Weak convergence of Galerkin finite element approximations of stochastic evolution equations with additive noise

Mihály Kovács

D.Sc. dissertation

Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden, 2018

dc_1541_18

Contents

dc_1541_18

Introduction

Stochastic partial differential equations (SPDEs) have been used in a variety of applied contexts to model systems with inherent or structural (e.g. complexity-related) randomness [**31, 88, 91**]. The study of SPDEs is a rather active field of research which gathered further momentum when one of the leading figures of the field, Martin Hairer, was awareded a Fields medal in 2014 for his work on SPDEs. This dissertation concerns numerical approximation of a class of SPDEs. Depending on the context, a particular functional of the solution of an SPDE can bear physical meaning, for example, energy stored in the system. We consider linear and semilinear SPDEs driven by additive Wiener or Lévy noise of pure jump type and investigate the accuracy of the numerical approximation of a functional of their solution. Due to its versatility the Galerkin finite element method is a popular and successful numerical method to discretise PDEs in space in science and engineering. Therefore, we consider this method when discretizing equations in space, however, spectral approximation would also fit our framework. While there are various approaches to SPDEs the one especially fitting the finite element analysis is the operator semigroup approach, or more generally, the evolution equation approach, of Da Parto and Zabczyk [**31**] for Wiener noise driven SPDEs and Peszat and Zabczyk [**88**] for Lévy noise driven SPDEs. We treat both parabolic and hyperbolic equations as well as Volterra integro-differential equations in this setting. In the Wiener case we also consider semilinear equations of parabolic type and Volterra integro-differential equations. For the model hyperbolic equation, which is the stochastic wave equation, we only consider linear equations both in the Wiener case and in the Lévy noise case. Time discretization is tailored to the specific problem: backward Euler method for parabolic problems together with a convolution quadrature for Volterra integro-differential equations, and rational approximation schemes for the wave equation, such as the Crank–Nicolson method.

Strong error estimates; that is, error estimates in the root-mean-squared sense (or more generally, in the root-mean-p-sense) has been extensively studied by many authors since the 1990's, early 2000's, see, for example, the seminal papers by I. Gyöngy and his collaborators [**45, 46, 47, 48**]. It would impossible to compile a list of all the work that has been done in this regard, and since the focus of the dissertation is also different, we refer to [**55**] for a comprehensive overview of the state-of-the-art of this kind of error analysis at the time the paper was published. Very often, the main stochastic tool in the analysis is Itô's Isometry (or more generally the Burkholder-Davis-Gundy inequality) and, at least when the nonlinear term in the equation has some sort of global Lipschitz bound, the analysis does not require sophisticated tools from stochastic analysis. The problems concerning the strong error analysis may, in this case, be usually reduced to finding appropriate deterministic error estimates, see for example, Theorems 1.3.14, 1.5.18, 1.5.24, 1.5.28, 2.3.2 and Remarks 3.3.1 and 3.5.1 in this dissertation.

The picture changes when one considers the so-called weak error; that is, when comparing the expected value of a functional g of the solution at a prescribed time horizon $T > 0$, where g is real valued and is defined on the state space of the solution (a Hilbert space in this dissertation), to the expected value of the same functional of its numerical approximation. By a change of variables one sees that this amounts to approximating the law of the solution in the weak sense of probability measures. Weak-order estimates are an important tool which allows one to express and handle the algorithmic uncertainty caused by numerical approximation schemes, and they are crucial for the analysis of Monte Carlo methods. When g is globally Lipschitz continuous, then it easily follows that the weak rate is bounded by the strong rate. However, it is a general phenomenon that for non-smooth noise, the rate of weak convergence is twice that of the strong convergence. To show this for various SPDEs, the analysis requires heavier probabilistic tools. We detail two approaches to tackle this problem and with it we relate the results of the dissertation to other research results in the field.

Methods. The first approach, which is used in Chapter 1 for equations driven by Wiener noise and in Chapter 3 driven by Lévy noise, uses a Kolmogorov's backward equation together with Itô's formula to represent the weak error. This approach was introduced by D. Talay in [**102**] for stochastic ordinary differential equations driven by Wiener noise. The use of Kolmogorov's equation for weak error analysis for Wiener noise driven SPDEs; that is, in the infinite dimensional setting, first appeared in [**100**] for a spectral Galerkin approximation for the stochastic heat equation when the covariance operator of the driving infinite dimensional Wiener process has the same eigenbasis as the linear operator in the equation (so-called commutative noise). This approach was later generalized to Galerkin finite elements for the linear stochastic heat equation with additive non-commutative noise in [**34**] and, around the same time, by the author in [**43**] albeit with a more restricted class of functionals g, see also [**61, 62**] and Remark 1.4.4. Later this was further extended to various type of semilinear stochastic heat equations in [**3, 33, 106, 108**] with the additional tool being Malliavin calculus, in particular, Malliavin integration-by-parts in most of these works, where the latter tool first appeared in [**49**] for a very restricted class of test functions. All these works consider parabolic type equations, in particular the stochastic heat equation, except for some results on the linear stochastic Cahn–Hilliard equation from [**61, 62**] which are presented in Section 1.4.

For hyperbolic equations; such as the stochastic wave equation or the stochastic Schrödinger equation, much less is known. The first weak approximation result appears in [**32**] for the time discretization of the stochastic Schrödinger equation in one spatial dimension and still in one spatial dimension in [**52**] for a finite difference scheme for the semilinear stochastic wave equation for a rather special class of test functions. Then, in [**61**] the linear stochastic wave equation was considered in several spatial dimensions and the weak error for a Galerkin finite element semidiscretization was analyzed. This was extended to fully discrete schemes in [**62**]. These results; that is, results from [**61, 62**] are presented in Chapter 2 of this dissertation. The methodology was further extended for a trigonometric integrator scheme in [**107**] for a semilinear wave equation.

Numerical methods for stochastic Volterra type integro-differential equations driven by Wiener noise were first considered in [**57**] without any error analysis. The first rigorous error analysis for such equations in the strong sense was carried out [**69**] followed by a weak error analysis in [**70**]. These results are presented in Section 1.5 with some improvements in the strong error analysis using regularity results from [**6**], see the Summary for more details on the nature of these improvements. Due to the presence of the memory kernel the solution to stochastic Volterra type integro-differential equations is not a Markov process hence there is no natural Kolmogorov's equation associated with the solution. However, in the linear case one might consider another (Itô) process which has the important property that at the prescribed time of interest it has the same distribution as the solution. Then one might use the Kolmogorov's equation associated with the auxiliary process in the error analysis. This trick was first used for a different purpose for the stochastic heat equation in [**34**], namely, to get rid of the drift in the Kolmogorov's equation and with it the unbounded differential operator. In the case of Volterra integro-differential equations the trick does even more: it allows Markovian tools in the analysis of a non-Markovian problem.

The same, Kolmogorov based approach is exploited in Chapter 3 for Lévy noise driven linear SPDEs. However, the extension of the arguments from the Wiener noise case is not straightforward. One of the difficulties in the general Lévy case (in contrast to the Gaussian case) is that there are no readily available, sufficiently general results on Kolmogorov's backward equation to suit our analysis. We remedy this situation in Proposition 3.2.10. Another complication arises from the fact that we use tools from the theory of stochastic integration based on two different settings. One, where we integrate operator-valued processes w.r.t. a Hilbert space-valued Lévy process, promoted in the monographs [**88**, Chapter 8], [**83, 84**], and another one where we integrate Hilbert space valued integrands w.r.t. a Poisson random measure [**80, 90**]. The problem occurs because our setting for stochastic differential equations is based on the first approach while the proof of the error representation formula in Theorem 3.2.6 is based on an Itô formula which appears in [**80**, Theorem 3.6]; the latter form is well suited for our purposes, but it is formulated using the second approach for stochastic integration. Therefore, we also link the two stochastic integrals in Lemma 3.2.4 so that we can use the results from both theories. We remark that weak error estimates for approximations of Lévy-driven stochastic ordinary differential equations have been considered by various authors, see, e.g. [**54, 82, 89, 92**] and the references therein. There also exists a series of papers on strong error estimates for approximations of SPDEs driven by Lévy processes or Poisson random measures, see, for example [**17, 18, 19, 37, 50, 51, 71**] and compare also with Remarks 3.3.1 and 3.5.1. However, to the best of our knowledge, the first steps in a weak error analysis for Lévy-driven SPDEs have been done only recently in [**75**] and an extension of these results in [**68**] on which Chapter 3 is based on. Very recently, the result from Theorem 3.3.4 on the stochastic heat equation driven by Lévy noise was reproved in [**20**], at least for a semidiscretization in space and with very minor improvements on the class of functionals considered, using a Malliavin integration-by-parts approach for Poisson random measures. Such an approach may pave the way for the weak error analysis of more general Lévy noise driven SPDEs. The results of the stochastic wave equation in the Lévy case is still state-of-the art.

The second approach, developed in Chapter 2, is based on [**1**] and it uses a duality argument with a suitable Gelfand triple of random variable spaces. In the triple, instead of the classical Sobolev–Malliavin spaces we use a refined version of these introduced in [**2**] for the purpose of weak error analysis. The use of these spaces allows us to exploit the smoothing property of the solution operator of the corresponding linear deterministic problem: the analytic semigroup generated by the differential operator in the equation in case of the stochastic heat equation and the resolvent family in case of stochastic Volterra integro-differential equations. The core idea is simple: use stability bounds in a stronger norm and error estimates in a weaker norm to double the strong rate of strong convergence. The motivation for using a relatively involved machinery stems from the problem that for semilinear stochastic Volterra integro-differential equations the trick used in the linear case does not work: it does not seem possible to use an auxiliary Markov process which has the same distribution as the solution of the equation at a prescribed time. Hence a Kolmogorov's equation based approach does not seem feasible, at least in the natural state space of the equation. One of course could try to consider the problem in a state space involving the memory as well to use a Markovian setting but then the key smoothing property of the corresponding deterministic solution operator would be lost. The approach taken here has two advantages. First, it allows us to treat the stochastic heat equation and a class of stochastic Volterra integro-differential equations at once in the same framework. Second, we may consider path-dependent functionals of the solution, albeit not in the full generality but general enough to cover, for example, approximation of covariances and higher order statistics, see Corollary 2.3.8. As far as we know this was the first, and to date almost the only, result on weak approximation of path dependent functionals for parabolic SPDEs. The recent paper [**22**] considers more general path dependent functionals for Wiener noise driven semilinear SPDEs but the noise have to be commutative and the discretization is a rather simple spectral Galerkin spatial-semidiscretization. It is not clear whether these arguments can be extended to time discretization and to more sophisticated space discretization methods such as the Galerkin finite element method.

We mention a third recent approach in the Markovian setting, not used in this dissertation, which is based on the so-called mild Itô formula from [**30**], see, for example [**29, 35**].

It is clear from the above that weak approximation of SPDEs are far from being fully understood. There are virtually no results for nonlinear equations without some sort of global Lipschitz condition such as the stochastic Allen–Cahn, Cahn–Hilliard–Cook or stochastic Navier–Stokes equations. For these equations even the strong error analysis is far from being complete. The weak error analysis of Lévy noise driven SPDEs are also in their infancy. Weak error analysis of path dependent functionals is also an area where future research is likely to grow quickly.

Although the main focus of the dissertation is to derive weak-order estimates for various SPDEs, to be able to compare the weak rate with the strong rate of the approximation schemes we also prove strong convergence rates in case it was established by the author, see Theorems 1.3.14, 1.5.18, 1.5.24, 1.5.28, 2.3.2 and Remarks 3.3.1 and 3.5.1.

Outline of the dissertation. The dissertation is based on the papers [**6, 61, 62, 66, 67, 69, 70**] for Chapter 1, [**1**] for Chapter 2 and [**68**] for Chapter 3. Unless otherwise stated, the results in the dissertation are due to the author and are taken from one of the aforementioned papers.

The dissertation is organized as follows. In Chapter 1 we first consider a general abstract framework to represent the weak error between two infinite dimensional drift-free Wiener-Itô integral processes. While the solution of the equations and their approximation we consider in this chapter are not of Itô form, as they are convolution processes, they indeed agree with an Itô process, perhaps after suitable interpolation, at a prescribed time $T > 0$. The importance of the weak error representation formula, presented in Theorem 1.2.1, lies in the fact that it allows us to study weak approximations of a wide range of SPDEs, parabolic, hyperbolic and even Volterra type, in a common framework. This weak error representation formula is then first applied to general abstract approximation schemes of the stochastic wave equation in Theorem 1.3.6, where we show that appropriate deterministic error estimates yield weak error estimates for the stochastic problem. We then apply Theorem 1.3.6 to a family of time semidiscretization schemes in Theorem 1.3.9, and to a full discretization scheme, where the space discretization is based on a Galerkin finite element method, in Theorem 1.3.13. For comparison, the strong rate of convergence of the same scheme is presented in Theorem 1.3.14 showing half the rate of that of the weak convergence. In Section 1.4 we apply the general error representation formula to parabolic SPDEs, first to the linear Cahn–Hilliard–Cook equation in Theorem 1.4.1 and then, without giving all the details to the stochastic heat equation in Remark 1.4.4. In Section 1.5 we consider a class of stochastic Volterra integro-differential equations, where the assumptions on the convolution kernel are typical in the linear theory of viscoleasticity. We first establish the key smoothing properties of the solution operator of the deterministic problem in Propositions 1.5.6 and 1.5.9 and then the regularity of the stochastic problem in 1.5.10. Then, in Theorem 1.5.18, we consider the strong rate of convergence of a time-semidiscretization scheme, the backward Euler scheme combined with a convolution quadrature based on the backward Euler scheme, under a rather general parabolicity condition on the convolution kernel. The strong rate of convergence of a the space semidiscretization via a Galerkin finite element method is presented in Theorem 1.5.24 for even slightly more general kernels. Finally, the strong rate of convergence of the fully discrete scheme is shown in Theorem 1.5.28 under the same condition on the kernel as for the time semidiscretization. To establish the weak rate of convergence, we have to impose a further analyticity condition on the convolution kernel, Assumption 1.5.30, which is satisfied, for example, for a family of tempered and untempered Riesz kernels. The weak rate of convergence for a fully discrete scheme is then stated in Theorem 1.5.33, based on the the general error representation formula fromTheorem 1.2.1, showing twice the rate of strong convergence established in Theorem 1.5.28, for a suitable class of memory kernels. The key deterministic error estimates of Section 1.5 are Theorem 1.5.13, Corollary 1.5.16, Proposition 1.5.22, Proposition 1.5.23, Lemma 1.5.27 and Theorem 1.5.32.

In Chapter 2 we consider mild solutions of semilinear parabolic type stochastic equations with additive Wiener noise, in particular, mild solutions of the semilinear heat equation and a class of semilinear Volterra integro-differential equations. We treat these equations in a common framework as their mild solutions satisfy the same integral equation, equation (2.1.2), with the only difference being the different degree of smoothing property, specified in equation (2.1.1), of the deterministic evolution operator appearing in the equation. As explained earlier in the Introduction, the approach of Chapter 1 does not work in this case and hence we first introduce some additional stochastic tools from Malliavin calculus in Subsection 2.1.2. In Section 2.2 we show the spatial and Hölder time-regularity of the solution of (2.1.2) both in classical L^p-norms of random variables and in Sobolev–Malliavin norms stated in Proposition 2.2.1. The regularity naturally depends on the smoothing property of the deterministic evolution family appearing in (2.1.2). Section 2.3 contains the main result of this Chapter, which is Theorem 2.3.7. In (2.3.6), we first introduce an abstract discrete version of (2.1.2) where we assume that the discrete deterministic time stepping-family appearing in the equation satisfies certain deterministic error bounds. We then prove, in Theorem 2.3.2, strong L^p-convergence rates for this abstract scheme. While this result is interesting on its own and one can compare this with the weak approximation rate, it is also used in proving Malliavin regularity of the solution of (2.1.2) in Proposition 2.3.4. The key result for proving the main weak convergence result of this chapter is Lemma 2.3.6 which establishes a strong convergence rate in a dual Sobolev–Malliavin norm which is twice the strong L^p-convergence rate. It is then used together with the Fundamental Theorem of Calculus, the earlier established Sobolev–Malliavin regularity of the solutions of (2.1.2) and (2.3.6), and a duality argument based on a well-chosen Gelfand triple of Sobolev–Malliavin and L^p-spaces to prove the desired weak convergence rate in Theorem 2.3.7 which is twice that of the strong convergence. The functionals considered in Theorem 2.3.7 may even depend on the paths of the solution in a special way and therefore we immediately get rates of approximations of covariances and higher order statistics of the solution of (2.1.2) in Corollary 2.3.8. Finally, in Section 2.4, we verify the abstract assumptions (2.1.1) on the deterministic evolution family and the abstract deterministic error bounds assumed in Subsection 2.3.1 for a suitable fully discrete scheme, first for a semilinear stochastic heat equation, in Subsection 2.4.1, and then for a class of semilinear stochastic Volterra integro-differential equations in Subsection 2.4.2.

In Chapter 3 we consider linear equations driven by additive square integrable Lévy noise of pure jump type. We develop a representation formula for the weak error between two infinite dimensional drift-free Lévy-Itô processes, stated in Theorem 3.2.6 and, in a slightly different form, in Corollary 3.2.8. In order to do so we have to prove additional preparatory results, interesting in their own right. First, in Subsection 3.2.1, we compare stochastic integrals of operator-valued processes w.r.t. a Hilbert space-valued Lévy process and integrals of Hilbert space valued integrands w.r.t. Poisson random measures. In Lemma 3.2.4 we show that these are equivalent using appropriate identifications of the integrands from Lemma 3.2.1. This allows us to use results from both integration theories. Second, in Proposition 3.2.10, we introduce the main technical result needed in the proof of the error representation in Theorem 3.2.6, a backward Kolmogorov equation associated with an infinite dimensional drift-free Lévy-Itô process given by (3.2.3). The proof of Theorem 3.2.6 is then presented in Subsection 3.2.3. In Remark 3.2.9 we also comment on the case when the Gaussian part of the Lévy process does not vanish connecting Chapter 3 to Chapter 1. The abstract weak error representation formulae from Theorem 3.2.6 and Corollary 3.2.8 allows us to study the weak error of space-time discretizations of stochastic equations with rather different properties in a common framework: parabolic equations, such as the stochastic heat equation and a stochastic Volterra integro-differential equation, and a hyperbolic equation, the stochastic wave equation. As space discretization we employ a standard continuous finite element method. As time discretization, similarly to the Gaussian case, we use the backward Euler method for the stochastic heat equation, the backward Euler method combined with a convolution quadrature for Volterra integro-differential equations, and an I-stable rational approximation of the exponential function, such as the Crank–Nicolson scheme, for the stochastic wave equation. For the stochastic heat equation, we show in Theorem 3.3.4 that for twice continuously differentiable test functions with bounded second derivatives the rate of weak convergence is essentially twice that of strong convergence. This extends the corresponding result from [**75**], where the analysis is restricted to so-called impulsive cylindrical processes on $L^2(\mathcal{D})$ as driving noise. Moreover, there is a serious restriction on the jump size intensity measure in [**75**, Section 6] which renders the sample paths of the process to be of bounded variation on finite time intervals. Here, the only restriction we have on the Lévy process is that it is square-integrable, non-Gaussian and has mean zero. Furthermore, we also remove the boundedness assumption on the test functions and their first derivatives. In Subsection 3.4 we briefly discuss a stochastic Volterra-type integro-differential equation and obtain a weak rate of convergence in Theorem 3.4.1 under the same conditions as in the Gaussian case in Theorem 1.5.33. For the stochastic wave equation we first prove appropriate deterministic estimates in Proposition 3.5.2 which is then used, together with the general error representation formula, to prove the order of weak convergence in Theorem 3.5.3. At the end of the chapter we discuss some examples where the conditions of Theorem 3.5.3, in particular (3.5.15) and (3.5.16), are satisfied.

Acknowledgements. I would like to thank my early career mentors, István Faragó, Frank Neubrander, Stig Larsson and Mark Meerschaert for their guidance, all my collaborators, especially Boris Baeumer, for the inspirational joint work and also my students Fredrik, Harish and Kristin for motivating and challenging me. I also would like to thank my wife Luca, daughter Lili, my mom, dad, sister and brother for their encouragement and support. Finally, I would like to thank Petra Csomós for carefully reading my dissertation summary.

dc_1541_18

CHAPTER 1

Linear stochastic PDEs driven by additive Wiener noise

1.1. Preliminaries

Here we collect some background material from infinite-dimensional stochastic analysis and stochastic PDEs driven by Wiener noise and introduce some notation. We use the semigroup approach of DaPrato and Zabczyk and we refer to the monograph [**31**] for details and proofs. Notation introduced here will be used throughout the dissertation.

Let U and H be real separable Hilbert spaces; we often denote both their norms and scalar products by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ when the meaning is clear from the context. We denote the space of bounded linear operators from U to H by $\mathcal{L}(U,H)$ with operator norm $\|\cdot\|_{\mathcal{L}(U,H)}$. The p:th Schatten class of operators from U to H is denoted by $\mathcal{L}_p(U,H)$. They are Banach spaces for all integers $p \geqslant 1$ and we will denote their norms by $\|\cdot\|_{\mathcal{L}_p(U,H)}$. The operators in $\mathcal{L}_1(U,H)$ are also refered to as trace class operators and operators in $\mathcal{L}_2(U,H)$ as Hilbert-Schmidt operators. The space $\mathcal{L}_2(U,H)$ is a Hilbert space with inner product denoted $\langle\cdot,\cdot\rangle_{\mathcal{L}_2(U,H)}$. When the underlying Hilbert spaces are understood from the context we will write $\|\cdot\| = \|\cdot\|_{\mathcal{L}(U,H)}$, $\|\cdot\|_{\mathrm{Tr}} = \|\cdot\|_{\mathcal{L}_1(U,H)}$, $\|\cdot\|_{\mathrm{HS}} = \|\cdot\|_{\mathcal{L}_2(U,H)}$ and $\langle\cdot,\cdot\rangle_{\mathrm{HS}} = \langle\cdot,\cdot\rangle_{\mathcal{L}_2(U,H)}$ in order to – we hope – increase the readability of the dissertation.

In case $H = U$ we write $\mathcal{L}(U) = \mathcal{L}(U,U)$ and $\mathcal{L}_p(U) = \mathcal{L}_p(U,U)$ for short. If $T \in \mathcal{L}_1(U)$ and $\{e_k\}_{k=1}^{\infty}$ is an orthonormal basis of U, then the trace of T,

$$\mathrm{Tr}(T) := \sum_{k=1}^{\infty} \langle Te_k, e_k\rangle_U,$$

is a well defined number, independent of the choice of orthonormal basis. Below we state a number of properties of Schatten class operators. For proofs and definitions we refer to, for example, [**31**, Appendix C], [**73**] and [**109**].

If $T \in \mathcal{L}_p(U,H)$, then its adjoint $T^* \in \mathcal{L}_p(H,U)$ and

$$\|T\|_{\mathcal{L}_p(U,H)} = \|T^*\|_{\mathcal{L}_p(H,U)}. \tag{1.1.1}$$

If $U = H$ and $p = 1$, then also

$$\mathrm{Tr}(T) = \mathrm{Tr}(T^*) \tag{1.1.2}$$

and

$$|\mathrm{Tr}(T)| \leqslant \|T\|_{\mathrm{Tr}}. \tag{1.1.3}$$

Further, if T is selfadjoint and positive semidefinite, then $\mathrm{Tr}(T) \geq 0$ and (1.1.3) holds with equality.

If U_1, U_2, and H are separable Hilbert spaces and $T \in \mathcal{L}_p(U_2, H)$ and if $S_1 \in \mathcal{L}(U_1, U_2)$ and $S_2 \in \mathcal{L}(H, U_1)$, then

$$\begin{aligned} \|TS_1\|_{\mathcal{L}_p(U_1,H)} &\le \|T\|_{\mathcal{L}_p(U_2,H)} \|S_1\|_{\mathcal{L}(U_1,U_2)}, \\ \|S_2T\|_{\mathcal{L}_p(U_2,U_1)} &\le \|T\|_{\mathcal{L}_p(U_2,H)} \|S_2\|_{\mathcal{L}(H,U_1)}. \end{aligned} \tag{1.1.4}$$

If $S \in \mathcal{L}(H, U)$ and $T \in \mathcal{L}_1(U, H)$, then we also have

$$\mathrm{Tr}(TS) = \mathrm{Tr}(ST). \tag{1.1.5}$$

Moreover, if $T\colon U \to H$ and $T^*T \in \mathcal{L}_1(U)$, then $T \in \mathcal{L}_2(U, H)$, $TT^* \in \mathcal{L}_1(H)$ and

$$\begin{aligned} \|T^*T\|_{\mathrm{Tr}} &= \mathrm{Tr}(T^*T) = \|T\|_{\mathrm{HS}}^2 = \|T^*\|_{\mathrm{HS}}^2 \\ &= \mathrm{Tr}(TT^*) = \|TT^*\|_{\mathrm{Tr}}. \end{aligned} \tag{1.1.6}$$

Finally, we note that if $T \in \mathcal{L}_2(U, H)$ and $S \in \mathcal{L}_2(H, U)$, then $TS \in \mathcal{L}_1(H)$ and

$$\|TS\|_{\mathrm{Tr}} \leqslant \|T\|_{\mathrm{HS}} \|S\|_{\mathrm{HS}} = (\mathrm{Tr}(TT^*)\mathrm{Tr}(SS^*))^{1/2}. \tag{1.1.7}$$

To be able to compare various assumptions on the regularity of the noise, where the regularity usually is measured in the trace or Hilbert-Schmidt norms, we have will use the following result.

THEOREM 1.1.1. *Assume that $Q \in \mathcal{L}(H)$ is selfadjoint, positive semidefinite and that A is a densely defined, unbounded, selfadjoint, positive definite, linear operator on H with an orthonormal basis of eigenvectors. Then the following inequalities hold, for $s \in \mathbb{R}$, $\alpha > 0$,*

$$\|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 \le \|A^s Q\|_{\mathrm{Tr}} \le \|A^{s+\alpha} Q\|_{\mathcal{B}(H)} \|\|A^{-\alpha}\|_{\mathrm{Tr}}, \tag{1.1.8}$$

$$\|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 \le \|A^{s+\frac12} Q A^{-\frac12}\|_{\mathrm{Tr}}, \tag{1.1.9}$$

provided that the respective norms are finite. Furthermore, if A and Q have a common basis of eigenvectors, in particular, if $Q = I$, then

$$\|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 = \|A^s Q\|_{\mathrm{Tr}} = \|A^{s+\frac12} Q A^{-\frac12}\|_{\mathrm{Tr}}. \tag{1.1.10}$$

PROOF. If $\{(\lambda_k, \phi_k)\}_{k=1}^\infty$ denotes a set of eigenpairs of A with orthonormal eigenvectors, then we define

$$A^s x = \sum_{k=1}^\infty \lambda_k^s \langle x, \phi_k \rangle \phi_k.$$

Although $[A^{\frac{s}{2}} Q^{\frac12}]^*$ is not equal to $Q^{\frac12} A^{\frac{s}{2}}$ in general, we do have $[A^{\frac{s}{2}} Q^{\frac12}]^* \phi_k = Q^{\frac12} A^{\frac{s}{2}} \phi_k$, and we compute using (1.1.1), (1.1.3), and (1.1.4),

$$\begin{aligned} \|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 &= \|[A^{\frac{s}{2}} Q^{\frac12}]^*\|_{\mathrm{HS}}^2 = \sum_{k=1}^\infty \|[A^{\frac{s}{2}} Q^{\frac12}]^* \phi_k\|^2 = \sum_{k=1}^\infty \|Q^{\frac12} A^{\frac{s}{2}} \phi_k\|^2 \\ &= \sum_{k=1}^\infty \lambda_k^s \|Q^{\frac12} \phi_k\|^2 = \sum_{k=1}^\infty \lambda_k^s \langle Q\phi_k, \phi_k \rangle = \sum_{k=1}^\infty \langle Q\phi_k, A^s \phi_k \rangle \\ &= \sum_{k=1}^\infty \langle A^s Q \phi_k, \phi_k \rangle = \mathrm{Tr}(A^s Q) \le \|A^s Q\|_{\mathrm{Tr}} \le \|A^{s+\alpha} Q\|_{\mathcal{B}(H)} \|A^{-\alpha}\|_{\mathrm{Tr}}. \end{aligned}$$

This is (1.1.8). Similarly, (1.1.9) is proved by

$$\begin{aligned}\|A^{\frac{s}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 &= \sum_{k=1}^{\infty}\lambda_k^s\langle Q\phi_k,\phi_k\rangle = \sum_{k=1}^{\infty}\langle Q\lambda_k^{-\frac{1}{2}}\phi_k,\lambda_k^{s+\frac{1}{2}}\phi_k\rangle\\ &= \sum_{k=1}^{\infty}\langle A^{s+\frac{1}{2}}QA^{-\frac{1}{2}}\phi_k,\phi_k\rangle = \mathrm{Tr}(A^{s+\frac{1}{2}}QA^{-\frac{1}{2}}) \le \|A^{s+\frac{1}{2}}QA^{-\frac{1}{2}}\|_{\mathrm{Tr}}.\end{aligned}$$

To show (1.1.10) we assume that Q has the same eigenvectors ϕ_k with eigenvalues γ_k. Then

$$A^sQx = \sum_{k=1}^{\infty}\lambda_k^s\gamma_k\langle x,\phi_k\rangle\phi_k,$$

and hence

$$\|A^sQ\|_{\mathrm{Tr}} \le \sum_{k=1}^{\infty}\lambda_k^s\gamma_k = \sum_{k=1}^{\infty}\|A^{\frac{s}{2}}Q^{\frac{1}{2}}\phi_k\|^2 = \|A^{\frac{s}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2,$$

which shows the first equality in (1.1.10) in view of (1.1.8). The second equality in (1.1.10) can be shown in a similar fashion. □

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space and let $L^p(\Omega;H)$ denote the space of random variables $X\colon(\Omega,\mathcal{F})\to(H,\mathcal{B}(H))$; that is, $\mathcal{F}/\mathcal{B}(H)$-measurable mappings $X:\Omega\to H$, where $\mathcal{B}(H)$ denotes the Borel σ-algebra of the separable Hilbert space H, such that

$$\|X\|_{L^p(\Omega;H)}^p = \mathbb{E}\big(\|X\|_H^p\big) = \int_\Omega \|X(\omega)\|_H^p\,\mathrm{d}\mathbb{P}(\omega) < \infty.$$

In the case $H=\mathbb{R}$ we write $L^p(\Omega)=L^p(\Omega;\mathbb{R})$. If X is a random variable on the probability space $(\Omega,\mathcal{F},\mathbb{P})$, we denote by $\mathbb{P}(X\in\cdot):=\mathbb{P}\circ X^{-1}$ the law of X under $\mathbb{P}$.

More generally, given a measure space $(M,\mathcal{M},\mu)$ and $1\leqslant p<\infty$, we denote by $L^p(M;H)=L^p(M,\mathcal{M},\mu;H)$ the space of all $\mathcal{M}/\mathcal{B}(H)$-measurable mappings $f:M\to H$ with finite norm $\|f\|_{L^p(M;H)}=(\int_M\|f\|_H^p\mathrm{d}\mu)^{1/p}$.

Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P})$ be a stochastic basis, or filtered probability space, satisfying the usual conditions. Let U be a separable Hilbert space and $Q\in\mathcal{L}(U)$ with $Q\ge0$ (selfadjoint, positive semidefinite). Let $W=(W(t))_{t\ge0}$ be a U-valued stochastic process on $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\ge0},\mathbb{P})$. We say that W is a *Q-Wiener process in U* if

(i) $W(0)=0$,
(ii) W has continuous trajectories (almost surely),
(iii) W has independent increments,
(iv) $W(t)-W(s)$ is a U-valued Gaussian random variable with zero mean and covariance operator $(t-s)Q$ for $0\le s\le t$.

Here Q is the unique operator defined by

$$\mathbb{E}\Big(\big\langle(W(t)-W(s)),x\big\rangle\big\langle(W(t)-W(s)),y\big\rangle\Big) = (t-s)\langle Qx,y\rangle \quad x,y\in U. \tag{1.1.11}$$

Condition (iv) implies that $\mathrm{Tr}(Q)<\infty$ because the covariance operator of a Gaussian random variable is necessarily of trace class, see [**31**, Proposition 2.15]. Therefore, W is also called a nuclear Wiener process.

A nuclear Wiener process can be constructed starting from its covariance operator Q and the construction extends to the case when $\mathrm{Tr}(Q) = \infty$ in the following way. Let $Q \in \mathcal{L}(U)$ with $Q \geq 0$. The *Cameron-Martin space* is defined as $U_0 := Q^{\frac{1}{2}}U$ endowed with the inner product $\langle x, y\rangle_0 := \langle Q^{-\frac{1}{2}}x, Q^{-\frac{1}{2}}y\rangle$, where Q^{-1} is understood as the pseudo-inverse if Q is not injective. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis for U_0, let $\{\beta_j\}_{j=1}^{\infty}$ be mutually independent real-valued Brownian motions on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$. If $\mathrm{Tr}(Q) < \infty$, then the series

$$W(t) := \sum_{k=1}^{\infty} \beta_k(t) e_k \tag{1.1.12}$$

converges in $L^2(\Omega; U)$ to a U-valued stochastic process, which has a version that is a nuclear Q-Wiener process, see [**31**, Section 4] and [**91**, Section 2].

If $\mathrm{Tr}(Q) = \infty$, then the series (1.1.12) does not converge in $L^2(\Omega; U)$. However, it converges in $L^2(\Omega; U_1)$ for a suitable (usually larger) space U_1 (see [**31**, Section 4.3.1]) to a U_1-valued stochastic process, which has a version that is a U_1-valued nuclear Wiener process. The constructed process, still denoted by W, is called a *cylindrical Q-Wiener process in U*. Also, it is easy to see that

$$W_x(t) = \sum_{k=1}^{\infty} \beta_k(t)\langle e_k, x\rangle, \quad x \in U,$$

exists in $L^2(\Omega; \mathbb{R})$ and defines a real-valued Wiener process (Brownian motion) satisfying

$$\mathbb{E}\Big(W_x(t) W_y(t)\Big) = t\langle Qx, y\rangle, \quad x, y \in U,$$

cf. (1.1.11). Hence, we may write formally $\langle W(t), x\rangle = W_x(t)$ although the process $W(t)$ constructed from (1.1.12) takes values in U_1.

In either case, $\mathrm{Tr}(Q) < \infty$ or $\mathrm{Tr}(Q) = \infty$, we denote by $W(t)$, $t \geq 0$, the series in (1.1.12), which is formal in case $\mathrm{Tr}(Q) = \infty$, and call it a *Q-Wiener process* in U.

Remark 1.1.2. It is often the case that there is an orthonormal basis $\{f_k\}_{k=1}^{\infty}$ in U consisting of eigenvectors of Q with corresponding non-negative eigenvalues $\{\gamma_k\}_{k=1}^{\infty}$. Then $e_k = Q^{1/2} f_k = \gamma_k^{1/2} f_k$ is an orthonormal basis for U_0 and, in particular, (1.1.12) becomes

$$W(t) = \sum_{k=1}^{\infty} \gamma_k^{1/2} \beta_k(t) f_k.$$

However, we prefer to avoid the eigenvector expansion of $W(t)$.

Finally we say that W is a Q-Wiener process in U with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if $W(t)$ is adapted to $\mathcal{F}_t$ for all $t \geq 0$ and $W(t) - W(s)$ is independent of $\mathcal{F}_s$ for all $0 \leqslant s \leqslant t$.

In what follows we need a simplified case of the stochastic integral, the Wiener integral, namely where the integrand is deterministic. In this case the class of integrands can be easily described. Let $F\colon [0, \infty) \to \mathcal{L}_2(U_0, H)$ be a measurable function, where $\mathcal{L}_2(U_0, H)$ is regarded as a Hilbert space endowed with its Borel

dc_1541_18

sigma algebra, and assume that F is square integrable,

(1.1.13)
$$\int_0^t \|F(s)\|^2_{\mathcal{L}_2(U_0,H)}\,\mathrm{d}s = \int_0^t \|F(s)Q^{\frac{1}{2}}\|^2_{\mathrm{HS}}\,\mathrm{d}s = \int_0^T \mathrm{Tr}(F(t)QF^*(t))\,\mathrm{d}t < \infty.$$

Then the stochastic integral $\int_0^t F(s)\,\mathrm{d}W(s)$ is a well defined Gaussian random variable with covariance operator

$$Q_F(t)x = \int_0^t F(s)QF^*(s)x\,\mathrm{d}s, \quad x \in H,$$

and the Itô isometry,

$$\Big\|\int_0^t F(s)\,\mathrm{d}W(s)\Big\|^2_{L^2(\Omega;H)} = \int_0^t \|F(s)Q^{\frac{1}{2}}\|^2_{\mathrm{HS}}\,\mathrm{d}s = \mathrm{Tr}(Q_F(t)) \tag{1.1.14}$$

holds, see [**31**, Chapter 4] and [**91**, Chapter 2].

For bounded functions G the next result can be found in, for example, [**31**, Proposition 1.12], which we extend to allow polynomial growth.

LEMMA 1.1.3. *Let $G : H \to \mathbb{R}$ be measurable such that $|G(x)| \le p_N(\|x\|)$ where p_N is a real polynomial of degree N. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G} \subset \mathcal{F}$ is a sub sigma-algebra of $\mathcal{F}$. Let $\xi_1, \xi_2 \in L^N(\Omega; H)$ be H-valued random variables such that ξ_1 is $\mathcal{G}$-measurable and ξ_2 is independent of $\mathcal{G}$. If we define $u : H \to \mathbb{R}$ by $u(x) = \mathbb{E}(G(x + \xi_2))$, $x \in H$, then, almost surely, $u(\xi_1) = \mathbb{E}(G(\xi_1 + \xi_2)|\mathcal{G})$.*

PROOF. Define $G_n(x) = G(\xi_{B_n(0)}(x)x)$ where $\xi_{B_n(0)}$ is the characteristic function of the closed unit ball around 0 with radius n. We clearly have that $G_n(x) \to G(x)$ for all $x \in H$. Furthermore, $|G_n(x)| \le p_N(\|x\|)$ for all $n \in \mathbb{N}$ and $x \in H$. Therefore, if $\eta \in L^N(\Omega, H)$, then by the dominated convergence theorem $G_n(\eta) \to G(\eta)$ in $L^1(\Omega; \mathbb{R})$. Let $x \in H$ and define $u(x) := \mathbb{E}(G(x + \xi_2))$ and $u_n(x) := \mathbb{E}(G_n(x + \xi_2))$. If we take $\eta := x + \xi_2$, then, for all $x \in H$,

$$|u_n(x) - u(x)| \le |\mathbb{E}(G_n(\eta) - G(\eta))| \le \|G_n(\eta) - G(\eta)\|_{L^1(\Omega;\mathbb{R})} \to 0$$

as $n \to \infty$. We also have that

$$\begin{aligned}|u_n(x)| &\le \mathbb{E}|(G_n(x + \xi_2))| \le \mathbb{E}(p_N(\|x + \xi_2\|)) \\ &\le C(p_N(\|x\|) + \mathbb{E}(p_N(\|\xi_2\|))) \le C(p_N(\|x\|) + \|\xi_2\|_{L^N(\Omega;H)}),\end{aligned}$$

and hence

$$|u_n(\xi_1)| \le C(p_N(\|\xi_1\|) + \|\xi_2\|_{L^N(\Omega;H)}) \in L^1(\Omega; \mathbb{R}).$$

Therefore,

$$u_n(\xi_1) \to u(\xi_1) \text{ in } L^1(\Omega; \mathbb{R}) \tag{1.1.15}$$

as $n \to \infty$ by dominated convergence. Since G_n is a bounded and measurable function it follows from [**31**, Proposition 1.12] that $u_n(\xi_1) = \mathbb{E}(G(\xi_1 + \xi_2)|\mathcal{G})$. By taking $\eta = \xi_1 + \xi_2$ it follows as above that $G_n(\xi_1 + \xi_2) \to G(\xi_1 + \xi_2)$ in $L^1(\Omega; \mathbb{R})$ and thus by the dominated convergence theorem for conditional expectations we conclude that

$$u_n(\xi_1) = \mathbb{E}(G_n(\xi_1 + \xi_2)|\mathcal{G}) \to \mathbb{E}(G(\xi_1 + \xi_2)|\mathcal{G}) \text{ in } L^1(\Omega; H)$$

as $n \to \infty$ which finishes the proof in view of (1.1.15). □

By $C^n(H,\mathbb{R})$ we denote the space of all n-times continuously Fréchet differentiable functions $f : H \to \mathbb{R}$, $x \mapsto f(x)$. By $C^n_{\mathrm{b}}(H,\mathbb{R})$ we denote the subspace of functions from $C^n(H,\mathbb{R})$ which are bounded together with their derivatives. Identifying H and $\mathcal{L}(H,\mathbb{R})$ via the Riesz isomorphism, we consider for fixed $x \in H$ the first derivative $f'(x)$ as an element of H. Similarly, the second derivative $f''(x)$ is considered as an element of $\mathcal{L}(H)$. We also write f_x and f_{xx} instead of f' and f''.

Suppose that $G : H \to \mathbb{R}$ is a functional such that

$$G \in C^2(H,\mathbb{R}) \text{ and } G'' \in C_b(H,\mathcal{L}(H)). \tag{1.1.16}$$

Then, it follows from Taylor's Formula that

$$|G(x) - G(y)| \leq \|G'(y)\| \cdot \|x - y\| + C\|x - y\|^2, \tag{1.1.17}$$

where $C = \sup_{x\in H} \|G''(x)\|_{\mathcal{L}(H)}$ and that

$$\|G'(x)\| \leq K(1 + \|x\|) \tag{1.1.18}$$

where $K = \max\{C, \|G'(0)\|\}$. Then, by (1.1.17) and (1.1.18), G has quadratic growth:

$$|G(x)| \leqslant L(1 + \|x\|^2), \tag{1.1.19}$$

for some $L > 0$. Let $(E(t))_{t\in[0,T]} \subset \mathcal{L}(H)$ be a strongly continuous family, $B \in \mathcal{L}(U,H)$ and W be a Q-Wiener process in U with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. Suppose that $EBQ^{1/2} \in L^2((0,T);\mathcal{L}_2(U,H))$. For any $x \in H$ and $t \in [0,T]$, we define

$$Z(T,t,x) := x + \int_t^T E(T - s)B \,\mathrm{d}W(s). \tag{1.1.20}$$

The above stochastic integral makes sense by (1.1.13). Let G satisfy (1.1.16), and by (1.1.19), we may define

$$u(t,x) := \mathbb{E}(G(Z(T,t,x))),\ x \in H,\ t \in [0,T]. \tag{1.1.21}$$

Since $EB \in L^2((0,T);\mathcal{L}_2(U_0,H)))$ is equivalent to

$$\mathrm{Tr}(E(T - \cdot)BQ[E(T - \cdot)B]^*) \in L^1(0,T)$$

and G satisfy (1.1.16), it is well known that u is a solution of the following backward Kolmogorov equation

$$u_t(t,x) + \frac{1}{2}\,\mathrm{Tr}\Big(u_{xx}(t,x)E(T - t)BQ[E(T - t)B]^*\Big) = 0, \quad x \in H,\ t \in [0,T), \tag{1.1.22}$$

with terminal condition $u(T,x) = G(x)$, $x \in H$. It is not hard to see that it follows from (1.1.20) and (1.1.21) that the partial derivatives of u are given by

$$\begin{aligned} u_x(t,x) &= \mathbf{E}\big(G'(Z(T,t,x))\big),\\ u_{xx}(t,x) &= \mathbf{E}\big(G''(Z(T,t,x))\big). \end{aligned} \tag{1.1.23}$$

Corollary 1.1.4. *Let* $\xi \in L^2(\Omega,\mathcal{F}_t,\mathbb{P};H)$ *and let* G *satisfy* (1.1.16). *If* u *defined by (1.1.21), then*

$$u(t,\xi) = \mathbb{E}(G(Z(T,t,\xi))|\mathcal{F}_t),\ t \in [0,T].$$

PROOF. The statement follows from Lemma 1.1.3 with $\xi_1 = \xi$ and $\xi_2 = \int_t^T S(T-s)\,\mathrm{d}W(s)$ noting that $\xi_2 \in L^2(\Omega; H)$ as, by Itô's Isometry,

$$\mathbb{E}\|\xi_2\|^2 = \int_t^T \|S(T-s)Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2\,\mathrm{d}s \le \int_0^T \|S(t)Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2\,\mathrm{d}t < \infty.$$

□

We quote the following Itô's formula from [**24**], see also, [**25**].

PROPOSITION 1.1.5 (Itô's formula). *Let $f : [c,d) \times H \to \mathbb{R}$, $0 \le c < d \le \infty$, such that f, f_t, f_x and f_{xx} are continuous on $[c,d)\times H$ with values in the appropriate spaces. Let $a \in L^1_{loc}(\Omega \times (c,d); H)$ and $\xi Q^{1/2} \in L^2_{loc}(\Omega \times (c,d); \mathrm{HS})$ and*

$$X(t) = X(c) + \int_c^t a(s)\,\mathrm{d}s + \int_c^t \xi(s)\,\mathrm{d}W(s),\ \ t \in [c,d).$$

Then, for all $t \in [c,d)$, almost surely,

$$f(t,X(t)) - f(c,X(c)) = \int_c^t f_t(s,X(s))\,\mathrm{d}s + \int_c^t (f_x(s,X(s)),a(s))\,\mathrm{d}s$$
$$+ \int_c^t (f_x(s,X(s)),\xi(s)\,\mathrm{d}W(s)) + \frac{1}{2}\int_c^t \mathrm{Tr}(f_{xx}(s,X(s))\xi(s)Q\xi^*(s))\,\mathrm{d}s.$$

1.2. An error representation formula

The proof of the main approximation results of this chapter relies on the ability to compare the laws of two different Itô processes of the form

$$Y(t) := Y(0) + \int_0^t E(T-s)B\,\mathrm{d}W(s), \quad t \in [0,T],$$

and

$$\tilde{Y}(t) := \tilde{Y}(0) + \int_0^t \tilde{E}(T-s)\tilde{B}\,\mathrm{d}W(s), \quad t \in [0,T]; \tag{1.2.1}$$

that is to bound the quantity

$$e(T) = \mathbb{E}\left(G(\tilde{Y}(T)) - G(Y(T))\right). \tag{1.2.2}$$

for a class of functions $G : H \to \mathbb{R}$. Here $(\tilde{E}(t))_{t\in[0,T]} \subset \mathcal{L}(H)$ denotes another family of bounded operators on H such that $t \mapsto \tilde{E}(t)\tilde{B}$ is a measurable mapping from $[0,T]$ to $\mathcal{L}_2(U_0,H)$, $\tilde{B} \in \mathcal{L}(U,H)$, $\tilde{E}\tilde{B} \in L^2((0,T);\mathcal{L}_2(U_0,H))$ and $Y(0), \tilde{Y}(0)$ are H-valued and $\mathcal{F}_0$-measurable.

THEOREM 1.2.1. *Let $T > 0$ and $(E(t))_{t\in[0,T]}$ and $(\tilde{E}(t))_{t\in[0,T]}$ be two families of bounded operators on H such that $(E(t))_{t\in[0,T]}$ is strongly continuous, $t \mapsto \tilde{E}(t)\tilde{B}$ is a measurable mapping from $[0,T]$ to $\mathcal{L}_2(U_0,H)$, $B, \tilde{B} \in \mathcal{L}(U,H)$ and $EB, \tilde{E}\tilde{B} \in L^2((0,T);\mathcal{L}_2(U_0,H))$. If G satisfies (1.1.16) and $Y(0), \tilde{Y}(0) \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$, then Y and $\tilde{Y}$ are well-defined and the weak error $e(T)$ in (1.2.2) has the representation*

$$e(T) = \mathbb{E}\big(u(0,\tilde{Y}(0)) - u(0,Y(0))\big) + \tfrac{1}{2}\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(t,\tilde{Y}(t))\mathcal{O}(t)\Big)\,\mathrm{d}t, \tag{1.2.3}$$

where

$$\mathcal{O}(t) = \big(\tilde{E}(T-t)\tilde{B} + E(T-t)B\big)Q\big(\tilde{E}(T-t)\tilde{B} - E(T-t)\tilde{B}\big)^*, \tag{1.2.4}$$

or

$$\mathcal{O}(t) = \big(\tilde{E}(T-t)\tilde{B} - E(T-t)B\big)Q\big(\tilde{E}(T-t)\tilde{B} + E(T-t)B\big)^*. \tag{1.2.5}$$

Proof. By the law of double expectation,

$$\mathbb{E}\Big(u(t,\xi)\Big) = \mathbb{E}\Big(\mathbb{E}\Big(G(Z(T,t,\xi))\Big|\mathcal{F}_t\Big)\Big) = \mathbb{E}\Big(G(Z(T,t,\xi))\Big).$$

Therefore, it follows that

$$\mathbb{E}\Big(G(Y(T))\Big) = \mathbb{E}\Big(G(Z(T,0,Y(0))\Big) = \mathbb{E}\Big(u(0,Y(0))\Big)$$

and that

$$\mathbb{E}\Big(G(\tilde{Y}(T))\Big) = \mathbb{E}\Big(G(Z(T,T,\tilde{Y}(T))\Big) = \mathbb{E}\Big(u(T,\tilde{Y}(T))\Big).$$

Hence,

$$\begin{aligned} e(T) &= \mathbb{E}\Big(G(\tilde{Y}(T)) - G(Y(T))\Big) = \mathbb{E}\Big(u(T,\tilde{Y}(T)) - u(0,Y(0))\Big) \\ &= \mathbb{E}\Big(u(0,\tilde{Y}(0)) - u(0,Y(0))\Big) + \mathbb{E}\Big(u(T,\tilde{Y}(T)) - u(0,\tilde{Y}(0))\Big). \end{aligned}$$

For the second term, we use Itô's formula from Proposition 1.1.5 for $u(t,\tilde{Y}(t))$ on $[0,T-\varepsilon]$ and then passing to the limit $\varepsilon \to 0+$ using the continuity of u on $[0,T]\times H$ and the continuity of the paths of $\tilde{Y}(t)$ on $[0,T]$. Thus, taking also Kolmogorov's equation (1.1.22) into account, we get

(1.2.6)

$$\begin{aligned} &\mathbb{E}\Big(u(T,\tilde{Y}(T)) - u(0,\tilde{Y}(0))\Big) \\ &\quad = \mathbb{E}\int_0^T \Big\{u_t(t,\tilde{Y}(t)) + \tfrac{1}{2}\mathrm{Tr}\Big(u_{xx}(t,\tilde{Y}(t))[\tilde{E}(T-t)\tilde{B}]Q[\tilde{E}(T-t)\tilde{B}]^*\Big)\Big\}\,\mathrm{d}t \\ &\quad = \tfrac{1}{2}\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(t,\tilde{Y}(t))\big\{[\tilde{E}(T-t)\tilde{B}]Q[\tilde{E}(T-t)\tilde{B}]^* \\ &\qquad - [E(T-t)B]Q[E(T-t)B]^*\big\}\Big)\,\mathrm{d}t. \end{aligned}$$

The operator $u_{xx}(r,\xi)$ is bounded for every ξ and r and both $\tilde{E}(s)\tilde{B}Q[\tilde{E}(s)\tilde{B}]^*$ and $E(s)BQ[E(s)B]^*$ are of trace class for almost every s by assumption. Hence, the trace above is well defined for almost every t since by (1.1.4) with $p=1$,

$$\begin{aligned} \|u_{xx}(r,\xi)[E(s)B]Q[E(s)B]^*\|_{\mathrm{Tr}} &\leqslant \|u_{xx}(r,\xi)\|_{\mathcal{L}(H)}\,\|[E(s)B]Q[E(s)B]^*\|_{\mathrm{Tr}} \\ &= \|u_{xx}(r,\xi)\|_{\mathcal{L}(H)}\,\mathrm{Tr}\big([E(s)B]Q[E(s)B]^*\big), \end{aligned}$$

where the last step is (1.1.3) with equality, which holds since $[E(s)B]Q[E(s)B]^*$ is selfadjoint and positive semidefinite. The same computations can be made with $[\tilde{E}(s)\tilde{B}]Q[\tilde{E}(s)\tilde{B}]^*$. Furthermore, the operator $u_{xx}(r,\xi)[E(s)B]Q[\tilde{E}(s)\tilde{B}]^*$ is also of trace class for almost every s, since, by (1.1.1), (1.1.4), and (1.1.7),

$$\begin{aligned} &\|u_{xx}(r,\xi)[E(s)B]Q[\tilde{E}(s)\tilde{B}]^*\|_{\mathrm{Tr}} \\ &\quad \le \|u_{xx}(r,\xi)\|_{\mathcal{L}(H)}\,\|[E(s)B]Q[\tilde{E}(s)\tilde{B}]^*\|_{\mathrm{Tr}} \\ &\quad \leqslant \|u_{xx}(r,\xi)\|_{\mathcal{L}(H)}\,\|[E(s)B]Q^{1/2}\|_{\mathrm{HS}}\,\|Q^{1/2}[\tilde{E}(s)\tilde{B}]^*\|_{\mathrm{HS}} \\ &\quad = \|u_{xx}(r,\xi)\|_{\mathcal{L}(H)}\,\|[E(s)B]Q^{1/2}\|_{\mathrm{HS}}\,\|[\tilde{E}(s)\tilde{B}]Q^{1/2}\|_{\mathrm{HS}} \\ &\quad = \|u_{xx}(r,\xi)\|_{\mathcal{L}(H)}\Big(\mathrm{Tr}\big([E(s)B]Q[E(s)B]^*\big)\mathrm{Tr}\big([\tilde{E}(s)\tilde{B}]Q[\tilde{E}(s)\tilde{B}]^*\big)\Big)^{1/2}. \end{aligned}$$

Therefore we may rewrite the operator in the trace in (1.2.6) by adding and subtracting $u_{xx}(r,\xi)[E(s)B]Q[\tilde{E}(s)\tilde{B}]^*$ to get

$$\begin{aligned} &u_{xx}(r,\xi)\big\{[\tilde{E}(s)\tilde{B}]Q[\tilde{E}(s)\tilde{B}]^* - [E(s)B]Q[E(s)B]^*\big\} \\ &= u_{xx}(r,\xi)[\tilde{E}(s)\tilde{B} - E(s)B]Q[\tilde{E}(s)\tilde{B}]^* \\ &\quad + u_{xx}(r,\xi)[E(s)B]Q[\tilde{E}(s)\tilde{B} - E(s)B]^* \\ &=: O_1 + O_2. \end{aligned}$$

Further, using (1.1.2), (1.1.5), and that Q and $u_{xx}(r,\xi)$ are selfadjoint, we obtain

$$\begin{aligned} \operatorname{Tr}(O_1+O_2) &= \operatorname{Tr}(O_1) + \operatorname{Tr}(O_2) = \operatorname{Tr}(O_1) + \operatorname{Tr}(O_2^*) \\ &= \operatorname{Tr}(O_1) + \operatorname{Tr}([\tilde{E}(s)\tilde{B} - E(s)B]Q[E(s)B]^* u_{xx}(r,\xi)) \\ &= \operatorname{Tr}(O_1) + \operatorname{Tr}(u_{xx}(r,\xi)[\tilde{E}(s)\tilde{B} - E(s)B]Q[E(s)B]^*) \\ &= \operatorname{Tr}\Big(u_{xx}(r,\xi)[\tilde{E}(s)\tilde{B} - E(s)B]Q[\tilde{E}(s)\tilde{B} + E(s)B]^*\Big) \qquad (1.2.7) \\ &= \operatorname{Tr}\Big([\tilde{E}(s)\tilde{B} + E(s)B]Q[\tilde{E}(s)\tilde{B} - E(s)B]^* u_{xx}(r,\xi)\Big) \\ &= \operatorname{Tr}\Big(u_{xx}(r,\xi)[\tilde{E}(s)\tilde{B} + E(s)B]Q[\tilde{E}(s)\tilde{B} - E(s)B]^*\Big). \qquad (1.2.8) \end{aligned}$$

Finally, by inserting (1.2.7) or (1.2.8) into (1.2.6) the proof is complete. □

1.3. The stochastic wave equation

1.3.1. A general error formula. In this subsection we apply the general result from Section 1.2 to the numerical approximation of the stochastic wave equation

$$\mathrm{d}\dot{U}(t) - \Delta U(t)\,\mathrm{d}t = \mathrm{d}W(t),\ U(t)|_{\partial\mathcal{D}} = 0,\ t>0; \quad U(0)=U_0,\ \dot{U}(0)=V_0, \qquad (1.3.1)$$

where the solution process $(U(t))_{t\geq 0}$ and the Wiener process $(W(t))_{t\geq 0}$ take values in $U = L^2(\mathcal{D})$, where $\mathcal{D}$ denote a convex bounded domain in $\mathbb{R}^d$ with boundary $\partial\mathcal{D}$. In the sequel we rewrite this equation in an abstract Itô form that fits the semigroup framework for SPDEs, see [**31**, Example 5.8]. At the same time we introduce a framework for measuring the regularity of the solution and to perform a careful error analysis. To this aim equip $L^2(\mathcal{D})$ with the usual norm $\|\cdot\|_{L^2(\mathcal{D})}$ and inner product $\langle\cdot,\cdot\rangle_{L^2(\mathcal{D})}$. Let $\Lambda := -\Delta = -\sum_{j=1}^d \partial^2/\partial\xi_j^2$ be the Laplace operator on $L^2(\mathcal{D})$ with zero-Dirichlet boundary condition, i.e., with domain $D(\Lambda) := \{v \in H_0^1(\mathcal{D}):\ \Lambda u \in L^2(\mathcal{D})\}$, where Λu is understood in the distributional sense, see [**5**, Example 3.4.7]. As usual, $H^n(\mathcal{D})$ denotes the L^2-Sobolev space of order $n \in \mathbb{N}_0$ on $\mathcal{D}$ and $H_0^1(\mathcal{D})$ is the $H^1(\mathcal{D})$-closure of the space $C_c^\infty(\mathcal{D})$ of compactly supported test functions. In the sequel, we use the smoothness spaces $\dot{H}^\alpha$, $\alpha \in \mathbb{R}$, defined by

$$\dot{H}^\alpha := D(\Lambda^{\alpha/2}) := \Big\{v = \sum_{k=1}^\infty v_k\varphi_k : (v_k)_{k\in\mathbb{N}} \subset \mathbb{R},\ \|v\|_{\dot{H}^\alpha} := \|\Lambda^{\alpha/2}v\|_{L^2(\mathcal{D})} = \Big(\sum_{k=1}^\infty \lambda_k^\alpha v_k^2\Big)^{1/2} < \infty\Big\},$$

where $(\varphi_k)_{k\in\mathbb{N}} \subset D(\Lambda)$ is an orthonormal basis of $L^2(\mathcal{D})$ consisting of eigenfunctions of Λ and $(\lambda_k)_{k\in\mathbb{N}} \subset (0,\infty)$ is the corresponding sequence of eigenvalues; compare [**103**, Chapters 3 and 19]. They are Hilbert spaces and one has the identities $\dot{H}^0 = H = L^2(\mathcal{D})$, $\dot{H}^1 = H_0^1(\mathcal{D})$ and $\dot{H}^2 = D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$, where the natural norms of the respective spaces are equivalent. The latter equality is a consequence of the elliptic regularity estimate

$$\|v\|_{H^2} \leq C\|v\|_{\dot{H}^2}, \quad v \in \dot{H}^2, \tag{1.3.2}$$

for bounded convex domains, see [**40**, Corollary 1]. Without the convexity assumption one can still define the $\dot{H}^\alpha$ spaces as above but one does not obtain a characterization of $D(\Lambda) = \dot{H}^2$ in terms of classical Sobolev spaces. For negative α, the elements of $\dot{H}^\alpha$ are formal sums and we identify them with elements of $L^2(\mathcal{D})$ if $\sum_{k=1}^\infty v_k^2 < \infty$, so that $\dot{H}^\alpha$ is the closure of $L^2(\mathcal{D})$ w.r.t. the $\|\cdot\|_{\dot{H}^\alpha}$-norm.

Remark 1.3.1. The spaces $\dot{H}^\alpha$, $\alpha \in \mathbb{R}$, can be obtained by both real and complex interpolation: For $\alpha = (1-\theta)\alpha_0 + \theta\alpha_2$, $\theta \in (0,1)$, one has $\dot{H}^\alpha = (\dot{H}^{\alpha_0}, \dot{H}^{\alpha_1})_{\theta,2} = [\dot{H}^{\alpha_0}, \dot{H}^{\alpha_1}]_\theta$ with equivalent norms, where $(\cdot,\cdot)_{\theta,2}$ and $[\cdot,\cdot]_\theta$ denotes real interpolation with summability parameter $q = 2$ and complex interpolation, respectively. This follows, e.g., from [**104**, Theorem 1.18.5] and the fact that the spaces $\dot{H}^\alpha$, $\alpha \in \mathbb{R}$, are isometrically isomorphic to weighted ℓ^2-spaces. We will frequently use the corresponding interpolation inequalities throughout the dissertation.

In addition we define the product space

$$\mathcal{H}^\alpha := \dot{H}^\alpha \times \dot{H}^{\alpha-1}, \quad \alpha \in \mathbb{R}, \tag{1.3.3}$$

with inner product

$$\langle v, w\rangle_{\mathcal{H}^\alpha} = \langle v_1, w_1\rangle_{\dot{H}^\alpha} + \langle v_2, w_2\rangle_{\dot{H}^{\alpha-1}},$$

where $v = [v_1, v_2]^T$ and $w = [w_1, w_2]^T$. The corresponding norms are

$$\|v\|^2_{\mathcal{H}^\alpha} := \|v_1\|^2_{\dot{H}^\alpha} + \|v_2\|^2_{\dot{H}^{\alpha-1}}.$$

We take H to be the special case of (1.3.3) when $\alpha = 0$ with norm $\|\cdot\| = \|\cdot\|_{\mathcal{H}^0}$ and inner product $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{L^2(\mathcal{D})} + \langle\Lambda^{-1/2}\cdot, \Lambda^{-1/2}\cdot\rangle_{L^2(\mathcal{D})}$. We now regard Λ as an operator from $\dot{H}^1 \to \dot{H}^{-1}$ defined by $\langle\Lambda x, y\rangle_{\dot{H}^{-1}\times\dot{H}^1} = \langle\nabla x, \nabla y\rangle_{L_2(D)}$ and we let $A : D(A) \subset H \to H$ and $B : \dot{H}^{-1} \to H$ be defined by

$$A := \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ I \end{bmatrix}, \tag{1.3.4}$$

where the domain of A is given by

$$D(A) = \left\{x \in H : Ax = \begin{bmatrix} x_2 \\ -\Lambda x_1 \end{bmatrix} \in H = \dot{H}^0 \times \dot{H}^{-1}\right\} = \mathcal{H}^1 = \dot{H}^1 \times \dot{H}^0.$$

We have some freedom in defining U but a natural choice is $U = \dot{H}^0 = L^2(\mathcal{D})$. Thus, we consider the process $(W(t))_{t\geqslant 0}$ to be a Q-Wiener process in $U = L^2(\mathcal{D})$ and thus Q to be bounded, selfadjoint and positive semidefinite on $U = L^2(\mathcal{D})$. Note that $\dot{H}^0 \hookrightarrow \dot{H}^{-1}$, and therefore $B \in \mathcal{L}(U, H)$. It is well known that the operator $-A$ is the generator of a strongly continuous semigroup $E(t) = \mathrm{e}^{-tA}$ on H, in fact, a unitary group, that can be written as

$$E(t) = \mathrm{e}^{-tA} = \begin{bmatrix} C(t) & \Lambda^{-1/2}S(t) \\ -\Lambda^{1/2}S(t) & C(t) \end{bmatrix}, \tag{1.3.5}$$

where $C(t) = \cos(t\Lambda^{1/2})$ and $S(t) = \sin(t\Lambda^{1/2})$ are defined via the spectral calculus of Λ.

With these definitions (1.3.1) can be written in the abstract Itô form

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t),\ t \in [0,T]; \quad X(0) = X_0, \tag{1.3.6}$$

with H-valued weak solution $X = \big[X_1(t), X_2(t)\big]^T$ given by

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}W(s), \tag{1.3.7}$$

if also the initial value $X_0 = \big[X_{0,1}, X_{0,2}\big]^T$ is H-valued and $\mathcal{F}_0$-measurable and if

$$\int_0^T \|E(t)BQ^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t = \mathrm{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) < \infty.$$

We have the following a technical result regarding the finiteness of the above trace.

Lemma 1.3.2. *If $E(t)$ is given by* (1.3.5)*, then the following four statements are equivalent.*

(i) $\mathrm{Tr}(\Lambda^{-1/2}Q\Lambda^{-1/2}) < \infty$.
(ii) $\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$.
(iii) $\|\Lambda^{-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$.
(iv) $\mathrm{Tr}(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t) < \infty$ *for some, hence all,* $T > 0$.

If either of them holds, then

$$\int_0^T \|E(t)BQ^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t = \mathrm{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) = T\,\mathrm{Tr}(\Lambda^{-1/2}Q\Lambda^{-1/2}). \tag{1.3.8}$$

Proof. The operator $\Lambda^{-1/2}Q\Lambda^{-1/2}$ is selfadjoint and positive semidefinite on $\dot H^0$. Hence, if the trace is finite it equals the trace norm and the other way around. This implies that (i) $\Leftrightarrow$ (iii). We know assume that (iv) holds. By monotone convergence,

$$\mathrm{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) = \int_0^T \mathrm{Tr}\Big(E(t)BQB^*E(t)^*\Big)\,\mathrm{d}t \tag{1.3.9}$$

and by (1.1.6) and since $E(t)^* = E(t)^{-1}$ we have

$$\begin{aligned}
\mathrm{Tr}_H(E(t)BQB^*E(t)^*) &= \mathrm{Tr}_{\dot H^0}(Q^{1/2}B^*E(t)^*E(t)BQ^{1/2}) = \mathrm{Tr}_{\dot H^0}(Q^{1/2}B^*BQ^{1/2})\\
&= \|BQ^{1/2}\|_{\mathcal{L}_2(\dot H^0,H)}^2 = \|[0,Q^{1/2}]^T\|_{\mathcal{L}_2(\dot H^0,H)}^2 = \|Q^{1/2}\|_{\mathcal{L}_2(\dot H^0,\dot H^{-1})}^2\\
&= \|\Lambda^{-1/2}Q^{1/2}\|_{\mathcal{L}_2(\dot H^0)}^2 = \mathrm{Tr}_{\dot H^0}(Q^{1/2}\Lambda^{-1}Q^{1/2}) = \mathrm{Tr}_{\dot H^0}(\Lambda^{-1/2}Q\Lambda^{-1/2}).
\end{aligned}$$

Thus, the integrand in the right-hand side of (1.3.9) is constant and is equal to $\mathrm{Tr}_{\dot H^0}(\Lambda^{-1/2}Q\Lambda^{-1/2})$, which implies (1.3.8). Therefore (i) must be true. This argument is reversible so (i) $\Leftrightarrow$ (iv). But from this computation it is also evident that $\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(\Lambda^{-1/2}Q\Lambda^{-1/2})$; i.e., that (i) $\Leftrightarrow$ (ii). □

Next we describe the spatial regularity of the solution of (1.3.6) given the regularity of Q and the initial data.

Theorem 1.3.3. *If* $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}} < \infty$, $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathcal{H}^\beta)$ *for some* $\beta \geqslant 0$ *and* A, B *as in* (1.3.4)*, then there is a unique weak solution of* (1.3.6) *given by* (1.3.7)*. Furthermore, it holds that*

$$\|X(t)\|_{L^2(\Omega;\mathcal{H}^\beta)} \le C\big(\|X_0\|_{L^2(\Omega;\mathcal{H}^\beta)} + t^{1/2}\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}\big) \quad t \in [0,T]. \tag{1.3.10}$$

PROOF. As, by (1.1.4), we have that

$$\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}} \leqslant \|\Lambda^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}$$

it follows from [**31**, Theorem 5.4] that (1.3.7) is the unique weak solution by Lemma 1.3.2. To show the required regularity let $(e_k)_{k=1}^{\infty}$, an arbitrary ON-basis in U. Then, and for any $\beta \geq 0$, we have

$$\begin{aligned}\int_0^t \|E(s)BQ^{1/2}\|^2_{\mathcal{L}_2(U,\mathcal{H}^\beta)}\,\mathrm{d}s &= \int_0^t\sum_{k=1}^{\infty}\|E(s)BQ^{1/2}e_k\|^2_{\mathcal{H}^\beta}\,\mathrm{d}s\\ &= \int_0^t\sum_{k=1}^{\infty}\big\{\big\|\Lambda^{-1/2}S(s)Q^{1/2}e_k\big\|^2_{\dot H^\beta} + \big\|C(s)Q^{1/2}e_k\big\|^2_{\dot H^{\beta-1}}\big\}\,\mathrm{d}s \qquad (1.3.11)\\ &= \int_0^t\big\{\big\|\Lambda^{(\beta-1)/2}S(s)Q^{1/2}\big\|^2_{\mathrm{HS}} + \big\|\Lambda^{(\beta-1)/2}C(s)Q^{1/2}\big\|^2_{\mathrm{HS}}\big\}\,\mathrm{d}s\\ &\leq 2t\|\Lambda^{(\beta-1)/2}Q^{1/2}\|^2_{\mathrm{HS}},\end{aligned}$$

where, for the last inequality, we used the fact that the Λ commutes with $C(s)$, $S(s)$ and (1.1.4) together with the uniform boundedness of $C(s)$, $S(s)$ in $\mathcal{L}(\dot H^0)$. Therefore, (1.3.10) follows from (1.3.7), the boundedness of $E(t)$ in $\mathcal{H}^\beta$, the Itô isometry (1.1.13), and (1.3.11):

$$\begin{aligned}\|X(t)\|^2_{L^2(\Omega;\mathcal{H}^\beta)} &\leq 2\Big(\|E(t)X_0\|^2_{L^2(\Omega;\mathcal{H}^\beta)} + \Big\|\int_0^t E(t-s)B\,\mathrm{d}W(s)\Big\|^2_{L^2(\Omega;\mathcal{H}^\beta)}\Big)\\ &\leq 2\Big(\|X_0\|^2_{L^2(\Omega;\mathcal{H}^\beta)} + \int_0^t \|E(s)BQ^{1/2}\|^2_{\mathcal{L}_2(U,\mathcal{H}^\beta)}\,\mathrm{d}s\Big).\end{aligned}$$

This completes the proof. □

COROLLARY 1.3.4. *If* $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$, $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathcal{H}^\beta)$ *for some* $\beta \geqslant 0$ *and* A, B *as in* (1.3.4)*, then there is a unique weak solution of* (1.3.6) *given by* (1.3.7)*. Furthermore, it holds that*

$$\|X(t)\|_{L^2(\Omega;\mathcal{H}^\beta)} \leq C\big(\|X_0\|_{L^2(\Omega;\mathcal{H}^\beta)} + t^{1/2}\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}\big) \quad t\in[0,T].$$

PROOF. Since $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|^2_{\mathrm{HS}} \leq \|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}}$ by (1.1.9), the statement follows immediately from Theorem 1.3.3. □

We will need the following result on the Hölder continuity of $E(t)$. It will put an ultimate limit on the convergence rate that one can achieve with respect to time-stepping.

LEMMA 1.3.5. *If* $(E(t))_{t\geqslant 0}$ *is the semigroup in* (1.3.5)*, then*

$$\|(E(t)-E(s))x\| \leqslant C|t-s|^\alpha\|x\|_{\mathcal{H}^\alpha}, \quad x\in\mathcal{H}^\alpha,\ t,s\geqslant 0,\ \alpha\in[0,1].$$

PROOF. The operator $E(t)$ is bounded on $\mathcal{H}$ so the statement is true for $\alpha = 0$. Let $\alpha = 1$ and $x = [x_1, x_2]^T \in \mathcal{H}^1$. Then

$$\begin{aligned}\|(E(t)-E(s))x\|^2 &= \|(C(t)-C(s))x_1 + (S(t)-S(s))\Lambda^{-1/2}x_2\|^2_{\dot H^0}\\ &\quad + \|(S(s)-S(t))\Lambda^{1/2}x_1 + (C(t)-C(s))x_2\|^2_{\dot H^{-1}} =: A_1 + A_2.\end{aligned}$$

By the triangle inequality and the definition of the norm on $\dot{H}^{-1}$, we have for the last term that

$$\begin{aligned} A_2 &\leqslant 2\|\Lambda^{-1/2}(S(t)-S(s))\Lambda^{1/2}x_1\|_{\dot{H}^0}^2 + 2\|\Lambda^{-1/2}(C(t)-C(s))x_2\|_{\dot{H}^0}^2 \\ &= 2\|(S(t)-S(s))x_1\|_{\dot{H}^0}^2 + 2\|(C(t)-C(s))\Lambda^{-1/2}x_2\|_{\dot{H}^0}^2. \end{aligned}$$

Since $x_2 \in \dot{H}^0$ it follows that $\Lambda^{-1/2}x_2 \in \dot{H}^1$. Hence we only need to investigate the Hölder continuity of C and S as functions from $[0,T]$ to $\mathcal{L}(\dot{H}^1, \dot{H}^0)$. To this aim we note that for real y the inequality

$$|\sin(ty) - \sin(sy)| \leqslant |t-s||y| \tag{1.3.12}$$

holds. It follows that for $\xi \in \dot{H}^1$ we have $\|(S(t)-S(s))\xi\|_{\dot{H}^0} \leqslant |t-s|\|\xi\|_{\dot{H}^1}$. Indeed,

$$\begin{aligned} \|(S(t)-S(s))\xi\|_{\dot{H}^0}^2 &= \sum_{j=1}^{\infty} \langle (S(t)-S(s))\xi, \phi_j\rangle_{\dot{H}^0}^2 \\ &= \sum_{j=1}^{\infty} (\sin(t\lambda_j^{1/2}) - \sin(s\lambda_j^{1/2}))^2 \langle \xi, \phi_j\rangle_{\dot{H}^0}^2 \\ &\leqslant \sum_{j=1}^{\infty} (t-s)^2 \lambda_j \langle \xi, \phi_j\rangle_{\dot{H}^0}^2 = (t-s)^2 \|\xi\|_{\dot{H}^1}^2. \end{aligned}$$

The inequality (1.3.12) holds also with sin replaced by cos. Thus the statement of the lemma holds also for $\alpha = 1$. The intermediate case follows by interpolation. □

We are now ready to prove a weak error bound for perturbations of the stochastic wave equation by assuming as little as possible about the perturbation $\tilde{X}$ in order to accommodate various approximation scenarios. We let V be a real separable Hilbert space with norm $\|\cdot\|_V$ and inner product $\langle\cdot,\cdot\rangle_V$.

THEOREM 1.3.6. *Assume that* $(X(t))_{t\in[0,T]}$ *is the weak solution given by* (1.3.7) *of the stochastic wave equation* (1.3.6) *with* A, B *as in* (1.3.4). *Suppose further that* $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ *and* $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathcal{H}^{2\beta})$ *for some* $\beta \geqslant 0$. *Assume further that* $\tilde{X}(T)$ *can be represented as* $\tilde{X}(T) = \tilde{Y}(T)$, *where* $\tilde{Y}(t)$ *is given by* (1.2.1) *with* $\tilde{Y}(0) = \tilde{E}(T)\tilde{P}X_0$, $\tilde{P} \in \mathcal{L}(H)$ *and* $\tilde{B} = \tilde{P}B$ *such that* $t \mapsto \tilde{E}(t)\tilde{B}$ *is a measurable mapping from* $[0,T]$ *to* $\mathcal{L}_2(U_0,H)$ *with* $\tilde{E}\tilde{B} \in L^2((0,T);\mathcal{L}_2(U_0,H))$. *Let* $g \in C^2(V,\mathbb{R})$ *with* $g'' \in C_b(V, \mathcal{L}(V))$ *and* $L \in \mathcal{L}(H,V)$. *Define*

$$\begin{aligned} K_1 &:= \sup_{t\in[0,T]} \|\tilde{E}(t)\tilde{P}\|_H, \\ K_2 &:= \|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}}. \end{aligned} \tag{1.3.13}$$

Then there is $C = C(T, \|L\|_{\mathcal{L}(H,V)}, X_0, g, K_1, K_2)$ *such that*

$$\left|\mathbb{E}\left(g(L\tilde{X}(T)) - g(LX(T))\right)\right| \leqslant C \sup_{t\in[0,T]} \|L(\tilde{E}(t)\tilde{P} - E(t))\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}. \tag{1.3.14}$$

We want to emphasise that this theorem reduces the problem of proving weak error estimates for the stochastic wave equation to proving deterministic error estimates for the approximation of the semigroup or, to be precise, to find a bound for

$$\sup_{t\in[0,T]} \|L(\tilde{E}(t)\tilde{P} - E(t))\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}.$$

PROOF. First note that (1.3.13) implies that $EB \in L^2((0,T);\mathcal{L}_2(U_0,H))$ using Lemma 1.3.2 since, as noted above,

(1.3.15)
$$\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}} \leqslant \|\Lambda^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}} \leqslant \|\Lambda^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}}.$$

Therefore, we may use Theorem 1.2.1 with $G(X) = g(LX)$ and $Y(0) = E(T)X_0$ noting that $X(T) = Y(T)$ in this case. To this aim we note that $G'(X) = L^*g'(LX)$ and $G''(X) = L^*g''(LX)L$. The terms in (1.2.3) will be estimated in order of appearance with $\mathcal{O}$ as in (1.2.5). For the first term, by (1.1.23), Corollary 1.1.4 and the fact that both $Y(0)$ and $\tilde Y(0)$ are $\mathcal{F}_0$-measurable, we have, using also the linear growth of g' and Theorem 1.3.3 with $\beta = 0$, that

$$\begin{aligned}
&\Big|\mathbb{E}\Big(\int_0^1 \Big\langle u_x\big(0, Y(0) + s(\tilde Y(0) - Y(0))\big), \tilde Y(0) - Y(0)\Big\rangle_H \,\mathrm{d}s\Big)\Big| \\
&= \Big|\mathbb{E}\Big(\int_0^1 \Big\langle \mathbb{E}\Big(g'\big(LZ(T,0,Y(0) + s(\tilde Y(0) - Y(0)))\big)\Big|\mathcal{F}_0\Big), L(\tilde Y(0) - Y(0))\Big\rangle_V \,\mathrm{d}s\Big)\Big| \\
&\leqslant \int_0^1 \big\|g'\big(LZ\big(T,0,Y(0) + \theta(\tilde Y(0) - Y(0))\big)\big)\big\|_{L^2(\Omega;V)}\,\mathrm{d}\theta \\
&\qquad \times \|L(\tilde E(T) - E(T))X_0\|_{L^2(\Omega,V)} \\
&\leqslant C\big(1 + \int_0^1 \big\|Z\big(T,0,Y(0) + \theta(\tilde Y(0) - Y(0))\big)\big\|_{L^2(\Omega;H)}\,\mathrm{d}\theta\big) \\
&\qquad \times \|L(\tilde E(T) - E(T))\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}\,\|X_0\|_{L^2(\Omega;\mathcal{H}^{2\beta})} \\
&\leqslant C\big(1 + \|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}} + \|X_0\|_{L^2(\Omega;H)}\big) \\
&\qquad \times \|L(\tilde E(T) - E(T))\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}\,\|X_0\|_{L^2(\Omega;\mathcal{H}^{2\beta})} \\
&\leqslant C\big(1 + \|\Lambda^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} + \|X_0\|_{L^2(\Omega;H)}\big) \\
&\qquad \times \|L(\tilde E(T) - E(T))\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}\,\|X_0\|_{L^2(\Omega;\mathcal{H}^{2\beta})}.
\end{aligned}$$

For the second term we note that with

$$\begin{aligned}
\mathcal{O}^-(t) &= \tilde E(T-t)\tilde B - E(T-t)B = (\tilde E(T-t)\tilde P - E(T-t))B, \\
\mathcal{O}^+(t) &= \tilde E(T-t)\tilde B + E(T-t)B = (\tilde E(T-t)\tilde P + E(T-t))B,
\end{aligned}$$

and by (1.1.23) and Corollary 1.1.4 we may write

$$\begin{aligned}
&\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}\big(t,\tilde Y(t)\big)\mathcal{O}^+(t)Q\mathcal{O}^-(t)^*\Big)\,\mathrm{d}t \\
&\qquad = \mathbb{E}\int_0^T \mathrm{Tr}\Big(\mathbb{E}\big(L^*g''\big(LZ(T,t,\tilde Y(t))\big)L\big|\mathcal{F}_t\big)\mathcal{O}^+(t)Q\mathcal{O}^-(t)^*\Big)\,\mathrm{d}t.
\end{aligned}$$

Using (1.1.5), (1.1.3) and (1.1.4) we bound the integrand above as follows:

$$\begin{aligned}
&\Big|\mathrm{Tr}\Big(\mathbb{E}\big(L^*g''(LZ(T,t,\tilde Y(t)))L\big|\mathcal{F}_t\big)\mathcal{O}^+(t)Q\mathcal{O}^-(t)^*\Big)\Big| \\
&\quad = \Big|\mathrm{Tr}\Big(\mathbb{E}\big(g''(LZ(T,t,\tilde Y(t)))\big|\mathcal{F}_t\big)L\mathcal{O}^+(t)Q\mathcal{O}^-(t)^*L^*\Big)\Big| \\
&\quad \leqslant \sup_{x\in V}\|g''(x)\|_{\mathcal{L}(V)}\|L\mathcal{O}^+(t)Q\mathcal{O}^-(t)^*L^*\|_{\mathrm{Tr}}.
\end{aligned}$$

Here we have, by assumption, that $\sup_{x\in V} \|g''(x)\|_{\mathcal{L}(V)} < \infty$. Furthermore, by (1.1.1) and (1.1.4),

$$\begin{aligned}
\|L\mathcal{O}^+(t)Q\mathcal{O}^-(t)^*L^*\|_{\mathrm{Tr}} &= \|L\mathcal{O}^+(t)\Lambda^{1/2}\Lambda^{-1/2}Q\mathcal{O}^-(t)^*L^*\|_{\mathrm{Tr}} \\
&\le \|L\mathcal{O}^+(t)\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,V)}\, \|\Lambda^{-1/2}Q\mathcal{O}^-(t)^*L^*\|_{\mathrm{Tr}} \\
&= \|L\mathcal{O}^+(t)\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,V)}\, \|L\mathcal{O}^-(t)Q\Lambda^{-1/2}\|_{\mathrm{Tr}} \\
&= \|L\mathcal{O}^+(t)\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,V)}\, \|L\mathcal{O}^-(t)\Lambda^{1/2-\beta}\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} \\
&\le \|L\mathcal{O}^+(t)\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,V)}\, \|L\mathcal{O}^-(t)\Lambda^{1/2-\beta}\, \|_{\mathcal{L}(\dot H^0,V)}\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}}.
\end{aligned}$$

The first factor can be estimated as

$$\begin{aligned}
\|L\mathcal{O}^+(t)\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,V)} &= \|L(\tilde E(T-t)\tilde P + E(T-t))B\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,V)} \\
&\le \|L\|_{\mathcal{L}(H,V)}\big(\|\tilde E(T-t)\tilde P\|_{\mathcal{L}(H,H)} + \|E(T-t)\|_{\mathcal{L}(H,H)}\big)\|B\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,H)} \\
&\le \|L\|_{\mathcal{L}(H,V)}(K_1+1),
\end{aligned}$$

because $\|E(s)\|_{\mathcal{L}(H,H)} = 1 = \|B\Lambda^{1/2}\|_{\mathcal{L}(\dot H^0,H)}$. Similarly, the middle factor may be bounded by

$$\big\|L\big(\tilde E(T-t)\tilde P - E(T-t)\big)\big\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}\, \|B\Lambda^{1/2-\beta}\|_{\mathcal{L}(\dot H^0,\mathcal{H}^{2\beta})},$$

where $\|B\Lambda^{1/2-\beta}\|_{\mathcal{L}(\dot H^0,\mathcal{H}^{2\beta})} = 1$. The third term is K_2. Thus, we conclude that

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}\big(t,\tilde Y(t)\big)\mathcal{O}^-(t)Q\mathcal{O}^+(t)^*\Big)\,\mathrm{d}t\Big| \\
&\qquad\leqslant C\int_0^T \big\|L\big(\tilde E(T-t)\tilde P - E(T-t)\big)\big\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)}\,\mathrm{d}t \\
&\qquad\leqslant C\,T \sup_{t\in[0,T]} \big\|L\big(\tilde E(t)\tilde P - E(t)\big)\big\|_{\mathcal{L}(\mathcal{H}^{2\beta},V)},
\end{aligned}$$

and the proof is complete. □

1.3.2. Weak convergence of temporally semidiscrete schemes. We begin by applying Theorem 1.3.6 to semidiscrete approximation schemes where the discretization is with respect to time. We will use results on so-called I-stable rational approximations considered in [**23**]. An I-stable rational approximation of order p is a rational function R such that

$$\begin{aligned}
|R(iy) - e^{-iy}| &\leqslant C|y|^{p+1}, \quad |y| \leqslant b, \\
|R(iy)| &\leqslant 1, \qquad\qquad\quad y\in\mathbb{R},
\end{aligned} \tag{1.3.16}$$

for some positive integer p and some $b > 0$. A class of such functions are constructed in [**16**] and analyzed further in connection to oscillation equations in [**86**]. It contains the implicit Euler method ($p = 1$) and, which is important since it preserves energy for the wave equation, the Crank-Nicolson method ($p = 2$).

For these functions the operators $E_{\Delta t} = R(\Delta t A)$, $\Delta t > 0$, are well defined on H and they are contractions, and hence stable, as $-A$ generates a unitary group. Here $\Delta t = T/N, N \in \mathbb{N}$, is the time step. We can approximate the solution of (1.3.6) on the uniform grid $t_j = j\Delta t,\ j = 0,\dots,N$, by the solution of the difference equation

$$X^j_{\Delta t} = E_{\Delta t}(X^{j-1}_{\Delta t} + B\Delta W^j),\ j = 1,\dots,N; \quad X^0_{\Delta t} = X_0, \tag{1.3.17}$$

dc_1541_18

given by

$$X_{\Delta t}^n = E_{\Delta t}^n X_0 + \sum_{j=1}^{n} E_{\Delta t}^{n-j+1} B\Delta W^j, \quad n \geqslant 1,$$

where $\Delta W^j = W(t_j) - W(t_{j-1})$. To We first define a new discrete process

$$Y_{\Delta t}^n = E_{\Delta t}^{N-n} X_{\Delta t}^n = E_{\Delta t}^N X_0 + \sum_{j=1}^{n} E_{\Delta t}^{N-j+1} B\Delta W^j. \tag{1.3.18}$$

Clearly $Y_{\Delta t}^N = X_{\Delta t}^N$. In order to make a piecewise constant time interpolation of (1.3.18) we introduce the time intervals $I_j = [t_{j-1}, t_j)$ for $j = 1, \ldots, N$ and $I_{N+1} = \{t_N\} = \{T\}$. With χ being the indicator function we then write

$$\tilde{E}_{\Delta t}(T-t) = \sum_{j=1}^{N+1} E_{\Delta t}^{N-j+1} \chi_{I_j}(t).$$

It may easily be checked that this corresponds to writing

$$\tilde{E}_{\Delta t}(t) = \sum_{j=0}^{N} E_{\Delta t}^{j} \chi_{\widehat{I}_j}(t)$$

with $\widehat{I}_0 = \{t_0\} = \{0\}$ and $\widehat{I}_j = (t_{j-1}, t_j]$ for $j = 1, \ldots, N$. We finally define

$$\tilde{Y}_{\Delta t}(t) := \tilde{E}_{\Delta t}(T) X_0 + \int_0^t \tilde{E}_{\Delta t}(T-s) B \,\mathrm{d}W(s). \tag{1.3.19}$$

The process $\tilde{Y}_{\Delta t}$ has the desired properties; that is, it is of the form (1.2.1) with $\tilde{Y}(0) = \tilde{E}_{\Delta t}(T)X_0$, and, in addition, $\tilde{Y}_{\Delta t}(t_n) = Y_{\Delta t}^n$, in particular for $n = N$. To apply Theorem 1.3.6 it remains to show that $\tilde{E}\tilde{B} \in L^2((0,T); \mathcal{L}_2(U_0, H))$ with $\tilde{E}(t) = \tilde{E}_{\Delta t}(t)$ and $\tilde{P} = I$ (hence $\tilde{B} = B$). This is indeed the case as soon as we are guaranteed a weak solution of (1.3.6), as stated in the following lemma.

Lemma 1.3.7. *If $\tilde{E}(t) = \tilde{E}_{\Delta t}(t)$ and $\tilde{B} = B$, then any of the equivalent conditions (i)-(iv) of Lemma 1.3.2 imply that $\tilde{E}_{\Delta t}\tilde{B} \in L^2((0,T); \mathcal{L}_2(U_0, H))$.*

Proof. If $\mathrm{Tr}(\Lambda^{-1/2} Q \Lambda^{-1/2}) < \infty$ then, for all $t \geq 0$, using (1.1.3) and (1.1.4), it follows that

$$\begin{aligned}\mathrm{Tr}(\tilde{E}_{\Delta t}(t) BQB^* \tilde{E}_{\Delta t}^*(t)) &\leqslant \|\tilde{E}_{\Delta t}(t)\|_{\mathcal{L}(H)}^2 \|BQB^*\|_{\mathrm{Tr}} \\ &\leq \|BQB^*\|_{\mathrm{Tr}} = \mathrm{Tr}(BQB^*) = \mathrm{Tr}(\Lambda^{-1/2} Q \Lambda^{-1/2}) < \infty,\end{aligned}$$

where the last equality is shown in the proof of Lemma 1.3.2 as $\mathrm{Tr}(BQB^*) = \mathrm{Tr}(Q^{1/2} B^* B Q^{1/2})$ by (1.1.6). The statement of the lemma follows by the monotone convergence theorem again as in the proof of Lemma 1.3.2. □

In order to make use of Theorem 1.3.6 we need a bound on $E(t) - \tilde{E}_{\Delta t}(t)$. The results in [**23**] are concerned with the difference at the grid points. With our notation their conclusion reads that, for $x \in \mathcal{H}^{p+1}$,

$$\|(E(t_n) - E_{\Delta t}^n)x\| \leqslant C t_n \Delta t^p \|x\|_{\mathcal{H}^{p+1}}. \tag{1.3.20}$$

As already mentioned, the conditions in (1.3.16) ensures that the operator $E_{\Delta t}^n$ is a contraction on H for any $n \geqslant 0$, so (1.3.20) can be extended to fractional order by interpolation; that is,

$$\|(E(t_n) - E_{\Delta t}^n)x\| \leqslant C t_n \Delta t^{\alpha \frac{p}{p+1}} \|x\|_{\mathcal{H}^\alpha}, \quad \alpha \in [0, p+1]. \tag{1.3.21}$$

For our purposes, it is not enough to consider only the grid points, but fortunately a global error estimate follows easily.

LEMMA 1.3.8. *For the operators $\tilde{E}_{\Delta t}(t)$ and $E(t)$ defined above, we have that*

$$\sup_{t\in[0,T]} \|\tilde{E}_{\Delta t}(t) - E(t)\|_{\mathcal{L}(\mathcal{H}^\alpha,H)} \leqslant C(T)\Delta t^{\min(\alpha\frac{p}{p+1},1)},\ k>0,$$

where p is a nonnegative integer as in (1.3.16) *and $\alpha \geqslant 0$.*

PROOF. The statement of the lemma follows from (1.3.21) and Lemma 1.3.5. Indeed, for $t \in \widehat{I}_j$, we have

$$\begin{aligned}\tilde{E}_{\Delta t}(t) - E(t) &= \big(\tilde{E}_{\Delta t}(t_j) - E(t_j)\big) + \big(E(t_j) - E(t)\big)\\ &= \big(E^j_{\Delta t} - E(t_j)\big) + \big(E(t_j) - E(t)\big).\end{aligned}$$

Hence, with $I = [0,T]$ and $\mathcal{I} = \{0,1,\dots,N\}$,

$$\begin{aligned}&\sup_{t\in I}\|\tilde{E}_{\Delta t}(t) - E(t)\|_{\mathcal{L}(\mathcal{H}^\alpha,H)}\\ &\quad\leqslant \sup_{j\in\mathcal{I}}\big(\|E^j_{\Delta t} - E(t_j)\|_{\mathcal{L}(\mathcal{H}^\alpha,H)} + \sup_{t\in\widehat{I}_j}\|E(t_j) - E(t)\|_{\mathcal{L}(\mathcal{H}^\alpha,H)}\big)\\ &\quad\leqslant C(T)(\Delta t^{\min(\alpha\frac{p}{p+1},p)} + \Delta t^{\min(\alpha,1)}) \leqslant C(T)\Delta t^{\min(\alpha\frac{p}{p+1},1)},\ \Delta t\leq 1.\end{aligned}$$

Finally, for $\Delta t > 1$, the statement follows by stability. ☐

We are now ready to prove a bound for the weak error of the pure time-discretization via (1.3.17) of the stochastic wave equation.

THEOREM 1.3.9. *Assume that $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$, $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathcal{H}^{2\beta})$ for some $\beta \geqslant 0$ and $g \in C^2(H,\mathbb{R})$ with $g'' \in C_b(H,\mathcal{L}(H))$. Then the weak error of the rational approximation algorithm* (1.3.17) *of the stochastic wave equation described above is bounded by*

$$\big|\mathbb{E}\left(g(X^N_{\Delta t}) - g(X(T))\right)\big| \leqslant C\Delta t^{\min(2\beta\frac{p}{p+1},1)},\ \Delta t > 0.$$

PROOF. We may use Theorem 1.3.6 with $\tilde{E}(t) = \tilde{E}_{\Delta t}(t)$, $\tilde{P} = I$, $V = H$ and $L = I$, because $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ implies $\mathrm{Tr}(\Lambda^{-1/2}Q\Lambda^{-1/2}) < \infty$ by (1.3.15) and hence $\tilde{E}_{\Delta t}\tilde{B} \in L^2((0,T);\mathcal{L}_2(U_0,H))$ by Lemma 1.3.7. Thus, our claim follows by applying Lemma 1.3.8 with $\alpha = 2\beta$ to (1.3.14). ☐

1.3.3. Weak convergence of fully discrete schemes. In this section we will present an error estimate for a fully discrete scheme. We will borrow the setting from [**16**], where estimates for the deterministic wave equation are proved. The spatial discretization can be performed by a Galerkin finite element method and the time discretization, as above, by I-stable rational approximations of the exponential function. We briefly describe this method and state the error estimates from [**16**]. We let $\{V^r_{h,0}\}_{0<h\leqslant 1} \subset H^1_0(\mathcal{D}) = \dot{H}^1$, $r \in \mathbb{N}, r \geq 2$, be a family of finite dimensional subspaces. Unless otherwise stated, we endow the spaces V_h with the inner product $\langle\cdot,\cdot\rangle_{L^2(\mathcal{D})}$ and the norm $\|\cdot\|_{L^2(\mathcal{D})}$. We assume that the error estimate

$$\|R_h v - v\| \leqslant Ch^\beta\|v\|_{\dot{H}^\beta},\quad v\in\dot{H}^\beta,\ \beta\in[1,r], \tag{1.3.22}$$

holds, where the Ritz projection $R_h\colon \dot{H}^1 \to V^r_{h,0}$ is defined by

$$\langle\nabla R_h v, \nabla\chi\rangle = \langle\nabla v,\nabla\chi\rangle,\quad \forall v\in\dot{H}^1, \chi\in V^r_{h,0}.$$

Remark 1.3.10. For $r = 2$ the subspaces $V^r_{h,0}$ can be chosen to be the space of continuous piecewise linear functions on a triangulation of $\mathcal{D}$ with maximal mesh-size h and the estimate in (1.3.22) requires no further assumption on the domain $\mathcal{D}$ (other than convexity) This follows from, for example, [**103**, Lemma 1.1], together with elliptic regularity (1.3.2). For $r > 2$, if $V^r_{h,0}$ is the space of continuous piecewise polynomials of degree at most $r-1$ on a triangulation of $\mathcal{D}$ with maximal mesh-size h the estimate (1.3.22) firstly requires extra assumptions on the domain $\mathcal{D}$, namely small enough interior angles in case of a polygon, or smooth enough boundary in case of a curved boundary. The reason is that to achieve $r > 2$ one needs an elliptic regularity estimate $\|u\|_{H^r} \le C\|\Lambda u\|_{H^{r-2}}$ with $r > 2$, where H^r denotes the standard Sobolev space of order r with norm $\|\cdot\|_{H^r}$, instead of the simplest elliptic regularity estimate corresponding to $r = 2$, namely, (1.3.2). Furthermore, if the boundary is curved, then the triangulation is not exact and one has to be more precise while approximating near the boundary and hence one needs special elements. For example, (1.3.22) holds for $r = 4$ for bounded convex domains with smooth boundary and $V^r_{h,0}$ consisting of continuous piecewise cubic polynomials using special, so-called isoparametric elements, near the boundary, see [**103**, Chapter 1].

Furthermore, we define the discrete Laplacian $\Lambda_h \colon V^r_{h,0} \to V^r_{h,0}$ by

$$\langle \Lambda_h \eta, \chi\rangle = \langle \nabla\eta, \nabla\chi\rangle, \quad \forall \eta, \chi \in V^r_{h,0}. \tag{1.3.23}$$

The homogeneous spatially semidiscrete wave equation with solution

$$u_h(t) := [u_{h,1}(t), u_{h,2}(t)]^T \in V^r_{h,0} \times V^r_{h,0}$$

reads as

$$\begin{bmatrix} \dot u_{h,1} \\ \dot u_{h,2} \end{bmatrix} + \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix} \begin{bmatrix} u_{h,1}(t) \\ u_{h,2}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},\ t > 0; \quad \begin{bmatrix} u_{h,1}(0) \\ u_{h,2}(0) \end{bmatrix} = \begin{bmatrix} P_{h,1}u_{0,1} \\ P_{h,2}u_{0,2} \end{bmatrix}. \tag{1.3.24}$$

Here $P_{h,1} \colon \dot H^0 \to V^r_{h,0}$ and $P_{h,2} \colon \dot H^{-1} \to V^r_{h,0}$ are the orthogonal projectors defined by $\langle P_{h,1}f, \chi\rangle = \langle f, \chi\rangle$, $\forall \chi \in V^r_{h,0}$, for $f \in \dot H^0$ and $\langle P_{h,2}f, \chi\rangle = \langle f, \chi\rangle_{\dot H^{-1}\times \dot H^1}$, $\forall \chi \in V^r_{h,0}$, for $f \in \dot H^{-1}$.

It is well known that Λ_h has eigenpairs $((\phi_{h,j}, \lambda_{h,j}))_{j=1}^{N_h}$, where $(\lambda_{h,j})_{j=1}^{N_h}$ is a positive, nondecreasing sequence and $(\phi_{h,j})_{j=1}^{N_h}$ an $\dot H^0$-orthonormal basis of $V^r_{h,0}$. If we write

$$A_h := \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix} \tag{1.3.25}$$

and if $P_h = [P_{h,1}, P_{h,2}]^T$ and $u_0 := [u_{0,1}, v_{0,2}]^T$, then (1.3.24) may be written

$$\dot u_h + A_h u_h = 0,\ t > 0; \quad u_h(0) = P_h u_0. \tag{1.3.26}$$

The operator $-A_h$ is the infinitesimal generator of a strongly continuous semigroup $E_h(t)$ and the solution of (1.3.26) is given by

$$u_h(t) = E_h(t)P_h u_0.$$

Similarly to (1.3.5) the operator $E_h(t)$ has a representation

$$E_h(t) = \begin{bmatrix} C_h(t) & \Lambda_h^{-1/2}S_h(t) \\ -\Lambda_h^{1/2}S_h(t) & C_h(t) \end{bmatrix} \tag{1.3.27}$$

with $S_h(t) = \sin(t\Lambda_h^{1/2})$ and $C_h(t) = \cos(t\Lambda_h^{1/2})$ defined by the spectral calculus of Λ_h.

The time discretization, as in the previous subsection, is performed by I-stable rational single step schemes; i.e., schemes where the rational function R fullfills (1.3.16) for some positive integer p. The fully discrete problem on the same uniform grid as in Subsection 1.3.2 then reads

$$v^n_{h,\Delta t} = R(\Delta t A_h) v^{n-1}_{h,\Delta t},\ n = 1, \dots, N; \quad v^0_{h,\Delta t} = P_h u_0. \tag{1.3.28}$$

We will henceforth write $E_{h,\Delta t} = R(\Delta t A_h)$ and the solution of (1.3.28) may then be written as

$$v^n_{h,\Delta t} = E^n_{h,\Delta t} P_h u_0.$$

The error estimate proved in [**16**] is as follows. It provides only a bound for the first component in u, which we express by means of a projector P^1.

Theorem 1.3.11. *If $P^1 \colon H \to \dot{H}^0$ is defined as $P^1 x = x_1$ for $x = [x_1, x_2]^T \in H$, then*

$$\|P^1(E^n_{h,\Delta t} P_h - E(t_n))u_0\|_{\dot{H}^0} \leqslant C(t_n)\big(h^r \|u_0\|_{\mathcal{H}^{r+1}} + \Delta t^p \|u_0\|_{\mathcal{H}^{p+1}}\big), \qquad t_n = n\Delta t \geq 0.$$

We note that Theorem 1.3.11 in [**16**] is proved for domains with smooth boundary and subspaces $V^r_{h,0}$ satisfying certain standard approximation properties. The proofs, however, remain valid under abstract assumption (1.3.22) on the Ritz projection and hence we posed it as our abstract simple assumption on the spaces $V^r_{h,0}$ although does not appear explicitly in our error analysis.

Using the stability of $E(t)$ and $E^n_{h,\Delta t}$ and a standard interpolation argument, this results in the following bound on the error operator.

Corollary 1.3.12. *Under the assumptions of Theorem 1.3.11 we have, for $\beta \geqslant 0$,*

$$\|P^1(E^n_{h,\Delta t} P_h - E(t_n))\|_{\mathcal{L}(\mathcal{H}^\beta, \dot{H}^0)} \leqslant C(t_n)\big(h^{\min(\beta\frac{r}{r+1}, r)} + \Delta t^{\min(\beta\frac{p}{p+1}, p)}\big), \qquad t_n = n\Delta t \geq 0.$$

We return to the stochastic wave equation whose fully discrete version now reads, with $B_h := P_h B = [0, P_{h,2}]^T$, as

$$X^j_{h,\Delta t} = E_{h,\Delta t}(X^{j-1}_{h,\Delta t} + B_h \Delta W^j),\ j = 1, \dots, N; \quad X^0_{h,\Delta t} = P_h X_0. \tag{1.3.29}$$

The solution is given by

$$X^n_{h,\Delta t} = E^n_{h,\Delta t} P_h X_0 + \sum_{j=1}^{n} E^{n-j+1}_{h,\Delta t} B_h \Delta W^j. \tag{1.3.30}$$

As in the previous section we multiply by $E^{N-n}_{h,\Delta t}$ and arrive at the drift free version

$$Y^n_{h,\Delta t} = E^N_{h,\Delta t} P_h X_0 + \sum_{j=1}^{n} E^{N-j+1}_{h,\Delta t} B_h W^j$$

and with piecewice constant interpolation

$$\tilde{Y}_{h,\Delta t}(t) = \tilde{E}_{h,\Delta t}(T) P_h X_0 + \int_0^t \tilde{E}_{h,\Delta t}(T-s) B_h \,\mathrm{d}W(s) \tag{1.3.31}$$

in exact analogy with the temporally semidiscrete case in (1.3.19).

Next we bound the weak error for fully discrete schemes given by (1.3.29). We only prove a result for the first component in X.

THEOREM 1.3.13. *Assume that* $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ *and suppose that the initial data satisfies* $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathcal{H}^{2\beta})$ *for some* $\beta \geqslant 0$. *If* $X^n_{h,\Delta t}$ *is given by* (1.3.29) *and* $(X(t))_{t\geqslant 0}$ *is the weak solution given by* (1.3.7) *of* (1.3.6) *with* A, B *as in* (1.3.4)*, then for* $g \in C^2(\dot H^0,\mathbb{R})$ *with* $g'' \in C_b(\dot H^0, \mathcal{L}(\dot H^0))$*, we have*

$$\big|\mathbb{E}\big(g(X^N_{h,\Delta t,1}) - g(X_1(T))\big)\big| \leqslant C(T)\big(h^{\min(2\beta\frac{r}{r+1},r)} + \Delta t^{\min(2\beta\frac{p}{p+1},1)}\big).$$

PROOF. The process in (1.3.31) is clearly of the form (1.2.1) with $\tilde Y(0) = \tilde E_{h,\Delta t}(T)P_hX_0$ and $\tilde Y_{h,k}(T) = X^N_{h,k}$. Furthermore,

$$\mathrm{Tr}(\tilde E_{h,\Delta t}(t)B_hQB_h^*\tilde E_{h,\Delta t}(t)^*) \leq \mathrm{Tr}(B_hQB_h) < \infty,$$

as B_hQB_h is a bounded operator with finite-dimensional range and hence it is of trace class. Therefore, $\tilde E_{h,\Delta t}B_h \in L^2((0,T);\mathcal{L}_2(U_0,H))$ holds and Theorem 1.3.6 can be applied with $V = \dot H^0$, $L = P^1$ (as defined in Theorem 1.3.11), $\tilde E(t) = \tilde E_{h,k}(t)$, $\tilde B = B_h$ and $\tilde P = P_h$. From Corollary 1.3.12 and Lemma 1.3.5, as in the proof of Lemma 1.3.8, it follows that

$$\sup_{t\in[0,T]} \|P^1(\tilde E_{h,\Delta t}(t)P_h - E(t))\|_{\mathcal{L}(\mathcal{H}^{2\beta},\dot H^0)} \leqslant C(T)\left(h^{\min(2\beta\frac{r}{r+1},r)} + \Delta t^{\min(2\beta\frac{p}{p+1},1)}\right).$$

Finally the statement of the theorem is immediate by inserting this into (1.3.14). □

1.3.4. Strong convergence of fully discrete schemes. It is a general phenomenon that the order of weak convergence is twice the strong order under the same regularity assumption on the noise. This essentially turns out to be the case also for the stochastic wave equation discretized by the method described in the previous section. We remark that the strong convergence is studied for a spatially semidiscrete finite element method in [**66**], for a fully discrete leap-frog scheme in one spatial dimension in [**105**], and for a spatially semidiscrete scheme in one dimension in [**97**].

First we form the strong error by taking the difference of (1.3.30) and (1.3.7), projecting onto the first component, and taking norms:

$$\begin{aligned}\mathbb{E}\Big(\|P^1(X^N_{h,\Delta t} - X(T))\|^2_{\dot H^0}\Big) &\leqslant C\mathbb{E}\Big(\|P^1(E^N_{h,\Delta t}P_h - E(T))X_0\|^2_{\dot H^0}\Big)\\ &+ C\mathbb{E}\Big(\Big\|P^1\int_0^T (\tilde E_{h,\Delta t}(T-s)P_h - E(T-s))B\,\mathrm{d}W(s)\Big\|^2_{\dot H^0}\Big) =: I_1 + I_2.\end{aligned}$$

If $X_0 \in L^2(\Omega;\mathcal{H}^\beta)$, then

$$\begin{aligned} I_1 &\leqslant C\|P^1(E_{h,\Delta t}P_h - E(T))\|^2_{\mathcal{L}(\mathcal{H}^\beta,\dot H^0)}\mathbb{E}\big(\|X_0\|^2_{\mathcal{H}^\beta}\big)\\ &\leqslant C(T)\big(h^{\min(\beta\frac{r}{r+1},r)} + \Delta t^{\min(\beta\frac{p}{p+1},p)}\big)^2\|X_0\|^2_{L^2(\Omega;\mathcal{H}^\beta)}\end{aligned}$$

by Corollary 1.3.12. For I_2 we use Itô's isometry (1.1.14) to get

$$
\begin{aligned}
I_2 &= \mathbb{E}\Big(\Big\| \int_0^T P^1(\tilde{E}_{h,\Delta t}(T-s)P_h - E(T-s))B\,\mathrm{d}W(s)\Big\|_{\dot{H}^0}^2\Big) \\
&= \int_0^T \|P^1(\tilde{E}_{h,\Delta t}(T-s)P_h - E(T-s))BQ^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s \\
&= \int_0^T \|P^1(\tilde{E}_{h,\Delta t}(T-s)P_h - E(T-s))B\Lambda^{\frac{1-\beta}{2}}\Lambda^{\frac{\beta-1}{2}}Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s \qquad (1.3.32)\\
&\leqslant \int_0^T \|P^1(\tilde{E}_{h,\Delta t}(T-s)P_h - E(T-s))\|_{\mathcal{L}(\mathcal{H}^\beta,\dot{H}^0)}^2 \|\Lambda^{\frac{\beta-1}{2}}Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s \\
&\leqslant T \sup_{t\in[0,T]} \big(\|P^1(\tilde{E}_{h,\Delta t}(t)P_h - E(t))\|_{\mathcal{L}(\mathcal{H}^\beta,\dot{H}^0)}^2\big)\|\Lambda^{\frac{\beta-1}{2}}Q^{1/2}\|_{\mathrm{HS}}^2 \\
&\leqslant C(T)\|\Lambda^{\frac{\beta-1}{2}}Q^{1/2}\|_{\mathrm{HS}}^2\big(h^{\min(\beta\frac{r}{r+1},r)} + \Delta t^{\min(\beta\frac{p}{p+1},1)}\big)^2,
\end{aligned}
$$

where the first inequality follows from the fact that $\|B\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot{H}^0,\mathcal{H}^\beta)} = 1$ combined with (1.1.4), and the last inequality from Corollary 1.3.12 and Lemma 1.3.5 as in the proof of Lemma 1.3.8. Combining the bounds for I_1 and I_2 and taking square roots, we have shown the following result.

THEOREM 1.3.14. *Let* $\|\Lambda^{\frac{\beta-1}{2}}Q^{1/2}\|_{\mathrm{HS}}^2 < \infty$ *and let* $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\mathcal{H}^\beta)$ *for some* $\beta \geq 0$. *Then the strong error of the approximation* $X_{h,\Delta t,1}^N = P^1X_{h,\Delta t}^N$ *of the displacement* $X_1(T) = P^1X(T)$ *in the stochastic wave equation is bounded by*

$$\|X_{h,\Delta t,1}^N - X_1(T)\|_{L^2(\Omega;\dot{H}^0)} \leqslant C(T)\big(h^{\min(\beta\frac{r}{r+1},r)} + \Delta t^{\min(\beta\frac{p}{p+1},1)}\big).$$

The regularity assumption on Q in Theorem 1.3.13 implies the assumption in Theorem 1.3.14, see Theorem 1.1.1 with $s = \beta - 1$ in (1.1.9). Thus the claim that the weak rate is essentially twice the strong rate is justified (if β is not too large) by comparing Theorems 1.3.13 and 1.3.14. Note also that the mean-square regularity of the solution is of order β according to Theorem 1.3.3.

1.4. Stochastic equations of parabolic type

Here, we give a detailed weak error analysis of a fully discrete scheme for the linearized Cahn–Hilliard–Cook (CHC) equation and also comment on the linear stochastic heat equation.

The linearized CHC equation, see [**72**], is

$$\mathrm{d}X + \Lambda^2 X\,\mathrm{d}t = \mathrm{d}W,\ t > 0; \quad X(0) = X_0, \qquad (1.4.1)$$

where now $\Lambda = -\Delta$ is the Laplacian together with homogeneous Neumann boundary conditions. To write the CHC equation in the form (1.3.6) we therefore set $H = \{f \in L^2(\mathcal{D}) : \langle f, 1\rangle = 0\}$, $A = \Lambda^2$ with $D(\Lambda) = \{f \in H^2(\mathcal{D}) \cap H : \frac{\partial f}{\partial n} = 0\}$, where $\mathcal{D}$ is a bounded convex domain, and we take $U = H$ and $B = I$. We further define the the spaces $\dot{H}^\alpha = D(\Lambda^{\alpha/2})$ in analogy with Section 1.3. Thus, $H = \dot{H}^0$, $D(A) = \dot{H}^4$ and $-A$ is known to be the infinitesimal generator of the analytic semigroup $E(t) = \mathrm{e}^{-tA} = \mathrm{e}^{-t\Lambda^2}$ on H.

Let $V_h^r \subset H^1(\mathcal{D})$, $r = 2, 3$, be finite dimensional subspaces and set $\dot{V}_h^r = \{v \in V_h^r : \langle v, 1\rangle = 0\}$. We now define the discrete Laplacian $\Lambda_h \colon \dot{V}_h^r \to \dot{V}_h^r$ and the Ritz projector $R_h \colon \dot{H}^1 \to \dot{V}_h^r$ in an analogous way as in Subection 1.3.3 and we assume an error bound of the same form as in (1.3.22). We set $A_h = \Lambda_h^2$ and note that $-A_h$ is the generator of an analytic semigroup $E_h(t)$ on $\dot{V}_h^r$. We consider only the backward Euler time-stepping and therefore introduce $E_{h,\Delta t} = (1+\Delta t A_h)^{-1}$ and define $\tilde{E}_{h,\Delta t}(t)$ in an analogous fashion to the case of the wave equation, see (1.3.19).

We need error bounds for the approximation of the semigroup. We claim that, for all $v \in H$,

$$\|(E_{h,\Delta t}^n P_h - E(t_n))v\| \leqslant C(h^\alpha + \Delta t^{\alpha/4}) t_n^{-\alpha/4}\|v\|, \quad t_n = n\Delta t,\ \alpha \in [0,r], \tag{1.4.2}$$

where $P_h \colon H \to \dot{V}_h^r$ denotes the $L^2(\mathcal{D})$-orthogonal projection to $\dot{V}_h^r$. To see this we write

$$E_{h,\Delta t}^n P_h v - E(t_n)v = \big(E_{h,\Delta t}^n P_h v - E_h(t_n)P_h v\big) + \big(E_h(t_n)P_h v - E(t_n)v\big).$$

It is well known and follows by a simple spectral argument, as A_h is self-adjoint positive semidefinite on $\dot{S}_h^r$, that the estimate

$$\|E_{h,\Delta t}^n P_h v - E_h(t_n)P_h v\| \leqslant C\Delta t^\gamma t_n^{-\gamma}\|v\|, \quad \gamma \in [0,1], \tag{1.4.3}$$

holds for the backward Euler method [**74**]. It follows from the stability of the Galerkin approximation and [**39**, Corollary 5.3] that

$$\|E_h(t)P_h v - E(t)v\| \leqslant C h^\gamma t^{-\gamma/4}\|v\|, \quad \gamma \in [0,r]. \tag{1.4.4}$$

Thus, with $\gamma = \alpha/4 \le r/4 \le 1$ in (1.4.3) and $\gamma = \alpha$ in (1.4.4), the estimate (1.4.2) follows.

It is also well known (see, for example, [**87**, Theorem 6.13]) that

$$\|(E(t) - E(s))A^{-\gamma}v\| \leqslant |t-s|^\gamma \|v\|, \quad \gamma \in [0,1], \tag{1.4.5}$$

and therefore, taking also (1.4.2) into account, it follows that

$$\|(\tilde{E}_{h,\Delta t}(t)P_h - E(t))v\| \leqslant C(h^\alpha + \Delta t^{\alpha/4})t^{-\alpha/4}\|v\|, \quad \alpha \in [0,r]. \tag{1.4.6}$$

Indeed, for $t \in (t_{j-1}, t_j]$ we have that

$$\begin{aligned}\|(\tilde{E}_{h,\Delta t}(t)P_h - E(t))v\| &= \|(E_{h,\Delta t}^j P_h - E(t))v\| \\ &\leqslant \|(E_{h,\Delta t}^j P_h - E(t_j))v\| + \|(E(t_j) - E(t))v\|.\end{aligned}$$

For the first term (1.4.2) applies and for the second term we use (1.4.5) to get

$$\begin{aligned}\|(E(t_j) - E(t))v\| &= \|A^{\alpha/4}E(t)(E(t_j - t) - I)A^{-\alpha/4}v\| \\ &\leqslant \|A^{\alpha/4}E(t)\|\|(E(t_j-t) - I)A^{-\alpha/4}v\| \leqslant C\Delta t^{\alpha/4}t^{-\alpha/4}\|v\|.\end{aligned}$$

Finally, we recall the smoothing property of the backward Euler scheme. It follows from [**103**, Lemma 7.3] by stability and interpolation that for $t \in (t_{j-1}, t_j]$,

$$\|A_h^\alpha \tilde{E}_{h,\Delta t}(t)P_h v\| = \|A_h^\alpha E_{h,\Delta t}^j P_h v\| \leqslant C t_j^{-\alpha}\|P_h v\| \leqslant C t^{-\alpha}\|v\|, \quad \alpha \ge 0.$$

Therefore,

$$\|A_h^\alpha \tilde{E}_{h,\Delta t}(t)P_h v\| \leqslant C t^{-\alpha}\|v\|, \quad \alpha \ge 0,\ t > 0. \tag{1.4.7}$$

We are now in the position to prove the following estimate for the weak error in case of the linearized CHC equation. As it was the case for the wave equation, the weak convergence rate is twice that of the strong convergence rate [**72**] (up to a logarithmic factor) under essentially the same regularity requirements on A and Q.

THEOREM 1.4.1. *Let X be the solution of* (1.3.6) *and $X^n_{h,\Delta t}$ be given by* (1.3.30) *with spaces and operators described above and $B_h = P_h$. Assume that $g \in C^2(H,\mathbb{R})$ with $g'' \in C_b(H,\mathcal{L}(H))$, $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$ and that*

$$\begin{aligned} \|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}} &= \|A^{(\beta-2)/2}Q\|_{\mathrm{Tr}} \leqslant K, \\ \|\Lambda_h^{\beta-2}P_hQ\|_{\mathrm{Tr}} &= \|A_h^{(\beta-2)/2}P_hQ\|_{\mathrm{Tr}} \leqslant K, \end{aligned} \tag{1.4.8}$$

for some $\beta \in (0,\frac{r}{2}]$ and $K>0$. Then there is $C>0$ depending on T, K, X_0, and g such that, for $N\Delta t = T$, $h^4 + \Delta t < T$,

$$\big|\mathbb{E}(g(X^N_{h,\Delta t}) - g(X(T)))\big| \leqslant C\big(h^{2\beta} + \Delta t^{\beta/2}\big)\log(\tfrac{T}{h^4+\Delta t}). \tag{1.4.9}$$

PROOF. Assumption (1.4.8) guarantees that $\|A^{\frac{\beta-2}{4}}Q^{\frac12}\|_{\mathrm{HS}} < \infty$ in view of Theorem 1.1.1. This in its turn implies that the weak solution X of (1.4.1) exists, as shown in [**72**] and is given by the variation of constants formula

$$X(t) = E(t)X_0 + \int_0^t E(t-s)\,\mathrm{d}W(s).$$

In particlular, $E \in L^2((0,T);\mathcal{L}_2(U_0,H))$ and

$$\|X(t)\|_{L^2(\Omega;H)} \leqslant C(\|X_0\|_{L^2(\Omega;H)} + \|A^{-\frac12}Q^{\frac12}\|_{\mathrm{HS}}), \quad t \geq 0. \tag{1.4.10}$$

We will use Theorem 1.2.1 with $\mathcal{O}$ as in (1.2.4) and $G := g$. Furthermore, let $\tilde{B} = P_h, \tilde{X}_0 = P_hX_0$, $\tilde{E}(t) = \tilde{E}_{h,\Delta t}(t)$ and $\tilde{Y}_{h,\Delta t}(t)$ be defined as in (1.3.31), whence $\tilde{Y}_{h,\Delta t}(T) = X^N_{h,\Delta t}$. The use of Theorem 1.2.1 is justified since the operator $\tilde{E}_{h,\Delta t}(t)P_hQ[\tilde{E}_{h,\Delta t}(t)P_h]^*$ is uniformly bounded in the operator norm and is of finite rank so that $\tilde{E}\tilde{B} \in L^2((0,T);\mathcal{L}_2(U_0,H))$.

We write $\tilde{F}_{h,\Delta t}(t) = \tilde{E}_{h,\Delta t}(t)P_h - E(t)$ and recall (1.4.6). For the first term in (1.2.3) we use that $\tilde{Y}_{h,\Delta t}(0) - Y(0) = \tilde{F}_{h,\Delta t}(T)X_0$, (1.4.6) and (1.4.10) to get, in a similar fashion as in the proof of Theorem 1.3.6,

$$\begin{aligned} &\Big|\mathbb{E}\int_0^1 \Big\langle u_x\big(0, Y(0) + \theta(\tilde{Y}_{h,\Delta t}(0) - Y(0))\big), \tilde{Y}_{h,\Delta t}(0) - Y(0)\Big\rangle\,\mathrm{d}\theta\Big| \\ &\leqslant \int_0^1 \big\|g'\big(Z\big(T,0,Y(0) + \theta(\tilde{Y}_{h,\Delta t}(0) - Y(0))\big)\big)\big\|_{L^2(\Omega;H)}\mathrm{d}\theta\,\|\tilde{F}_{h,\Delta t}(T)X_0\|_{L^2(\Omega;H)} \\ &\leqslant C\big(1 + \int_0^1 \big\|Z\big(T,0,Y(0) + \theta(\tilde{Y}_{h,\Delta t}(0) - Y(0))\big)\big\|_{L^2(\Omega;H)}\mathrm{d}\theta\big) \\ &\qquad \times (h^{2\beta} + \Delta t^{\beta/2})\,T^{-\beta/2}\,\|X_0\|_{L^2(\Omega;H)} \\ &\leqslant C\big(1 + \|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}} + \|X_0\|_{L^2(\Omega;H)}\big)\,\|X_0\|_{L^2(\Omega;H)}\,T^{-\beta/2}\,(h^{2\beta} + \Delta t^{\beta/2}). \end{aligned} \tag{1.4.11}$$

For the second term of (1.2.3) we have by (1.1.3) and repeated use of (1.1.4) that

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}\big(t,\tilde Y_{h,\Delta t}(t)\big)(\tilde E_{h,\Delta t}(T-t)P_h + E(T-t))Q\tilde F_{h,\Delta t}(T-t)^*\Big)\,\mathrm{d}t\Big|\\
&\quad\leqslant \mathbb{E}\Big(\int_0^T \|u_{xx}(t,\tilde Y_{h,\Delta t}(t))(\tilde E_{h,\Delta t}(T-t)P_h\\
&\qquad\qquad + E(T-t))Q\|_{\mathrm{Tr}}\|\tilde F_{h,\Delta t}(T-t)\|_{\mathcal{L}(\mathcal{H})}\,\mathrm{d}t\Big)\\
&\quad\leqslant \sup_{(x,t)\in H\times[0,T]}\|u_{xx}(t,x)\|_{\mathcal{L}(H)}\\
&\qquad\times\int_0^T \|(A_h^{-(\beta-2)/2}\tilde E_{h,\Delta t}(t)A_h^{(\beta-2)/2}P_h + A^{-(\beta-2)/2}E(t)A^{(\beta-2)/2})Q\|_{\mathrm{Tr}}\\
&\qquad\times\|\tilde F_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t\\
&\quad\leqslant \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\int_0^T\Big(\|A_h^{-(\beta-2)/2}\tilde E_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\|A_h^{(\beta-2)/2}P_hQ\|_{\mathrm{Tr}}\\
&\qquad+\|A^{-(\beta-2)/2}E(t)\|_{\mathcal{L}(H)}\,\|A^{(\beta-2)/2}Q\|_{\mathrm{Tr}}\Big)\,\|\tilde F_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t\\
&\quad\leqslant \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\Big(\|A_h^{(\beta-2)/2}P_hQ\|_{\mathrm{Tr}}+\|A^{(\beta-2)/2}Q\|_{\mathrm{Tr}}\Big)\\
&\qquad\times\int_0^T\Big(\|A_h^{-(\beta-2)/2}\tilde E_{h,\Delta t}(t)\|_{\mathcal{L}(H)}+\|A^{-(\beta-2)/2}E(t)\|_{\mathcal{L}(H)}\Big)\|\tilde F_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t.
\end{aligned}$$

By (1.4.8) the factors in front of the integral are bounded by

$$2K\sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}.$$

We proceed by splitting the integral in two as $\int_0^T = \int_0^{h^4+\Delta t} + \int_{h^4+\Delta t}^T$. For the first integral we notice that the last factor of the integrand is uniformly bounded and hence, by the analyticity of $E(t)$ and (1.4.7) with $\alpha = -(\beta-2)/2$,

$$\begin{aligned}
&\int_0^{h^4+\Delta t}\Big(\|A_h^{-(\beta-2)/2}\tilde E_{h,\Delta t}(t)\|_{\mathcal{L}(H)}+\|A^{-(\beta-2)/2}E(t)\|_{\mathcal{L}(H)}\Big)\|\tilde F_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t\\
&\qquad\leqslant C\int_0^{h^4+\Delta t} t^{(\beta-2)/2}\,\mathrm{d}t = C(h^4+\Delta t)^{\beta/2}\le C(h^{2\beta}+\Delta t^{\beta/2}).
\end{aligned}$$

For the second part we use again the analyticity of $E(t)$, (1.4.6) with $\alpha = 2\beta$ and (1.4.7) to get

$$\begin{aligned}
&\int_{h^4+\Delta t}^T\Big(\|A_h^{-(\beta-2)/2}\tilde E_{h,\Delta t}(t)\|_{\mathcal{L}(H)}+\|A^{-(\beta-2)/2}E(t)\|_{\mathcal{L}(H)}\Big)\|\tilde F_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t\\
&\qquad\leqslant C\int_{h^4+\Delta t}^T\big(t^{(\beta-2)/2}+t^{(\beta-2)/2}\big)\big(h^{2\beta}+\Delta t^{\beta/2}\big)t^{-\beta/2}\,\mathrm{d}t\\
&\qquad= C(h^{2\beta}+\Delta t^{\beta/2})\int_{h^4+\Delta t}^T t^{-1}\,\mathrm{d}t = C\log\big(\tfrac{T}{h^4+\Delta t}\big)(h^{2\beta}+\Delta t^{\beta/2}).
\end{aligned}$$

□

REMARK 1.4.2. The dependence on T of C in (1.4.9) can be removed if we assume that $X_0 \in L^2(\Omega; \dot{H}^{2\beta})$ by using the deterministic error estimate for smooth initial data from [**72**] in (1.4.11).

Next we discuss some h-independent conditions guaranteeing (1.4.8).

THEOREM 1.4.3. *The following conditions imply* (1.4.8).

(i) Assume that $Q = I$, $0 < \beta \le \frac{r}{2}$, *and* $\|\Lambda^{\beta-2}\|_{\mathrm{Tr}} < \infty$.
(ii) Assume that $r = 3$, $\beta = \frac{3}{2}$, *and* $\|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}} < \infty$.
(iii) Assume that V_h^r *is based on a quasi-uniform mesh family and that, for some* $\alpha > 0$, *we have* $0 < \beta \le \frac{r}{2}$, $0 \le \beta - 2 + \alpha \le 1$ *and*

$$\|\Lambda^{\beta-2+\alpha}Q\|_{\mathcal{L}(H)} < \infty, \quad \|\Lambda^{-\alpha}\|_{\mathrm{Tr}} < \infty. \tag{1.4.12}$$

PROOF. (i) The eigenvalues of Λ_h and Λ satisfy $\lambda_{h,j} \ge \lambda_j$ and $\|P_h\|_{\mathcal{L}(H)} \le 1$, so that

$$\|\Lambda_h^{-\alpha}P_h\|_{\mathrm{Tr}} \le \|\Lambda_h^{-\alpha}\|_{\mathrm{Tr}} = \sum_{j=1}^{N_h} \lambda_{h,j}^{-\alpha} \le \sum_{j=1}^{\infty} \lambda_j^{-\alpha} = \|\Lambda^{-\alpha}\|_{\mathrm{Tr}}, \quad \alpha \ge 0. \tag{1.4.13}$$

With $\alpha = 2 - \beta \ge 0$ and $Q = I$ we obtain

$$\|\Lambda_h^{\beta-2}P_hQ\|_{\mathrm{Tr}} \le \|\Lambda_h^{\beta-2}\|_{\mathrm{Tr}} \le \|\Lambda^{\beta-2}\|_{\mathrm{Tr}} = \|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}}, \quad 0 < h \le 1.$$

For case (ii) we use the fact that

$$\|\Lambda_h^{-\frac{1}{2}}P_h\Lambda^{\frac{1}{2}}\| \le C.$$

This follows by using $\|\Lambda_h^{\frac{1}{2}}w_h\| = \|\Lambda^{\frac{1}{2}}w_h\|$ for $w_h \in \dot{V}_h^r$ in the calculation

$$\begin{aligned}\|\Lambda_h^{-\frac{1}{2}}P_hf\| &= \sup_{v_h\in\dot{V}_h^r} \frac{|\langle \Lambda_h^{-\frac{1}{2}}P_hf, v_h\rangle|}{\|v_h\|} = \sup_{v_h\in\dot{V}_h^r} \frac{|\langle f, \Lambda_h^{-\frac{1}{2}}v_h\rangle|}{\|v_h\|} = \sup_{w_h\in\dot{V}_h^r} \frac{|\langle f, w_h\rangle|}{\|\Lambda_h^{\frac{1}{2}}w_h\|} \\ &= \sup_{w_h\in\dot{V}_h^r} \frac{|\langle f, w_h\rangle|}{\|\Lambda^{\frac{1}{2}}w_h\|} \le \sup_{v\in H} \frac{|\langle f, \Lambda^{-\frac{1}{2}}v\rangle|}{\|v\|} = \left\|\Lambda^{-\frac{1}{2}}f\right\|.\end{aligned}$$

Hence with $\beta = \frac{3}{2}$, we have

$$\begin{aligned}\|\Lambda_h^{\beta-2}P_hQ\|_{\mathrm{Tr}} &= \|\Lambda_h^{\beta-2}P_h\Lambda^{2-\beta}\Lambda^{\beta-2}Q\|_{\mathrm{Tr}} \\ &\le \|\Lambda_h^{-(2-\beta)}P_h\Lambda^{2-\beta}\|_{\mathcal{L}(H)}\|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}} \le C\|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}}.\end{aligned}$$

Finally, for case (iii) we first note that Theorem 1.1.1 shows that (1.4.12) implies that $\|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}} < \infty$. For quasi-uniform mesh families we have the inverse inequality $\|\nabla v_h\| \le Ch^{-1}\|v_h\|$, $v_h \in \dot{V}_h^r$, so that

$$\|\Lambda_h\|_{\mathcal{L}(H)} = \max_{1\le j\le N_h} \lambda_{h,j} = \max_{v_h\in\dot{V}_h^r} \frac{\|\nabla v_h\|^2}{\|v_h\|^2} \le Ch^{-2}.$$

Hence, using also $R_h = \Lambda_h^{-1}P_h\Lambda$ and (1.3.22), we get

$$\begin{aligned}\|\Lambda_hP_h\Lambda^{-1}f\| &= \|\Lambda_hP_h\Lambda^{-1}f - \Lambda_h\Lambda_h^{-1}P_h\Lambda\Lambda^{-1}f + P_hf\| \\ &\le \|\Lambda_hP_h(I - \Lambda_h^{-1}P_h\Lambda)\Lambda^{-1}f\| + \|P_hf\| \\ &\le Ch^{-2}\|(I - R_h)\Lambda^{-1}f\| + \|f\| \\ &\le Ch^{-2}Ch^2\|f\| + \|f\| \le C\|f\|.\end{aligned}$$

dc_1541_18

We conclude

$$\|\Lambda_h^{\delta} P_h \Lambda^{-\delta}\|_{\mathcal{L}(H)} \leq C, \quad 0 \leq \delta \leq 1.$$

With $\delta = \beta - 2 + \alpha \in [0,1]$ and (1.4.13) we obtain

$$\begin{aligned}\|\Lambda_h^{\beta-2} P_h Q\|_{\mathrm{Tr}} &\leq \|\Lambda_h^{-\alpha}\|_{\mathrm{Tr}} \|\Lambda_h^{\beta-2+\alpha} P_h \Lambda^{-(\beta-2+\alpha)}\|_{\mathcal{L}(H)} \|\Lambda^{\beta-2+\alpha} Q\|_{\mathcal{L}(H)} \\ &\leq C\|\Lambda^{-\alpha}\|_{\mathrm{Tr}} \|\Lambda^{\beta-2+\alpha} Q\|_{\mathcal{L}(H)}.\end{aligned}$$

This completes the proof. □ □

Finally, we comment on the cases (i) and (iii) in the previous theorem.

(i) The asymptotics $\lambda_j \sim j^{2/d}$, $j \to \infty$, shows that $\|\Lambda^{\beta-2}\|_{\mathrm{Tr}} < \infty$, if $\beta < 2 - \frac{d}{2}$.

(iii) As mentioned in (i) above, we have $\|\Lambda^{-\alpha}\|_{\mathrm{Tr}} < \infty$ if $\alpha > \frac{d}{2}$ and hence it is possible to choose $\beta \in (0, 3 - \frac{d}{2})$. In particular, we may have $\beta = 1$ for $d = 1,2,3$ and thus for $r = 2$ the (almost) optimal order can be achieved in this case.

Remark 1.4.4. The weak convergence of the finite element space discretization and backward Euler time discretization of stochastic heat equation with additive noise was considered in [**34**]. The results there can be recovered using the well-known deterministic estimates (see also Lemma 3.3.3 for a piecewise constant time-interpolated version)

$$\|E_{h,\Delta t}^n P_h - E(t_n)\|_{\mathcal{L}(H)} \leq C(h^2 + \Delta t) t_n^{-1}$$

and

$$\|A^{\delta}(E_{h,\Delta t}^n P_h \pm E(t_n))\|_{\mathcal{L}(H)} \leq C t_n^{-\delta}, \quad \delta \in [0, \tfrac{1}{2}],$$

together with Theorem 1.2.1 where $H = L^2(\mathcal{D})$, $B = I$, $(A, D(A)) := (\Lambda, D(\Lambda))$ as defined in Subsection 1.3.1, $A_h = \Lambda_h$ is defined by (1.3.23) with $r = 2$, $E(t) = e^{-tA}$ and $E_{h,\Delta t}^n = (I + \Delta t A_h)^{-n}$. The technicalities are the same as in the spatially semidiscrete case [**61**, Theorem 4.1], or as the proof of Theorem 1.5.33 with $\rho = 1$, under the same symmetric condition

$$\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty, \quad \beta \in (0,1].$$

We do not detail this here any further as it recovers a known result:

$$|\mathbb{E}\, g(X_{h,\Delta t}^N) - \mathbb{E}\, g(X(T))| \leqslant C \ln\left(\frac{T}{h^2 + \Delta t}\right)(\Delta t^{\beta} + h^{2\beta}), \quad \beta \in (0,1], \tag{1.4.14}$$

with $g \in C^2(H, \mathbb{R})$ such that $g'' \in C_b(H, \mathcal{L}(H))$ and $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, only with perhaps a more transparent proof. The rate in (1.4.14) is twice that of the strong rate under the same assumption on Q found in [**110**], see also Theorem 2.3.2.

1.5. Stochastic Volterra integro-differential equations

As before, H denotes a real separable Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\|\cdot\|$ and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is a stochastic basis satisfying the usual conditions. We consider the stochastic Volterra equation given in the abstract Itô form as

$$\mathrm{d}X + \left(\int_0^t b(t-s) A X(s)\, \mathrm{d}s\right) \mathrm{d}t = \mathrm{d}W, \quad t \in (0,T]; \quad X(0) = X_0, \tag{1.5.1}$$

dc_1541_18

where the process $(X(t))_{t\in[0,T]}$ is an H-valued stochastic process, A is a densely defined, nonnegative self-adjoint unbounded operator on H with compact inverse and domain of definition $D(A)$, and W is a Q-Wiener process in $U := H$ with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$. We take X_0 to be an $\mathcal{F}_0$-measurable, H-valued random variable. The weak solution of (1.5.1) is a mean-square continuous H-valued process satisfying

$$\langle X(t), \eta\rangle + \int_0^t \int_0^r b(r-s)\,\langle X(s), A^*\eta\rangle\,\mathrm{d}s\,\mathrm{d}r = \langle X_0, \eta\rangle + \int_0^t \langle \eta, \mathrm{d}W(s)\rangle\,,$$

for all $\eta \in D(A^*)$ almost surely for all $t\in[0,T]$.

Having applications in mind stemming from the linear theory of viscoelasticity, the Hilbert space H is typically a space of square integrable functions on a spatial domain $\mathcal{D}$, the operator $-A$ is an elliptic differential operator like the Laplace operator, the elasticity operator or the Stokes operator together with appropriate boundary conditions. We will consider a family of convolution kernels b accordingly.

Our analysis will also use the Laplace transform. Let $f:\mathbb{R}_+\to H$ be subexponential; i.e., that for any $\varepsilon>0$ the function $t\mapsto f(t)e^{-\varepsilon t}$ belongs to $L^1(\mathbb{R}_+;H)$. We define the Laplace transform of $\widehat{f}:\mathbb{C}_+\to H$ by

$$\widehat{f}(z) = \int_0^{+\infty} f(t)e^{-zt}\,dt, \quad \mathrm{Re}\,z>0,$$

where we have used the same notation H for the complexification of H. We denote by $\star$ the Laplace convolution product on $[0,t]$ of two locally integrable subexponential functions $f,g\in L^1_{loc}(\mathbb{R}_+,H)$ defined as

$$(f\star g)(t) = \int_0^t f(t-s)g(s)\,\mathrm{d}s.$$

It is well known that $f\star g\in L^1_{loc}(\mathbb{R}_+,H)$ is subexponential and

$$\widehat{f\star g}\,(z) = \widehat{f}(z)\,\widehat{g}(z), \quad \mathrm{Re}\,z>0.$$

We make state our main assumptions on the kernel b.

Assumption 1.5.1. The kernel $0\neq b\in L^1_{\mathrm{loc}}(\mathbb{R}_+)$, is 3-monotone and

$$\lim_{t\to\infty} b(t) = 0.$$

Furthermore,

$$\limsup_{t\to 0,\infty} \frac{\frac{1}{t}\int_0^t sb(s)\,\mathrm{d}s}{\int_0^t -s\dot{b}(s)\,\mathrm{d}s} < +\infty. \tag{1.5.2}$$

It follows from [**93**, Proposition 3.10], see also [**85**], that for 3-monotone and locally integrable kernels b, condition (1.5.2) is equivalent to

$$\rho := 1 + \frac{2}{\pi}\sup\{|\arg\widehat{b}(\lambda)|,\ \mathrm{Re}\,\lambda>0\}\in(1,2). \tag{1.5.3}$$

Also note that, by (1.5.3), we have that $\mathrm{Re}\,(\widehat{b}(\lambda))\geq 0$ for $\mathrm{Re}\,\lambda>0$ and hence b is a positive definite kernel is the sense that

$$\int_0^T\int_0^t b(t-s)f(s)f(t)\,\mathrm{d}s\,\mathrm{d}t \geqslant 0, \tag{1.5.4}$$

for any continuous function f on $[0,T]$. The parameter ρ in (1.5.3) plays a vital role throughout the analysis as it quantifies the smoothing effect of the linear deterministic problem, see Propositions 1.5.6 and 1.5.9. An important example is the kernel $b(t) = \frac{1}{\Gamma(\rho-1)} t^{\rho-2} e^{-\eta t}$, $1 < \rho < 2$ and $\eta \geq 0$. When $\eta = 0$, then the corresponding equation (1.5.1) can be viewed as a fractional-in-time stochastic equation.

When considering time discretisation we make a slightly stronger assumption on b.

Assumption 1.5.2. The kernel b satisfies Assumption 1.5.1 and, in addition, b is 4-monotone.

1.5.1. Resolvent family. Under the assumptions on A and b it follows from [**93**, Corollary 1.2] that there exists a strongly continuous family $(S(t))_{t\geq 0}$ such that the function $u(t) = S(t)u_0$, $u_0 \in H$, is the unique solution of

$$u(t) + A\int_0^t B(t-s)u(s)\,\mathrm{d}s = u_0, \quad t \geq 0, \tag{1.5.5}$$

with $B(t) = \int_0^t b(s)\mathrm{d}s$. In fact, [**93**, Corollary 3.3] even shows that $t \to u(t) = S(t)u_0$ is differentiable for $t > 0$ and $u_0 \in H$ and hence it is the unique solution of

$$\dot{u}(t) + A\int_0^t b(t-s)u(s)\,\mathrm{d}s = 0, \quad t > 0; \quad u(0) = u_0. \tag{1.5.6}$$

Note that the resolvent family does not satisfy the semi-group property because of the presence of a memory term. It is well known that such assumptions on A implies the existence of a sequence of nondecreasing positive real numbers $(\lambda_k)_{k\geqslant 1}$ and an orthonormal basis $(e_k)_{k\geqslant 1}$ of H such that

$$Ae_k = \lambda_k e_k, \quad \lim_{k\to+\infty} \lambda_k = +\infty.$$

Then, the resolvent family can be written explicitly as

$$S(t)x = \sum_{k=1}^{+\infty} s_k(t)\langle x, e_k\rangle e_k, \quad , x \in H, \tag{1.5.7}$$

where the functions $s_k(\cdot) := s_{\lambda_k}(\cdot)$ are the solutions of the ordinary differential equations

$$\dot{s}_\mu(t) + \mu\int_0^t b(t-s)s_\mu(s)\,\mathrm{d}s = 0, \quad s_\mu(0) = 1, \quad \mu > 0, \tag{1.5.8}$$

with $\mu = \lambda_k$.

Remark 1.5.3. The positivity of the kernel b defined in (1.5.4), together with the positivity of the operator A already allows for the construction of a unique solution to (1.5.6) using an energy argument, see [**93**, Corollary 1.2]. Assumption 1.5.1 gives further integrability and smoothing properties for $(S(t))_{t\geq 0}$ as will be discussed in the next subsection.

Remark 1.5.4. We note that since A is nonnegative self-adjoint, $-A$ generates an analytic contraction semigroup on H. Moreover, for any $\theta < \pi$, there exists $M_\theta \geqslant 1$ such that the following holds:

$$\|(zI + A)^{-1}\|_{\mathcal{L}(H)} \leqslant \frac{M_\theta}{|z|}, \quad \text{for any } z \in \Sigma_\theta, \tag{1.5.9}$$

where $\Sigma_\theta = \{z \in \mathbb{C}\backslash\{0\},\ |\arg(z)| < \theta\}$.

1.5.2. Smoothing of the deterministic problem. To measure regularity, we define, by means of the spectral decomposition of A, the fractional powers A^s of A for $s \in \mathbb{R}$. That is, for $s \geq 0$ we set

$$A^s x = \sum_{k\geq 1} \lambda_k^s \langle x, e_k \rangle e_k, \tag{1.5.10}$$

with domain $D(A^s)$ being all $x \in H$ for which the sum converges in H. In particular, $D(A^0) = H$. For $s < 0$ we define $A^s x$ as in (1.5.10) for all $x \in H$ and $D(A^s)$ to be the completion of H with respect to the norm of $\|x\|_{2s} := \|A^s x\|$.

The next proposition summarizes the main properties of the functions $\{s_k\}_{k\geqslant 1}$ when b satisfies Assumption 1.5.1.

LEMMA 1.5.5. *Suppose that b satisfies Assumption 1.5.1 and let $\rho \in (1,2)$ as defined in (1.5.3). Then $\lim_{r\to\infty} s_k(r) = 0$ for all $k \geq 1$ and there exists $C_0 > 0$ such that for any $k \geqslant 1$,*

$$\|s_k\|_{L^\infty(\mathbb{R}_+)} \leqslant 1, \tag{1.5.11}$$

$$\|\dot{s}_k\|_{L^1(\mathbb{R}_+)} \leqslant C_0, \tag{1.5.12}$$

$$\|t\dot{s}_k\|_{L^1(\mathbb{R}_+)} \leqslant C_0\, \lambda_k^{-1/\rho}, \tag{1.5.13}$$

$$\|s_k\|_{L^1(\mathbb{R}_+)} \leqslant C_0\, \lambda_k^{-1/\rho}. \tag{1.5.14}$$

PROOF. Estimate (1.5.11) follows from [**93**, Corollary 1.2], inequalities (1.5.12) and (1.5.13) can be found in [**85**, Proposition 6] and estimate (1.5.14) is shown in [**28**, Lemma 3.1] where also the fact $\lim_{r\to\infty} s_k(r) = 0$ for all $k \geq 1$ is shown in the proof of the lemma. □

Smoothing effects of the resolvent family $(S(t))_{t\geq 0}$ when b satisfies Assumption 1.5.1 can be now easily proved using Lemma 1.5.5.

PROPOSITION 1.5.6. *Let b and ρ as in Lemma 1.5.5. Then there exist constants $C_0, C_1 > 0$ such that for $0 \leqslant s \leq 2/\rho$ and $0 \leqslant s' \leq 2$,*

$$\|A^{s/2} S(t)\| \leqslant C_0\, t^{-s\rho/2}, \quad t > 0, \tag{1.5.15}$$

$$\|A^{-s'/2} \dot{S}(t)\| \leqslant C_1 \|b\|_{L^1(0,t)}^{s'/2}\, t^{s'/2-1}, \quad t > 0, \tag{1.5.16}$$

$$\int_0^t \|S(s)x\|^2 \,\mathrm{d}s \leqslant C\|A^{-\frac{1}{2\rho}} x\|, \quad t > 0. \tag{1.5.17}$$

PROOF. For any $\delta \in (0,1)$ and any $k \geqslant 1$, Hölder's inequality, (1.5.12) and (1.5.13) yields

$$\begin{aligned} \int_0^{+\infty} u^\delta |\dot{s}_k(u)| \,\mathrm{d}u &= \int_0^{+\infty} u^\delta |\dot{s}_k(u)|^\delta |\dot{s}_k(u)|^{1-\delta} \,\mathrm{d}u \\ &\leqslant \left(\int_0^{+\infty} u |\dot{s}_k(u)| \,\mathrm{d}u \right)^\delta \left(\int_0^{+\infty} |\dot{s}_k(u)| \,\mathrm{d}u \right)^{1-\delta} \\ &\leqslant C_0\, \lambda_k^{-\delta/\rho}. \end{aligned}$$

Note, that the previous final estimate also holds for $\delta = 0, 1$ by (1.5.12) and (1.5.13). Then, since $s_k(t) = -\int_t^{+\infty} u^{-\delta} u^\delta \dot{s}_k(u)\, du$ as $\lim_{r\to\infty} s_k(r) = 0$ for all $k \geq 1$ by

Lemma 1.5.5, we can conclude that

$$|s_k(t)| \leqslant C_0\, t^{-\delta} \lambda_k^{-\delta/\rho}, \quad t > 0, \quad \delta \in [0,1]. \tag{1.5.18}$$

Thus, for any $s \in [0, 2/\rho]$ and $x \in H$, (1.5.18) with $0 \le \delta = \rho s/2 \le 1$ implies

$$\|A^{s/2} S(t) x\|^2 \;=\; \sum_{k\geqslant 1} \lambda_k^s\, s_k(t)^2 \langle x, e_k\rangle^2 \leqslant C_0\, t^{-\rho s/2} \|x\|^2,$$

which is (1.5.15). To show (1.5.16), we use [**93**, Corollary 3.3] which states that under Assumption 1.5.1 and since 0 belongs to the resolvent set of A, there is $M > 0$ such that

$$\|\dot S(t) x\| \leqslant M t^{-1} \|x\|, \quad x \in H, \quad t > 0. \tag{1.5.19}$$

On the other hand, we can bound $\dot S(t)x$ for $x \in D(A)$ as follows:

$$\begin{aligned} \|\dot S(t)x\|^2 &= \sum_{k\geq 1} (\dot s_k(t))^2 \langle x, e_k\rangle^2 \\ &= \sum_{k\geq 1} \lambda_k^2 \left(\int_0^t b(t-s) s_k(s) ds \right)^2 \langle x, e_k\rangle^2 \leqslant \|b\|^2_{L^1(0,t)} \|Ax\|^2, \end{aligned} \tag{1.5.20}$$

where we have used (1.5.8) and (1.5.11). Interpolation between (1.5.19) and (1.5.20) yields (1.5.16). Finally, (1.5.17) follows by (1.5.11) and (1.5.14) as

$$\begin{aligned} &\int_0^t \|S(s)x\|^2\, \mathrm{d}s = \sum_{k=1}^{\infty} \int_0^t s_k^2(s)\, \mathrm{d}s\, \langle x, e_k\rangle^2 \\ &\le \sum_{k=1}^{\infty} \|s_k\|_{L^\infty(\mathbb{R}_+)} \|s_k\|_{L^1(\mathbb{R}_+)} \langle x, e_k\rangle^2 \le C_0 \sum_{k=1}^{\infty} \lambda_k^{-1/\rho} \langle x, e_k\rangle^2 = C_0 \|A^{-\frac{1}{2\rho}} x\|^2. \end{aligned}$$

□

When the kernel b satisfies the slightly stronger Assumption 1.5.2 we do get an additional smoothing effect of $\dot S$. First, recall the following important result regarding L^1 bounds for Laplace transforms which is essentially follows from [**53**, Theorem 4.3].

LEMMA 1.5.7. *Let r be a bounded analytic function in the right halfplane with boundary function $g(x) = \lim_{\varepsilon\to 0} r(\varepsilon + ix)$ for almost all $x \in \mathbb{R}$. If $g \in L^p(\mathbb{R})$ for some $1 \le p < \infty$ and $r' \in H^1(\mathbb{C}_+)$ where $H^1(\mathbb{C}_+)$ denotes the Hardy space, then there exists $f \in L^1(\mathbb{R}_+)$ with $\int_0^\infty e^{-\lambda t} f(t)\, dt = r(\lambda)$ and*

$$\|f\|_{L^1(\mathbb{R}_+)} \le \frac{1}{2} \|r'\|_{H^1(\mathbb{R})}.$$

LEMMA 1.5.8. *Suppose that the kernel b satisfies Assumption 1.5.2. Then there exists $C_0 > 0$ such that*

$$\|t \ddot s_\mu\|_{L^1(\mathbb{R}_+)} \leqslant C_0, \qquad \mu > 0, \tag{1.5.21}$$

$$\|t^2 \ddot s_\mu\|_{L^1(\mathbb{R}_+)} \leqslant C_0 \mu^{-1/\rho}, \qquad \mu > 0. \tag{1.5.22}$$

PROOF. We follow the idea of the proof of [**85**, Lemma 6]. To be more precise, we show that the Laplace transforms $f^1_\mu = \widehat{t\ddot s_\mu}$ and $f^2_\mu = \widehat{t^2 \ddot s_\mu}$ of $t\ddot s_\mu$ and $t^2 \ddot s_\mu$, respectively, are bounded in $\mathbb{C}_+$, $f^1_\mu(i\cdot), f^1_\mu(i\cdot) \in L^p(\mathbb{R})$ for some $1 \le p < \infty$ and their derivatives are in the Hardy space $H^1(\mathbb{C}_+)$ with a suitable estimate. It then follows from Proposition 1.5.7 that

$$\|t\ddot{s}_\mu\|_{L^1(\mathbb{R}_+)} \leqslant \frac{1}{2}\|(f^1_\mu)'\|_{H^1(\mathbb{C}_+)} \text{ and } \|t^2\ddot{s}_\mu\|_{L^1(\mathbb{R}_+)} \leqslant \frac{1}{2}\|(f^2_\mu)'\|_{H^1(\mathbb{C}_+)}.$$

In order to match notation with the proof of [**85**, Lemma 6], we define g and r_μ by $\dot{s}_\mu = -\mu r_\mu$ for $\mu > 0$ and $g = \widehat{b}$.

Thanks to $r_\mu(0) = 0$, using integration by parts, we find

$$f^1_\mu(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda t} t\mu\dot{r}_\mu(t)\mathrm{d}t = \int_0^\infty (\lambda t - 1)\,\mathrm{e}^{-\lambda t}\mu r_\mu(t)\mathrm{d}t = -\mu\lambda h'_\mu(\lambda) - \mu h_\mu(\lambda),$$

$$f^2_\mu(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda t} t^2\mu\dot{r}_\mu(t)\mathrm{d}t = \int_0^\infty \left(\lambda t^2 - 2t\right)\mathrm{e}^{-\lambda t}\mu r_\mu(t)\mathrm{d}t = -\mu\lambda h''_\mu(\lambda) + 2\mu h'_\mu(\lambda),$$

where we have set $h_\mu = \widehat{r}_\mu$. Note that the Laplace transform of the solution of equation (1.5.8) is given by

$$\widehat{s}_\mu(\lambda) = \frac{1}{\lambda + \mu\widehat{b}(\lambda)} = \frac{1}{\lambda + \mu g(\lambda)}, \quad \operatorname{Re}\lambda > 0,\ \mu > 0.$$

By definition $s_\mu(0) = 1$, and by the operational properties of the Laplace transform, we thus obtain

$$h_\mu(\lambda) = \widehat{r}_\mu(\lambda) = -\frac{1}{\mu}\widehat{\dot{s}_\mu}(\lambda) = -\frac{1}{\mu}(\lambda\widehat{s}_\mu(\lambda) - 1) = \frac{g(\lambda)}{\lambda + \mu g(\lambda)}, \quad \operatorname{Re}\lambda > 0,\ \mu > 0.$$

Therefore, an elementary calculation shows that

$$|f^1_\mu(\lambda)| = \left|\mu\frac{g'(\lambda)\lambda^2 + \mu(g(\lambda))^2}{(\lambda + \mu g(\lambda))^2}\right|,$$

$$|f^2_\mu(\lambda)| = \left|\mu\frac{\mu\lambda^2 g''(\lambda)g(\lambda) - 2\mu\lambda^2(g'(\lambda))^2 + \lambda^3 g''(\lambda) - 4\lambda^2 g'(\lambda)}{(\lambda + \mu g(\lambda))^3} + \mu\frac{2\mu g(\lambda)^2 + 4\lambda g(\lambda)}{(\lambda + \mu g(\lambda))^3}\right|.$$

Since a 4-monotone kernel is 3-regular it follows that

$$|\lambda^n g^{(n)}(\lambda)| \leqslant C|g(\lambda)|, \quad \operatorname{Re}\lambda \geqslant 0,\ \mu > 0, n = 0, 1, 2, 3. \tag{1.5.23}$$

Moreover, the estimate

$$|\lambda + \mu g(\lambda)| \geqslant C|\lambda| + \mu|g(\lambda)|, \quad \operatorname{Re}\lambda \geqslant 0,\ \mu > 0, \tag{1.5.24}$$

holds, see [**85**, (5.16)].

We thus obtain

$$|f^1_\mu(\lambda)| \leqslant C\frac{|\mu g(\lambda)\lambda| + |\mu^2(g(\lambda))^2|}{(|\lambda| + \mu|g(\lambda)|)^2} \leqslant C, \qquad \operatorname{Re}\lambda \geqslant 0,\ \mu > 0,$$

$$|f^2_\mu(\lambda)| \leqslant C\frac{|\mu^2 g^2(\lambda)| + |\mu\lambda g(\lambda)|}{(|\lambda| + \mu|g(\lambda)|)^3} \leqslant \frac{C}{|\lambda| + \mu|g(\lambda)|} \leqslant C, \qquad \operatorname{Re}\lambda \geqslant 0,\ \mu > 0,$$

i.e. f^1_μ and f^2_μ are bounded in $\mathbb{C}_+$. Furthermore, by [**93**, Proposition 3.9],

$$2\int_0^{\frac{1}{|\rho|}} b(t)\,\mathrm{d}t \geqslant |g(i\rho)| \geqslant (2\sqrt{2})^{-1}\int_0^{\frac{1}{|\rho|}} b(t)\,\mathrm{d}t, \quad \rho \neq 0,$$

and hence $f^1_\mu(i\cdot), f^2_\mu(i\cdot) \in L^p(\mathbb{R})$ for all $p > 1$. Next, the product rule yields

$$(f^1_\mu)' = -\,\mu h'_\mu - \mu\lambda h''_\mu - \mu h'_\mu = -2\mu h'_\mu - \mu\lambda h''_\mu,$$
$$(f^2_\mu)' = -\,\mu h''_\mu - \lambda\mu h'''_\mu + 2\mu h''_\mu = \mu h''_\mu - \lambda\mu h'''_\mu.$$

We first consider $(f^1_\mu)'$. Since $\|\mu h'_\mu\|_{H^1(\mathbb{C}_+)} \leqslant C$, $\mu > 0$ has already been established in [**85**, Lemma 6], it is enough to consider the term $\lambda\mu h''_\mu$. By (1.5.23) and (1.5.24), we obtain

$$\begin{aligned}&|\lambda\mu h''_\mu(\lambda)|\\ &= \left|\mu\lambda\frac{\mu\lambda g''(\lambda)g(\lambda) - 2\mu\lambda(g'(\lambda))^2 + \lambda^2 g''(\lambda) + 2\mu g'(\lambda)g(\lambda) - 2\lambda g'(\lambda) + 2g(\lambda)}{(\lambda+\mu g(\lambda))^3}\right|\\ &\leqslant C\frac{\mu|g(\lambda)|}{(|\lambda| + \mu|g(\lambda)|)^2}, \quad \operatorname{Re}\lambda \geqslant 0,\ \mu > 0.\end{aligned}$$

Following the estimates for h'_μ in the proof of [**85**, Lemma 6], we end up with $\|\lambda\mu h''_\mu\|_{H^1(\mathbb{C}_+)} \leqslant C$, $\mu > 0$ and (1.5.21) follows.

Next, we turn our attention to $(f^2_\mu)'$. Since $\|\mu h''_\mu\|_{H^1(\mathbb{C}_+)} \leqslant C|\mu|^{-1/\rho}$, $\mu > 0$ is already shown in [**85**, Lemma 6], it is enough to consider the term $\lambda\mu h'''_\mu$. An elementary but rather long calculation gives

$$\begin{aligned}h'''_\mu(\lambda) =& \frac{g'''(\lambda)}{\lambda+\mu g(\lambda)} - \frac{3g''(\lambda)(1+\mu g'(\lambda)) + 3\mu g'(\lambda))g''(\lambda) + \mu g(\lambda)g'''(\lambda)}{(\lambda+\mu g(\lambda))^2}\\ &+ \frac{6g'(\lambda)(1+\mu g'(\lambda))^2 + 6\mu g(\lambda))(1+\mu g'(\lambda))g''(\lambda)}{(\lambda+\mu g(\lambda))^3} - \frac{6g(\lambda)(1+\mu g'(\lambda))^3}{(\lambda+\mu g(\lambda))^4}.\end{aligned}$$

Using the estimates (1.5.23), (1.5.24) and the identity

$$\frac{g'''(\lambda)}{\lambda+\mu g(\lambda)} - \frac{\mu g(\lambda)g'''(\lambda)}{(\lambda+\mu g(\lambda))^2} = \frac{\lambda g'''(\lambda)}{(\lambda+\mu g(\lambda))^2},$$

we obtain

$$|\lambda\mu h'''_\mu(\lambda)| \leqslant C\frac{|\mu g'(\lambda)| + |\mu\lambda g''(\lambda)| + 1 + |\lambda^2\mu g'''(\lambda)|}{|\lambda+\mu g(\lambda)|^2}, \quad \operatorname{Re}\lambda \geqslant 0,\ \mu > 0.$$

It follows from [**93**, Proposition 3.8] that

$$|t^{n-1}g^{(n)}(it)| \leqslant C|t|^{n-1}\int_0^{1/|t|}\tau^n b(\tau)\,d\tau \le C\Psi(1/|t|), \quad t \in \mathbb{R}\setminus\{0\},\ n = 1,2,3,$$

where $\Psi(s) = \int_0^s tb(t)\mathrm{d}t$, $s > 0$. We thus arrive at

$$|it\mu h'''_\mu(it)| \leqslant C\frac{\mu\Psi(1/|t|) + 1}{|t| + \mu g(|t|)|^2}, \quad t \in \mathbb{R}\setminus\{0\},\ \mu > 0,$$

and the proof of the estimate for h''_μ given in [**85**, Lemma 6] shows that

$$\|\lambda\mu h'''_\mu(\lambda)\|_{H^1(\mathbb{C}_+)} \leqslant C|\mu|^{-1/\rho}, \quad \mu > 0.$$

The proof is complete. □

The next result specifies the key smoothing property of the derivative of the solution operator of the linear, homogeneous, deterministic problem with 4-monotone kernels.

PROPOSITION 1.5.9. *Suppose that the kernel b satisfies Assumption 1.5.2. Then*

$$\|A^s\dot{S}(t)\| \leqslant Ct^{-s\rho-1}, \quad t>0,\ s\in[0,1/\rho], \tag{1.5.25}$$

PROOF. To show (1.5.25) first note that $s=0$ is included in (1.5.16). Let $\mu>0$. Next we use Lemma 1.5.8 and Hölder's inequality to conclude that, for $0<\delta<1$,

$$\begin{aligned}\int_0^\infty t^{\delta+1}|\ddot{s}_\mu(t)|\,\mathrm{d}t &= \int_0^\infty |t^2\ddot{s}_\mu(t)|^\delta |t\ddot{s}_\mu(t)|^{1-\delta}\,\mathrm{d}t\\ &\le \left(\int_0^\infty |t^2\ddot{s}_\mu(t)|\,\mathrm{d}t\right)^\delta \left(\int_0^\infty |t\ddot{s}_\mu(t)|\,\mathrm{d}t\right)^{1-\delta} \le C_0\mu^{-\frac{\delta}{\rho}}.\end{aligned}$$

Note that this estimate also holds for $\delta=1$, by (1.5.22). Therefore, taking into account that $\lim_{t\to\infty}\dot{s}_\mu(t)=0$ by (1.5.16) with $s'=0$, we have that

$$|\dot{s}_\mu(t)| = \left|\int_t^\infty \ddot{s}_\mu(r)\,\mathrm{d}r\right| = \left|\int_t^\infty r^{-\delta-1}r^{\delta+1}\ddot{s}_\mu(r)\,\mathrm{d}r\right| \le C_0 t^{-\delta-1}\mu^{-\frac{\delta}{\rho}}.$$

That is,

$$|\mu^s\dot{s}_\mu(t)| \le C_0 t^{-1-\rho s},\ 0<s\le\frac{1}{\rho}.$$

Hence,

$$\|A^s\dot{S}(t)x\|^2 = \sum_{k=1}^\infty \left(\lambda_k^s\dot{s}_k(t)\right)^2\langle x,e_k\rangle^2 \le C_0 t^{-2-2\rho s}\|x\|^2$$

and (1.5.25) follows. □

1.5.3. Regularity of the continuous stochastic problem. Next we recall an existence result for the problem (1.5.1) and, for the sake of completeness, we indicate a proof (see [**28**, Theorem 2.1] and we refer to [**101**] for more general noise).

PROPOSITION 1.5.10. *Suppose that* $\|A^{(\beta-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathrm{HS}}<+\infty$ *for some* $\beta\ge 0$, X_0 *is* $\mathcal{F}_0$*-measurable with* $X_0\in L^2(\Omega,\mathcal{F}_0,\mathbb{P};D(A^{\frac{\beta}{2}}))$ *and let b satisfy Assumption 1.5.1. Then there exists an unique H-valued weak solution* $X\in L^2(\Omega;D(A^{\frac{\beta}{2}}))$ *of* (1.5.1) *given by the variation of constants formula*

$$X(t) = S(t)X_0 + \int_0^t S(t-s)\,dW(s). \tag{1.5.26}$$

Furthermore, if $\beta>0$, *then the stochastic convolution term has a version whose trajectories are a.s.* θ*-Hölder continuous for any* $\theta<\min(\frac{1}{2},\frac{\rho\beta}{2})$.

PROOF. Analogously to [**31**, Theorem 5.4], it is sufficient to show that the stochastic convolution is well-defined and has the required regularity. It follows by Itô's Isometry and the fact that $\|S(t)\|\le 1$ that

$$\mathbb{E}\|X(t)\|_\beta^2 \le 2\mathbb{E}\|X_0\|_\beta^2 + 2\int_0^t \|A^{\beta/2}S(s)Q^{1/2}\|_{\mathrm{HS}}^2\mathrm{d}s.$$

Let (e_k, λ_k) be the eigenpairs of A. Then, by monotone convergence, the self-adjointness of A and S, and Lemma 1.5.5, it follows that

$$\begin{aligned}
&\int_0^t \|A^{\beta/2}S(s)Q^{1/2}\|_{\mathrm{HS}}^2 \mathrm{d}s = \sum_{k=1}^{\infty}\int_0^t \|A^{\beta/2}S(s)Q^{1/2}e_k\|^2\,\mathrm{d}s \\
&= \sum_{j,k=1}^{\infty}\int_0^t \langle A^{\beta/2}S(s)Q^{1/2}e_k, e_j\rangle^2\,\mathrm{d}s = \sum_{j,k=1}^{\infty}\int_0^t \langle Q^{1/2}e_k, S(s)A^{\beta/2}e_j\rangle^2\,\mathrm{d}s \\
&= \sum_{j,k=1}^{\infty}\langle Q^{1/2}e_k, \lambda_j^{\beta/2}e_j\rangle^2 \int_0^t s_j^2(s)\,\mathrm{d}s \\
&\le \sum_{j,k=1}^{\infty}\langle Q^{1/2}e_k, \lambda_j^{\beta/2}e_j\rangle^2 \|s_j\|_{L^\infty(\mathbb{R}_+)}\|s_j\|_{L^1(\mathbb{R}_+)} \\
&\le C_0\sum_{j,k=1}^{\infty}\langle Q^{1/2}e_k, \lambda_j^{\beta/2}e_j\rangle^2\lambda_j^{-1/\rho} = C_0\sum_{j,k=1}^{\infty}\langle Q^{1/2}e_k, \lambda_j^{\beta/2-\frac{1}{2\rho}}e_j\rangle^2 \\
&= C_0\|A^{(\beta-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathrm{HS}}^2.
\end{aligned}$$

Finally, the proof of the Hölder regularity in time of X uses similar techniques and is omitted. □

1.5.4. Strong convergence of a semidiscretization in time. Time discretization of (1.5.1) is achieved via the classical backward Euler scheme and, concerning the convolution in time, via a quadrature rule introduced in [**76, 77**]. Let $\Delta t > 0$ and we set $t_n = n\,\Delta t$ for any integer $n \ge 0$. We consider an approximation $X_{\Delta t}^n$ of $X(t_n)$ defined by the recurrence

$$X_{\Delta t}^n - X_{\Delta t}^{n-1} + \Delta t\left(\sum_{k=1}^{n}\omega_{n-k}\,AX_{\Delta t}^k\right) = W(t_n) - W(t_{n-1}), \quad n \geqslant 1, \tag{1.5.27}$$

with initial condition $X_{\Delta t}^0 = X_0$. The coefficients $(\omega_k)_{k\geqslant 0}$ of the quadrature are chosen such that

$$\sum_{k=0}^{+\infty}\omega_k z^k = \widehat{b}\left(\frac{1-z}{\Delta t}\right), \quad |z| < 1. \tag{1.5.28}$$

In the sequel we derive a discrete mild formulation (variation of constants formula) for (1.5.27). This formulation can not be made easily explicit as a function of the operators A, Q and the kernel b, because of the memory effect in the drift. First consider the deterministic algorithm

$$v^n - v^{n-1} + \Delta t\left(\sum_{k=1}^{n}\omega_{n-k}\,Av^k\right) = 0, \quad n \geqslant 1; \quad v^0 = x. \tag{1.5.29}$$

Taking the z-transform, using the notation

$$\hat{V}(z) = \sum_{k=0}^{\infty} v^k z^k \text{ and } \hat{\omega}(z) = \sum_{k=0}^{\infty}\omega_k z^k,$$

we get

$$\hat{V}(z) - x - z\hat{V}(z) + \Delta t\hat{\omega}(z)A(\hat{V}(z) - x) = 0.$$

Thus,

$$\hat{V}(z) = (I + \Delta t\hat{\omega}(z)A)((1-z)I + \Delta t\hat{\omega}(z)A)^{-1}x := \hat{B}(z)x,$$

where

$$\hat{B}(z)x = \sum_{k=0}^{\infty} B_{\Delta t}^k x z^k.$$

This means that $v^k = B_{\Delta t}^k x$, $k = 0, 1, \ldots$ Note that $B_{\Delta t}^0 = \hat{B}(0) = I$. For the stochastic equation it will be useful to rewrite $\hat{B}(z)x$ as

(1.5.30)

$$\begin{aligned}\hat{B}(z)x =& ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}(I + \hat{\omega}(z)\Delta tA)x \\ &= ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x + \hat{\omega}(z)\Delta tA((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x \\ &= ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x - (1-z)((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}x + x \\ &= (z((1-z)I + \hat{\omega}(z)\Delta tA)^{-1} + I)x.\end{aligned}$$

Now, we consider the stochastic case (1.5.27) which reads, after taking the z-transform, rearranging, and using the notation $w^n = W(t_n) - W(t_{n-1})$ for $n \geq 1$, $w^0 = 0$, and

$$\hat{w}(z) = \sum_{k=0}^{\infty} w^k z^n \text{ and } \hat{U}(z) = \sum_{k=0}^{\infty} X_{\Delta t}^k z^k,$$

as

$$\begin{aligned}\hat{U}(z) &= \hat{B}(z)x + ((1-z)I + \hat{\omega}(z)\Delta tA)^{-1}\hat{w}(z) \\ &= \hat{B}(z)x + \frac{\hat{B}(z) - I}{z}\hat{w}(z) = \hat{B}(z)x + \hat{B}(z)\frac{\hat{w}(z)}{z} - \frac{1}{z}\hat{w}(z),\end{aligned}$$

where we also used (1.5.30) to rewrite the stochastic term in the previous calculation. This yields the discrete variation of constants formula, taking into account that $w^0 = 0$ and that $B_{\Delta t}^0 = I$,

$$X_{\Delta t}^n = B_{\Delta t}^n X_0 + \sum_{k=0}^{n} B_{\Delta t}^{n-k} w^{k+1} - w^{n+1} = B_{\Delta t}^n X_0 + \sum_{k=0}^{n-1} B_{\Delta t}^{n-k} w^{k+1}. \tag{1.5.31}$$

The importance of this formula lies in the fact that it connects the deterministic case to the stochastic case with the deterministic time-discrete solution operators $B_{\Delta t}^n$ explicitly appearing in the formula. We have the following representation of $B_{\Delta t}^n$.

Lemma 1.5.11. *Suppose that he resolvent family* $(S(t))_{t\geq 0}$ *of* (1.5.6) *is strongly continuous and uniformly bounded in* t. *Then,* $B_{\Delta t}^0 = I$ *and*

$$B_{\Delta t}^k x = \int_0^{\infty} S(\Delta ts)xf_k(s)\,\mathrm{d}s \text{ for } k \geq 1, \tag{1.5.32}$$

where

$$f_k(s) := \frac{e^{-s}s^{k-1}}{(k-1)!}, \quad k \geq 1.$$

Proof. The Laplace Transform of $(S(t))_{t\geq 0}$ is given by

$$\hat{S}(z)x = (zI + \hat{b}(z)A)^{-1}x.$$

Using (1.5.28) and (1.5.30) we see that the z-transform $\hat{B}x$ of $\{B^n_{\Delta t}x\}_n$ is given by

$$\begin{aligned}\hat{B}(z) &= z\frac{1}{\Delta t}\hat{S}(\frac{1-z}{\Delta t})x + x = x + z\int_0^\infty S(\Delta ts)e^{-s}e^{zs}\,\mathrm{d}s\\ &= x + \sum_{k=1}^\infty z^k \int_0^\infty S(\Delta ts)x\frac{e^{-s}s^{k-1}}{(k-1)!}\,\mathrm{d}s.\end{aligned}$$

Therefore, we conclude that $B^0_{\Delta t} = I$ and that

$$B^k_{\Delta t}x = \int_0^\infty S(\Delta ts)x\frac{e^{-s}s^{k-1}}{(k-1)!}\,\mathrm{d}s \text{ for } k \geq 1.$$

□

1.5.4.1. *Deterministic estimates: stability, smoothing and rate of convergence of Lubich's convolution quadrature of order 1.* The results of this section are interesting in their own right. They will be later on used to show that Lubich's convolution quadrature based on the backward Euler scheme have remarkable qualitative properties: it preserves the L^p-norm of the orbits and admits smooth and non-smooth data estimates under mild assumptions on the family $(S(t))_{t\geq 0}$. In case S is a strongly continuous semigroup of bounded linear operators, the operator $B^k_{\Delta t}$ corresponds to k backward Euler steps when approximating the corresponding Cauchy problem.

We state and prove the results in Banach spaces as the proofs do not use Hilbert space techniques. We start with a result on stability.

Lemma 1.5.12. *Let X be a Banach space with norm $\|\cdot\|$ and let v be a continuous X-valued function. If*

$$v \in L^p((0,\infty);X)$$

for some $1 \leq p \leq \infty$ and

$$b^k_{\Delta t} = \int_0^\infty v(\Delta ts)\frac{e^{-s}s^{k-1}}{(k-1)!}\,\mathrm{d}s \text{ for } k \geq 1,\ \Delta t > 0, \tag{1.5.33}$$

then

$$\Delta t\sum_{k=1}^n \|b^k_{\Delta t}\|^p \leq \int_0^\infty \|v(t)\|^p\,\mathrm{d}t, \quad 1 \leq p < \infty,$$

and

$$\sup_{k\geq 1}\|b^k_{\Delta t}\| \leq \|v(\cdot)\|_{L^\infty(\mathbb{R}_+)}.$$

Proof. Note that $f_k \geq 0$, $\|f_k\|_{L^1(\mathbb{R}_+)} = 1$ and hence a probability density. Therefore, if $p = \infty$, we immediately obtain from (1.5.33) that

$$\sup_{k\geq 1}\|b^k_{\Delta t}\| \leq \|v(\cdot)\|_{L^\infty(\mathbb{R}_+)}.$$

If $1 \leq p < \infty$, then we use Jensen's inequality in (1.5.32) to get

$$\begin{aligned}\Delta t\sum_{k=1}^n \|b^k_{\Delta t}\|^p &\leq \sum_{k=1}^n \Delta t\int_0^\infty \|v(\Delta ts)\|^p f_k(s)\,\mathrm{d}s\\ &= \int_0^\infty \|v(t)x\|^p \sum_{k=1}^n f_k(\frac{t}{\Delta t})\,\mathrm{d}t \leq \sup_{t>0}\sum_{k=1}^\infty f_k(t)\int_0^\infty \|v(t)\|^p\,\mathrm{d}t.\end{aligned}$$

Finally, the observation that $\sum_{k=1}^{\infty} f_k \equiv 1$ completes the proof. □

Next we state and prove the main error estimate for the sequence $\{b^k_{\Delta t}\}$. We note that the method of proof of Theorem 1.5.13 (a) is essentially contained in [**21**] in a probabilistic language (M. Haase 2014, personal communication).

THEOREM 1.5.13. *Let X be a Banach space with norm $\|\cdot\|$ and suppose that v is a twice strongly continuously differentiable X-valued function for $t > 0$.*

(a) If $\sup_{s>0} \|s^2 \ddot{v}(s)\| < \infty$, then

$$\|b^n_{\Delta t} - v(t_n)\| \leq \frac{1}{n-1} \sup_{s>0} \|s^2 \ddot{v}(s)\|, \quad n \geq 2,\ \Delta t > 0.$$

(b) If $\sup_{s>0} \|s \ddot{v}(s)\| < \infty$, then,

$$\|b^n_{\Delta t} - v(t_n)\| \leqslant \Delta t \sup_{s>0} \|s \ddot{v}(s)\|, \quad \Delta t > 0.$$

PROOF. As v is twice strongly continuously differentiable, we have, by Taylor's theorem, that

$$v(s) - v(t) = (s-t)\dot{v}(t) + \int_t^s (s-r)\ddot{v}(r) x \,\mathrm{d}r.$$

Note that

$$\int_0^\infty s f_n(s)\,\mathrm{d}s = n. \tag{1.5.34}$$

Therefore, using that $t_n = n\Delta t$,

$$\begin{aligned} b^n_{\Delta t} - v(t_n) &= \int_0^\infty v(s\Delta t) f_n(s)\,\mathrm{d}s - \int_0^\infty v(t_n) f_n(s)\,\mathrm{d}s \\ &= \int_0^\infty \int_{t_n}^{s\Delta t} (s\Delta t - r)\ddot{v}(r)\,\mathrm{d}r f_n(s)\,\mathrm{d}s. \end{aligned} \tag{1.5.35}$$

(a) Suppose that $\sup_{s>0} \|s^2 \ddot{v}(t)\| < \infty$. We have, by (1.5.35), that

$$\|b^n_{\Delta t} - v(t_n)\| \leqslant \int_0^\infty \left\| \int_{t_n}^{s\Delta t} \frac{s\Delta t - r}{r^2} r^2 \ddot{v}(r)\,\mathrm{d}r \right\| f_n(s)\,\mathrm{d}s. \tag{1.5.36}$$

Next, we write

$$\begin{aligned} &\left\| \int_{t_n}^{s\Delta t} \frac{s\Delta t - r}{r^2} r^2 \ddot{v}(r)\,\mathrm{d}r \right\| \\ &= (s\Delta t - t_n)^2 \\ &\quad \times \left\| \int_0^1 \frac{1-\tau}{(t_n + \tau(s\Delta t - t_n))^2} (t_n + \tau(s\Delta t - t_n))^2 \ddot{v}(t_n + \tau(s\Delta t - t_n))\,\mathrm{d}\tau \right\| \\ &\leqslant \sup_{s>0} \|s^2 \ddot{v}(s)\| (s\Delta t - t_n)^2 \int_0^1 \frac{1-\tau}{(t_n + \tau(s\Delta t - t_n))^2}\,\mathrm{d}\tau \\ &\leqslant \sup_{s>0} \|s^2 \ddot{v}(s)\| (s\Delta t - t_n)^2 \int_0^1 \frac{1}{(t_n + \tau(s\Delta t - t_n))^2}\,\mathrm{d}\tau \\ &= \sup_{s>0} \|s^2 \ddot{v}(s)\| (s\Delta t - t_n) \left(\frac{1}{t_n} - \frac{1}{s\Delta t} \right). \end{aligned}$$

Therefore, using (1.5.34) again, we conclude from (1.5.36) that

$$\|b^n_{\Delta t} - v(t_n)\| \leqslant \sup_{s>0} \|s^2 \ddot{v}(s)\| \left(\int_0^\infty \frac{t_n}{s\Delta t} f_n(s)\,\mathrm{d}s - 1 \right). \tag{1.5.37}$$

Finally, we have that

$$\begin{aligned} \int_0^\infty \frac{t_n}{s\Delta t} f_n(s)\,\mathrm{d}s &= n \int_0^\infty \frac{f_n(s)}{s} \mathrm{d}s = n \int_0^\infty \frac{e^{-s} s^{n-2}}{(n-1)!}\,\mathrm{d}s \\ &= \frac{n}{n-1} \int_0^\infty f_{n-1}(s)\,\mathrm{d}s = \frac{n}{n-1}. \end{aligned}$$

Thus, the claim in (a) follows from (1.5.37).

(b) Suppose now that $\sup_{s>0} \|s\ddot{v}(s)\|\,\mathrm{d}s < \infty$. Then, by (1.5.35),

$$\|b^n_{\Delta t} x - v(t_n)\| \leqslant \int_0^\infty \left\| \int_{t_n}^{s\Delta t} (s\Delta t - r)\ddot{v}(r)\,\mathrm{d}r \right\| f_n(s)\,\mathrm{d}s. \tag{1.5.38}$$

Next, by noticing that $t_n + \tau(s\Delta t - t_n) = t_n(1-\tau) + \tau s\Delta t \geq 0$ for $0 \leqslant \tau \leqslant 1$, we obtain

(1.5.39)

$$\begin{aligned} &\left\| \int_{t_n}^{s\Delta t} (s\Delta t - r)\ddot{v}(r)\,\mathrm{d}r \right\| = \left\| \int_{t_n}^{s\Delta t} \frac{s\Delta t - r}{r} r\ddot{v}(r)\,\mathrm{d}r \right\| \\ &= (s\Delta t - t_n)^2 \left\| \int_0^1 \frac{1-\tau}{t_n + \tau(s\Delta t - t_n)} (t_n + \tau(s\Delta t - t_n))\ddot{v}(t_n + \tau(s\Delta t - t_n))\,\mathrm{d}\tau \right\| \\ &\leqslant \sup_{s>0} \|s\ddot{v}(s)\|(s\Delta t - t_n)^2 \int_0^1 \frac{1-\tau}{t_n + \tau(s\Delta t - t_n)}\,\mathrm{d}\tau \\ &= \sup_{s>0} \|s\ddot{v}(s)\|(s\Delta t - t_n)^2 \int_0^1 \frac{1-\tau}{t_n(1-\tau) + \tau s\Delta t}\,\mathrm{d}\tau \leqslant \sup_{s>0} \|s\ddot{v}(s)\|(s\Delta t - t_n)^2 \frac{1}{t_n}. \end{aligned}$$

Thus, inserting (1.5.39) into (1.5.38) we obtain

$$\|b^n_{\Delta t} - v(t_n)\| \leqslant \sup_{s>0} \|s\ddot{v}(s)\| \frac{1}{t_n} \int_0^\infty (s\Delta t - t_n)^2 f_n(s)\,\mathrm{d}s. \tag{1.5.40}$$

Finally, similarly as above, we may explicitly calculate

$$\begin{aligned} &\int_0^\infty (s\Delta t - t_n)^2 f_n(s)\,\mathrm{d}s = \Delta t^2 \int_0^\infty s^2 f_n(s)\mathrm{d}s - 2t_n\Delta t \int_0^\infty s^2 f_n(s)\mathrm{d}s + t_n^2 \\ &= \Delta t^2 (n+1)n - 2t_n \Delta t n + t_n^2 = n\Delta t^2 = t_n \Delta t, \end{aligned}$$

and the proof of the claim in (b) is complete in view of (1.5.40). □

1.5.4.2. *Deterministic estimates: convergence rates for deterministic Volterra equations.* In this subsection we apply the results from the previous subsection to the deterministic time stepping scheme (1.5.29)

We start with stability.

Corollary 1.5.14. *If b satisfies Assumption 1.5.1, then, for all $x \in H$,*

$$\sup_{k\geq 1} \|B^k_{\Delta t} x\| \leq \|x\| \text{ and } \Delta t \sum_{k=1}^n \|B^k_{\Delta t} x\|^2 \leq C\|A^{-\frac{1}{2\rho}} x\|^2,\ n \geq 1.$$

PROOF. The statement follows from Lemma 1.5.12 applied to $t \mapsto v(t) = S(t)x$ taking into account the integral representation of $B^k_{\Delta t}$ from Lemma 1.5.11 together with (1.5.17) and the fact that $\|S(t)\| \leq 1$ for $t \geq 0$. □

Next, we derive error estimates. We first need the following regularity result.

PROPOSITION 1.5.15. *Suppose b satisfies Assumption 1.5.2. Then*

$$\sup_{s>0} \|s^2 \ddot{S}(s)\| < \infty, \tag{1.5.41}$$

and for any $\varepsilon > 0$ there is $C_\varepsilon > 0$ such that

$$\sup_{s>0} \|s\ddot{S}(s)x\| \leq C_\varepsilon \|A^{1+\varepsilon}x\|. \tag{1.5.42}$$

PROOF. As b satisfies the sector condition (1.5.3) and A satisfies (1.5.9) it follows that (1.5.5) is parabolic by [**93**, Proposition 3.1]. Furthermore, b is 4-monotone and thus 3-regular and so is $B(t) = \int_0^t b(s)\,\mathrm{d}s$. Therefore, (1.5.41) follows immediately from [**93**, Theorem 3.1]. To show (1.5.42), first note that by (1.5.41) we have that

$$\|s\ddot{S}(s)x\| \leqslant Cs^{-1}\|x\| \leqslant Ks^{-1}\|A^{1+\varepsilon}x\|, \tag{1.5.43}$$

whence

$$\sup_{s\geq 1} \|s\ddot{S}(s)x\| \leqslant K\|A^{1+\varepsilon}x\|.$$

If $x \in \mathcal{D}(A)$, then $u(t) = S(t)x$ is a strong solution of (1.5.6), see [**93**, Proposition 1.2]; that is, $u(t) = S(t)x$ satisfies (1.5.6) with $f \equiv 0$ for all $t > 0$. Then,

$$\ddot{S}(t)x + \int_0^t b(t-s)A\dot{S}(s)x\,\mathrm{d}s + b(t)Ax = 0, \ t > 0.$$

As b is non-negative and non-increasing we have that

$$|tb(t)| \leqslant \int_0^t b(s)\mathrm{d}s = \|b\|_{L^1(0,t)}. \tag{1.5.44}$$

Therefore,

$$\|t\ddot{S}(t)x\| \leqslant t\int_0^t b(t-s)\|\dot{S}(s)Ax\|\,\mathrm{d}s + \|b\|_{L^1(0,t)}\|Ax\|. \tag{1.5.45}$$

We have, by (1.5.44), that

$$\begin{aligned}\int_0^t b(t-s)\|\dot{S}(s)Ax\|\,\mathrm{d}s &= \int_0^{\frac{t}{2}} b(t-s)\|\dot{S}(s)Ax\|\,\mathrm{d}s + \int_{\frac{t}{2}}^t b(t-s)\|\dot{S}(s)Ax\|\,\mathrm{d}s \\ &\leqslant b(\frac{t}{2})\int_0^{\frac{t}{2}} \|\dot{S}(s)Ax\|\,\mathrm{d}s + \sup_{s\in[\frac{t}{2},t]}(\|\dot{S}(s)Ax\|)\int_{\frac{t}{2}}^t b(t-s)\,\mathrm{d}s \\ &\leqslant 2t^{-1}\|b\|_{L^1(0,t)}\int_0^t \|\dot{S}(s)Ax\|\,\mathrm{d}s + \|b\|_{L^1(0,t)}\sup_{s\in[\frac{t}{2},t]}\|\dot{S}(s)Ax\|.\end{aligned}$$

Furthermore, for $0 < s < 1$,

$$\|A^{-\varepsilon}\dot{S}(s)x\| \leqslant Cs^{-1+\varepsilon}\|x\|,$$

and hence

$$\int_0^t b(t-s)\|\dot{S}(s)Ax\|\,\mathrm{d}s \leqslant C_\varepsilon t^{-1+\varepsilon}\|b\|_{L^1(0,t)}\|A^{1+\varepsilon}x\|.$$

Inserting this to (1.5.45) we obtain that

(1.5.46) $$\|t\ddot{S}(t)x\| \leqslant \tilde{C}_\varepsilon t^\varepsilon \|b\|_{L^1(0,t)}\|A^{1+\varepsilon}x\| \leqslant \tilde{K}_\varepsilon \|A^{1+\varepsilon}x\|, \quad 0 < t < 1.$$

Finally, (1.5.42) follows from (1.5.43) and (1.5.46). □

COROLLARY 1.5.16. *Suppose that b satisfies Assumption 1.5.2. Then the following statements hold.*

(a) For every $\varepsilon > 0$, there exits, $C = C(\varepsilon)$

(1.5.47) $$\|B^n_{\Delta t}x - S(t_n)x\| \leqslant C\|A^{1+\varepsilon}x\|\Delta t, \quad \Delta t > 0, x \in \mathcal{D}(A^{1+\varepsilon}).$$

(b) There exists $C > 0$ such that

(1.5.48) $$\|B^n_{\Delta t}x - S(t_n)x\| \leqslant \frac{C}{t_n}\Delta t\,\|x\|, \quad n \geqslant 1.$$

(c) There exists $C = C(T, \gamma, \rho)$ such that

(1.5.49) $$\left(\Delta t \sum_{k=0}^{n} \|B^k_{\Delta t}x - S(t_k)x\|^2\right)^{1/2} \leq C\Delta t^\gamma \|A^{(s-\frac{1}{\rho})/2}x\|, \quad n\Delta t = T,$$

for all $0 \leqslant \gamma < \frac{\rho s}{2}$ where $0 < s \leq \frac{1}{\rho}$.

PROOF. To show (1.5.47) we apply Theorem 1.5.13 (b) to $t \mapsto v(t) = S(t)x$ taking into account the integral representation of $B^k_{\Delta t}$ from Lemma 1.5.11 and (1.5.42) from Proposition 1.5.15. The bound in (1.5.48), for $n \geq 2$, follows from Theorem 1.5.13 (a) applied to $t \mapsto v(t) = S(t)x$ taking into account the integral representation of $B^k_{\Delta t}$ from Lemma 1.5.11 and (1.5.41) from Proposition 1.5.15. For $n = 1$, it follows from Corollary 1.5.14 and the fact that $\|S(t)\| \leq 1$ for $t \geq 0$. To show (1.5.49), note that it follows from (1.5.15), with $s = 1/\rho - \varepsilon$, and Corollary 1.5.14 that there is a constant $C = C(\varepsilon, T)$ such that, for $0 < \varepsilon \leq \frac{1}{\rho}$,

$$\left(\Delta t \sum_{k=0}^{n} \|B^k_{\Delta t}x - S(t_k)x\|^2\right)^{1/2} \leq C\|A^{(\varepsilon-\frac{1}{\rho})/2}x\|, \; n\Delta t = T, \quad \varepsilon > 0,$$

where we also used the fact that $B^0_{\Delta t} = S(0) = I$. Furthermore, since

$$\|S(t_k) - B^k_{\Delta t}\| \leq 2$$

by Corollary 1.5.14, it follows from (1.5.48) that

$$\|B^k_{\Delta t}x - S(t_k)x\| \leqslant C\Delta t^{\frac{1}{2}-\varepsilon} t_k^{\varepsilon-\frac{1}{2}}\|x\|, \quad k \geq 1,$$

and thus, for some $C = C(\varepsilon, T, \rho)$,

$$\left(\Delta t \sum_{k=0}^{n} \|B^k_{\Delta t}x - S(t_k)x\|^2\right)^{1/2} \leq C\Delta t^{\frac{1}{2}-\varepsilon}\|x\|.$$

Interpolation finishes the proof. □

Finally we will need a Hölder type estimate on the resolvent family $(S(t))_{t\geq 0}$.

LEMMA 1.5.17. *If b satisfies Assumption 1.5.1, then there is $C = C(T, \gamma) > 0$ such that*

$$\left(\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|(S(t_n - t_{k-1}) - S(t_n - s))x\|^2 \,\mathrm{d}s\right)^{1/2} \leq C\Delta t^\gamma \|A^{(s-\frac{1}{\rho})/2}x\|, \; n\Delta t = T,$$

for all $0 \leqslant \gamma < \frac{\rho s}{2}$ *where* $0 < s \leq \frac{1}{\rho}$.

PROOF. It follows from (1.5.15), with $s = \frac{1}{\rho} - \varepsilon$, that there is a constant $C = C(\varepsilon, T)$ such that, for $0 < \varepsilon \leq \frac{1}{\rho}$,

$$\left(\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|(S(t_n - t_{k-1}) - S(t_n - s))x\|^2 \, \mathrm{d}s\right)^{1/2} \leq C\|A^{(\varepsilon - \frac{1}{\rho})/2}x\|, \; n\Delta t = t_n = T.$$

Next, it follows from Lemma 1.5.5 that

$$\begin{aligned} &\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \|(S(t_n - t_{k-1}) - S(t_n - s))x\|^2 \, \mathrm{d}s \\ &\qquad = \sum_{i=1}^{\infty} (x, e_i)^2 \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} (s_i(t_n - t_{k-1}) - s_i(t_n - s))^2 \, \mathrm{d}s \\ &\qquad \leq 2 \sum_{i=1}^{\infty} (x, e_i)^2 \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} |s_i(t_n - t_{k-1}) - s_i(t_n - s)| \, \mathrm{d}s \\ &\qquad \leq 2 \sum_{i=1}^{\infty} (x, e_i)^2 \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \int_{t_n - s}^{t_n - t_{k-1}} |\dot{s}_i(t)| \, \mathrm{d}t \, \mathrm{d}s \\ &\qquad \leq 2 \sum_{i=1}^{\infty} (x, e_i)^2 \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \int_{t_n - t_k}^{t_n - t_{k-1}} |\dot{s}_i(t)| \, \mathrm{d}t \, \mathrm{d}s \\ &\qquad \leq 2\Delta t \|x\|^2 \sup_{i \geq 1} \|\dot{s}_i\|_{L^1(\mathbb{R}_+)} \leq C\Delta t \|x\|^2. \end{aligned}$$

Finally, interpolation gives the desired result. □

1.5.4.3. *Strong error estimate for the stochastic problem.* We can now state and prove a strong error bound for the semidiscretization in time of (1.5.1).

THEOREM 1.5.18. *Suppose that* $\|A^{(\beta - \frac{1}{\rho})/2} Q^{1/2}\|_{\mathrm{HS}} < +\infty$ *for some* $0 < \beta \leqslant \frac{1}{\rho}$, $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, *and let* b *satisfy Assumption 1.5.2. For* $T > 0$, *let* $(X(t))_{t \in [0,T]}$ *be the unique weak solution of* (1.5.1) *and let* $X_{\Delta t}^N$ *be the solution of the scheme* (1.5.27) *with* $T = N\Delta t = t_N$. *Then for any* $\gamma < \frac{\rho\beta}{2}$, *there is* $C > 0$ *and* $K = K(T, \gamma, \rho) > 0$ *such that*

$$\|X_{\Delta t}^N - X(T)\|_{L^2(\Omega;H)} \leqslant CT^{-1}\Delta t \|X_0\|_{L^2(\Omega;H)} + K\Delta t^{\gamma}, \quad t_n = n\Delta t = T. \tag{1.5.50}$$

PROOF. If $e_N = X_{\Delta t}^N - X(T) = X_{\Delta t}^N - X(t_N)$, then (1.5.26) and (1.5.31) yields

$$e_N = (B_{\Delta t}^N - S(t_N))X_0 + \sum_{k=1}^{N} \left[\int_{t_{k-1}}^{t_k} (B_{\Delta t}^{N-k+1} - S(t_N - s)) \, \mathrm{d}W(s) \right].$$

Taking the expectation of the square of the H-norm of e_n gives, by independence and Itô's isometry,

$$\mathbb{E}\|e_N\|^2 \leq 2(a + b),$$

where a denotes the deterministic part of the error:

$$a = \mathbb{E}\|(B_{\Delta t}^N - S(t_N))X_0\|^2,$$

dc_1541_18

and b the stochastic part

$$\begin{aligned} b &= \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \|(B_{\Delta t}^{N-k+1} - S(t_N - s))Q^{1/2}\|_{\mathrm{HS}}^2 \,\mathrm{d}s \\ &= \sum_{i=1}^{+\infty} \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \|(B_{\Delta t}^{N-k+1} - S(t_N - s))Q^{1/2}e_i\|^2 \,\mathrm{d}s. \end{aligned}$$

Owing to (1.5.48), a can be bounded as

$$a \leqslant \frac{C}{t_N^2} \Delta t^2 \mathbb{E}\|X_0\|^2, \quad n \geqslant 1.$$

Finally, we use Corollary 1.5.16 and Lemma 1.5.17 to bound b as

$$\begin{aligned} b &\leqslant 2 \sum_{i=1}^{\infty} \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \|(S(t_N - t_{k-1}) - S(t_N - s))Q^{1/2}e_i\|^2 \,\mathrm{d}s \\ &\quad + 2 \sum_{i=1}^{\infty} \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \|(B_{\Delta t}^{N-k+1} - S(t_N - t_{k-1}))Q^{1/2}e_i\|^2 \,\mathrm{d}s \\ &\leq C\Delta t^{2\gamma} \sum_{i=1}^{\infty} \|A^{(\beta - \frac{1}{\rho})/2} Q^{1/2} e_i\| = C\Delta t^{2\gamma} \|A^{(\beta-\frac{1}{\rho})/2} Q^{1/2}\|_{\mathrm{HS}}^2. \end{aligned}$$

for all $\gamma < \frac{\rho\beta}{2}$. □

REMARK 1.5.19. In particular, if $Q = I$ then $d = 1$ and $\beta < \frac{1}{\rho} - \frac{1}{2}$ whence $\gamma < 1/2 - \frac{\rho}{4}$. For trace class noise; that is, when $\mathrm{Tr}(Q) < \infty$, we can take $s = \frac{1}{\rho}$ and hence $\gamma < 1/2$. Remarkably, this is the same rate as for the heat equation [**110**] independently of the value of ρ.

REMARK 1.5.20. If the initial data satisfies $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; D(A^{s(1+\varepsilon)}))$, for some $\varepsilon > 0$ and $0 < s \leqslant 1$, then the deterministic part of the estimate in (1.5.50); that is, the first term on the right hand side, modifies to $C\Delta t^s \|A^{s(1+\varepsilon)} X_0\|_{L^2(\Omega;H)}$ with no singularity at $T = 0$ using the smooth data estimate (1.5.47), stability, and interpolation.

1.5.5. Strong convergence of a Galerkin finite element semidiscretization for a model problem. The theory in the previous sections covers a whole class of problems where the linear operator A is specified via abstract spectral properties. To demonstrate ideas behind the convergence analysis of a semidiscrete Galerkin finite element approximation scheme we consider a concrete setting for A and H but the error analysis can be carried out along the same lines for more complicated operators and spaces as well. To this end, we let $\mathcal{D}$ be a bounded convex domain in $\mathbb{R}^d$, $H := L^2(\mathcal{D})$ and let $A := -\Delta$ be the Laplace operator with $D(A) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$; that is, $(A, D(A)) := (\Lambda, D(\Lambda))$, where Λ is defined precisely in Subsection 1.3.1. Below, we will use some notation introduced in Subsection 1.3.1, in particular the definition of the smoothness spaces $\dot{H}^\beta$ and the corresponding norms.

We discretize (1.5.1) in space by a Galerkin finite element method. We shall derive strong error estimates for the spatially semidiscrete problem for smooth initial data only imposing Assumption 1.5.1 on b. Let $\{V_h\}_{0<h<1}$ be a family of finite dimensional subspaces of $H_0^1(\mathcal{D})$. As in Subsection 1.3.3, the deterministic error analysis is based on the Ritz projection

$$R_h : H_0^1(\mathcal{D}) \to V_h, \quad \langle \nabla R_h v, \nabla \chi \rangle = \langle \nabla v, \nabla \chi \rangle, \; v \in H_0^1(\mathcal{D}), \; \chi \in V_h.$$

We assume that R_h satisfies the error bound

$$\|R_h v - v\| \le C h^\gamma \|v\|_{\dot{H}^\gamma}, \; v \in \dot{H}^\gamma, \; 1 \le \gamma \le 2, \tag{1.5.51}$$

which coincides with (1.3.22) for $r = 2$. Recalling Remark 1.3.10, the subspaces V_h can be chosen, for example, to be the space of continuous piecewise linear functions on a triangulation of $\mathcal{D}$ with maximal mesh-size h and the estimate in (1.5.51) holds as $\mathcal{D}$ is a bounded convex domain.

In order to derive the finite element formulation we look for a V_h-valued process X_h such that

$$\begin{cases} \langle \mathrm{d}X_h(t), \chi \rangle + \displaystyle\int_0^t b(t-s) \langle \nabla X_h(t), \nabla \chi \rangle \,\mathrm{d}s\,\mathrm{d}t = \langle \mathrm{d}W(t), \chi \rangle, \; \chi \in V_h, \; t > 0, \\ \langle X_h(0), \chi \rangle = \langle X_0, \chi \rangle. \end{cases}$$

As before, we introduce the discrete Laplacian

$$A_h : V_h \to V_h, \quad \langle A_h \psi, \chi \rangle = \langle \nabla \psi, \nabla \chi \rangle, \quad \psi, \chi \in V_h, \tag{1.5.52}$$

and the orthogonal projector

$$P_h : H \to V_h, \quad \langle P_h f, \chi \rangle = \langle f, \chi \rangle, \quad \chi \in V_h.$$

It is clear that the operator A_h is a positive definite bounded operator on V_h. Let us note also that using the definition (1.5.52) of A_h, the following uniform inequality can be easily derived

$$\|A_h^{-1/2} P_h x\| \leqslant \|A^{-1/2} x\|, \quad x \in H,$$

see the proof of Theorem 1.4.3. Then, using the L_2-stability of P_h and some interpolation theory, we also have that

$$\|A_h^{-\delta} P_h x\| \le \|A^{-\delta} x\|, \quad \delta \in [0, \frac{1}{2}], \quad x \in H. \tag{1.5.53}$$

Similarly to $-A$, the operator $-A_h$ generates an analytic contraction semigroup on V_h and satisfies the uniform resolvent estimate

$$\|z(z + A_h)^{-1} P_h\| = \|z R(z, A_h) P_h\| \le M_\phi,$$

for $z \in \Sigma_\phi = \{z \in \mathbb{C} : |\cdot(z)| < \phi < \pi\}$. Since $A_h R(z, A_h) = I - z R(z, A_h)$, it follows that

$$\|A_h R(z, A_h) P_h\|_{\mathcal{L}(H)} \leqslant M_\phi + 1, \quad z \in \Sigma_\phi. \tag{1.5.54}$$

Then, we can rewrite the spatially semidiscrete problem in the same form as the original one as

$$\begin{cases} \mathrm{d}X_h + \left(\displaystyle\int_0^t b(t-s) A_h X_h(s) \,\mathrm{d}s \right) \mathrm{d}t = P_h \,\mathrm{d}W(t), \quad t > 0, \\ X_h(0) = P_h X_0, \end{cases}$$

with weak solution given by

$$X_h(t) = S_h(t) P_h X_0 + \int_0^t S_h(t-s) P_h \,\mathrm{d}W(s).$$

The resolvent family $(S_h(t))_{t\geq 0}$ can be written explicitly, with $\dim(V_h) = N_h$, as

$$S_h(t)P_h u_0 = \sum_{k=1}^{N_h} s_{h,k}(t)\langle u_0, e_{h,k}\rangle e_{h,k}. \tag{1.5.55}$$

Here $(\lambda_{h,k}, e_{h,k})$ are the eigenpairs of A_h and $s_{h,k}(t)$ are the solutions of the ODEs

$$\dot{s}_{h,k}(t) + \lambda_{h,k}\int_0^t b(t-s)s_{h,k}(s)\,\mathrm{d}s = 0,\ \ s_{h,k}(0) = 1.$$

We have the following stability result.

Lemma 1.5.21. *If b satisfies Assumption 1.5.1, then for some $C > 0$,*

$$\int_0^t \|S_h(s)P_h x\|^2\,\mathrm{d}s \leq C\|x\|^2_{\dot{H}^{-\frac{1}{\rho}}},\ \ t > 0,\ \ h > 0.$$

Proof. As the constants in (1.5.11) and (1.5.14) do not depend on λ_k, similarly to the proof of (1.5.17), we obtain

$$\int_0^t \|S_h(s)P_h x\|^2\,\mathrm{d}s \leq C_0\|A_h^{-1/2\rho}P_h x\|^2.$$

Since $-1/2 < -1/2\rho < -1/4$, using (1.5.53) with $\delta = 1/(2\rho)$, the statement follows. □

Next we prove an $L^2((0,\infty);H)$ error estimate for the space semidiscretization of the deterministic problem. It is an extension of the result in [**27**] where the special kernel $b(t) = \frac{1}{\Gamma(\beta)}e^{-t}t^{\beta-1}$ was considered.

Proposition 1.5.22. *If b satisfies Assumption 1.5.1, then there is $C > 0$ such that*

$$\int_0^\infty \|S_h(t)P_h x - S(t)x\|^2\,\mathrm{d}t \leq Ch^{2s}\|x\|^2_{\dot{H}^{s-\frac{1}{\rho}}},\ \ 0 \leq s \leq 2.$$

Proof. It follows from (1.5.17) and Lemma 1.5.21 that

$$\int_0^t \|(S_h(s)P_h - S(s))x\|^2\,\mathrm{d}s \leq 2\int_0^t \|S_h(s)P_h x\|^2 + \|S(s)x\|^2\,\mathrm{d}s \leq C\|x\|^2_{\dot{H}^{-\frac{1}{\rho}}}. \tag{1.5.56}$$

To prove an error estimate of optimal order we set

$$\begin{aligned} e(t) := S_h(t)P_h x - S(t)x := v_h(t) - v(t) \\ = v_h(t) - P_h v(t) + P_h v(t) - v(t) := \theta(t) + \rho(t). \end{aligned}$$

For ρ, using the best approximation property of P_h, we obtain by (1.5.17) and (1.5.51) that

$$\int_0^\infty \|\rho(t)\|^2\,\mathrm{d}t \leq \int_0^\infty \|(R_h - I)v(t)\|^2\,\mathrm{d}t \leq Ch^4\|x\|^2_{\dot{H}^{2-\frac{1}{\rho}}}. \tag{1.5.57}$$

In a standard way one derives an equation for θ which reads as

$$\begin{cases} \dot{\theta}(t) + \displaystyle\int_0^t b(t-s)A_h\theta(s)\,\mathrm{d}s = A_h P_h \int_0^t b(t-s)(I - R_h)v(s)\,\mathrm{d}s,\ \ t > 0, \\ \theta(0) = 0. \end{cases}$$

Taking Laplace transforms of both sides yields

$$z\widehat{\theta}(z) + \widehat{b}(z)A_h\widehat{\theta}(z) = A_hP_h(I - R_h)\widehat{v}(z)\widehat{b}(z).$$

Therefore,

$$\widehat{\theta}(z) = A_hR(\frac{z}{\widehat{b}(z)}, A_h)P_h(I - R_h)\widehat{v}(z). \tag{1.5.58}$$

It can be shown that $\widehat{b}$ extends continuously to $i\mathbb{R} \setminus \{0\}$, see, for example, [**85**]. Therefore, using (1.5.3), it follows that $\frac{ik}{\widehat{b}(ik)} \in \Sigma_\phi$, $k \in \mathbb{R} \setminus \{0\}$, with $\phi < \pi$. Thus, $\|A_hR(\frac{ik}{\widehat{b}(ik)}, A_h)P_h\|_{\mathcal{L}(H)} \leq (M_\phi + 1)$ by (1.5.54). Therefore, setting $z = ik$, $k \in \mathbb{R} \setminus \{0\}$, in (1.5.58) and using the isometry property of the Fourier transform we obtain, by (1.5.17) and (1.5.51), that

$$\int_0^\infty \|\theta(t)\|^2 \,\mathrm{d}t \leq (M_\phi + 1)\int_0^\infty \|(I - R_h)v(t)\|^2 \,\mathrm{d}t \leq Ch^4\|x\|^2_{\dot{H}^{2-\frac{1}{\rho}}}. \tag{1.5.59}$$

Interpolation using (1.5.56), (1.5.57), and (1.5.59) yields

$$\int_0^\infty \|e(t)\|^2 \,\mathrm{d}t \leq 2\int_0^\infty \left(\|\rho(t)\|^2 + \|\theta(t)\|^2\right) \,\mathrm{d}t \leq Ch^{2s}\|x\|^2_{\dot{H}^{s-\frac{1}{\rho}}},\ 0 \leq s \leq 2.$$

□

Next, using the error analysis from [**81**], noting that the proofs of that paper remain valid under (1.5.51) on R_h, we have the following pointwise smooth data estimate for the spatially semidiscrete scheme.

PROPOSITION 1.5.23. *If b satisfies Assumption 1.5.1, then for every $\varepsilon, T > 0$ there is $C = C(T, \varepsilon)$ such that*

$$\|S_h(t)P_hx - S(t)x\| \leq Ch^s\|x\|_{\dot{H}^{s(1+\varepsilon)}},\ 0 \leq s \leq 2,\ t \in [0, T].$$

Furthermore, if b satisfies Assumption 1.5.2, then $C = C(\varepsilon)$ is independent of T.

PROOF. As already observed, Assumption 1.5.1 (and hence also Assumption 1.5.2) implies that b is a positive definite kernel. Therefore, the proof of [**81**, Theorem 2.1] yields under assumption (1.5.51) on R_h, using also the best approximation property of P_h together with (1.5.51) with $\beta = 2$, that

$$\|S_h(t)P_hx - S(t)x\| \leq Ch^2\left(\|x\|_{\dot{H}^2} + \int_0^t \|\dot{S}(s)x\|_{\dot{H}^2} \,\mathrm{d}s\right).$$

If b satisfies Assumption 1.5.1, then Proposition 1.5.6 implies that

$$\int_0^t \|\dot{S}(s)x\|_{\dot{H}^2} \,\mathrm{d}s = \int_0^t \|A^{-\varepsilon}\dot{S}(s)A^{1+\varepsilon}x\| \,\mathrm{d}s \leq C(T, \varepsilon)\|x\|_{\dot{H}^{2+2\varepsilon}}. \tag{1.5.60}$$

Furthermore, if b satisfies Assumption 1.5.2 then, by Proposition 1.5.9 we also have that

$$\|\dot{S}(t)x\|_{\dot{H}^2} \leqslant Ct^{-\rho\varepsilon-1}\|A^{1-\varepsilon}x\| \leqslant C_\varepsilon t^{-\rho\varepsilon-1}\|x\|_{\dot{H}^{2+2\varepsilon}},$$

and hence the bound in (1.5.60) can be chosen to be independent of T. Finally, since $\|S(t) - S_h(t)P_h\|_{\mathcal{L}(H)} \leq 2$, interpolation finishes the proof. □

THEOREM 1.5.24. *Suppose that* $\|A^{(\beta-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathrm{HS}} < +\infty$ *for some* $0 < \beta \leq 2$, $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \dot{H}^{\beta(1+\varepsilon)})$ *for some* $\varepsilon > 0$, *and let* b *satisfy Assumption 1.5.1. Then, there is* $C = C(T, \varepsilon, \beta)$ *such that*

$$\|X_h(t) - X(t)\|_{L^2(\Omega;H)} \leq Ch^\beta, \ t \in [0, T].$$

PROOF. By the variation of constants formula,

$$X_h(t) - X(t) = S_h(t)P_hX_0 - S(t)X_0 + \int_0^t (S_h(t-s)P_h - S(t-s))\,\mathrm{d}W(s).$$

Thus,

$$\begin{aligned}\mathbb{E}\|X_h(t) - X(t)\|^2 &\leq 2\mathbb{E}\|S_h(t)P_hX_0 - S(t)X_0\|^2 \\ &\quad + 2\mathbb{E}\left\|\int_0^t (S_h(t-s)P_h - S(t-s))\,\mathrm{d}W(s)\right\|^2 := e_1 + e_2.\end{aligned}$$

It follows from Proposition 1.5.23 that

$$e_1 \leq Ch^{2\beta}\mathbb{E}\|X_0\|^2_{\dot{H}^{\beta(1+\varepsilon)}}.$$

Finally, to bound e_2 we use Itô's Isometry and Proposition 1.5.22 to obtain

$$\begin{aligned} e_2 &= \mathbb{E}\left\|\int_0^t (S_h(t-s)P_h - S(t-s))\,\mathrm{d}W(s)\right\|^2 \\ &= \int_0^t \|(S_h(t-s)P_h - S(t-s))Q^{1/2}\|^2_{\mathrm{HS}}\,\mathrm{d}s \\ &= \sum_{k=1}^\infty \int_0^t \|(S_h(t-s)P_h - S(t-s))Q^{1/2}e_k\|^2\,\mathrm{d}s \\ &\leq Ch^{2\beta}\sum_{k=1}^\infty \|A^{(\beta-\frac{1}{\rho})/2}Q^{1/2}e_k\|^2 = Ch^{2\beta}\|A^{(\beta-\frac{1}{\rho})/2}Q^{1/2}\|^2_{\mathrm{HS}}. \end{aligned}$$

□

REMARK 1.5.25. In particular, if $Q = I$, then $d = 1$ and $\beta < \frac{1}{\rho} - \frac{1}{2}$. For trace class noise; that is, when $\mathrm{Tr}(Q) < \infty$, we can take $\beta = \frac{1}{\rho}$.

1.5.6. Strong convergence of a fully discrete scheme for a model problem. In this section we derive strong error estimates for a fully discrete scheme for (1.5.1) with smooth initial data in the specific setting described in the beginning of Section 1.5.5. As the fully discrete scheme, similarly to the time semidiscretization (1.5.27), we consider the recurrence
(1.5.61)

$$X^n_{h,\Delta t} - X^{n-1}_{h,\Delta t} + \Delta t\left(\sum_{k=1}^n \omega_{n-k}\, A_hX^k_{h,\Delta t}\right) = P_h(W(t_n) - W(t_{n-1})), \quad n \geqslant 1,$$

with $X^0_{h,\Delta t} = P_hX_0$. Again, analogously to (1.5.31), the solution is given by the discrete variation of constants formula

$$X^n_{h,\Delta t} = B^n_{h,\Delta t}P_hX_0 + \sum_{k=0}^{n-1} B^{n-k}_{h.\Delta t}P_hw^{k+1},$$

where $\Delta w^{k+1} = W(t_{k+1}) - W(t_k)$ and $(B^k_{h,\Delta t})_{k\geqslant 0}$ is a family of linear bounded operators on V_h with $B^0_{h.\Delta t} = I$ and representation, similarly to formula (1.5.32) representing $B^k_{\Delta t}$ from Lemma 1.5.11,

$$B^k_{h,\Delta t} v_h = \int_0^\infty S_h(\Delta t s) v_h f_k(s)\,\mathrm{d}s \text{ for } k \geq 1, \quad v_h \in V_h, \tag{1.5.62}$$

where

$$f_k(s) := \frac{e^{-s} s^{k-1}}{(k-1)!}, \quad k \geq 1.$$

In order to use the analysis directly from the previous sections, we need the following lemmata.

LEMMA 1.5.26. *Let $(S(t))_{t\geq 0}$ and $(S_h(t))_{t\geq 0}$ be the resolvent families given by* (1.5.7)*, respectively,* (1.5.55) *with corresponding time stepping operators $(B^n_{\Delta t})_{n\in\mathbb{N}}$ and $(B^n_{h,\Delta t})_{n\in\mathbb{N}}$ given by* (1.5.32)*, respectively,* (1.5.62)*. Then, for $v \in H$,*

$$\Delta t \sum_{k=1}^n \|B^k_{h,\Delta t} P_h v - B^k_{\Delta t} v\|^2 \leqslant \int_0^\infty \|S_h(t) P_h v - S(t) v\|^2\,\mathrm{d}t, \quad n \geq 1 \tag{1.5.63}$$

and

$$\|B^n_{h,\Delta t} P_h v - B^n_{\Delta t} v\| \leqslant \sup_{t>0} \|S_h(t) P_h v - S(t) v\|, \quad n \geq 1. \tag{1.5.64}$$

PROOF. The statement follows immediately from Lemma 1.5.12 applied to $t \mapsto S_h(t) P_h v - S(t) v$. □

LEMMA 1.5.27. *Suppose that b satisfies Assumption 1.5.2. If $(B^n_{h,\Delta t})_{n\in\mathbb{N}}$ is given by* (1.5.62)*, and $x \in \dot H^{2s(1+\varepsilon)}$ for some $0 \leqslant s \leqslant 1$, then*

$$\|B^n_{h,\Delta t} P_h x - S(t_n) x\| \leq C(\Delta t^s + h^{2s}),\ n = 0, 1, ...$$

PROOF. Let $n \geq 1$. We have that

$$B^n_{h,\Delta t} P_h x - S(t_n) x = B^n_{h,\Delta t} P_h x - B^n_{\Delta t} x + B^n_{\Delta t} x - S(t_n) x = e_1 + e_2.$$

To bound e_1 we use (1.5.64) from Lemma 1.5.26 together with Proposition 1.5.23 to conclude that that there is $C = C(\varepsilon) > 0$ such that

$$\|e_1\| \leq C h^{2s} \|x\|_{\dot H^{2s(1+\varepsilon)}}.$$

It follows from (1.5.47), stability and interpolation that that there is $C = C(\varepsilon) > 0$ such that

$$\|e_2\| \leq C \Delta t^s \|x\|_{\dot H^{2s(1+\varepsilon)}}.$$

Finally, for $n = 0$, using the best approximation property of P_h together with (1.5.51) with $\gamma = 2$, stability and interpolation we get

$$\|B^0_{h,\Delta t} P_h x - S(t_0) x\| = \|(P_h - I) x\| \leqslant C h^{2s} \|x\|_{\dot H^{2s}},$$

which concludes the proof. □

THEOREM 1.5.28. *Suppose that $\|A^{(\beta-\frac{1}{\rho})/2} Q^{1/2}\|_{\mathrm{HS}} < +\infty$ for some $0 < \beta \leqslant \frac{1}{\rho}$, $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \dot H^{2s(1+\varepsilon)})$, for some $\varepsilon > 0$ and $0 \leqslant s \leqslant 1$, and let b satisfy Assumption 1.5.2. For $T > 0$, let $(X(t))_{t\in[0,T]}$ be the unique weak solution of* (1.5.1) *and let $X^N_{h,\Delta t}$ be the solution of the scheme* (1.5.61) *with $T = N\Delta t = t_N$. Then there is $C > 0$ and $K = K(T, \gamma, \rho) > 0$ such that*

$$\|X^N_{h,\Delta t} - X(T)\|_{L^2(\Omega;H)} \leq C(\Delta t^s + h^{2s}) \|X_0\|_{L^2(\Omega;\dot H^{2s(1+\varepsilon)})} + K(\Delta t^\gamma + h^\beta),$$

where $\gamma < \frac{\rho\beta}{2}$.

Proof. We decompose the error as

$$\begin{aligned}
X^N_{h,\Delta t} - X(T) &= B^N_{h,\Delta t} P_h X_0 - S(T)X_0 \\
&\quad + \sum_{k=1}^{N} \left[\int_{t_{k-1}}^{t_k} (B^{N-k+1}_{h,\Delta t} P_h - B^{N-k+1}_{\Delta t}) \,\mathrm{d}W(s) \right] \\
&\quad + \sum_{k=1}^{N} \left[\int_{t_{k-1}}^{t_k} (B^{N-k+1}_{\Delta t} - S(t_N - s)) \,\mathrm{d}W(s) \right] \\
&:= e_1 + e_2 + e_3.
\end{aligned}$$

We bound e_1 using Lemma 1.5.27 as

$$(\mathbb{E}\|e_1\|^2)^{1/2} \le C(\Delta t^s + h^{2s})(\mathbb{E}\|X_0\|^2_{\dot{H}^{2s(1+\varepsilon)}})^{1/2}.$$

Next, by Itô's isometry and independence,

$$\begin{aligned}
\mathbb{E}\|e_2\|^2 &= \sum_{k=1}^{N} \int_{t_{k-1}}^{t_k} \|(B^{N-k+1}_{h,\Delta t} P_h - B^{N-k+1}_{\Delta t}) Q^{1/2}\|^2_{\mathrm{HS}} \,\mathrm{d}s \\
&= \sum_{i=1}^{+\infty} \Delta t \sum_{k=1}^{N} \|(B^{N-k+1}_{h,\Delta t} P_h - B^{N-k+1}_{\Delta t}) Q^{1/2} e_i\|^2 \\
&\leqslant \sum_{i=1}^{+\infty} \int_0^{\infty} \|(S_h(t)P_h - S(t))Q^{1/2} e_i\|^2 \\
&\le C h^{2\beta} \sum_{k=1}^{\infty} \|A^{(\beta - \frac{1}{\rho})/2} Q^{1/2} e_k\|^2 = C h^{2\beta} \|A^{(\beta - \frac{1}{\rho})/2} Q^{1/2}\|^2_{\mathrm{HS}},
\end{aligned}$$

where we have used (1.5.63) from Lemma 1.5.26 and the deterministic error estimate from Proposition 1.5.22. Finally, e_3 can be bounded using Theorem 1.5.18 with $X_0 = 0$, by

$$(\mathbb{E}\|e_3\|^2)^{1/2} \le K \Delta t^{\gamma}$$

for all $\gamma < \frac{\rho\beta}{2}$, where $K = K(T, \gamma, \rho) > 0$, and the proof is complete. □

Remark 1.5.29. We would like to highlight three important special cases. Firstly, if $Q = I$ then $d = 1$ whence $\beta < \frac{1}{\rho} - \frac{1}{2}$ and $\gamma < 1/2 - \frac{\rho}{4}$. If $\mathrm{Tr}(Q) < \infty$, then we may set $\beta = \frac{1}{\rho}$. Thus, the time order is almost 1/2, the same as for the heat equation with trace class noise, but the space order is less than 1, which is the space order for the heat equation, see [**110**]. Finally, suppose that there exist some real numbers κ and $\alpha > 0$ such that $A^{\kappa} Q \in \mathcal{L}(H)$, $\mathrm{Tr}(A^{-\alpha}) < \infty$ and $\alpha - 1/\rho < \kappa \leqslant \alpha$. Then, since

$$\|A^{\frac{\beta - 1/\rho}{2}} Q^{\frac{1}{2}}\|^2_{\mathrm{HS}} \leqslant \|A^{\kappa} Q\|_{\mathcal{L}(H)} \mathrm{Tr}(A^{\beta - 1/\rho - \kappa}),$$

we recover a space order $\beta < 1/\rho - (\alpha - \kappa)$ and a time order $\gamma < \frac{1}{2}(1 - \rho(\alpha - \kappa))$.

1.5.7. Weak convergence of a fully discrete scheme for a model problem. In order to give a weak error estimate of optimal order we have to impose another assumption on b. This kind of assumption; that is, the existence of an analytic extension of $\hat{b}$ to a sector beyond the left halfplane, is fairly standard in the existing deterministic literature, see, for example, [**38, 76, 78, 79**], but it clearly represents a major restriction compared to Assumption 1.5.1. However, it allows us to use the deterministic nonsmooth data estimate [**78**, Theorems 2.1 and 3.2,] in the proof of Theorem 1.5.32 below.

Assumption 1.5.30. *The Laplace transform $\widehat{b}$ of b can be extended to an analytic function in a sector Σ_θ with $\theta > \pi/2$ and $|\widehat{b}^{(k)}(z)| \le C|z|^{1-\rho-k}$, $k = 0, 1$, $z \in \Sigma_\theta$.*

An important example of a family of kernels satisfying both Assumptions 1.5.1 and 1.5.30 is given by $b(t) = Ct^{\beta-1}e^{-\eta t}$, $0 < \beta < 1$ and $\eta \ge 0$.

Next we state the smoothing properties of the resolvent family $(S_h(t))_{t\ge0}$ similar to that of $(S(t))_{t\ge0}$ from Proposition 1.5.6.

Lemma 1.5.31. *Suppose that the kernel b satisfies Assumption 1.5.1 and let $(S_h(t))_{t\ge0}$ be the resolvent familiy given by* (1.5.55). *Then, there exist a constant $C > 0$, independent of $t > 0$ and $h > 0$, such that for any $0 \leqslant s \le 2/\rho$,*

$$\|A_h^{s/2} S_h(t)\| \leqslant C\, t^{-s\rho/2}, \quad t > 0. \tag{1.5.65}$$

Proof. The proof is completely analogous to that of Proposition 1.5.6 and therefore it is ommited. □

Let $\sigma(t) := \lceil \frac{t}{\Delta t} \rceil$ and define the piecewise constant operator function

$$\tilde{B}_{h,\Delta t}(t) := B_{h,\Delta t}^{\sigma(t)} P_h, \quad 0 \le t \le T = N\Delta t. \tag{1.5.66}$$

Theorem 1.5.32. *If b satisfies Assumptions 1.5.1 and 1.5.30, then the following estimates hold for some $C > 0$ where $\tilde{F}_{h,\Delta t}(t) = \tilde{B}_{h,\Delta t}(t) - S(t)$, $N\Delta t = T$ and $h > 0$:*

$$\|S(t) - S(s)\| \le C\, s^{-\alpha} |t-s|^\alpha, \;\; 0 \le \alpha \le 1, \;\; 0 < s \le t; \tag{1.5.67}$$

$$\|A^{\frac{\nu}{2}} \tilde{B}_{h,\Delta t}(t)\| \le C\, t^{-\rho\nu/2}, \;\; 0 \le \nu \le \frac{1}{\rho}, \;\; 0 < t \le T; \tag{1.5.68}$$

$$\|\tilde{F}_{h,\Delta t}(t)\| \le C\, t^{-\rho\nu} (\Delta t^{\rho\nu} + h^{2\nu}), \;\; 0 \le \nu \le \frac{1}{\rho}, \;\; 0 < t \le T; \tag{1.5.69}$$

$$\|A^{\frac{1/\rho-\nu}{2}} \tilde{F}_{h,\Delta t}(t)\| \le C\, t^{-\frac{1}{2}-\frac{\rho\nu}{2}} (\Delta t^{\rho\nu} + h^{2\nu}), \;\; 0 \le \nu \le \frac{1}{\rho}, \;\; 0 < t \le T. \tag{1.5.70}$$

Proof. It follows from [**93**, Corollary 3.3] that $\|\dot{S}(t)x\| \le Ct^{-1}\|x\|$ for all $x \in H$ and $t > 0$. Thus, for $0 < s \le t$, we have

$$\|S(t)x - S(s)x\| \le \int_s^t \|\dot{S}(r)x\| \,\mathrm{d}r \le C\|x\| \int_s^t r^{-1} \,\mathrm{d}r \le C\|x\| s^{-1}|t-s|.$$

Since we also have that $\|S(t)x - S(s)x\| \le 2\|x\|$, the inequality in (1.5.67) follows.

To show (1.5.68), first note that it follows from Lemma 1.5.12 applied to $t \mapsto S_h(t)P_h x$ and the fact that $\|S_h(t)\| \leqslant 1$ for all $t \ge 0$ taking into account representation (1.5.62) that

$$\|B_{h,\Delta t}^k\| \le 1, \;\; \text{for all } k \ge 1, \;\; h > 0. \tag{1.5.71}$$

dc_1541_18

From (1.5.62) and (1.5.65) $s = 2/\rho$ we conclude that, for $k \geq 2$ and $h > 0$,
(1.5.72)

$$\|A_h^{1/\rho} B_{h,\Delta t}^k P_h x\| \leq C\|x\|(\Delta t)^{-1} \int_0^\infty \frac{e^{-s} s^{k-2}}{(k-1)!} \,\mathrm{d}s$$
$$= C\|x\|((k-1)\Delta t)^{-1} \int_0^\infty \frac{e^{-s} s^{k-2}}{(k-2)!} \,\mathrm{d}s = C\|x\| t_{k-1}^{-1} = C\|x\| \frac{k}{k-1} t_k^{-1} \leq C\|x\| t_k^{-1}.$$

For $k = 1$, by (1.5.65), we have

$$\begin{aligned} \|A_h^\gamma B_{h,\Delta t}^1 P_h x\| &\leq \int_0^\infty \|A_h^\gamma S_h(\Delta t s) x\| e^{-s} \,\mathrm{d}s \\ &\leq C \Delta t^{-\rho\gamma} \int_0^\infty s^{-\rho\gamma} \,\mathrm{d}s \|x\| = C_\gamma (\Delta t)^{-\rho\gamma} \|x\|, \ \gamma < 1/\rho. \end{aligned} \tag{1.5.73}$$

By interpolation, using (1.5.71), (1.5.73) and (1.5.72) we conclude that

$$\|A_h^{\frac{\nu}{2}} B_{h,\Delta t}^k P_h\| \leq C t_k^{-\frac{\rho\nu}{2}}, \ 0 \leq \nu \leq \frac{1}{\rho}, \ k \geq 1, \ h > 0.$$

Since for $\delta \in [0, 1/2]$ and $v_h \in V_h$ we have that $\|A^\delta v_h\| \leq \|A_h^\delta v_h\|$ it also follows that

$$\|A^{\frac{\nu}{2}} B_{h,\Delta t}^k P_h\| \leq C t_k^{-\frac{\rho\nu}{2}}, \ 0 \leq \nu \leq \frac{1}{\rho}, \ k \geq 1, \ h > 0.$$

Finally, for $t \in (t_{j-1}, t_j]$, $j \geq 1$, we see that

$$\|A^{\frac{\nu}{2}} \tilde{B}_{h,\Delta t}(t)\| = \|A^{\frac{\nu}{2}} B_{h,\Delta t}^j P_h\| \leq C t_j^{-\frac{\rho\nu}{2}} \leq C t^{-\frac{\rho\nu}{2}}, \ 0 \leq \nu \leq \frac{1}{\rho}, \ h > 0,$$

and the proof of (1.5.68) is complete.

Next we prove (1.5.69). First, we write

$$\|B_{h,\Delta t}^k P_h - S(t_k)\| \leq \|B_{h,\Delta t}^k - S_h(t_k)\| + \|S_h(t_k) - S(t_k)\| := e_1 + e_2.$$

It follows from in [**78**, Theorems 2.1 and 3.2] that if b satisfies Assumptions 1.5.1 and 1.5.30, then

$$e_1 \leq C t_k^{-1} \Delta t \text{ and } e_2 \leq C t_k^{-\rho} h^2, \ k \geq 1, \ h > 0.$$

Furthermore, we also have that $\max\{e_1, e_2\} \leq 2$, and thus

$$\|B_{h,\Delta t}^k P_h - S(t_k)\| \leq C t_k^{-\rho\nu} (\Delta t^{\rho\nu} + h^{2\nu}), \ 0 \leq \nu \leq 1/\rho, \ k \geq 1, \ h > 0. \tag{1.5.74}$$

Next, for $t \in (t_{k-1}, t_k]$, $k \geq 1$, we have by (1.5.67) and (1.5.74), that

$$\begin{aligned} \|\tilde{F}_{h,\Delta t}(t)\| &\leq \|B_{h,\Delta t}^k P_h - S(t_k)\| + \|S(t_k) - S(t)\| \\ &\leq C t_k^{-\rho\nu} (\Delta t^{\rho\nu} + h^{2\nu}) + C t_k^{-\rho\nu} \Delta t^{\rho\nu} \\ &\leq C t^{-\rho\nu} (\Delta t^{\rho\nu} + h^{2\nu}), \ 0 \leq \nu \leq 1/\rho, \ k \geq 1, \ h > 0, \end{aligned}$$

which finishes the proof of (1.5.69).

Finally, by interpolation, for $0 \leq \alpha \leq 1/(2\rho)$ have that

$$\begin{aligned} \|A^\alpha \tilde{F}_{h,\Delta t}(t)\| &\leq \|\tilde{F}_{h,\Delta t}(t)\|^{1-2\rho\alpha} \|A^{1/(2\rho)} \tilde{F}_{h,\Delta t}(t)\|^{2\rho\alpha} \\ &\leq \|\tilde{F}_{h,\Delta t}(t)\|^{1-2\rho\alpha} \left(\|A^{1/(2\rho)} S(t)\|^{2\rho\alpha} + \|A^{1/2\rho} \tilde{B}_{h,\Delta t}(t)\|^{2\rho\alpha} \right). \end{aligned}$$

Setting $\alpha = \frac{1/\rho - \nu}{2}$, $0 \leq \nu \leq 1/\rho$, and using (1.5.15) from Proposition 1.5.6, (1.5.68) and (1.5.69) all with $\nu = 1/\rho$ the estimate in (1.5.70) follows. □

THEOREM 1.5.33. *Suppose that b satisfies Assumptions 1.5.1 and 1.5.30. For $T>0$, let $(X(t))_{t\in[0,T]}$ be the unique weak solution of* (1.5.1) *and let $X^N_{h,\Delta t}$ be the solution of the scheme* (1.5.61) *with $T=N\Delta t=t_N$. Let $g\in C^2(H,\mathbb{R})$ such that $g''\in C_b(H,\mathcal{L}(H))$ and suppose that $X_0\in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$. If $\|A^{(\beta-\frac{1}{\rho})/2}Q^{1/2}\|_{\mathrm{HS}}<\infty$, $0<\beta\le 1/\rho$, then there exists a constant $C=C(T,\beta,g,X_0)>0$, which does not depend on h and N, such that for $h^{2/\rho}+\Delta t<T$,*

$$|\mathbb{E}g(X^N_{h,\Delta t})-\mathbb{E}g(X(T))|\leqslant C\ln\left(\frac{T}{h^{2/\rho}+\Delta t}\right)(\Delta t^{\rho\beta}+h^{2\beta}). \tag{1.5.75}$$

PROOF. We use Theorem 1.2.1 with $U=H$, $G=g$, $Y(0)=S(T)X_0$, $E(t)=S(t)$, $\tilde Y(0)=\tilde B_{h,\Delta t}(T)X_0$, $B=\tilde B=I$ and $\tilde S(t)=\tilde B_{h,\Delta t}(t)$. Using (1.1.4) and (1.5.68) with $\nu=0$, we have that

$$\|\tilde B_{h,\Delta t}(t)Q^{\frac12}\|_{\mathrm{HS}}=\|P_h\tilde B_{h,\Delta t}(t)Q^{\frac12}\|_{\mathrm{HS}}\le C\|P_h\|_{\mathrm{HS}},\ t\in(0,T).$$

Therefore, as V_h is finite dimensional, it follows that $\tilde S\in L^2((0,T),\mathcal{L}_2(U_0,H))$. Furthermore,

$$\int_0^T\|S(t)Q^{\frac12}\|^2_{\mathrm{HS}}\,\mathrm{d}t<\infty,$$

as shown in Proposition 1.5.10. Thus, Theorem 1.2.1 is applicable.

We estimate the first term in Theorem 1.2.1 in a similar fashion as in the proof of Theorem 1.3.6, using Proposition 1.5.10 and (1.5.69),

(1.5.76)

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^1\Big\langle u_x\big(0,Y(0)+\theta(\tilde Y(0)-Y(0))\big),\tilde Y(0)-Y(0)\Big\rangle\,\mathrm{d}\theta\Big|\\
&\leqslant\int_0^1\big\|g'\big(Z\big(T,0,Y(0)+\theta(\tilde Y(0)-Y(0))\big)\big)\big\|_{L^2(\Omega;H)}\mathrm{d}\theta\,\|\tilde F_{h,\Delta t}(T)X_0\|_{L^2(\Omega;H)}\\
&\leqslant C\big(1+\int_0^1\big\|Z\big(T,0,Y(0)+\theta(\tilde Y(0)-Y(0))\big)\big\|_{L^2(\Omega;H)}\mathrm{d}\theta\big)\\
&\qquad\times T^{-\rho\beta}(\Delta t^{\rho\beta}+h^{2\beta})\,\|X_0\|_{L^2(\Omega;H)}\\
&\leqslant C\big(1+\|A^{-1/2\rho}Q^{1/2}\|_{\mathrm{HS}}+\|X_0\|_{L^2(\Omega;H)}\big)\,\|X_0\|_{L^2(\Omega;H)}\,T^{-\rho\beta}(\Delta t^{\rho\beta}+h^{2\beta}).
\end{aligned}$$

Next we estimate the trace term. Using that the operators $\tilde B_{h,\Delta t}(r)$, and $S(r)$, $r\in[0,T]$, are self-adjoint, and taking inequalities (1.1.3), (1.1.4) and (1.1.7) into account, we have

$$
\begin{aligned}
&\Big|\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(t,\tilde{Y}(t)) \\
&\qquad \times [\tilde{B}_{h,\Delta t}(T-t)+S(T-t)]Q[\tilde{B}_{h,\Delta t}(T-t)-S(T-t)]^*\Big)\,\mathrm{d}t\Big| \\
&\quad = \Big|\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(t,\tilde{Y}(t))[\tilde{B}_{h,\Delta t}(T-t)+S(T-t)]^* \\
&\qquad \times A^{\frac{1/\rho-\beta}{2}}A^{\frac{\beta-1/\rho}{2}}Q^{\frac12}Q^{\frac12}A^{\frac{\beta-1/\rho}{2}}A^{\frac{1/\rho-\beta}{2}}E_{h,N}(T-t)\Big)\,\mathrm{d}t\Big| \\
&\quad = \Big|\mathbb{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(t,\tilde{Y}(t))(A^{\frac{1/\rho-\beta}{2}}[\tilde{B}_{h,\Delta t}(T-t)+S(T-t)])^* \\
&\qquad \times A^{\frac{\beta-1/\rho}{2}}Q^{\frac12}Q^{\frac12}A^{\frac{\beta-1/\rho}{2}}A^{\frac{1/\rho-\beta}{2}}\tilde{F}_{h,\Delta t}(T-t)\Big)\,\mathrm{d}t\Big| \\
&\quad \le \mathbb{E}\int_0^T \|u_{xx}(t,\tilde{Y}(t))(A^{\frac{1/\rho-\beta}{2}}[\tilde{B}_{h,\Delta t}(T-t)+S(T-t)])^*A^{\frac{\beta-1/\rho}{2}}Q^{\frac12}\|_{\mathrm{HS}} \\
&\qquad \times \|Q^{\frac12}A^{\frac{\beta-1/\rho}{2}}A^{\frac{1/\rho-\beta}{2}}\tilde{F}_{h,\Delta t}(T-t)\|_{\mathrm{HS}}\,\mathrm{d}t \\
&\quad \le \sup_{(t,x)\in[0,T]\times H}\|u_{xx}(t,x)\|_{\mathcal{L}(H)}\|A^{\frac{\beta-1/\rho}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 \\
&\qquad \times \int_0^T \|A^{\frac{1/\rho-\beta}{2}}(\tilde{B}_{h,\Delta t}(t)+S(t))\|_{\mathcal{L}(H)}\|A^{\frac{1/\rho-\beta}{2}}\tilde{F}_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t.
\end{aligned}
$$

Next we split the integral from 0 to $\Delta t + h^{2/\rho}$ and from $h^{2/\rho}+\Delta t$ to T. Then, using (1.5.15) and (1.5.68),

$$
\begin{aligned}
&\int_0^{\Delta t+h^{2/\rho}} \|A^{\frac{1/\rho-\beta}{2}}(\tilde{B}_{h,\Delta t}(t)+S(t))\|_{\mathcal{L}(H)}\|A^{\frac{1/\rho-\beta}{2}}\tilde{F}_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\quad \le 2\int_0^{\Delta t+h^{2/\rho}} \Big(\|A^{\frac{1/\rho-\beta}{2}}\tilde{B}_{h,\Delta t}(t)\|_{\mathcal{L}(H)}^2+\|A^{\frac{1/\rho-\beta}{2}}S(t)\|_{\mathcal{L}(H)}^2\Big)\,\mathrm{d}t \\
&\quad \le C\int_0^{\Delta t+h^{2/\rho}} t^{-1+\rho\beta}\,\mathrm{d}t \le C(\Delta t^{\rho\beta}+h^{2\beta}).
\end{aligned}
$$

Furthermore, by (1.5.15), (1.5.68) and (1.5.70), it follows that

$$
\begin{aligned}
&\int_{\Delta t+h^{2/\rho}}^{T} \|A^{\frac{1/\rho-\beta}{2}}(\tilde{B}_{h,\Delta t}(t)+S(t))\|_{\mathcal{L}(H)}\|A^{\frac{1/\rho-\beta}{2}}\tilde{F}_{h,\Delta t}(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\quad \le C\int_{\Delta t+h^{2/\rho}}^{T} t^{-1/2+\rho\beta/2}(\Delta t^{\rho\beta}+h^{2\beta})t^{-1/2-\rho\beta/2}\,\mathrm{d}t \\
&\quad = C\ln\Big(\frac{T}{\Delta t+h^{2/\rho}}\Big)(\Delta t^{\rho\beta}+h^{2\beta}).
\end{aligned}
$$

This finishes the proof considering that

$$
\sup_{(t,x)\in[0,T]\times H}\|u_{xx}(t,x)\|_{\mathcal{L}(H)} \le \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}.
$$

□

REMARK 1.5.34. The dependence on T of C in (1.5.75) can be removed if we assume, for example, that $X_0 \in L^2(\Omega,\mathcal{F}_0,\mathbb{P};\dot{H}^{2(1+\varepsilon)})$ and that b satisfies the slightly stronger Assumption 1.5.2 instead of Assumption 1.5.1 using the deterministic estimate for smooth initial data from Lemma 1.5.27 in (1.5.76).

REMARK 1.5.35. Below we give examples of the rate of convergence obtained in Theorem 1.5.33 in some typical cases.
(i) If $Q = I$ (white noise), then we must have $d = 1$ and the rate of weak convergence in time is $(1 - \rho/2)_-$ and in space it is $(2/\rho - 1)_-$.
(ii) If Q is of trace class, then we may take $\beta = 1/\rho$ and almost recover the deterministic order; that is, 1_- in time and an order $2/\rho_-$ in space.
(iii) Suppose that there exist some real numbers κ and $\alpha > 0$ such that $A^\kappa Q \in \mathcal{L}(H)$, $\mathrm{Tr}(A^{-\alpha}) < \infty$ and $\alpha - 1/\rho < \kappa \leqslant \alpha$. Then, since

$$\|A^{\frac{\beta-1/\rho}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \leqslant \|A^\kappa Q\|_{\mathcal{L}(H)} \mathrm{Tr}(A^{\beta-1/\rho-\kappa}),$$

we recover a space weak order of convergence $(2/\rho - 2(\alpha - \kappa))_-$ and a time weak order of $(1 - \rho(\alpha - \kappa)_-$.

All these are twice the strong orders (modulo the logarithmic term and for smooth initial data) respectively in space and in time found in Theorem 1.5.28, c.f., Remark 1.5.29.

dc_1541_18

CHAPTER 2

Semilinear stochastic PDEs driven by additive Wiener noise

2.1. Preliminaries

Let $(S(t))_{t\in[0,T]}$ be a strongly continuous evolution family of bounded, self-adjoint, linear operators on a separable Hilbert space $(H, \|\cdot\|, \langle\cdot,\cdot\rangle)$, not necessarily enjoying the semigroup property. Related to $(S(t))_{t\in[0,T]}$ is a densely defined, linear, self-adjoint, positive definite operator $A\colon D(A)\subset H\to H$ with compact inverse. Let $(A^\alpha)_{\alpha\in\mathbb{R}}$ denote the fractional powers of A, which are well defined via spectral calculus, let $(\dot H^\alpha)_{\alpha\in R}$ denote the spaces $\dot H^\alpha = D(A^{\alpha/2})$ for $\alpha\geqslant 0$ with dual spaces $\dot H^{-\alpha} = (\dot H^\alpha)^*$. We assume that $(S(t))_{t\in[0,T]}$ is strongly differentiable on $(0,T]$ with derivative $(\dot S(t))_{t\in(0,T]}$ and that there exist $\rho\in[1,2)$ and constants $(L_s)_{s\in[0,2]}$ so that

(2.1.1)
$$\big\|A^{\frac{\min(1,s)}{\rho}} S(t)x\big\| + \big\|A^{\frac{s-1}{\rho}}\dot S(t)x\big\| \leqslant L_s t^{-s}\|x\|, \quad t\in(0,T],\ x\in H,\ s\in[0,2].$$

If $(S(t))_{t\in[0,T]}$ is the analytic semigroup generated by $-A$, then (2.1.1) holds with $\rho = 1$. If $(S(t))_{t\in[0,T]}$ is the solution operator of the Volterra equation,

$$\dot u(t) + \int_0^t b(t-s)Au(s)\,\mathrm{d}s = 0, \quad t\in(0,T]; \quad u(0) = x;$$

that is, $u(t) = S(t)x$, where $b\colon(0,\infty)\to\mathbb{R}$ is the Riesz kernel $b(t) = t^{\rho-2}/\Gamma(\rho-1)$ for some $\rho\in(1,2)$, then $(S(t))_{t\in[0,T]}$ satisfies (2.1.1). In Subsection 2.4.2 we verify (2.1.1) more general kernels b.

In this chapter we will consider the stochastic evolution equation

$$X(t) = S(t)x_0 + \int_0^t S(t-s)f(X(s))\,\mathrm{d}s + \int_0^t S(t-s)\,\mathrm{d}W(s), \quad t\in[0,T], \tag{2.1.2}$$

where W is a Q-Wiener process in H on a stochastic basis $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]}\,\mathbb{P})$ with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and $f: H\to H$ is a nonlinear mapping. The initial value x_0 is deterministic and satisfies

$$x_0\in\dot H^3 := D(A^{\frac{3}{2}}). \tag{2.1.3}$$

For Hilbert spaces U, V the space $\mathcal{G}_{\mathrm{b}}^k(U,V)$ consists of all, not necessarily bounded, functions $\phi\colon U\to V$, whose Gâteaux derivatives of orders $1,\dots,k$ are bounded, symmetric and strongly continuous, see Subsection 2.1.1. The non-linear drift $f\colon H\to H$ is assumed to satisfy, for some $\delta\in[0,2/\rho)$,

$$f\in\mathcal{G}_{\mathrm{b}}^1(H,H)\cap\mathcal{G}_{\mathrm{b}}^2(H,\dot H^{-\delta}). \tag{2.1.4}$$

This assumption includes interesting cases where $f \notin \mathcal{G}_{\mathrm{b}}^2(H,H)$, e.g., Nemytskii operators on $H = L^2(\mathcal{D})$ for a spatial domain $\mathcal{D} \subset \mathbb{R}^d$, with $\delta > d/2$.

The regularity of the noise is measured by a parameter $\beta \in (0, 1/\rho]$, by assuming

$$\|A^{\frac{\beta-1/\rho}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty. \tag{2.1.5}$$

The smoothest case $\beta = 1/\rho$ corresponds to trace class noise as (2.1.5) reduces to $\|Q^{\frac{1}{2}}\|_{\mathrm{HS}} = \sqrt{\mathrm{Tr}(Q)} < \infty$. The motivation to consider (2.1.2) is that semilinear parabolic SPDEs and stochastic integro-differential equations of Volterra type with additive noise have a mild formulation given by (2.1.2) and hence we can study these type of equations simultaneously.

2.1.1. Spaces of functions and operators. Let $(U, \|\cdot\|_U, \langle\cdot,\cdot\rangle_U)$, $(V, \|\cdot\|_V, \langle\cdot,\cdot\rangle_V)$ be real separable Hilbert spaces. For $k \geqslant 1$, let $\mathcal{L}^{[k]}(U,V)$ be the Banach space of all bounded multilinear mappings $b\colon U^k \to V$, equipped with the norm

$$\|b\|_{\mathcal{L}^{[k]}(U,V)} = \sup_{u_1,\ldots,u_k \in U} \frac{\|b(u_1,\ldots,u_k)\|_V}{\|u_1\|_U \cdots \|u_k\|_U}.$$

It is clear that $\mathcal{L}^{[1]}(U,V) = \mathcal{L}(U,V)$. As before, we denote by $C(U,V)$ the space of all continuous mappings $U \to V$ and further by $C_{\mathrm{str}}(U, \mathcal{L}^{[k]}(U,V))$ the space of strongly continuous mappings $U \to \mathcal{L}^{[k]}(U,V)$, i.e., mappings $B\colon U \to \mathcal{L}^{[k]}(U,V)$ such that for $u_1,\ldots,u_k \in U$, the mapping

$$U \ni x \mapsto B(x)(u_1,\ldots,u_k) \in V,$$

is continuous. A function $\phi\colon U \to V$ is said to be k times Gâteaux differentiable if the recursively defined (directional) derivatives, $\phi^{(l)}\colon U^{l+1} \to V$, $l \in \{1,\ldots,k\}$,

$$\begin{aligned}&\phi^{(l)}(x)(u_1,\ldots,u_l)\\&\quad= \lim_{\varepsilon\to 0} \frac{\phi^{(l-1)}(x+\varepsilon u_l)(u_1,\ldots,u_{l-1}) - \phi^{(l-1)}(x)(u_1,\ldots,u_{l-1})}{\varepsilon},\end{aligned}$$

exist for $u_1,\ldots,u_l, x \in U$, $l \in \{1,\ldots,k\}$, as limits in V, where $\phi^{(0)} = \phi$. This class of functions is large and fails to have natural properties, e.g., Gâteaux differentiability does not imply continuity and the mapping $\phi^{(l)}(x)$ may not be multilinear and symmetric. We therefore introduce a smaller class, with useful properties. For $k \geqslant 1$, let $\mathcal{G}^k(U,V) \subset C(U,V)$ be the subset of all k-times Gâteaux differentiable mappings $\phi \in C(U,V)$, whose Gâteaux derivatives $\phi^l(x)$, $l \in \{1,\ldots,k\}$, are multilinear, bounded and symmetric at every $x \in U$ and $\phi^{(l)} \in C_{\mathrm{str}}(U, \mathcal{L}^{[l]}(U,V))$. This is a weaker assumption than requiring $\phi^{(l)} \in C(U, \mathcal{L}^{[l]}(U,V))$, $l \in \{1,\ldots,k\}$, which would imply (continuous) Fréchet differentiability. For integers $k \in \{0,\ldots,m\}$ and $\phi \in \mathcal{G}^k(U,V)$, let

$$|\phi|_{\mathcal{G}_{\mathrm{p}}^{k,m}(U,V)} = \sup_{u\in U} \frac{\|\phi^{(k)}(u)\|_{\mathcal{L}^{[k]}(U,V)}}{(1+\|u\|_U^{m-k})},$$

and let $\mathcal{G}_{\mathrm{p}}^{k,m}(U,V)$ be the space of $\phi \in \mathcal{G}^k(U,V)$ such that $|\phi|_{\mathcal{G}_{\mathrm{p}}^{l,m}(U,V)} < \infty$ for $l \in \{1,\ldots,k\}$. Let $\mathcal{G}_{\mathrm{p}}^{\infty}(U,V)$ be the space of all infinitely often differentiable mappings $\phi\colon U \to V$ such that ϕ and all its derivatives satisfy a polynomial bound. Let $\mathcal{G}_{\mathrm{b}}^k(U,V)$ denote the space of $\phi \in \mathcal{G}^k(U,V)$ such that

$$|\phi|_{\mathcal{G}_{\mathrm{b}}^l(U,V)} = \sup_{u\in U} \|\phi^{(l)}(u)\|_{\mathcal{L}^{[l]}(U,V)} < \infty, \quad l \in \{1,\ldots,k\}.$$

For $\phi \in \mathcal{G}^1(U,\mathbb{R})$ we can identify the derivative with the gradient $\phi'(u) \in U^* = U$, by the Riesz Representation Theorem. For $m \geqslant 1$, $\phi \in \mathcal{G}_{\mathrm{p}}^{1,m}(U,V)$, the map $[0,1] \ni \lambda \mapsto \phi'(y+\lambda(x-y)) \cdot (x-y) \in V$ is continuous and Bochner integrable and therefore

$$\phi(x) = \phi(y) + \int_0^1 \phi'(y+\lambda(x-y)) \cdot (x-y)\,\mathrm{d}\lambda, \quad x,y \in U. \tag{2.1.6}$$

By $\mathcal{M}_T$ we denote the space of all finite Borel measures on the interval $[0,T]$. For $\nu \in \mathcal{M}_T$ we write $|\nu| = \nu([0,T])$ and for a Banach space V we let $L^p_\nu((0,T);V))$ be the Bochner space of ν-measurable mappings $Z\colon [0,T] \to V$ such that

$$\|Z\|_{L^p_\nu(0,T;V)} = \Big(\int_0^T \|Z(t)\|_V^p \,\mathrm{d}\nu_t\Big)^{\frac{1}{p}} < \infty,$$

with the usual modification for $p = \infty$. When ν is Lebesgue measure we write $L^p((0,T);V)$.

The next lemma is used in the proof of Malliavin regularity by a limiting procedure in Proposition 2.3.4.

Lemma 2.1.1. *Let $\mathcal{X}$, $\mathcal{Y}$ be separable Hilbert spaces such that the embedding $\mathcal{X} \subset \mathcal{Y}$ is continuous. If $x \in \mathcal{Y}$ and $(x_n)_{n\in\mathbb{N}} \subset \mathcal{X}$ satisfies $x_n \to x$ weakly in $\mathcal{Y}$ as $n \to \infty$ and $\sup_{n\in\mathbb{N}} \|x_n\|_{\mathcal{X}} < \infty$, then $x \in \mathcal{X}$.*

Proof. Any closed ball in $\mathcal{X}$ is weakly compact and since $(x_n)_{n\in\mathbb{N}}$ is a bounded sequence in $\mathcal{X}$, there exists a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ and $\tilde{x} \in \mathcal{X}$ such that $x_{n_k} \to \tilde{x}$ weakly in $\mathcal{X}$. Therefore, $x_{n_k} \to \tilde{x}$ also in the weak topology of $\mathcal{Y}$ because $\mathcal{Y}^* \subset \mathcal{X}^*$ is continuous. By assumption $x_n \to x$ weakly in $\mathcal{Y}$, as $n \to \infty$, so $x = \tilde{x} \in \mathcal{X}$. ☐

We need a generalized Gronwall, see [**36**, Theorem 6.1]. Here, and later, we use the convention that an empty sum equals 0.

Lemma 2.1.2. *Let $T > 0$, $N \in \mathbb{N}$, $\Delta t = T/N$, and $t_n = n\Delta t$ for $0 \leqslant n \leqslant N$. If $\varphi_0, \ldots, \varphi_N \geqslant 0$ satisfy for some $M_0, M_1 \geq 0$ and $0 < \nu < 1$ the inequality*

$$\varphi_n \leqslant M_0 + M_1\,\Delta t \sum_{j=0}^{n-1} t_{n-j}^{-1+\nu} \varphi_j, \quad 0 \leqslant n \leqslant N,$$

then there exists a constant $M_2 = M_2(\nu, M_1, T)$ such that

$$\varphi_n \leqslant M_0 M_2, \quad 1 \leqslant n \leqslant N.$$

2.1.2. Malliavin calculus. Let W be a Q-Wiener process in H on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and let $H_0 = Q^{\frac{1}{2}}(H)$ be the Cameron–Martin space endowed with inner product $\langle u,v\rangle_{H_0} = \langle Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}v\rangle$, where $Q^{-\frac{1}{2}}$ denotes the pseudo-inverse of $Q^{\frac{1}{2}}$ if it is not injective. By $\mathcal{L}_2^0 = \mathcal{L}_2(H_0,H)$ we denote the space of Hilbert–Schmidt operators $H_0 \to H$.

For $\Phi \in L^2((0,T);\mathcal{L}_2^0)$ the Wiener integral $\int_0^T \Phi(t)\,\mathrm{d}W(t)$, is well defined and is a random variable in $L^p(\Omega;H)$, $p \in [2,\infty)$. Furhermore, the following consequence of the Burkholder inequality for deterministic integrands holds for $p \geqslant 2$,

$$\Big\|\int_0^T \Phi(t)\,\mathrm{d}W(t)\Big\|_{L^p(\Omega;H)} \leqslant \frac{p(p-1)}{2}\|\Phi\|_{L^2((0,T);\mathcal{L}_2^0)}, \quad \Phi \in L^2((0,T);\mathcal{L}_2^0), \tag{2.1.7}$$

see, for example, [**31**, Lemma 7.2]. By taking $H = \mathbb{R}$ and noting the isomorphisms $H_0 \cong H_0^* \cong \mathcal{L}_2(H_0, \mathbb{R})$ we see that a function $\phi \in L^2((0,T); H_0)$ defines an integrand in $L^2((0,T); \mathcal{L}_2(H_0, \mathbb{R}))$ for the stochastic integral and the integral $\int_0^T \phi(t)\,\mathrm{d}W(t) \in L^2(\Omega)$ is real-valued. As $L^p((0,T); H_0) \subset L^2((0,T); H_0)$ for $p \geqslant 2$ the stochastic integral is well defined for $\phi \in L^p((0,T); H_0)$.

We now recall some concepts from Malliavin calculus introduced in [**2**]. For $q \in [2, \infty]$ let $\mathcal{S}^q(\mathbb{R})$ be the class of smooth cylindrical random variables of the form

$$F = f\Big(\int_0^T \phi_1(s)\,\mathrm{d}W(s), \dots, \int_0^T \phi_n(s)\,\mathrm{d}W(s)\Big),$$
$$f \in \mathcal{G}_{\mathrm{p}}^\infty(\mathbb{R}^n, \mathbb{R}),\ (\phi_k)_{k=1}^n \subset L^q((0,T); H_0),\ n \in \mathbb{N}.$$

For $F \in \mathcal{S}^q(\mathbb{R})$ with the above representation, we define the Malliavin derivative

$$\big(D_t F\big)_{t\in[0,T]} = \left(\sum_{j=1}^n \partial_j f\Big(\int_0^T \phi_1(s)\,\mathrm{d}W(s), \dots, \int_0^T \phi_n(s)\,\mathrm{d}W(s)\Big) \otimes \phi_j(t)\right)_{t\in[0,T]}.$$

Let V be a separable Hilbert space. We define $\mathcal{S}^q(V)$ to be the space of all V-valued random variables of the form $Y = \sum_{i=1}^m v_i \otimes F_i$ with $(v_i)_{i=1}^m \subset V$, $(F_i)_{i=1}^m \subset \mathcal{S}^q(\mathbb{R})$, $m \in \mathbb{N}$. The Malliavin derivative of $Y \in \mathcal{S}^q(V)$ of the above form is given by $D_t Y = \sum_{i=1}^m v_i \otimes D_t F_i$. As $(D_t F_i)_{t\in[0,T]}$ is an H_0-valued process, $(D_t Y)_{t\in[0,T]}$ is a $V \otimes H_0 = \mathcal{L}_2(H_0, V)$-valued process.

For $p \in [2, \infty)$, $q \in [2, \infty]$, $\mathcal{S}^q(V) \subset L^p(\Omega; V)$ is dense by [**2**, Lemma 3.1] and the operator $D\colon \mathcal{S}^q(V) \to L^p(\Omega; L^q((0,T); \mathcal{L}_2(H_0, V)))$ is closable by [**2**, Lemma 3.2]. Let $\mathbf{M}^{1,p,q}(V)$ denote the closure of $\mathcal{S}^q(V)$ with respect to the norm

$$\|Y\|_{\mathbf{M}^{1,p,q}(V)} = \Big(\|Y\|_{L^p(\Omega;V)}^p + \|DY\|_{L^p(\Omega;L^q((0,T);\mathcal{L}_2(H_0,V)))}^p\Big)^{\frac{1}{p}}.$$

We also use the corresponding seminorm

$$|Y|_{\mathbf{M}^{1,p,q}(V)} = \|DY\|_{L^p(\Omega;L^q((0,T);\mathcal{L}_2(H_0,V)))}.$$

The spaces $\mathbf{M}^{1,p,q}(V)$ are Banach spaces, densely and continuously embedded into $L^2(\Omega; V)$. Thus, $\mathbf{M}^{1,p,q}(V) \subset L^2(\Omega; V) \subset \mathbf{M}^{1,p,q}(V)^*$ is a Gelfand triple. By [**2**, Theorem 3.5] the following inequality holds for $p \in [2, \infty)$, $q \in [2, \infty]$ with $\frac{1}{q} + \frac{1}{q'} = 1$:

(2.1.8)
$$\Big\|\int_0^T \Phi(t)\,\mathrm{d}W(t)\Big\|_{\mathbf{M}^{1,p,q}(V)^*} \leqslant \big\|\Phi\big\|_{L^{q'}((0,T);\mathcal{L}_2(H_0,V))}, \quad \Phi \in L^2((0,T); \mathcal{L}_2(H_0, V)).$$

What makes this duality theory useful is the possibility of taking q' close to 1, c.f., (2.1.7) where the exponent is 2. Following [**2**] we refer to $\mathbf{M}^{1,p,q}(H)$ for $q > 2$ as refined Sobolev–Malliavin spaces. The spaces $\mathbf{M}^{1,p,2}(V)$ are the classical Sobolev–Malliavin spaces, often denoted $\mathbf{D}^{1,p}(V)$. For $p = q$ we write $\mathbf{M}^{1,p}(V) := \mathbf{M}^{1,p,p}(V)$.

We next state a modified version of [**2**, Lemma 3.10]. It provides a local Lipschitz bound that enables us to prove an error estimate in the $\mathbf{M}^{1,p}(H)^*$-norm by a Gronwall argument in Lemma 2.3.6 below. More precisely, [**2**, Lemma 3.10] gives a local Lipschitz bound from $\mathbf{G}^{1,p}(U)^*$ to $\mathbf{G}^{1,p}(V)^*$ for mappings $\sigma \in \mathcal{G}_{\mathrm{b}}^2(U, V)$, where $\mathbf{G}^{1,p}(U) = \mathbf{M}^{1,p}(U) \cap L^{2p}(\Omega; U)$. The Lipschitz constant depends on the $\mathbf{M}^{1,2p,p}(U)$-norms of the random variables. By restriction to random variables in

$\mathbf{M}^{1,p}(U)$ with Malliavin derivative bounded over Ω, Lemma 2.1.3 provides a more natural bound, obviating the need for the spaces $\mathbf{G}^{1,p}(V)$. The Lipschitz constant now depends on the $\mathbf{M}^{1,\infty,p}(U)$-seminorm. It is proved in the same way as [**2**, Lemma 3.10], by application of a modified version of [**2**, Lemma 3.8], based on Hölder's inequality with exponents 1, ∞ instead of 2, 2. We omit the details. In the subsequent Lemma 2.1.4 we cite parts of [**2**, Lemma 3.9].

Lemma 2.1.3. *Let U, V be separable Hilbert spaces, $\sigma \in \mathcal{G}^2_{\mathrm{b}}(U,V)$, and $p \in [2,\infty)$. For $Y^1, Y^2 \in \mathbf{M}^{1,p}(U)$ with $DY^1, DY^2 \in L^\infty(\Omega; L^p((0,T);\mathcal{L}(H_0,U)))$, it holds that*

$$\begin{aligned}\big\|\sigma(Y^1) - \sigma(Y^2)\big\|_{\mathbf{M}^{1,p}(V)^*} &\leqslant \max\big(|\sigma|_{\mathcal{G}^1_{\mathrm{b}}(U,V)}, |\sigma|_{\mathcal{G}^2_{\mathrm{b}}(U,V)}\big)\\ &\qquad\times\Big(1 + \sum_{i=1}^{2} \big|Y^i\big|_{\mathbf{M}^{1,\infty,p}(U)}\Big)\big\|Y^1 - Y^2\big\|_{\mathbf{M}^{1,p}(U)^*}.\end{aligned}$$

Lemma 2.1.4. *Let $p \in [2,\infty)$, $q \in [2,\infty]$. Then for all $S \in \mathcal{L}(H)$, $Y \in L^2(\Omega;H)$ it holds that $\|SY\|_{\mathbf{M}^{1,p,q}(H)^*} \leqslant \|S\|_{\mathcal{L}(H)}\|Y\|_{\mathbf{M}^{1,p,q}(H)^*}$.*

2.2. Existence, uniqueness and regularity

Throughout this section we assume that (2.1.1) and (2.1.5)–(2.1.3) hold with $\rho \in [1,2)$, $\beta \in (0, 1/\rho]$. We begin by proving existence, uniqueness, and Malliavin regularity of the solution of (2.1.2). Recall that two stochastic processes X^1, X^2 are modifications of each other if for all $t \in [0,T]$ it holds that $\mathbb{P}(X^1(t) \neq X^2(t)) = 0$.

Proposition 2.2.1. *There exists an, up to modification, unique stochastic process $X\colon [0,T]\times\Omega \to H$ such that $X \in \mathbb{C}((0,T), L^p(\Omega;H))$ for $p \in [2,\infty)$, $X \in \mathbb{C}((0,T), \mathbf{M}^{1,p,q}(H))$ for $p \in [2,\infty)$, $q \in [2, \frac{2}{1-\rho\beta})$, and which satisfies equation* (2.1.2) $\mathbb{P}$*-a.s..*

Proof. Existence is proved by a standard application of Banach's Fixed Point Theorem, see, e.g., [**56**, Theorem 1] or [**6**, Theorem 3.3]. We note that for proving existence and uniqueness in $C((0,T), L^p(\Omega;H))$ it is not crucial whether $(S(t))_{t\in[0,T]}$ is a semigroup or not. For the $C((0,T), \mathbf{M}^{1,p,q}(H))$ regularity, see Proposition 2.3.4 below. □

The next proposition states the temporal Hölder regularity of X in the $L^p(\Omega;H)$ and $\mathbf{M}^{1,p,q}(H)^*$ norms. Note that the Hölder exponent in the $\mathbf{M}^{1,p,q}(H)^*$ norm is twice that in the $L^p(\Omega;H)$ norm.

Proposition 2.2.2. *Let X be the solution to* (2.1.2)*. For $\gamma \in (0,\beta)$, $p \geqslant 2$, $q = \frac{2}{1-\rho\gamma}$, there exists $C > 0$ such that*

$$\begin{aligned}\big\|X(t_2) - X(t_1)\big\|_{L^p(\Omega;H)} &\leqslant C\big|t_2 - t_1\big|^{\frac{\rho\gamma}{2}}, \quad t_1, t_2 \in [0,T],\\ \big\|X(t_2) - X(t_1)\big\|_{\mathbf{M}^{1,p,q}(H)^*} &\leqslant C\big|t_2 - t_1\big|^{\rho\gamma}, \quad t_1, t_2 \in [0,T].\end{aligned}$$

Proof. Fix $\gamma \in (0,\beta)$, $p \geqslant 2$. In order to treat both cases simultaneously we define $V_2 = L^p(\Omega;H)$, $c_{p,2} = p(p-1)/2$, and $V_r = \mathbf{M}^{1,p,r}(H)^*$, $c_{p,r} = 1$ for $r \in (2,\infty]$. In view of (2.1.7) and (2.1.8) it holds that

(2.2.1)
$$\Big\|\int_0^T \Phi(t)\,\mathrm{d}W(t)\Big\|_{V_r} \leqslant c_{p,r}\big\|\Phi\big\|_{L^{r'}((0,T);\mathcal{L}^0_2)}, \quad \Phi \in L^2((0,T);\mathcal{L}^0_2),\ r \in [2,\infty],$$

where $\frac{1}{r}+\frac{1}{r'}=1$. Let $t_2>t_1$. The difference $X(t_2)-X(t_1)$ can be written in the form

$$\begin{aligned}X(t_2)-X(t_1) &= \big(S(t_2)-S(t_1)\big)x_0+\int_0^{t_1}\big(S(t_2-s)-S(t_1-s)\big)f(X(s))\,\mathrm{d}s\\ &\quad+\int_{t_1}^{t_2}S(t_2-s)f(X(s))\,\mathrm{d}s\\ &\quad+\int_0^{t_1}\big(S(t_2-s)-S(t_1-s)\big)\,\mathrm{d}W(s)+\int_{t_1}^{t_2}S(t_2-s)\,\mathrm{d}W(s).\end{aligned}$$

Taking V_r-norms, using the continuous embeddings $H\subset L^p(\Omega;H)\subset L^2(\Omega;H)\subset \mathbf{M}^{1,p,r}(H)^*$, yields

$$\begin{aligned}&\big\|X(t_2)-X(t_1)\big\|_{V_r}\leqslant \big\|\big(S(t_2)-S(t_1)\big)x_0\big\|\\ &\quad+\Big\|\int_0^{t_1}\big(S(t_2-s)-S(t_1-s)\big)f(X(s))\,\mathrm{d}s\Big\|_{L^p(\Omega;H)}\\ &\quad+\Big\|\int_{t_1}^{t_2}S(t_2-s)f(X(s))\,\mathrm{d}s\Big\|_{L^p(\Omega;H)}\\ &\quad+\Big\|\int_0^{t_1}\big(S(t_2-s)-S(t_1-s)\big)\,\mathrm{d}W(s)\Big\|_{V_r}+\Big\|\int_{t_1}^{t_2}S(t_2-s)\,\mathrm{d}W(s)\Big\|_{V_r}.\end{aligned}$$

First, by (2.1.1) and (2.1.3), we obtain

$$\big\|\big(S(t_2)-S(t_1)\big)x_0\big\|=\Big\|\int_{t_1}^{t_2}\dot S(t)A^{-\frac{1}{\rho}}A^{\frac{1}{\rho}}x_0\,\mathrm{d}t\Big\|\leqslant L_0\big\|A^{\frac{1}{\rho}}x_0\big\|(t_2-t_1).$$

It is straightforward to show that the terms containing f are bounded up to a constant by $|t_2-t_1|^{1-\varepsilon}$, and $|t_2-t_1|$ respectively, for every $\varepsilon\in(0,1)$. For the case $\rho=1$ see the proof of [**2**, Proposition 3.11].

By (2.2.1), (2.1.5), and (2.1.1) we get

$$\begin{aligned}&\Big\|\int_0^{t_1}\big(S(t_2-s)-S(t_1-s)\big)\,\mathrm{d}W(s)\Big\|_{V_r}\\ &\quad\leqslant c_{p,r}\Big(\int_0^{t_1}\big\|\big(S(t_2-s)-S(t_1-s)\big)A^{\frac{1-\beta\rho}{2\rho}}\big\|_{\mathcal{L}}^{r'}\big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}}\\ &\quad\leqslant c_{p,r}\big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}\Big(\int_0^{t_1}\Big(\int_{t_1}^{t_2}\|\dot S(t-s)A^{\frac{(3-\beta\rho)/2-1}{\rho}}\|_{\mathcal{L}}\,\mathrm{d}t\Big)^{r'}\mathrm{d}s\Big)^{\frac{1}{r'}}\\ &\quad\leqslant c_{p,r}\big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}L_{\frac{3-\beta\rho}{2}}\Big(\int_0^{t_1}\Big(\int_{t_1}^{t_2}(t-s)^{-\frac{3-\beta\rho}{2}}\,\mathrm{d}t\Big)^{r'}\mathrm{d}s\Big)^{\frac{1}{r'}}.\end{aligned}$$

Bounding the integrals yields, for $\eta\in(0,1/\rho)$ to be chosen,

$$\begin{aligned}&\Big(\int_0^{t_1}\Big(\int_{t_1}^{t_2}(t-s)^{-\frac{3-\beta\rho}{2}}\,\mathrm{d}t\Big)^{r'}\mathrm{d}s\Big)^{\frac{1}{r'}}\\ &\quad\leqslant\Big(\int_0^{t_1}\Big((t_1-s)^{-\frac{1-(\beta-2\eta)\rho}{2}}\int_{t_1}^{t_2}(t-t_1)^{-1+\eta\rho}\,\mathrm{d}t\Big)^{r'}\mathrm{d}s\Big)^{\frac{1}{r'}}\\ &\quad=\frac{(t_2-t_1)^{\eta\rho}}{\eta\rho}\Big(\int_0^{t_1}(t_1-s)^{-\frac{r}{r-1}\frac{1-(\beta-2\eta)\rho}{2}}\,\mathrm{d}s\Big)^{\frac{r-1}{r}}.\end{aligned}$$

dc_1541_18

For $r = q = 2/(1-\gamma\rho)$ and $\eta < (\beta+\gamma)/2$, the exponent is

$$\frac{r}{r-1}\frac{1-(\beta-2\eta)\rho}{2} = \frac{1-\beta\rho+2\eta\rho}{1+\rho\gamma} < 1.$$

In particular, we can take $\eta = \gamma$ as required since $\gamma < \beta$. For $r = 2$, the analogous condition is $\eta < \beta/2$ and we can take $\eta = \gamma/2$. Next, similarly,

$$\begin{aligned}\Big\| \int_{t_1}^{t_2} S(t_2-s)\,\mathrm{d}W(s) \Big\|_{V_r} &\leqslant c_{p,r} \Big(\int_{t_1}^{t_2} \big\|S(t_2-s)A^{\frac{1-\beta\rho}{2\rho}}\big\|_{\mathcal{L}}^{r'} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}^{r'}\,\mathrm{d}s\Big)^{\frac{1}{r'}} \\ &\leqslant c_{p,r} L_{\frac{1-\beta\rho}{2}} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}^{r'} \Big(\int_{t_1}^{t_2} (t_2-s)^{-\frac{r}{r-1}\frac{1-\beta\rho}{2}}\,\mathrm{d}s\Big)^{\frac{r-1}{r}} \\ &\leqslant c_{p,r} L_{\frac{1-\beta\rho}{2}} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}^{r'} (t_2-t_1)^{\frac{r-1}{r}-\frac{1-\beta\rho}{2}}.\end{aligned}$$

For $r = q = 2/(1-\gamma\rho)$ we have the Hölder exponent

$$\frac{r-1}{r} - \frac{1-\beta\rho}{2} = \frac{\rho(\beta+\gamma)}{2} > \gamma\rho,$$

and for $r = 2$ the Hölder exponents equals $\beta\rho/2 > \gamma\rho/2$. □

2.3. Weak and strong convergence

This section contains the main result of the chapter and its proof. Theorem 2.3.7 states a weak error estimate for abstractly defined approximations of quantities of the form $\mathbb{E}[\Phi(X)] = \mathbb{E}[\prod_{i=1}^K \varphi_i(\int_0^T X(t)\,\mathrm{d}\nu_t^i)]$ for $(\nu^i)_{i=1}^K \subset \mathcal{M}_T$, $(\varphi_i)_{i=1}^K \subset \mathcal{G}_{\mathrm{p}}^{2,m}(H,\mathbb{R})$, $m \geqslant 2$, and X being the solution to (2.1.2). Theorem 2.3.2 provides a strong error estimate for approximations of X. For parabolic problems, weak convergence, more precisely, convergence of approximations of $\mathbb{E}[\varphi(X(t))]$ for fixed $t \in [0,T]$ has been considered [**2**], and for Volterra equations in Subsection 1.5.7 but only in the linear case $f = 0$. The rate of convergence for $\mathbb{E}[\Phi(X)]$ is twice the strong rate as expected. We begin by presenting a family of abstractly defined approximations.

2.3.1. Approximation. Assume that (2.1.1), (2.1.5)–(2.1.3) hold. Suppose that $(V_h)_{h\in(0,1)}$ is a family of finite-dimensional subspaces of H and let $P_h\colon H \to V_h$ be the orthogonal projector. Let $\Delta t \in (0,1)$ and $t_n = n\Delta t$, $n = 0,\dots,N$, where $t_N < T \leqslant t_N + \Delta t$. Let $(B_{h,\Delta t})_{h,\Delta t\in(0,1)}$ be a family of operator-valued functions $B_{h,\Delta t}\colon \{0,\dots,N\} \to \mathcal{L}(H,V_h)$ such that $B_{h,\Delta t}^n = B_{h,\Delta t}^n P_h$, and let $(A_h)_{h\in(0,1)}$ be a collection of linear operators $A_h\colon V_h \to V_h$ such that for $n = 1,\dots,N$ it holds that

$$\big\|A_h^{\frac{s}{\rho}} B_{h,\Delta t}^n x\big\| \leqslant L_s t_n^{-s}\|x\|, \quad x \in H,\ 0 \leqslant s < 1, \tag{2.3.1}$$

with the same constants $(L_s)_{s\in[0,1]}$ as in (2.1.1). For other constants $(K_\varepsilon)_{\varepsilon\in(0,\infty)}$ and $(R_s)_{s\in[0,1]}$, let the corresponding error operator $(F_{h,\Delta t})_{h,\Delta t\in(0,1)}$, given by $F_{h,\Delta t}^n = B_{h,\Delta t}^n - S(t_n)$ for $n = 0,\dots,N$, satisfy the smooth data error estimate

$$\big\|F_{h,\Delta t}^n x\big\| \leqslant K_\varepsilon\big(h^\sigma + \Delta t^{\frac{\sigma}{2}}\big)\|x\|_{\dot H^{\sigma(1+\varepsilon)}}, \quad 0 \leqslant \sigma \leqslant 2,\ \varepsilon > 0, \tag{2.3.2}$$

and the non-smooth data error estimates, for $n = 1,\dots,N$, $t > 0$,

$$\big\|A^{\frac{s}{2\rho}} F_{h,\Delta t}^n x\big\| \leqslant R_s\big(h^{\frac{\sigma}{\rho}} + \Delta t^{\frac{\sigma}{2}}\big)t_n^{-\frac{\sigma+s}{2}}\|x\|, \quad 0 \leqslant \sigma \leqslant 2,\ 0 \leqslant s \leqslant 1-\sigma/2, \tag{2.3.3}$$

$$\big\|\big(e^{-tA_h}P_h - e^{-tA}\big)x\big\| \leqslant R_0 h^\sigma t^{-\frac{\sigma}{2}}\|x\|, \quad 0 \leqslant \sigma \leqslant 2, \tag{2.3.4}$$

where $(e^{-tA})_{t\geqslant 0}$ and $(e^{-tA_h})_{t\geqslant 0}$ are the analytic semigroups generated by $-A$ and $-A_h$, respectively. As before, we introduce the piecewise continuous operator function $\tilde{F}_{h,\Delta t}\colon [0,T] \to \mathcal{L}(H)$ given by $\tilde{F}_{h,\Delta t}(t) = B^n_{h,\Delta t} - S(t)$ for $t \in (t_{n-1}, t_n]$ and $n = 1, \dots N$ with $\tilde{F}_{h,\Delta t}(0) = F^0_{h,\Delta t}$. By (2.1.1) and (2.3.2) the family $(\tilde{F}_{h,\Delta t}(t))_{t\in[0,T]}$ satisfies for $t \in (0,T]$ the bound

$$\big\|A^{\frac{s}{2\rho}}\tilde{F}_{h,\Delta t}(t)\big\| \leqslant R_s\big(h^{\frac{\sigma}{\rho}} + \Delta t^{\frac{\sigma}{2}}\big)t^{-\frac{\sigma+s}{2}}, \quad 0 \leqslant \sigma \leqslant 2,\ 0 \leqslant s \leqslant 1 - \sigma/2. \tag{2.3.5}$$

The discrete and continuous stochastic convolutions are defined by

$$W_S(t) = \int_0^t S(t-s)\,\mathrm{d}W(s), \quad t \in [0,T]; \quad W^n_{B_{h,\Delta t}} = \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\,\mathrm{d}W(t),$$
$$n = 1, \dots, N.$$

We now define approximations of equation (2.1.2). For $h, \Delta t \in (0,1)$, let $(X^n_{h,\Delta t})^N_{n=0}$ be given by

$$X^n_{h,\Delta t} = B^n_{h,\Delta t}x_0 + \Delta t\sum_{j=0}^{n-1} B^{n-j}_{h,\Delta t}f(X^j_{h,\Delta t}) + W^n_{B_{h,\Delta t}}, \quad n = 1, \dots, N, \tag{2.3.6}$$

and $X^0_{h,\Delta t} = B^0_{h,\Delta t}x_0$.

2.3.2. Strong convergence. Boundedness in the $L^p(\Omega; H)$-sense of the approximate family $(X^n_{h,\Delta t})^N_{n=0}$ is stated in the next proposition. For a proof in the parabolic case, i.e., for $\rho = 1$, see [**2**, Proposition 3.15]. The general case is proved in the same way but using the different smoothing property in (2.3.1).

PROPOSITION 2.3.1. *Let the setting of Section 2.3.1 hold. For $p \geqslant 2$ it holds that*

$$\sup_{h,\Delta t\in(0,1)}\max_{n\in\{0,\dots,N\}}\big\|X^n_{h,\Delta t}\big\|_{L^p(\Omega;H)} < \infty.$$

We next prove strong convergence. This is interesting in itself, but it is also used in our proof of the Malliavin regularity of X in Proposition 2.3.4.

THEOREM 2.3.2. *Let the setting of Section 2.3.1 hold, let X be the solution of* (2.1.2) *and let $(X_{h,\Delta t})_{h,\Delta t\in(0,1]}$ be given by* (2.3.6). *For $\gamma \in [0,\beta)$, $p \in [2,\infty)$, there exists $C > 0$ such that*

$$\max_{n\in\{0,\dots,N\}}\big\|X^n_{h,\Delta t} - X(t_n)\big\|_{L^p(\Omega;H)} \leqslant C\big(h^\gamma + \Delta t^{\frac{\rho\gamma}{2}}\big), \quad h, \Delta t \in (0,1).$$

PROOF. For $n = 0$ the estimate holds by (2.3.2). For $n \geqslant 1$, we take the difference of (2.1.2) and (2.3.6) to obtain the equation for the error:

(2.3.7)

$$X^n_{h,\Delta t} - X(t_n) = \big(B^n_{h,\Delta t} - S(t_n)\big)x_0 + \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\big(B^{n-j}_{h,\Delta t} - S(t_n - t)\big)f(X(t))\,\mathrm{d}t$$
$$+ \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\,\mathrm{d}t + W^n_{B_{h,\Delta t}} - W_S(t_n).$$

The deterministic nature of the first two terms allows us to obtain twice the rate of convergence compared to the other terms. This will be used later in the proof of Lemma 2.3.6. Recall that

$$\tilde{F}_{h,\Delta t}(t) = B^k_{h,\Delta t} - S(t)$$

for $t \in (t_{k-1}, t_k]$ and $n = 1, \dots, N$. Then, we get

$$\begin{aligned}\big\|X^n_{h,\Delta t} - X(t_n)\big\|_{L^p(\Omega;H)} \leqslant \big\|F^n_{h,\Delta t}x_0\big\|_H &+ \Big\|\int_0^{t_n} \tilde{F}_{h,\Delta t}(t_n - t)f(X(t))\,\mathrm{d}t\Big\|_{L^p(\Omega;H)} \\ &+ \Big\|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\big(f(X^j_{h,k}) - f(X(t))\big)\,\mathrm{d}t\Big\|_{L^p(\Omega;H)} \\ &+ \big\|W^n_{B_{h,\Delta t}} - W_S(t_n)\big\|_{L^p(\Omega;H)}.\end{aligned}$$

Using (2.1.3), (2.3.2) with $\sigma = 2\rho\gamma$, $\varepsilon = (3 - 2\gamma\rho)/2\gamma\rho$ we obtain

$$\max_{n\in\{0,\dots,N\}} \big\|F^n_{h,\Delta t}x_0\big\| \leqslant K_{\frac{3-2\gamma\rho}{2\gamma\rho}}\big(h^{2\rho\gamma} + \Delta t^{\rho\gamma}\big)\|x_0\|_{\dot{H}^3}. \tag{2.3.8}$$

By Proposition 2.2.1, (2.1.4), (2.3.5) it holds that

$$\begin{aligned}&\Big\|\int_0^{t_n} \tilde{F}_{h,\Delta t}(t_n - t)f(X(t))\,\mathrm{d}t\Big\|_{L^p(\Omega;H)} \\ &\leqslant \int_0^{t_n} \big\|\tilde{F}_{h,\Delta t}(t_n - t)\big\|\big\|f(X(t))\big\|_{L^p(\Omega;H)}\,\mathrm{d}t \\ &\leqslant R_0\big(h^{2\gamma} + \Delta t^{\rho\gamma}\big)|f|_{\mathcal{G}^1_{\mathrm{b}}(H,H)}\Big(1 + \sup_{t\in[0,T]}\big\|X(t)\big\|_{L^p(\Omega;H)}\Big)\int_0^{t_n}(t_n - t)^{-\rho\gamma}\,\mathrm{d}t \\ &\leqslant C(h^{2\gamma} + \Delta t^{\rho\gamma}).\end{aligned} \tag{2.3.9}$$

Using (2.1.4), (2.1.6), (2.3.1), and Proposition 2.2.2 yields

$$\begin{aligned}&\Big\|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\,\mathrm{d}t\Big\|_{L^p(\Omega;H)} \\ &\leqslant |f|_{\mathcal{G}^1_{\mathrm{b}}(H,H)}\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \big\|B^{n-j}_{h,\Delta t}\big\|\big\|X^j_{h,\Delta t} - X(t)\big\|_{L^p(\Omega;H)}\,\mathrm{d}t \\ &\leqslant L_0|f|_{\mathcal{G}^1_{\mathrm{b}}(H,H)}\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\Big(\big\|X^j_{h,\Delta t} - X(t_j)\big\|_{L^p(\Omega;H)} \\ &\quad + \big\|X(t_j) - X(t)\big\|_{L^p(\Omega;H)}\Big)\,\mathrm{d}t \\ &\leqslant L_0|f|_{\mathcal{G}^1_{\mathrm{b}}(H,H)}\Big(CT\Delta t^{\frac{\rho\gamma}{2}} + \Delta t\sum_{j=0}^{n-1}\big\|X^j_{h,\Delta t} - X(t_j)\big\|_{L^p(\Omega;H)}\Big).\end{aligned}$$

For the error of the stochastic convolution we write the difference in the form

$$\begin{aligned}W^n_{B_{h,\Delta t}} - W_S(t_n) &= \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}}\big(B^{n-j}_{h,\Delta t} - S(t_n - t)\big)\,\mathrm{d}W(t) \\ &= \int_0^{t_n} \tilde{F}_{h,\Delta t}(t_n - t)\,\mathrm{d}W(t).\end{aligned} \tag{2.3.10}$$

By (2.1.7) and (2.3.5) with $\sigma = \gamma\rho$, and $s = 1 - \beta\rho$, we obtain the estimate

$$\begin{aligned}\big\|W^n_{B_{h,\Delta t}} - W_S(t_n)\big\|_{L^p(\Omega;H)} &\leqslant \Big(\frac{p(p-1)}{2}\int_0^{t_n} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|^2_{\mathcal{L}_2^0}\big\|A^{\frac{1-\beta\rho}{2\rho}}\tilde{F}_{h,\Delta t}(t)\big\|^2_{\mathcal{L}}\,\mathrm{d}t\Big)^{\frac12}\\ &\leqslant CR_{1-\beta\rho}\Big(\int_0^{t_n} t^{\rho(\beta-\gamma)-1}\,\mathrm{d}t\Big)^{\frac12}\big(h^\gamma + \Delta t^{\frac{\rho\gamma}{2}}\big) \leqslant C(h^\gamma + \Delta t^{\frac{\rho\gamma}{2}}).\end{aligned}$$

Collecting the estimates yields that, for all $n = 0, \ldots, N$,

$$\big\|X^n_{h,\Delta t} - X(t_n)\big\|_{L^p(\Omega;H)} \leqslant C\left(h^\gamma + \Delta t^{\frac{\rho\gamma}{2}} + \Delta t\sum_{j=0}^{n-1}\big\|X^j_{h,\Delta t} - X(t_j)\big\|_{L^p(\Omega;H)}\right).$$

The proof is concluded by using Gronwall's lemma. ☐

2.3.3. Regularity and weak convergence. Here we state and prove our main result on weak convergence. It is based on a strong error estimate in the $\mathbf{M}^{1,p}(H)^*$ norm combined with boundedness of X and $X_{h,\Delta t}$ in $\mathbf{M}^{1,p,q}(H)$ for suitable p, q. The methodology was introduced in [**2**], but here we exploit it further in a more general setting. We begin by proving the Malliavin differentiability of $X_{h,\Delta t}$.

PROPOSITION 2.3.3. *Let the setting of Section 2.3.1 hold, and let $X_{h,\Delta t}$ be given by* (2.3.6). *For $p \in [2,\infty)$, $q \in [2, \frac{2}{1-\rho\beta})$, it holds that*

$$\sup_{h,\Delta t\in(0,1)}\max_{n\in\{0,\ldots,N\}}\Big(\big\|X^n_{h,\Delta t}\big\|_{\mathbf{M}^{1,p,q}(H)} + \big|X^n_{h,\Delta t}\big|_{\mathbf{M}^{1,\infty,q}(H)}\Big) < \infty.$$

SKETCH OF PROOF. Note first that $DX^0_{h,\Delta t} = 0$ as $X^0_{h,\Delta t}$ is deterministic. Therefore, it follows inductively that $X^j_{h,\Delta t}$, $j = 0, \ldots, N$, are differentiable and the derivative satisfies the equation

$$D_r X^n_{h,\Delta t} = \Delta t\sum_{j=0}^{n-1} B^{n-j}_{h,\Delta t} f'(X^j_{h,\Delta t}) D_r X^j_{h,\Delta t} + \sum_{j=0}^{n-1}\chi_{[t_j,t_{j+1})}(r) B^{n-j}_{h,\Delta t}.$$

The proof is performed by straightforward analysis of this equation using the discrete Gronwall's lemma, see [**2**, Proposition 3.16] for details in the parabolic case $\rho = 1$. The general case is treated analogously. ☐

The Malliavin regularity of X is next obtained by a limiting procedure.

PROPOSITION 2.3.4. *Let the setting of Section 2.3.1 hold and let X be the solution to* (2.1.2)*. For $p \in [2,\infty)$, $q \in [2, \frac{2}{1-\rho\beta})$, it holds that $X \in C((0,T), \mathbf{M}^{1,p,q}(H))$. Furthermore, we have that*

$$\sup_{t\in[0,T]}\big|X(t)\big|_{\mathbf{M}^{1,\infty,q}(H)} < \infty.$$

PROOF. Let $\tilde{X}_{h,\Delta t}(t) = X^n_{h,\Delta t}$ for $t \in [t_n, t_{n+1})$, $n = 0, \ldots, N-1$, $h, \Delta t \in (0,1)$. By Proposition 2.3.3 it holds, in particular, that the family $(\tilde{X}_{h,\Delta t})_{h,\Delta t\in(0,1)}$ is bounded in the Hilbert space $\mathcal{X} = L^2((0,T); \mathbf{M}^{1,2,2}(H))$. From Proposition 2.2.2 and Theorem 2.3.2 it follows that $\tilde{X}_{h,\Delta t} \to X$ as $h, \Delta t \to 0$ in the Hilbert space $\mathcal{Y} = L^2((0,T); L^2(\Omega; H))$. Therefore, by Lemma 2.1.1, $X \in \mathcal{X} = L^2((0,T); \mathbf{M}^{1,2,2}(H))$.

dc_1541_18

By [**41**, Lemma 3.6] it holds that also

$$\int_0^{\cdot} S(\cdot - s) f(X(s))\,\mathrm{d}s \in L^2((0,T);\mathbf{M}^{1,2,2}(H))$$

with $D_r \int_0^t S(t-s)f(X(s))\,\mathrm{d}s = \int_r^t S(t-s)f'(X(s))D_rX(s)\,\mathrm{d}s$, for $0 \leqslant r \leqslant t \leqslant T$, and $\int_0^{\cdot} S(\cdot - s)\,\mathrm{d}W(s) \in L^2(0,T;\mathbf{M}^{1,2,2}(H))$ with $D_r \int_0^t S(t-s)\,\mathrm{d}W(s) = S(t-r)$, for $0 \leqslant r \leqslant t \leqslant T$. We remark that [**41**, Lemma 3.6] is formulated for semigroups, but the semigroup property is not used in the proof. We have thus proved that we can differentiate the equation for X term by term, and obtain the equation

$$D_rX(t) = \begin{cases} S(t-r) + \int_r^t S(t-s)f'(X(s))D_rX(s)\,\mathrm{d}s, & t \in (r,T], \\ 0, & t \in [0,r]. \end{cases}$$

A straightforward analysis of this equation, by a Gronwall argument, as in the proof of [**2**, Proposition 3.10] completes the proof. □

In the proof of [**2**, Lemma 4.6], which is the analogue of Lemma 2.3.6 below, a bound

$$\|A_h^{-\frac{\delta}{2}} P_h x\| \leqslant \|A_h^{\frac{\delta}{2}} P_h A^{-\frac{\delta}{2}}\| \|A^{-\frac{\delta}{2}} x\| \leqslant C\|A^{-\frac{\delta}{2}} x\|, \tag{2.3.11}$$

was used in the special case $\delta = 1$. This estimate is true for all $\delta \in [0,1]$ for both the finite element method and for spectral approximation. For $\delta > 1$ it holds only for spectral approximation. As we need $\delta \in [0, 2/\rho)$ we cannot rely on (2.3.11). In [**103**, Lemma 5.3] it is shown that for finite element discretization and for $\delta = 0,1,2$ it holds

$$\|A_h^{-\frac{\delta}{2}} P_h x\| \leqslant C\big(\|A^{-\frac{\delta}{2}} x\| + h^{\delta}\|x\|\big), \quad x \in H.$$

The next lemma is a generalization of this result, assuming the availability of a non-smooth data estimate of the form (2.3.4). It will be used in the proof of Lemma 2.3.6 below with $\mathcal{X} = \mathbf{M}^{1,p}(H)^*$ for a certain p. By using it we don't have to rely on (2.3.11) and this way we may include finite element discretizations under the same generality as spectral approximations.

Lemma 2.3.5. *Let the setting of Section 2.3.1 hold and let $\mathcal{X}$ be a Banach space such that the embedding $L^2(\Omega;H) \subset \mathcal{X}$ is continuous. For $\kappa \in [0,2)$, $\sigma \in [0,\kappa)$, there exists $C > 0$ such that for $Y \in L^2(\Omega;H)$ it holds that*

$$\big\|A_h^{-\frac{\kappa}{2}} P_h Y\big\|_{\mathcal{X}} \leqslant \big\|A^{-\frac{\kappa}{2}} Y\big\|_{\mathcal{X}} + Ch^{\sigma}\big\|Y\big\|_{L^2(\Omega;H)}, \quad h \in (0,1).$$

Proof. By the continuous embedding $L^2(\Omega;H) \subset \mathcal{X}$ we get that

$$\begin{aligned} \big\|A_h^{-\frac{\kappa}{2}} P_h Y\big\|_{\mathcal{X}} &\leqslant \big\|A^{-\frac{\kappa}{2}} Y\big\|_{\mathcal{X}} + \big\|\big(A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}}\big) Y\big\|_{\mathcal{X}} \\ &\leqslant \big\|A^{-\frac{\kappa}{2}} Y\big\|_{\mathcal{X}} + C\big\|A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}}\big\|_{\mathcal{L}(H)}\big\|Y\big\|_{L^2(\Omega;H)}. \end{aligned}$$

By [**87**, Chapter 2, (6.9)] we have that

$$A_h^{-\frac{\kappa}{2}} P_h - A^{-\frac{\kappa}{2}} = \frac{1}{\Gamma(\kappa/2)} \int_0^{\infty} t^{\frac{\kappa}{2}-1}\big(e^{-tA_h}P_h - e^{-tA}\big)\,\mathrm{d}t.$$

dc_1541_18

Therefore, by (2.3.4),

$$\begin{aligned}\|A_h^{-\frac{\kappa}{2}}P_h - A^{-\frac{\kappa}{2}}\|_{\mathcal{L}(H)} &\leqslant \frac{1}{\Gamma(\kappa/2)}\int_0^\infty t^{\frac{\kappa}{2}-1}\|e^{-tA_h}P_h - e^{-tA}\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&= \frac{1}{\Gamma(\kappa/2)}\int_0^{h^{-2}} t^{\frac{\kappa}{2}-1}\|e^{-tA_h}P_h - e^{-tA}\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\quad + \frac{1}{\Gamma(\kappa/2)}\int_{h^{-2}}^\infty t^{\frac{\kappa}{2}-1}\|e^{-tA_h}P_h - e^{-tA}\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\leqslant C\left(h^{\frac{\kappa+\sigma}{2}}\int_0^{h^{-2}} t^{\frac{\kappa-\sigma}{4}-1}\,\mathrm{d}t + h^2\int_{h^{-2}}^\infty t^{\frac{\kappa}{2}-2}\,\mathrm{d}t\right) = C\left(\frac{4h^\sigma}{\kappa-\sigma} + \frac{2}{2-\kappa}h^2h^{2-\frac{\kappa}{2}}\right) \\
&\leqslant Ch^\sigma.\end{aligned}$$

□

The next result is a strong error estimate in the $\mathbf{M}^{1,p}(H)^*$ norm. Together with the regularity stated in Propositions 2.3.3 and 2.3.4 it is the key to the proof of Theorem 2.3.7 below on weak convergence.

Lemma 2.3.6. *Let the setting of Section 2.3.1 hold, let X be the solution of* (2.1.2), *and let $X_{h,\Delta t}$ be given by* (2.3.6). *For $\gamma \in [0,\beta)$, $p = \frac{2}{1-\rho\gamma}$, there exists $C > 0$ such that*

$$\max_{n\in\{0,\dots,N\}} \|X^n_{h,\Delta t} - X(t_n)\|_{\mathbf{M}^{1,p}(H)^*} \leqslant C\big(h^{2\gamma} + \Delta t^{\rho\gamma}\big), \quad h, \Delta t \in (0,1).$$

Proof. The proof is similar to that of Theorem 2.3.2. First note that we have the continuous embeddings $H \subset L^p(\Omega;H) \subset L^2(\Omega;H) \subset \mathbf{M}^{1,p}(H)^*$. Therefore, for $n = 0$ the statement follows from (2.3.2). For $n = 1,\dots N$ it follows from (2.3.7) that

$$\begin{aligned}\|X^n_{h,\Delta t} - X(t_n)\|_{\mathbf{M}^{1,p}(H)^*} &\leqslant \|F^n_{h,\Delta t}x_0\|_H + \Big\|\int_0^{t_n}\tilde{F}_{h,\Delta t}(t_n-t)f(X(t))\,\mathrm{d}t\Big\|_{L^p(\Omega;H)} \\
&\quad + \Big\|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\,\mathrm{d}t\Big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\quad + \|W^n_{B_{h,\Delta t}} - W_S(t_n)\|_{\mathbf{M}^{1,p}(H)^*}.\end{aligned}$$

The first two terms was already estimated as desired in (2.3.8) and (2.3.9). Choose κ so that $\max(\delta, 2\gamma) < \kappa < 2/\rho$, where δ is the parameter in (2.1.4). Since $\rho\kappa < 2$, we have, by Lemma 2.1.4 and (2.3.1) with $s = \rho\kappa/2$, that

$$\begin{aligned}&\Big\|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\,\mathrm{d}t\Big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\quad\leqslant \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \|B^{n-j}_{h,\Delta t}A_h^{\frac{\kappa}{2}}P_h\|_{\mathcal{L}(H)}\|A_h^{-\frac{\kappa}{2}}P_h\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\|_{\mathbf{M}^{1,p}(H)^*}\,\mathrm{d}t \\
&\quad\leqslant L_{\frac{\kappa\rho}{2}}\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} t_{n-j}^{-\frac{\kappa\rho}{2}}\|A_h^{-\frac{\kappa}{2}}P_h\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\|_{\mathbf{M}^{1,p}(H)^*}\,\mathrm{d}t.\end{aligned}$$

dc_1541_18

Applying Lemma 2.3.5 with $\mathcal{X} = \mathbf{M}^{1,p}(H)^*$ and $\sigma = 2\gamma < \kappa$ yields

$$\begin{aligned}&\big\|A_h^{-\frac{\kappa}{2}} P_h\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\big\|_{\mathbf{M}^{1,p}(H)^*}\\&\quad\leqslant Ch^{2\gamma}\big\|f(X^j_{h,\Delta t}) - f(X(t))\big\|_{L^2(\Omega;H)} + \big\|A^{-\frac{\kappa}{2}}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\big\|_{\mathbf{M}^{1,p}(H)^*}.\end{aligned}$$

For the first term we get by (2.1.4), Propositions 2.2.1, and 2.3.1 that

$$\begin{aligned}&\sup_{t\in[0,T]}\max_{j\in\{0,\dots,N\}}\big\|f(X^j_{h,\Delta t}) - f(X(t))\big\|_{L^2(\Omega;H)}\\&\quad\leqslant |f|_{\mathcal{G}^1_{\mathrm{b}}(H,H)}\Big(\max_{j\in\{0,\dots,N\}}\big\|X^j_{h,\Delta t}\big\|_{L^2(\Omega;H)} + \sup_{t\in[0,T]}\big\|X(t)\big\|_{L^2(\Omega;H)}\Big) < \infty.\end{aligned}$$

By duality in the Gelfand triple $\mathbf{M}^{1,p}(\dot H^{-\delta}) \subset L^2(\Omega;\dot H^{-\delta}) \subset \mathbf{M}^{1,p}(\dot H^{-\delta})^*$ we compute that for $Y \in L^2(\Omega;\dot H^{-\delta})$,

$$\begin{aligned}\|Y\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})^*} &= \sup_{Z\in\mathbf{M}^{1,p}(\dot H^{-\delta})}\frac{\big\langle Z, Y\big\rangle_{L^2(\Omega;\dot H^{-\delta})}}{\|Z\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})}}\\&= \sup_{Z\in\mathbf{M}^{1,p}(\dot H^{-\delta})}\frac{\big\langle A^{-\frac{\delta}{2}}Z, A^{-\frac{\delta}{2}}Y\big\rangle_{L^2(\Omega;H)}}{\|Z\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})}}\\&= \sup_{Z\in\mathbf{M}^{1,p}(\dot H^{-\delta})}\frac{\big\langle Z, A^{-\frac{\delta}{2}}Y\big\rangle_{L^2(\Omega;H)}}{\|A^{\frac{\delta}{2}}Z\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})}}\\&= \sup_{Z\in\mathbf{M}^{1,p}(H)}\frac{\big\langle Z, A^{-\frac{\delta}{2}}Y\big\rangle}{\|Z\|_{\mathbf{M}^{1,p}(H)}} = \|A^{-\frac{\delta}{2}}Y\|_{\mathbf{M}^{1,p}(H)^*}.\end{aligned}$$

Therefore, by Lemma 2.1.4 and Lemma 2.1.3 applied with $U = H$, $V = \dot H^{-\delta}$, $\sigma = f$ we get

$$\begin{aligned}&\big\|A^{-\frac{\kappa}{2}}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\big\|_{\mathbf{M}^{1,p}(H)^*}\\&\quad\leqslant \big\|A^{-\frac{\kappa-\delta}{2}}\big\|_{\mathcal{L}(H)}\big\|A^{-\frac{\delta}{2}}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\big\|_{\mathbf{M}^{1,p}(H)^*}\\&\quad= \big\|A^{-\frac{\kappa-\delta}{2}}\big\|_{\mathcal{L}(H)}\big\|f(X^j_{h,\Delta t}) - f(X(t))\big\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})^*}\\&\quad\leqslant \big\|A^{-\frac{\kappa-\delta}{2}}\big\|_{\mathcal{L}(H)}\max\Big(|f|_{\mathcal{G}^1_{\mathrm{b}}(H,\dot H^{-\delta})}, |f|_{\mathcal{G}^2_{\mathrm{b}}(H,\dot H^{-\delta})}\Big)\\&\qquad\times\Big(1 + \sup_{j\in\{0,\dots,N\}}\big|X^j_{h,\Delta t}\big|_{\mathbf{M}^{1,\infty,p}(H)} + \sup_{t\in[0,T]}\big|X(t)\big|_{\mathbf{M}^{1,\infty,p}(H)}\Big)\\&\qquad\times\Big(\big\|X^j_{h,\Delta t} - X(t_j)\big\|_{\mathbf{M}^{1,p}(H)^*} + \big\|X(t_j) - X(t)\big\|_{\mathbf{M}^{1,p}(H)^*}\Big).\end{aligned}$$

By Propositions 2.2.2, 2.3.3 and 2.3.4, we conclude that

$$\big\|A^{-\frac{\kappa}{2}}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\big\|_{\mathbf{M}^{1,p}(H)^*} \leqslant C\Big(\Delta t^{\rho\gamma} + \big\|X^j_{h,\Delta t} - X(t_j)\big\|_{\mathbf{M}^{1,p}(H)^*}\Big).$$

Thus,

$$\begin{aligned}&\Big\|\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} B^{n-j}_{h,\Delta t}\big(f(X^j_{h,\Delta t}) - f(X(t))\big)\,\mathrm{d}t\Big\|_{\mathbf{M}^{1,p}(H)^*}\\&\quad\leqslant C\left(h^{2\gamma} + \Delta t^{\rho\gamma} + \Delta t\sum_{j=0}^{n-1} t_{n-j}^{-\frac{\kappa\rho}{2}}\big\|X^j_{h,\Delta t} - X(t_j)\big\|_{\mathbf{M}^{1,p}(\dot H^{-\delta})^*}\right).\end{aligned}$$

By (2.3.10), (2.1.8), and (2.3.5), with $s = 1 - \beta\rho$, $\sigma = 2\gamma\rho$, and since $p = \frac{2}{1-\rho\gamma}$ and $p' = \frac{2}{1+\rho\gamma}$, we get

$$\begin{aligned}
&\big\|W^n_{B_{h,\Delta t}} - W_S(t_n)\big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\leqslant \Big(\int_0^{t_n} \big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}^{\frac{2}{1+\rho\gamma}} \big\|A^{\frac{1-\beta\rho}{2\rho}}\tilde{F}_{h,\Delta t}(t)\big\|_{\mathcal{L}}^{\frac{2}{1+\rho\gamma}}\,\mathrm{d}t\Big)^{\frac{1+\rho\gamma}{2}} \\
&\leqslant R_{1-\beta\rho}\big\|A^{\frac{\beta\rho-1}{2\rho}}\big\|_{\mathcal{L}_2^0}\Big(\int_0^{t_n} t^{\frac{\rho(\beta-\gamma)}{1+\rho\gamma}-1}\,\mathrm{d}t\Big)^{\frac{1+\rho\gamma}{2}}\big(h^{2\gamma} + \Delta t^{\rho\gamma}\big).
\end{aligned}$$

Altogether we have that for every $n = 0, \dots, N$ it holds that

$$\begin{aligned}
&\big\|X^n_{h,\Delta t} - X(t_n)\big\|_{\mathbf{M}^{1,p}(H)^*} \\
&\qquad\leqslant C\left(h^{2\gamma} + \Delta t^{\rho\gamma} + \Delta t\sum_{j=0}^{n-1} t_{n-j}^{-\frac{\kappa\rho}{2}}\big\|X^j_{h,\Delta t} - X(t_j)\big\|_{\mathbf{M}^{1,p}(H)^*}\right).
\end{aligned}$$

Lemma 2.1.2 finishes the proof. □

We next state our main result on weak convergence. Note that this is a more general type of weak convergence result than the ones in the previous chapter which concern the convergence of $|\mathbb{E}[\varphi(X^n_{h,\Delta t}) - \varphi(X(t_n))]|$ for fixed $t_n \in [0,T]$. This is a special case of the following theorem.

THEOREM 2.3.7. *Let X be the solution of* (2.1.2) *and let $X_{h,\Delta t}$ be given by* (2.3.6). *Let $\tilde{X}_{h,\Delta t}(t) = X^n_{h,\Delta t}$, for $t \in [t_n, t_{n+1})$, $n \in \{0,\dots,N-1\}$ and $\tilde{X}_{h,\Delta t}(t) = X^N_{h,\Delta t}$, for $t \in [t_N, T]$. For $K \geqslant 1$, $m_1, \dots, m_K \geqslant 2$, $\varphi_i \in \mathcal{G}^{2,m_i}_{\mathrm{p}}(H,\mathbb{R})$, $\nu_i \in \mathcal{M}_T$, $i = 1,\dots,K$, $\Phi(Z) = \prod_{i=1}^K \varphi_i(\int_0^T Z(t)\,\mathrm{d}\nu_{i,t})$, $\gamma \in [0,\beta)$, there exists $C > 0$ such that*

$$\big|\mathbb{E}\big[\Phi(\tilde{X}_{h,\Delta t}) - \Phi(X)\big]\big| \leqslant C\big(h^{2\gamma} + \Delta t^{\rho\gamma}\big), \quad h, \Delta t \in (0,1).$$

PROOF. We start by observing that by (2.1.6) we have

$$\begin{aligned}
&\prod_{i=1}^K \varphi_i(x_i) - \prod_{i=1}^K \varphi_i(y_i) \\
&\quad= \sum_{l=1}^K \prod_{i=1}^{l-1}\varphi_i(x_i)\prod_{j=l+1}^K \varphi_j(y_j)\big(\varphi_l(x_l) - \varphi_l(y_l)\big) \\
&\quad= \sum_{l=1}^K \left\langle \prod_{i=1}^{l-1}\varphi_i(x_i)\prod_{j=l+1}^K \varphi_j(y_j)\int_0^1 \varphi_l'(y_l + \lambda(x_l - y_l))\,\mathrm{d}\lambda, x_l - y_l\right\rangle \\
&\quad=: \sum_{l=1}^K \langle \gamma_l(x_1,\dots,x_l,y_l,\dots,y_K), x_l - y_l\rangle.
\end{aligned}$$

Here we use the convention that an empty product equals 1. We get

$$\big|\mathbb{E}\big[\Phi(\tilde{X}_{h,\Delta t}) - \Phi(X)\big]\big| = \Big|\sum_{l=1}^K \Big\langle \gamma_l(Y^l_{h,\Delta t}), \int_0^T \big(\tilde{X}_{h,\Delta t}(t) - X(t)\big)\,\mathrm{d}\nu_{l,t}\Big\rangle_{L^2(\Omega;H)}\Big|,$$

where

$$Y^l_{h,\Delta t} = \Big(\int_0^T \tilde{X}_{h,\Delta t}(t)\,\mathrm{d}\nu_{l,t}, \dots, \int_0^T \tilde{X}_{h,\Delta t}(t)\,\mathrm{d}\nu_{K,t}, \\ \int_0^T X(t)\,\mathrm{d}\nu_{1,t}, \dots, \int_0^T X(t)\,\mathrm{d}\nu_{l,t}\Big).$$

By duality in the Gelfand triple $\mathbf{M}^{1,p}(H) \subset L^2(\Omega; H) \subset \mathbf{M}^{1,p}(H)^*$ we obtain

$$\begin{aligned} &\big|\mathbb{E}\big[\Phi(\tilde{X}_{h,\Delta t}) - \Phi(X)\big]\big| \\ &\quad \leqslant \sum_{l=1}^K \big\|\gamma_l(Y^l_{h,\Delta t})\big\|_{\mathbf{M}^{1,p}(H)} \Big\| \int_0^T \big(\tilde{X}_{h,\Delta t}(t) - X(t)\big)\,\mathrm{d}\nu_{l,t}\Big\|_{\mathbf{M}^{1,p}(H)^*} \\ &\quad \leqslant \sum_{l=1}^K \Big(\sup_{h,\Delta t \in (0,1)} \big\|\gamma_l(Y^l_{h,\Delta t})\big\|_{\mathbf{M}^{1,p}(H)}\Big) \big\|\tilde{X}_{h,\Delta t} - X\big\|_{L^1_{\nu_l}((0,T);\mathbf{M}^{1,p}(H)^*)}. \end{aligned}$$

Here $\gamma_l \in \mathcal{G}^{1,r}_{\mathrm{p}}(H^{K+1}; H)$ and $Y^l_{h,\Delta t} \in \mathbf{M}^{1,rp}(H^{K+1})$ with $r = \sum_{i=1}^K m_i - 1$. Therefore [**2**, Lemma 3.3] applied with $U = H^{K+1}$ and $V = H$ gives for $l \in \{1, \dots, K\}$ the bound

$$\sup_{h,\Delta t \in (0,1)} \big\|\gamma_l(Y^l_{h,\Delta t})\big\|_{\mathbf{M}^{1,p}(H)} \leqslant C_l \Big(1 + \sup_{h,\Delta t \in (0,1)} \big\|Y^l_{h,\Delta t}\big\|^r_{\mathbf{M}^{1,rp}(H^{K+1})}\Big).$$

Propositions 2.3.3 and 2.3.4 ensure that

$$\begin{aligned} &\sup_{h,\Delta t \in (0,1)} \big\|\gamma_l(Y^l_{h,\Delta t})\big\|_{\mathbf{M}^{1,p}(H)} \\ &\leqslant \tilde{C}_l \Big(1 + \sum_{i=1}^K \Big(\sup_{h,\Delta t \in (0,1)} \big\|\tilde{X}_{h,\Delta t}\big\|^r_{L^1_{\nu_i}((0,T);\mathbf{M}^{1,rp,p}(H))} + \big\|X\big\|^r_{L^1_{\nu_i}((0,T);\mathbf{M}^{1,rp,p}(H))}\Big)\Big) \\ &< \infty. \end{aligned}$$

Let $\tilde{X}$ be the process $\tilde{X}(t) = X(t_n)$ for $t \in [t_n, t_{n+1})$, $n \in \{0, \dots, N-1\}$. For $l \in \{1, \dots, K\}$, Proposition 2.2.2 and Lemma 2.3.6 yield

$$\begin{aligned} &\big\|\tilde{X}_{h,\Delta t} - X\big\|_{L^1_{\nu_l}((0,T);\mathbf{M}^{1,p}(H)^*)} \\ &\quad \leqslant \big\|\tilde{X}_{h,\Delta t} - \tilde{X}\big\|_{L^1_{\nu_l}((0,T);\mathbf{M}^{1,p}(H)^*)} + \big\|\tilde{X} - X\big\|_{L^1_{\nu_l}((0,T);\mathbf{M}^{1,p}(H)^*)} \\ &\quad \leqslant C\left(h^{2\gamma} + \Delta t^{\rho\gamma}\right). \end{aligned}$$

This completes the proof. □

Finally, we formulate a corollary of Theorem 2.3.7 that can be used to prove convergence of covariances and higher order statistics of approximate solutions. We demonstrate this for covariances; higher order statistics can be treated in a similar way.

Corollary 2.3.8. *Let X be the solution of* (2.1.2) *and let $X_{h,\Delta t}$ be given by* (2.3.6). *Let $\tilde{X}_{h,\Delta t}(t) = X^n_{h,\Delta t}$, for $t \in [t_n, t_{n+1})$, $n \in \{0, \dots, N-1\}$ and $\tilde{X}_{h,\Delta t}(t) = X^N_{h,\Delta t}$, for $t \in [t_N, T]$. Then, for $K \geqslant 1$, $\phi_1, \dots, \phi_K \in H$, $t_1, \dots, t_K \in (0, T]$, $\gamma \in [0, \beta)$, there exists $C > 0$ such that*

$$\Big|\mathbb{E}\Big[\prod_{i=1}^K \big\langle \tilde{X}_{h,\Delta t}(t_i), \phi_i\big\rangle - \prod_{i=1}^K \big\langle X(t_i), \phi_i\big\rangle\Big]\Big| \leqslant C\big(h^{2\gamma} + \Delta t^{\rho\gamma}\big), \quad h, \Delta t \in (0,1).$$

In particular, for $\phi_1, \phi_2 \in H$ and $t_1, t_2 \in (0,T]$, it holds that

$$\begin{aligned}&\big|\mathrm{Cov}\big(\big\langle \tilde X_{h,\Delta t}(t_1), \phi_1\big\rangle, \big\langle \tilde X_{h,\Delta t}(t_2), \phi_2\big\rangle\big) - \mathrm{Cov}\big(\big\langle X(t_1), \phi_1\big\rangle, \big\langle X(t_2), \phi_2\big\rangle\big)\big|\\ &\quad\leqslant C\big(h^{2\gamma} + \Delta t^{\rho\gamma}\big), \quad h, \Delta t \in (0,1).\end{aligned}$$

Proof. The first statement follows from Theorem 2.3.7 by setting $\varphi_i = \langle \phi_i, \cdot\rangle$, $\nu_i = \delta_{t_i}$, $i \in \{1, \dots, K\}$, where δ_{t_i} is the Dirac measure concentrated at t_i. The second is a consequence of the first and the fact that

$$\begin{aligned}&\mathrm{Cov}\big(\big\langle \tilde X_{h,\Delta t}(t_1), \phi_1\big\rangle, \big\langle \tilde X_{h,\Delta t}(t_2), \phi_2\big\rangle\big) - \mathrm{Cov}\big(\big\langle X(t_1), \phi_1\big\rangle, \big\langle X(t_2), \phi_2\big\rangle\big)\\ &\quad= \mathbb{E}\big[\big\langle \tilde X_{h,\Delta t}(t_1), \phi_1\big\rangle\big\langle \tilde X_{h,\Delta t}(t_2), \phi_2\big\rangle\big] - \mathbb{E}\big[\big\langle X(t_1), \phi_1\big\rangle\big\langle X(t_2), \phi_2\big\rangle\big]\\ &\qquad- \mathbb{E}\big[\big\langle \tilde X_{h,\Delta t}(t_1), \phi_1\big\rangle - \big\langle X(t_1), \phi_1\big\rangle\big]\mathbb{E}\big[\big\langle X(t_2), \phi_2\big\rangle\big]\\ &\qquad- \mathbb{E}\big[\big\langle \tilde X_{h,\Delta t}(t_1), \phi_1\big\rangle\big]\mathbb{E}\big[\big\langle \tilde X_{h,\Delta t}(t_2), \phi_2\big\rangle - \big\langle X(t_2), \phi_2\big\rangle\big].\end{aligned}$$

□

2.4. Examples

In this section we consider two different types of equations and write them in the abstract form of Section 2.1. We verify the abstract assumptions in both cases. Numerical approximation by the finite element method and suitable time discretization schemes are proved to satisfy the assumptions of Section 2.3. We start with parabolic stochastic partial differential equations and continue with Volterra equations in a separate subsection.

2.4.1. Stochastic parabolic partial differential equations. Let $\mathcal{D} \subset \mathbb{R}^d$ for $d = 1, 2, 3$ be a convex polygonal domain. Let $\Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$ be the Laplace operator and $g \in \mathcal{G}_{\mathrm{b}}^2(\mathbb{R}, \mathbb{R})$. We consider the stochastic partial differential equation:

$$\begin{aligned}&\dot u(t,x) = \Delta u(t,x) + g(u(t,x)) + \dot\eta(t,x), && (t,x) \in (0,T] \times \mathcal{D},\\ &u(t,x) = 0, && (t,x) \in (0,T] \times \partial\mathcal{D},\\ &u(0,x) = u_0(x), && x \in \mathcal{D}.\end{aligned} \tag{2.4.1}$$

The noise $\dot\eta$ is not well defined as a function, as it is written, but makes sense as a random measure. We will study this equation in the abstract framework of Section 2.1. Let $H = L^2(\mathcal{D})$, $A\colon D(A) \subset H \to H$ be given by $A = -\Delta$ with $D(A) = H_0^1(\mathcal{D}) \cap H^2(\mathcal{D})$. Let $(S(t))_{t\in[0,T]}$ denote the analytic semigroup $S(t) = e^{-tA}$ of bounded linear operators generated by $-A$. Assumption 2.1.1 is satisfied with $\rho = 1$, as is easily seen by a spectral argument. The drift $f\colon H \to H$ is the Nemytskii operator determined by the action $(f(r))(x) = g(r(x))$, $x \in \mathcal{D}$, $r \in H$. Condition (2.1.4) for f is verified in [**106**] for $\delta = \frac{d}{2} + \varepsilon$. With these definitions we may write (2.4.1) is the abstract Itô form

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = f(X(t))\,\mathrm{d}t + \mathrm{d}W(t),\ t \in (0,T]; \quad X(0) = x_0,$$

with mild solution given by (2.1.2).

Let $(\mathcal{T}_h)_{h\in(0,1)}$ denote a family of regular triangulations of $\mathcal{D}$ where h denotes the maximal mesh size. Let $(V_h)_{h\in[0,1]}$ be the finite element spaces of continuous piecewise linear functions with respect to $(\mathcal{T}_h)_{h\in(0,1)}$ and $P_h\colon H \to V_h$ be the orthogonal projector. Let the operators $A_h\colon V_h \to V_h$ be defined by (1.5.52).

REMARK 2.4.1. If the domain $\mathcal{D}$ is such that the pairs of eigenvalues and eigenfunctions $(\lambda_n, e_n)_{n\in\mathbb{N}}$ of A are known, e.g., $\mathcal{D} = [0,1]^d$, then instead of finite element discretization one can consider a spectral Galerkin approximation. Let the eigenvalues be ordered in increasing order so that $\lambda_n \leqslant \lambda_{n+1}$ for every $n \in \mathbb{N}$. Further, let $h = \lambda_{N+1}^{-\frac{1}{2}}$ and $V_h = \operatorname{span}\{\phi_n : n \leqslant N\}$. By $P_h \colon H \to V_h$ we denote the orthogonal projector and we define $A_h = AP_h = P_hA = P_hAP_h$.

We discretize in time by a semi-implicit Euler-Maruyama method. By defining $B^1_{h,\Delta t} = (I + \Delta t A_h)^{-1}P_h$ and $B^n_{h,\Delta t} = (B^1_{h,\Delta t})^n P_h$ for $n \geqslant 1$, the discrete solutions $(X^n_{h,\Delta t})^N_{n=0}$ are recursively given by

$$\begin{aligned} X^n_{h,\Delta t} &= B^1_{h,\Delta t}X^{n-1}_{h,\Delta t} + \Delta t B^1_{h,\Delta t} f(X^{n-1}_{h,\Delta t}) + \int_{t_{n-1}}^{t_n} B^1_{h,\Delta t}\,\mathrm{d}W(s), \quad n = 1,\dots,N,\\ X^0_{h,\Delta t} &= P_h x_0. \end{aligned}$$

Iterating the scheme gives the discrete variation of constants formula (2.3.6). For both finite element and spectral approximation the assumptions (2.3.1), (2.3.2), (2.3.3), (2.3.4), are valid, see, for example, [**103**].

2.4.2. Stochastic Volterra integro-differential equations. Consider the semi-linear stochastic Volterra type equation

(2.4.2)
$$\begin{aligned} \dot u(t,x) &= \int_0^t b(t-s)\Delta u(t,x)\,\mathrm{d}s + g(u(t,x)) + \dot\eta(t,x), && (t,x) \in (0,T]\times\mathcal{D},\\ u(t,x) &= 0, && (t,x)\in(0,T]\times\partial\mathcal{D},\\ u(0,x) &= u_0, && x\in\mathcal{D}. \end{aligned}$$

Suppose that b satisfies Assumptions 1.5.1 and 1.5.30. We write (2.4.2) in the abstract Itō form

(2.4.3)
$$\mathrm{d}X(t) + \Big(\int_0^t b(t-s)AX(s)\,\mathrm{d}s\Big)\,\mathrm{d}t = f(X(t))\,\mathrm{d}t + \mathrm{d}W(t),\ t\in(0,T];\ X(0) = x_0,$$

with A, f, W, x_0 as in Subsection 2.4.1. Here one needs $\delta = \frac{d}{2} + \varepsilon < \frac{2}{\rho}$ and this requires $\rho < \frac{4}{3}$ and ε small in the case $d = 3$ but causes no restrictions in the case $d = 1, 2$. Under the above assumptions there exist a resolvent family of operators $(S(t))_{t\in[0,T]}$ defined by (1.5.7) and the mild solution of (2.4.3) is given by (2.1.2). By Propositions 1.5.6 and 1.5.9 condition (2.1.1) holds for S. Thus, the setting of Section 2.1 is applicable.

Using the notation from Subsections 1.5.6 and 1.5.7 we now turn our attention to the numerical approximation via the semilinear variant of (1.5.61):

$$\begin{aligned} X^n_{h,\Delta t} - X^{n-1}_{h,\Delta t} + \Delta t\left(\sum_{k=1}^n \omega_{n-k}\, A_h X^k_{h,\Delta t}\right)&\\ = \Delta t P_h f(X^{n-1}_{h,k}) + \int_{t_{n-1}}^{t_n} P_h\,\mathrm{d}W(t),& \quad n = 1,\dots,N,\\ X^0_{h,k} = P_h x_0.& \end{aligned}$$

It is possible to write $(X_{h,k}^n)_{n=0}^N$ as a variation of constants formula (2.3.6) with $B_{h,\Delta t}^n$ defined by (1.5.62). The stability (2.3.1) holds by (1.5.71), (1.5.72), (1.5.73) and interpolation. The smooth data error estimate (2.3.2) was proved in Lemma 1.5.27. It remains to verify (2.3.3). By (1.5.74) there exist $\tilde{C}$ so that

$$\left\|F_{h,\Delta t}^n\right\| \leqslant \tilde{C} t_n^{-\frac{\delta}{2}} \big(h^{\frac{\delta}{\rho}} + \Delta t^{\frac{\delta}{2}}\big), \quad 0 \leqslant \delta \leqslant 2,\ n = 1, \ldots, N.$$

Let $0 \leqslant \delta \leqslant 2$. Interpolation with $0 \leqslant s \leqslant 1$ yields

$$\begin{aligned}
\left\|A^{\frac{s}{2\rho}} F_{h,\Delta t}^n\right\| &\leqslant \left\|F_{h,\Delta t}^n\right\|^{1-s} \left\|A^{\frac{1}{2\rho}} F_{h,\Delta t}^n\right\|^s \\
&\leqslant \left\|F_{h,\Delta t}^n\right\|^{1-s} \Big(\left\|A^{\frac{1}{2\rho}} S(t_n)\right\| + \left\|A^{\frac{1}{2\rho}} B_{h,\Delta t}^n\right\|\Big)^s \\
&\leqslant \Big(\tilde{C} t_n^{-\frac{\delta}{2}} \big(h^{\frac{\delta}{\rho}} + \Delta t^{\frac{\delta}{2}}\big)\Big)^{1-s} \big(2L_{\frac{1}{2}} t_n^{-\frac{1}{2}}\big)^s \\
&\leqslant \tilde{C}^{1-s} (2L_{\frac{1}{2}})^s t_n^{-\frac{\delta(1-s)+s}{2}} \big(h^{\frac{\delta(1-s)}{\rho}} + \Delta t^{\frac{\delta(1-s)}{2}}\big).
\end{aligned}$$

Setting $\sigma = \delta(1-s)$ and $R_s = \tilde{C}^{1-s}(2L_{\frac{1}{2}})^s$ yields the estimate

$$\left\|A^{\frac{s}{2\rho}} F_{h,\Delta t}^n\right\| \leqslant R_s t_n^{-\frac{\sigma+s}{2}} \big(h^{\frac{\sigma}{\rho}} + \Delta t^{\frac{\sigma}{2}}\big), \quad 0 \leqslant \sigma \leqslant 2,\ 0 \leqslant s \leqslant 1 - \frac{\sigma}{2},$$

for $n = 1, \ldots, N$. Therefore (2.3.3) holds.

CHAPTER 3

Linear stochastic PDEs driven by additive Lévy noise

3.1. Preliminaries

To describe the problem in an abstract setting, let H be a real separable Hilbert space and $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a stochastic basis satisfiying the usual conditions, $L = (L(t))_{t\geqslant 0}$ be a square-integrable cylindrical Lévy process in a real separable Hilbert space U with respect to the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$, taking values in a possibly larger Hilbert space $U_1 \supset U$, and $B : U \to H$ is a bounded linear operator. Consider an H-valued stochastic convolution process

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}L(s) \tag{3.1.1}$$

where $(E(t))_{t\in[0,T]}$ is a family of bounded linear operators on H and X_0 is an $\mathcal{F}_0$-measurable H-valued random variable. Without loss of generality, all Hilbert spaces are assumed to be infinite-dimensional. Important examples of such processes are weak solutions $(X(t))_{t\geqslant 0}$ of certain SPDEs driven by additive Lévy noise; these can be written as abstract Itô stochastic differential equations

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}L(t), \quad t \geqslant 0; \quad X(0) = X_0, \tag{3.1.2}$$

where $-A$ is the generator of a strongly continuous semigroup $(E(t))_{t\geqslant 0}$ on H. In particular, we consider the stochastic heat equation

$$\mathrm{d}X(t) + \Lambda X(t)\mathrm{d}t = \mathrm{d}L(t), \quad t \geqslant 0; \quad X(0) = X_0, \tag{3.1.3}$$

and the stochastic wave equation, written as a first order system,

$$\begin{aligned} \mathrm{d}X_1(t) - X_2(t)\mathrm{d}t &= 0, & t \geqslant 0; \quad X_1(0) = X_{0,1}, \\ \mathrm{d}X_2(t) + \Lambda X_1(t)\mathrm{d}t &= \mathrm{d}L(t), & t \geqslant 0; \quad X_2(0) = X_{0,2}. \end{aligned} \tag{3.1.4}$$

In both cases $\Lambda := -\Delta = -\sum_{j=1}^d \partial^2/\partial\xi_j^2$ is the Laplace operator on $L^2(\mathcal{D})$ where $\mathcal{D} \subset \mathbb{R}^d$ is a bounded domain supplemented by Dirichlet zero boundary conditions. For the precise abstract setup of these equations we refer to Sections 3.3 and 3.5. Similarly to the Gaussian case discussed in Chapters 1 and 2, we do not require that $(E(t))_{t\geqslant 0}$ enjoys the semigroup property so that the abstract framework can accommodate stochastic Volterra-type evolution equations as well, see Subsection 3.4 for details.

Consider an approximation $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ of the process $(X(t))_{t\in[0,T]}$ given by

$$\tilde{X}(t) = \tilde{E}(t)X_0 + \int_0^t \tilde{E}(t-s)B\,\mathrm{d}L(s), \tag{3.1.5}$$

where $(\tilde{E}(t))_{t\in[0,T]}$ is a family of bounded linear operators on H, which is again not necessarily (extendable to) an operator semigroup. For example, the family $(\tilde{E}(t))_{t\in[0,T]}$ may be a time-interpolated solution operator family of a space-time discretized stochastic evolution problem, when H is an L^2-space of some spatial domain $\mathcal{D}$. As is Chapter 1 we study the weak error

$$e(T) := \mathbb{E}\big(G(\tilde{X}(T)) - G(X(T))\big) \tag{3.1.6}$$

for suitable test functions $G : H \to \mathbb{R}$. At the heart of the chapter are the error representation formulae for $e(T)$, Theorem 3.2.6 and Corollary 3.2.8. The proof of Theorem 3.2.6 is based on Kolmogorov's backward equation for the martingale $Y(t) = E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}L(s)$, $t \in [0,T]$, which has the important property that $Y(T) = X(T)$. The introduction of such an auxiliary process Y is well-known for linear equations with additive Gaussian noise, and it was used extensively in Chapter 1.

3.1.1. The driving Lévy process L. The process $L = (L(t))_{t\geqslant 0}$ in (3.1.2) is a Lévy process with values in a real and separable Hilbert space U_1, defined on a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ satisfying the usual conditions (cf. [**88**]) if L is $(\mathcal{F}_t)$-adapted and for $t, h \geqslant 0$ the increment $L(t+h) - L(t)$ is independent of $\mathcal{F}_t$. We always consider a càdlàg (right continuous with left limits) modification of L, i.e., a modification such that $L(t) = \lim_{s\searrow t} L(s)$ for all $t \geqslant 0$ and $L(t-) := \lim_{s\nearrow t} L(s)$ exists for all $t > 0$, where the limits are pathwise limits in U_1. Our standard reference for Hilbert space-valued Lévy processes is [**88**].

We assume that L is square-integrable, i.e., $\mathbb{E}\|L(t)\|_{U_1}^2 < \infty$, and that it has mean zero, i.e., $\mathbb{E}L(t) = 0$ in U_1. Moreover, we assume that the Gaussian part of L vanishes; see Remark 3.2.9 below for a comment on how to treat the case of a non-vanishing Gaussian part. Let ν be the jump intensity measure (Lévy measure) of L. Note that the jump intensity measure ν of a general Lévy process in U_1 satisfies $\nu(\{0\}) = 0$ and $\int_{U_1} \min(1, \|y\|_{U_1}^2)\nu(\mathrm{d}y) < \infty$, cf. [**88**, Section 4]. Due to our assumptions we have

$$\int_{U_1} \|y\|_{U_1}^2 \nu(\mathrm{d}y) < \infty, \tag{3.1.7}$$

and the characteristic function of L is given by

$$\mathbb{E}e^{i\langle x, L(t)\rangle_{U_1}} = \exp\Big\{ -t\int_{U_1} \big(1 - e^{i\langle x,y\rangle_{U_1}} + i\langle x,y\rangle_{U_1}\big)\nu(\mathrm{d}y)\Big\}, \quad t \geqslant 0,\ x \in U_1. \tag{3.1.8}$$

Conversely, any U_1-valued Lévy process L satisfying (3.1.7) and (3.1.8) is square-integrable, with mean zero and vanishing Gaussian part.

Let $Q_1 \in \mathcal{L}_1(U_1)$ be the covariance operator of L. It is determined by the jump intensity measure ν via

$$\langle Q_1 x, y\rangle_{U_1} = \int_{U_1} \langle x, z\rangle_{U_1} \langle y, z\rangle_{U_1} \nu(\mathrm{d}z), \quad x, y \in U_1, \tag{3.1.9}$$

see [**88**, Theorem 4.47]. Further, let

$$(U_0, \langle\,\cdot\,,\,\cdot\,\rangle_{U_0}) := \big(Q_1^{1/2}(U_1), \langle Q_1^{-1/2}\,\cdot\,, Q_1^{-1/2}\,\cdot\,\rangle_{U_1}\big)$$

be the reproducing kernel Hilbert space, or Cameron–Martin space, of L, where $Q_1^{-1/2}$ denotes the pseudo-inverse of $Q_1^{1/2}$, see [**88**, Section 7]. Recall that the operator B in (3.1.2) is defined on the Hilbert space U. We assume that

$$U_0 \subset U \subset U_1, \tag{3.1.10}$$

and that the inclusions (3.1.10) define continuous embeddings. We denote the embedding of U_0 into U by $J_0 \in \mathcal{L}(U_0, U)$ and set

$$Q := J_0 J_0^* \in \mathcal{L}(U). \tag{3.1.11}$$

The nonnegative and symmetric operator Q is the covariance operator of L considered as a cylindrical process in U, cf. Remark 3.1.1 below. As a consequence of Douglas' theorem as stated in [**88**, Appendix A.4], compare also [**91**, Corollary C.0.6], the reproducing kernel Hilbert space of L has the alternative representation

$$(U_0, \langle\,\cdot\,,\,\cdot\,\rangle_{U_0}) = \big(Q^{1/2}(U), \langle Q^{-1/2}\,\cdot\,, Q^{-1/2}\,\cdot\,\rangle_U\big).$$

Remark 3.1.1. Suppose w.l.o.g. that U is dense in U_1, identify U and U^* via the Riesz isomorphism, and consider the Gelfand triple $U_1^* \subset U^* \equiv U \subset U_1$. Then it is not difficult to see that

$$\mathbb{E}\langle L(t), x\rangle\langle L(t), y\rangle = t\langle Qx, y\rangle_U, \quad t \geqslant 0,\ x, y \in U_1^*,$$

where $\langle\,\cdot\,,\,\cdot\,\rangle : U_1 \times U_1^* \to \mathbb{R}$ is the canonical dual pairing; compare [**88**, Proposition 7.7]. The unique continuous extensions of the linear mappings $U_1^* \ni x \mapsto \langle L(t), x\rangle \in L^2(\mathbb{P})$, $t \geqslant 0$, to the larger space U^* determine a 2-cylindrical U-process in the sense of [**84**], compare with also [**4**], [**94**], [**95**].

Remark 3.1.2. Note that, using (1.1.4), we have $\mathcal{L}(U_1, H) \subset \mathcal{L}_2(U_0, H)$ since

$$\|C\|_{\mathcal{L}_2(U_0,H)} = \|CQ_1^{1/2}\|_{\mathcal{L}_2(U_1,H)} \leqslant \|C\|_{\mathcal{L}(U_1,H)}\|Q_1^{1/2}\|_{\mathcal{L}_2(U_1)}$$

for all $C \in \mathcal{L}(U_1, H)$ and $\|Q_1^{1/2}\|_{\mathcal{L}_2(U_1)} = \operatorname{Tr} Q_1 = \|Q_1\|_{\mathcal{L}_1(U_1)} < \infty$.

Remark 3.1.3. Unlike in the case of a mean-zero (cylindrical) Q-Wiener process in U, the covariance operators $Q \in \mathcal{L}(U)$ and $Q_1 \in \mathcal{L}_1(U_1)$ do *not* determine the distribution of the Lévy process L, but the jump intensity measure ν does so according to (3.1.8). Note that the law of a general Lévy process is determined by its characteristics (Lévy triplet), cf. [**88**, Definition 4.28], and that the characteristics of L are $(-\int_{\{\|y\|_{U_1}\geqslant 1\}} y\,\nu(\mathrm{d}y), 0, \nu)$. Nevertheless, the operator Q in (3.1.11) will play an important role in our error analysis. Let us briefly make the connection of our setting to the construction of a cylindrical Q-Wiener process in U as described in [**31**], [**91**], see also, Chapter 1. To this end, let $(f_k)_{k\in\mathbb{N}}$ be an orthonormal basis of U_1 consisting of eigenvectors of Q_1 with eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$ and consider the orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of U_0 given by $e_k := \lambda_k^{1/2} f_k$. To simplify notation we suppose for the moment that all eigenvalues λ_k of Q_1 are strictly positive. Then, compare [**88**, Section 4.8], the real-valued Lévy processes $L_k = (L_k(t))_{t\geqslant 0}$, $k \in \mathbb{N}$, given by

$$L_k(t) := \lambda_k^{-1/2}\langle L(t), f_k\rangle_{U_1}$$

are uncorrelated, i.e., $\mathbb{E}L_k(t)L_j(s) = 0$ if $k \neq j$, they satisfy $\mathbb{E}(L_k^2(t)) = t$, and we have

$$L(t) = \sum_{k\in\mathbb{N}} L_k(t)e_k. \tag{3.1.12}$$

The infinite sum in (3.1.12) converges for all finite $T > 0$ in the space $\mathcal{M}_T^2(U_1)$ of càdlàg square-integrable U_1-valued $(\mathcal{F}_t)$-martingales $M = (M(t))_{t\in[0,T]}$ with norm $\|M\|_{\mathcal{M}_T^2(U_1)} = (\mathbb{E}\|M(T)\|_{U_1}^2)^{1/2}$. In contrast to the Gaussian case, where uncorrelated coordinates are always independent, the coordinate processes L_k, $k \in \mathbb{N}$, are in general only uncorrelated but *not* independent.

Conversely, suppose that we are given an arbitrary symmetric and nonnegative operator $Q \in \mathcal{L}(U)$, an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of $U_0 = Q^{1/2}(U)$, and a family L_k, $k \in \mathbb{N}$, of real-valued Lévy processes on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ that satisfy the following conditions:

- Each L_k is $(\mathcal{F}_t)$-adapted and for $t, h \geqslant 0$ the increment $L_k(t+h) - L_k(t)$ is independent of $\mathcal{F}_t$;
- each L_k is square-integrable with $\mathbb{E}L_k(t) = 0$ and $\mathbb{E}(L_k^2(t)) = t$;
- the processes L_k, $k \in \mathbb{N}$, are uncorrelated;
- for all $n \in \mathbb{N}$ the $\mathbb{R}^n$-valued process $((L_1(t), \ldots, L_n(t))^\top)_{t\geqslant 0}$ is a Lévy process;
- the Gaussian part of each L_k is zero.

Then, if U_1 is a Hilbert space containing U such that the natural embedding of $U_0 = Q^{1/2}(U)$ into U_1 is Hilbert-Schmidt, the infinite sum in (3.1.12) converges in $\mathcal{M}_T^2(U_1)$ and defines a Lévy process L with reproducing kernel Hilbert space U_0 that fits into our setting.

We end this subsection with some examples of Lévy processes L. We suppose that all processes are defined relative to the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ and that their increments on time intervals $[t, t+h]$ are independent of $\mathcal{F}_t$.

Example 3.1.4 (Subordinate cylindrical $\widetilde{Q}$-Wiener process). Let W be a cylindrical $\widetilde{Q}$-Wiener process in U in the sense of [**91**, Section 2.5.1], where $\widetilde{Q} \in \mathcal{L}(U)$ is a given nonnegative and symmetric operator. Assume that W takes values in a possibly larger Hilbert space $U_1 \supset U$ such that the natural embedding of U into U_1 is continuous. Let $\widetilde{Q}_1 \in \mathcal{L}_1(U_1)$ be the covariance operator of W considered as a Wiener process in U_1, i.e., $\mathbb{E}\langle W(t), x\rangle_{U_1}\langle W(s), y\rangle_{U_1} = \min(s,t)\langle \widetilde{Q}_1 x, y\rangle_{U_1}$ for x, $y \in U_1$, $s, t \geqslant 0$. Let $Z = (Z(t))_{t\geqslant 0}$ be a subordinator, i.e., a real-valued increasing Lévy process in the sense of [**98**, Definition 21.4], [**99**]. Assume that W and Z are independent, that the drift of Z is zero, and that the jump intensity measure ρ of Z satisfies

$$\int_0^\infty s\,\rho(\mathrm{d}s) < \infty. \tag{3.1.13}$$

The latter is equivalent to assuming that Z has first moments, $\mathbb{E}|Z(t)| < \infty$. According to [**98**, Remark 21.6], the Laplace tranform of $Z(t)$ is given by

$$\mathbb{E}(e^{-rZ(t)}) = \exp\big(-t\int_0^\infty (1-e^{-rs})\rho(\mathrm{d}s)\big), \quad r \geqslant 0. \tag{3.1.14}$$

In this situation, subordinate cylindrical Brownian motion

$$L(t) := W(Z(t)), \quad t \geqslant 0,$$

defines a U_1-valued Lévy process $L = (L(t))_{t\geqslant 0}$ that fits into the general framework described above. Indeed, L has stationary and independent increments. Moreover, the independence of W and Z, the identity $\mathbb{E}e^{i\langle x, W(s)\rangle_{U_1}} = e^{-s\frac{1}{2}\langle \widetilde{Q}_1 x, x\rangle_{U_1}}$, Eq. (3.1.14) and the symmetry of the distribution $\mathbb{P}(W(1) \in \cdot) = N(0, Q_1)$ imply that characteristic function of $L(t)$ is given by

$$\begin{aligned}\mathbb{E}e^{i\langle x,L(t)\rangle_{U_1}} &= \int_0^\infty e^{-s\frac{1}{2}\langle \widetilde{Q}_1 x,x\rangle_{U_1}}\,\mathbb{P}(Z(t)\in \mathrm{d}s)\\ &= \exp\Big[-t\int_0^\infty (1-e^{-\frac{1}{2}\langle \widetilde{Q}_1 x,x\rangle_{U_1} s})\rho(\mathrm{d}s)\Big]\\ &= \exp\Big[-t\int_0^\infty\int_{U_1}(1-e^{i\langle x,\sqrt{s}y\rangle_{U_1}}+i\langle x,\sqrt{s}y\rangle_{U_1})\\ &\qquad\qquad \mathbb{P}(W(1)\in \mathrm{d}y)\rho(\mathrm{d}s)\Big].\end{aligned}$$

As a consequence, (3.1.8) holds with

$$\nu = \big(\mathbb{P}(W(1)\in\cdot)\otimes\rho\big)\circ\kappa^{-1}, \tag{3.1.15}$$

where $\kappa : U_1\times(0,\infty)\to U_1$ is defined by $\kappa(y,s)=\sqrt{s}y$; compare [**95**, Lemma 4.8]. (Note that, by the scaling property of W, (3.1.15) is equivalent to the standard formula $\nu=\int_0^\infty \mathbb{P}(W(s)\in\cdot)\,\rho(\mathrm{d}s)$, where the measure-valued integral is defined in a weak sense, cf. [**98**, Section 30]). Moreover, (3.1.7) holds due to (3.1.13) as we have the equality $\int_{U_1}\|y\|^2_{U_1}\nu(\mathrm{d}y)=\int_0^\infty s\,\rho(\mathrm{d}s)\,\mathbb{E}(\|W(1)\|^2_{U_1})$ according to (3.1.15). It follows that L is a U_1-valued, square-integrable, mean-zero Lévy process with vanishing Gaussian part. It is also not difficult to show that the covariance operators $Q_1\in\mathcal{L}_1(U_1)$ and $Q\in\mathcal{L}(U)$ of L in (3.1.9) and (3.1.11) are given by $Q_1=\int_0^\infty s\,\rho(\mathrm{d}s)\,\widetilde{Q}_1$ and $Q=\int_0^\infty s\,\rho(\mathrm{d}s)\,\widetilde{Q}$. Subordinate cylindrical Wiener processes have been considered, e.g., in [**26**].

Example 3.1.5 (Independent one-dimensional Lévy processes). Let $Q\in\mathcal{L}(U)$ be symmetric, nonnegative and let $(e_k)_{k\in\mathbb{N}}$ be an orthonormal basis of $U_0:=Q^{1/2}(U)\subset U$. Let $L_k=(L(t))_{k\in\mathbb{N}}$, $k\in\mathbb{N}$, be independent real-valued square-integrable Lévy processes with vanishing Gaussian part and $\mathbb{E}L_k(t)=0$, $\mathbb{E}(L_k^2(t))=t$. Let $U_1\supset U$ be another Hilbert space such that the natural embedding of U_0 into U_1 is a Hilbert-Schmidt operator. Then, the series (3.1.12) converges for all $T\in(0,\infty)$ in the space $\mathcal{M}^2_T(U_1)$ and defines a Lévy process $L=(L(t))_{t\geqslant 0}$ satisfiying (3.1.7) and (3.1.8) with jump intensity measure

$$\nu=\sum_{k\in\mathbb{N}}\nu_k\circ\pi_k^{-1},$$

where ν_k is the Lévy measure of L_k and $\pi_k:\mathbb{R}\to U_1$ is defined by $\pi_k(\xi):=\xi e_k$; compare with [**88**, Section 4.8.1].

Example 3.1.6 (Impulsive cylindrical process). Let μ be a Lévy measure on $\mathbb{R}$ such that $\int_{\mathbb{R}}\sigma^2\mu(\mathrm{d}\sigma)<\infty$. Let $\mathcal{O}\subset\mathbb{R}^d$ be a bounded domain and $Z=(Z(t))_{t\geqslant 0}$ an impulsive cylindrical process on $U:=L^2(\mathcal{O})=L^2(\mathcal{O},\mathcal{B}(\mathcal{O}),\lambda^d)$ with jump size intensity μ in the sense of [**88**, Definition 7.23]. Here, λ^d denotes d-dimensional Lebesgue measure. The process Z is a measure-valued process defined, informally, by $Z(t,\mathrm{d}\xi)=\int_0^t\int_{\mathbb{R}}\sigma\hat{\pi}(\mathrm{d}s,\mathrm{d}\xi,\mathrm{d}\sigma)$, where $\hat{\pi}$ is a compensated Poisson random measure on $[0,\infty)\times\mathcal{O}\times\mathbb{R}$ with reference measure $\lambda^1\otimes\lambda^d\otimes\mu$; see [**88**, Section 7.2] for details. Let $\widetilde{Q}\in\mathcal{L}(U)$ be symmetric and nonnegative, $(b_k)_{k\in\mathbb{N}}$ an orthonormal basis of U, and $U_1\supset U$ a Hilbert space such that the natural embedding of

$U_0 = \tilde{Q}^{1/2}(U) \subset U$ into U_1 is Hilbert-Schmidt. Then the series

$$L(t) := \widetilde{Q}^{\frac{1}{2}} Z(t) := \sum_{k\in\mathbb{N}} \int_0^t \int_{\mathcal{O}} \int_{\mathbb{R}} \sigma b_k(\xi) \hat{\pi}(\mathrm{d}s, \mathrm{d}\xi, \mathrm{d}\sigma) \widetilde{Q}^{\frac{1}{2}} b_k, \quad t \geqslant 0, \tag{3.1.16}$$

converges for all $T \in (0, \infty)$ in $\mathcal{M}_T^2(U_1)$ and defines a Lévy process that fits into our general framework with $Q = \int_{\mathbb{R}} \sigma^2 \mu(\mathrm{d}\sigma) \widetilde{Q}$ and

$$\nu = (\lambda^d \otimes \mu) \circ \phi^{-1},$$

where $\phi \in L^2(\mathcal{O} \times \mathbb{R}, \lambda^d \otimes \mu; U_1)$ is defined by $\phi(\xi, \sigma) = \sum_{n\in\mathbb{N}} \sigma b_k(\xi) \widetilde{Q}^{\frac{1}{2}} b_k$ (convergence in $L^2(\mathcal{O} \times \mathbb{R}, \lambda^d \otimes \mu; U_1)$). In [**75**] weak approximation of the stochastic heat equation driven by an impulsive process of the form (3.1.16) was considered. The results in Section 3.3.1 improve the results of [**75**] in several aspects.

3.1.2. Linear stochastic evolution equations with additive noise. We are mainly interested in equations of the type (3.1.2), where $A : D(A) \subset H \to H$ is an unbounded linear operator such that $-A$ is the generator of a strongly continuous semigroup $(E(t))_{t\geqslant 0} \subset \mathcal{L}(H)$, $B \in \mathcal{L}(U, H)$, $L = (L(t))_{t\geqslant 0}$ is a square-integrable Lévy process with reproducing kernel Hilbert space $U_0 \subset U$ as described in Subsection 3.1.1, and $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. It is well known that if

$$\int_0^T \|E(t)B\|^2_{\mathcal{L}_2(U_0,H)} \mathrm{d}t < \infty \tag{3.1.17}$$

for some (and hence for all) $T > 0$, then there exists a unique weak solution $X = (X(t))_{t\geqslant 0}$ to (3.1.2) which is given by the variation-of-constants formula (3.1.1), see, e.g., [**88**, Chapter 9]. Similarly, if $(\tilde{E}(t))_{t\in[0,T]} \subset \mathcal{L}(H)$ is given by some approximation scheme such that $t \mapsto \tilde{E}(t)B$ is a measurable mapping from $[0, T]$ to $\mathcal{L}_2(U_0, H)$, then the condition

$$\int_0^T \|\tilde{E}(t)B\|^2_{\mathcal{L}_2(U_0,H)} \mathrm{d}t < \infty \tag{3.1.18}$$

ensures that the approximation $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ of $(X(t))_{t\in[0,T]}$ in (3.1.5) exists as a square-integrable H-valued process. We refer to [**88**, Chapter 8] for details on the construction and properties of the stochastic integral w.r.t. Hilbert space-valued Lévy processes.

It turns out that our general error representation formula for the weak error $e(T)$ in (3.1.6) does not require the semigroup property of the strongly continuous family of operators $(E(t))_{t\geqslant 0}$. This paves the way for analysing a more general class of Lévy-driven linear stochastic evolution equations, including, for example, stochastic Volterra-type integro-differential equations as considered in Section 1.5 for the Gaussian case, see Subsection 3.4 for details. For such equations, the weak solution still has the form (3.1.1) but the solution operator family $(E(t))_{t\geqslant 0} \subset \mathcal{L}(H)$ is not a semigroup anymore. Therefore, we weaken our abstract assumptions and summarize them as follows.

ASSUMPTION 3.1.7. We will use the following assumptions:

(i) H, U and U_1 are real and separable Hilbert spaces;

(ii) $L = (L(t))_{t\geqslant 0}$ is a U_1-valued Lévy process with vanishing Gaussian part on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$ with zero mean and finite second moments and reproducing kernel Hilbert space U_0 such that $U_0 \subset U \subset U_1$ as described in Subsection 3.1.1;

(iii) $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$;
(iv) $B \in \mathcal{L}(U, H)$ and $(E(t))_{t\in[0,T]} \subset \mathcal{L}(H)$ is a strongly continuous family of linear operators such that (3.1.17) holds;
(v) for all $\varepsilon > 0$ there exists $\Phi_\varepsilon \in \mathcal{L}_2(U_0, H)$ and $C_\varepsilon > 0$ such that

$$\|E(t)Bx\|_H \leqslant \|\Phi_\varepsilon x\|_H, \quad (t,x) \in [\varepsilon, T] \times U_0;$$

(vi) $(\tilde{E}(t))_{t\in[0,T]} \subset \mathcal{L}(H)$ is a family of linear operators such that $t \mapsto \tilde{E}(t)B$ is a measurable mapping from $[0,T]$ to $\mathcal{L}_2(U_0, H)$ and (3.1.18) holds;
(vii) $X = (X(t))_{t\in[0,T]}$ and $\tilde{X} = (\tilde{X}(t))_{t\in[0,T]}$ are H-valued stochastic processes given by (3.1.1) and (3.1.5).

Remark 3.1.8. If $(E(t))_{t\geqslant 0}$ is an operator semigroup, then 3.1.7(v) is a consequence of 3.1.7(iv). Indeed, if (3.1.17) holds then $E(t_0)B \in \mathcal{L}_2(U_0, H)$ for some $t_0 \in (0, \varepsilon)$ and hence

$$\|E(t)Bx\|_H \leq \|E(t_0)Bx\|_H \sup_{t\in[t_0,T]} \|E(t-t_0)\|_{\mathcal{L}(H)}, \quad (t,x) \in [\varepsilon, T] \times U_0.$$

Hence, one may take $\Phi_\varepsilon := cE(t_0)B$ where $c := \sup_{t\in[t_0,T]} \|E(t-t_0)\|_{\mathcal{L}(H)}$.

To fix notation, let us briefly recall the Itô isometry for stochastic integrals w.r.t. L. It has the same form as the Itô isometry for stochastic integrals w.r.t. Hilbert space-valued Wiener processes. We set $\Omega_T := \Omega \times [0,T]$ and $\mathbb{P}_T := \mathbb{P} \otimes \lambda$, where λ is Lebesgue measure on $[0,T]$. The predictable σ-algebra on Ω_T w.r.t. $(\mathcal{F}_t)_{t\in[0,T]}$ is denoted by $\mathcal{P}_T$. For operator-valued processes $\Phi = (\Phi(t))_{t\in[0,T]}$ in

$$L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H)) := L^2(\Omega_T, \mathcal{P}_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H)),$$

we have

$$\mathbb{E}\Big(\Big\|\int_0^t \Phi(s)\mathrm{d}L(s)\Big\|_H^2\Big) = \mathbb{E}\int_0^t \|\Phi(s)\|^2_{\mathcal{L}_2(U_0,H)}\mathrm{d}s, \quad t \in [0,T], \tag{3.1.19}$$

and the integral process $(\int_0^t \Phi(s)\mathrm{d}L(s))_{t\in[0,T]}$ belongs to the space $\mathcal{M}_T^2(H)$ of càdlàg square-integrable H-valued $(\mathcal{F}_t)$-martingales. The usual norm in $\mathcal{M}_T^2(H)$ is defined by $\|M\|_{\mathcal{M}_T^2(H)} = (\mathbb{E}\|M(T)\|_H^2)^{1/2}$, $M = (M(t))_{t\in[0,T]} \in \mathcal{M}_T^2(H)$. Note, however, that the integral processes given by the stochastic integrals in (3.1.1) and (3.1.5) are in general *not* martingales since the (deterministic) operator-valued integrands also depend on t.

3.2. An error representation formula

In this section, we state and prove a general representation formula for the weak approximation error $e(T)$ in (3.1.6) within the abstract setting described above.

For the formulation and the proof of the error representation formula, we introduce auxiliary drift-free Itô processes $Y = (Y(t))_{t\in[0,T]}$ and $\tilde{Y} = (\tilde{Y}(t))_{t\in[0,T]}$ such that

$$X(T) = Y(T), \quad \tilde{X}(T) = \tilde{Y}(T).$$

To this aim, we set

$$Y(t) := E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}L(s), \quad t \in [0,T], \tag{3.2.1}$$

and

$$\tilde{Y}(t) := \tilde{E}(T)X_0 + \int_0^t \tilde{E}(T-s)B\,\mathrm{d}L(s), \quad t \in [0,T]. \tag{3.2.2}$$

Moreover, we consider the auxiliary problem

$$\mathrm{d}Z(t) = E(T-t)B\,\mathrm{d}L(t), \quad t \in [\tau, T]; \qquad Z(\tau) = \xi,$$

where $\tau \in [0,T)$ and ξ is an H-valued $\mathcal{F}_\tau$-measurable random variable. Its solution is given by

$$Z(t,\tau,\xi) := \xi + \int_\tau^t E(T-s)B\,\mathrm{d}L(s), \quad t \in [\tau, T], \tag{3.2.3}$$

and we use it to define for $G \in C^2(H,\mathbb{R})$ with $\sup_{x\in H} \|G''(x)\|_{\mathcal{L}(H)} < \infty$ a function $u : [0,T] \times H \to \mathbb{R}$ by

$$u(t,x) := \mathbb{E}\, G(Z(T,t,x)), \quad (t,x) \in [0,T] \times H. \tag{3.2.4}$$

As noted in (1.1.18) and (1.1.19), the boundedness of G'' implies quadratic and linear growth of G and G', respectively. That is,

$$|G(x)| \leqslant C(1 + \|x\|_H^2), \quad \|G'(x)\|_H \leqslant C(1 + \|x\|_H) \tag{3.2.5}$$

for all $x \in H$ and a constant $C \in (0,\infty)$ that does not depend on x. It is also not difficult to see that u is twice Fréchet differentiable w.r.t. x and we have

$$u_x(t,x) = \mathbb{E}\, G'(Z(T,t,x)), \quad u_{xx}(t,x) = \mathbb{E}\, G''(Z(T,t,x)). \tag{3.2.6}$$

All expectations appearing in (3.2.4) and (3.2.6) make sense due to (3.2.5), the assumption $\sup_{x\in H} \|G''(x)\|_{\mathcal{L}(H)} < \infty$, (3.1.17) and Itô's isometry (3.1.19).

3.2.1. Poisson random measures and a comparison of stochastic integrals. Before stating the representation formula, we show in the following lemma how operators in $\mathcal{L}_2(U_0, H)$ can be identified with functions in

$$L^2(U_1, \nu; H) := L^2(U_1, \mathcal{B}(U_1), \nu; H)$$

and how processes in $L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$ can be identified with elements in

$$L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; H) := L^2(\Omega_T \times U_1, \mathcal{P}_T \otimes \mathcal{B}(U_1), \mathbb{P}_T \otimes \nu; H).$$

These identifications will be used throughout this chapter, see Remark 3.2.2 below. They also lead to a generic identification of integrals w.r.t. (cylindrical) Hilbert space-valued Lévy processes of jump type and integrals w.r.t. the associated Poisson random measures.

LEMMA 3.2.1. *Let $(f_k)_{k\in\mathbb{N}} \subset U_0$ be an orthonormal basis of U_1 consisting of eigenvectors of the covariance operator $Q_1 \in \mathcal{L}_1(U_1)$ of L and let $(\lambda_k)_{k\in\mathbb{N}} \subset [0,\infty)$ be the corresponding sequence of eigenvalues.*

(i) *Given $\Phi \in \mathcal{L}_2(U_0, H)$, the series*

$$\iota(\Phi) := \sum_{k\in\mathbb{N}, \lambda_k \neq 0} \langle \,\cdot\, , f_k \rangle_{U_1} \Phi f_k$$

converges in $L^2(U_1, \nu; H)$ and the linear mapping

$$\iota : \mathcal{L}_2(U_0, H) \to L^2(U_1, \nu; H),\ \Phi \mapsto \iota(\Phi)$$

is an isometric embedding.

(ii) *Given $\Phi \in L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$, the series*

$$\kappa(\Phi) := \sum_{k\in\mathbb{N}, \lambda_k \neq 0} \langle \,\cdot\, , f_k \rangle_{U_1} \Phi(\,\cdot\,) f_k$$

converges in $L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; H)$ *and the linear mapping*

$$\kappa : L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H)) \to L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; H),\ \Phi \mapsto \kappa(\Phi)$$

is an isometric embedding. For $\mathbb{P}_T$*-almost all* $(\omega, t) \in \Omega_T$ *we have* $\kappa(\Phi)(\omega, t, \cdot\,) = \iota(\Phi(\omega, t))$ *in* $L^2(U_1, \nu; H)$*, where* ι *is the embedding from* (i).

Proof. **(i)** W.l.o.g. all eigenvalues λ_k are strictly positive. Let $(e_k)_{k\in\mathbb{N}}$ be the orthonormal basis of U_0 given by $e_k := \lambda_k^{1/2} f_k$. For $m, n \in \mathbb{N}$ with $m \leqslant n$ we have

$$\begin{aligned}\Big\| \sum_{k=m}^{n} \langle\,\cdot\,, f_k\rangle_{U_1} \Phi f_k \Big\|^2_{L^2(U_1,\nu;H)} &= \int_{U_1} \Big\| \sum_{k=m}^{n} \langle x, f_k\rangle_{U_1} \Phi f_k \Big\|^2_H \nu(\mathrm{d}x) \\ &= \sum_{j,k=m}^{n} \lambda_j^{-1/2} \lambda_k^{-1/2} \int_{U_1} \langle x, f_j\rangle_{U_1} \langle x, f_k\rangle_{U_1} \nu(\mathrm{d}x)\, \langle \Phi e_j, \Phi e_k\rangle_H = \sum_{k=m}^{n} \|\Phi e_k\|_H^2;\end{aligned}$$

in the last step we used (3.1.9). Since $\sum_{k\in\mathbb{N}} \|\Phi e_k\|_H^2 = \|\Phi\|^2_{\mathcal{L}_2(U_0,H)} < \infty$, this shows that the partial sums $\sum_{k=1}^{n} \langle\,\cdot\,, f_k\rangle_{U_1} \Phi f_k$, $n \in \mathbb{N}$, are a Cauchy sequence in $L^2(U_1, \nu; H)$ and

$$\Big\| \sum_{k=1}^{\infty} \langle\,\cdot\,, f_k\rangle_{U_1} \Phi f_k \Big\|_{L^2(U_1,\nu;H)} = \|\Phi\|_{\mathcal{L}_2(U_0,H)}.$$

(ii) The first two assertions can be shown as in the proof of (i). The last assertion is due the fact that the iterated integral

$$\int_\Omega \int_0^T \int_{U_1} \|\iota(\Phi(\omega, t))(x) - \kappa(\Phi)(\omega, t, x)\|_H^2 \,\nu(\mathrm{d}x)\,\mathrm{d}t\,\mathbb{P}(\mathrm{d}\omega)$$

equals zero, which follows from an approximation argument. □

Remark 3.2.2. From now on we will identify operators $\Phi \in \mathcal{L}_2(U_0, H)$ with the corresponding mappings $\iota(\Phi) \in L^2(U_1, \nu; H)$ and write

$$\Phi x = \iota(\Phi)(x), \quad x \in U_1.$$

Analogously, we identify processes $\Phi \in L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$ with the corresponding mappings $\kappa(\Phi) \in L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; H)$ and write

$$\Phi(\omega, t)x = \kappa(\Phi)(\omega, t, x), \quad (\omega, t, x) \in \Omega_T \times U_1.$$

For processes $\Phi \in L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$ both identifications are compatible $\mathbb{P}\otimes\lambda$-almost everywhere on Ω_T in the sense that we have $\kappa(\Phi)(\omega, t, \cdot\,) = \iota(\Phi(\omega, t))$ in $L^2(U_1, \nu; H)$ for $\mathbb{P}\otimes\lambda$-almost all $(\omega, t) \in \Omega_T$.

The proof of Theorem 3.2.6 below is based on Itô's formula for Banach space-valued jump processes driven by Poisson random measures as presented in [**80**]. Alternatively, one could use Itô's formula as proved in [**44**], but the formula in [**80**] is more convenient in our setting. Therefore, we use Lemma 3.2.1 to relate our setting to the setting in [**80**] and state a suitable version of Itô's formula for integral processes $(\int_0^t \Phi(s)\,\mathrm{d}L(s))_{t\in[0,T]}$.

It is well-known that the jumps of a Lévy process determine a Poisson random measure on the product space of the underlying time interval and the state space.

We refer to [**88**, Section 6] for a definition and properties of Poisson random measures. For $(\omega, t) \in \Omega \times (0, \infty)$ we denote by $\Delta L(t)(\omega) := L(t)(\omega) - \lim_{s \nearrow t} L(s)(\omega) \in U_1$ the jump of a trajectory of L at time t. Setting

$$N(\omega) := \sum_{\Delta L(t)(\omega) \neq 0} \delta_{(t, \Delta L(t)(\omega))}, \quad \omega \in \Omega,$$

defines a Poisson random measure N on $([0,\infty) \times U_1, \mathcal{B}([0,\infty)) \otimes \mathcal{B}(U_1))$ with intensity measure (or compensator) $\lambda \otimes \nu$, where λ is Lebesgue measure on $[0,\infty)$ and ν is the jump intensity measure of L. This follows, e.g., from Theorem 6.5 in [**88**] together with Theorems 4.9, 4.15, 4.23 and Lemma 4.25 therein. We denote the compensated Poisson random measure by

$$q := N - \lambda \otimes \nu.$$

Let V be a (real and separable) Hilbert space. The stochastic integral with respect to q of functions in $L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; V) = L^2(\Omega_T \times U_1, \mathcal{P}_T \otimes \mathcal{B}(U_1), \mathbb{P}_T \otimes \nu; V)$ is constructed as a linear isometry

$$L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; V) \to \mathcal{M}_T^2(V), \; f \mapsto \Big(\int_0^t \int_{U_1} f(s,x)\, q(\mathrm{d}s, \mathrm{d}x) \Big)_{t \in [0,T]}.$$

In particular, the V-valued integral processes have càdlàg modifications; we will always work with such a càdlàg modification. Using a standard stopping procedure, the stochastic integral can be extendend to functions $f \in L^0(\Omega_T \times U_1, \mathcal{P}_T \otimes \mathcal{B}(U_1), \mathbb{P}_T \otimes \nu; V)$ such that

$$\mathbb{P}\Big(\int_0^T \int_{U_1} \|f(s,x)\|_V^2 \, \nu(\mathrm{d}x)\, \mathrm{d}s < \infty \Big) = 1.$$

We refer to [**80**], [**90**] and the references therein for details on stochastic integration w.r.t. Poisson random measures, compare also [**88**, Section 8.7].

Remark 3.2.3. Strictly speaking, in [**80**] the integrands f do not have to be predictable but only $\mathcal{F}_t \otimes \mathcal{B}(U_1)$-adapted and $\mathcal{F} \otimes \mathcal{B}([0,T]) \otimes \mathcal{B}(U_1)$-measurable. However, it is clear that in the case of predictable, i.e., $\mathcal{P}_T \otimes \mathcal{B}(U_1)$-measurable, and square integrable Hilbert space-valued integrands f the stochastic integral in [**80**] coincides with the stochastic integral considered in [**88**], [**90**]. See [**96**] for a detailed comparison of the different spaces of integrands.

Since $\mathbb{E} \int_0^T \int_{U_1} \|x\|_{U_1}^2 \, \nu(\mathrm{d}x) \mathrm{d}t$ is finite for all $T < \infty$, it follows that the integral process $(\int_0^t \int_{U_1} x\, q(\mathrm{d}s, \mathrm{d}x))_{t \geqslant 0}$ is uniquely determined (up to indistinguishability) as a U_1-valued square-integrable càdlàg martingale. Taking into account the assumptions on the Lévy process L, the Lévy-Khinchin decomposition [**88**, Theorem 4.23], the definition of q, and the construction of the stochastic integral w.r.t. q, it is not difficult to see that the processes L and $(\int_0^t \int_{U_1} x\, q(\mathrm{d}s, \mathrm{d}x))_{t \geqslant 0}$ are indistinguishable, i.e.,

$$\mathbb{P}\Big(L(t) = \int_0^t \int_{U_1} x\, q(\mathrm{d}s, \mathrm{d}x) \quad \forall\, t \geqslant 0 \Big) = 1. \tag{3.2.7}$$

Using Lemma 3.2.1, we are now able to identify stochastic integrals w.r.t. L and stochastic integrals w.r.t. the compensated Poisson random measure q. Recall from Remark 3.2.2 that we identify processes $\Phi \in L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$ with the corresponding functions $\kappa(\Phi) \in L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; H)$. Thus, for such Φ the integral process $(\int_0^t \int_{U_1} \Phi(s)x\, q(\mathrm{d}s, \mathrm{d}x))_{t\in[0,T]}$ is defined.

LEMMA 3.2.4. *Given $\Phi \in L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$, the H-valued càdlàg integral processes $(\int_0^t \Phi(s)\,\mathrm{d}L(s))_{t\in[0,T]}$ and $(\int_0^t \int_{U_1} \Phi(s)x\, q(\mathrm{d}s, \mathrm{d}x))_{t\in[0,T]}$ are indistinguishable. That is,*

$$\mathbb{P}\Big(\int_0^t \Phi(s)\,\mathrm{d}L(s) = \int_0^t \int_{U_1} \Phi(s)x\, q(\mathrm{d}s, \mathrm{d}x) \quad \forall\, t \in [0,T]\Big) = 1.$$

PROOF. We first assume that Φ is a simple $\mathcal{L}(U_1, H)$-valued process of the form

$$\Phi(s) = \sum_{k=0}^{m-1} \mathbb{1}_{F_k} \mathbb{1}_{(t_k, t_{k+1}]}(s)\Phi_k, \quad s \in [0,T],$$

with $0 \leqslant t_0 < t_1 < \cdots < t_m \leqslant T$, $m \in \mathbb{N}$, $F_k \in \mathcal{F}_{t_k}$ and $\Phi_k \in \mathcal{L}(U_1, H)$. Recall from Section 3.1.2 that $\mathcal{L}(U_1, H)$ is a subspace of $\mathcal{L}_2(U_0, H)$. Using (3.2.7) and applying standard arguments for the evaluation of stochastic integrals, we obtain for fixed $t \in [0,T]$, $\mathbb{P}$-almost surely,

$$\begin{aligned}\int_0^t \Phi(s)\,\mathrm{d}L(s) &= \sum_{k=0}^{m-1} \mathbb{1}_{F_k}\Phi_k(L(t_{k+1}\wedge t) - L(t_k \wedge t)) \\ &= \sum_{k=0}^{m-1} \mathbb{1}_{F_k}\Phi_k\Big(\int_0^T \int_{U_1} \mathbb{1}_{(t_k\wedge t, t_{k+1}, \wedge t]}(s)x\, q(\mathrm{d}s, \mathrm{d}x)\Big) \\ &= \sum_{k=0}^{m-1} \int_0^T \int_{U_1} \mathbb{1}_{F_k}\mathbb{1}_{(t_k\wedge t, t_{k+1}, \wedge t]}(s)\Phi_k x\, q(\mathrm{d}s, \mathrm{d}x) \\ &= \int_0^t \int_{U_1} \Phi(s)x\, q(\mathrm{d}s, \mathrm{d}x).\end{aligned}$$

Since both processes are right-continuous, we see that the processes

$$(\int_0^t \Phi(s)\,\mathrm{d}L(s))_{t\in[0,T]}$$

and $(\int_0^t \int_{U_1} \Phi(s)x\, q(\mathrm{d}s, \mathrm{d}x))_{t\in[0,T]}$ are indistinguishable.

For general $\Phi \in L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$, take a sequence $(\Phi_n)_{n\in\mathbb{N}}$ of simple $\mathcal{L}(U_1, H)$-valued process such that $\Phi_n \to \Phi$ in $L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$; see, e.g., [**91**, Proposition 2.3.8] for a proof of the existence of such a sequence. Then, the processes

$$\int_0^\cdot \Phi_n(s)\,\mathrm{d}L(s) = \Big(\int_0^t \Phi_n(s)\,\mathrm{d}L(s)\Big)_{t\in[0,T]}$$

and $\int_0^\cdot \int_{U_1} \Phi_n(s)x\, q(\mathrm{d}s, \mathrm{d}x) = (\int_0^t \int_{U_1} \Phi_n(s)x\, q(\mathrm{d}s, \mathrm{d}x))_{t\in[0,T]}$ are indistinguishable for all $n \in \mathbb{N}$, and we have the convergence $\int_0^\cdot \Phi_n(s)\,\mathrm{d}L(s) \to \int_0^\cdot \Phi(s)\,\mathrm{d}L(s)$ in $\mathcal{M}_T^2(H)$ by the construction of the stochastic integral w.r.t. L. According to Lemma 3.2.1, the convergence $\Phi_n \to \Phi$ in $L^2(\Omega_T, \mathbb{P}_T; \mathcal{L}_2(U_0, H))$ entails the convergence $\kappa(\Phi_n) \to \kappa(\Phi)$ in $L^2(\Omega_T \times U_1, \mathbb{P}_T \otimes \nu; H)$, so that we also have

$$\int_0^\cdot \int_{U_1} \Phi_n(s)x\, q(\mathrm{d}s, \mathrm{d}x) \to \int_0^\cdot \int_{U_1} \Phi(s)x\, q(\mathrm{d}s, \mathrm{d}x)$$

in $\mathcal{M}_T^2(H)$. Thus, $\int_0^\cdot \Phi(s)\,\mathrm{d}L(s) = \int_0^\cdot \int_{U_1} \Phi(s)x\,q(\mathrm{d}s,\mathrm{d}x)$ as an equality in $\mathcal{M}_T^2(H)$, which yields the assertion. □

As a direct consequence of Lemma 3.2.4 and [**80**, Theorem 3.6] we obtain the following version of Itô's formula. By $C_{\mathrm{b}}^{1,2}([0,T]\times H,\mathbb{R})$ we denote the space of continuous and bounded functions $F:[0,T]\times H\to\mathbb{R}$, $(s,x)\mapsto F(s,x)$ such that the Fréchet partial derivatives F_s, F_x and F_{xx} exist and are continuous and bounded on $[0,T]\times H$.

THEOREM 3.2.5. *Let $F\in C_{\mathrm{b}}^{1,2}([0,T]\times H,\mathbb{R})$, $\Phi\in L^2(\Omega_T,\mathbb{P}_T;\mathcal{L}_2(U_0,H))$ and $M=(M(t))_{t\in[0,T]}$ be the H-valued càdlàg process given by*

$$M(t) = M(0) + \int_0^t \Phi(s)\,\mathrm{d}L(s)$$

with $M(0)\in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$. Then, $\mathbb{P}$-almost surely,

$$\int_0^T \int_{U_1} \big|F(s,M(s)+\Phi(s)y) - F(s,M(s))\big|^2\,\nu(\mathrm{d}y)\,\mathrm{d}s < \infty,$$
$$\int_0^T \int_{U_1} \big|F(s,M(s)+\Phi(s)y) - F(s,M(s)) - \big\langle F_x(s,M(s)),\Phi(s)y\big\rangle_H\big|\,\nu(\mathrm{d}y)\,\mathrm{d}s < \infty,$$

and the following equality holds for all $t\in[0,T]$:

$$\begin{aligned} F(t,M(t)) \;&=\; F(0,M(0)) + \int_0^t F_s(s,M(s))\,\mathrm{d}s \\ &+ \int_0^t \int_{U_1} \big\{F(s,M(s-)+\Phi(s)y) - F(s,M(s-))\big\}\,q(\mathrm{d}s,\mathrm{d}y) \\ &+ \int_0^t \int_{U_1} \big\{F(s,M(s)+\Phi(s)y) - F(s,M(s)) - \big\langle F_x(s,M(s)),\Phi(s)y\big\rangle_H\big\}\,\nu(\mathrm{d}y)\,\mathrm{d}s. \end{aligned}$$

3.2.2. Formulation of the result. Next state the main result of this section.

THEOREM 3.2.6. *Let Assumption 3.1.7 hold and let $G\in C^2(H,\mathbb{R})$ such that $\sup_{x\in H}\|G''(x)\|_{\mathcal{L}(H)}<\infty$. Then, for the process $(\tilde Y(t))_{t\in[0,T]}$ from (3.2.2) and the function $u:[0,T]\times H\to\mathbb{R}$ from (3.2.4) it holds that*

$$\begin{aligned} \mathbb{E}\int_0^T \int_{U_1} \Big| & u\big(t,\tilde Y(t)+\tilde E(T-t)By\big) - u\big(t,\tilde Y(t)+E(T-t)By\big) \\ & - \big\langle u_x(t,\tilde Y(t)),\big(\tilde E(T-t)B - E(T-t)B\big)y\big\rangle_H\Big|\,\nu(\mathrm{d}y)\,\mathrm{d}t \;<\; \infty. \end{aligned} \tag{3.2.8}$$

Furthermore, the weak error $e(T)$ in (3.1.6) has the representation

$$\begin{aligned} e(T) = \mathbb{E}\big\{&u(0,\tilde E(T)X_0) - u(0,E(T)X_0)\big\} \\ &+ \mathbb{E}\int_0^T \int_{U_1} \Big\{u\big(t,\tilde Y(t)+\tilde E(T-t)By\big) - u\big(t,\tilde Y(t)+E(T-t)By\big) \\ &\qquad - \big\langle u_x(t,\tilde Y(t)),\big(\tilde E(T-t)B - E(T-t)B\big)y\big\rangle_H\Big\}\,\nu(\mathrm{d}y)\,\mathrm{d}t. \end{aligned} \tag{3.2.9}$$

Remark 3.2.7. The terms $E(T-t)By$ and $\tilde{E}(T-t)By$ appearing in (3.2.8) and (3.2.9) are defined for $\lambda\otimes\nu$-almost all $(t,y)\in[0,T]\times U_1$. This follows from (3.1.17), (3.1.18), Lemma 3.2.1 and Remark 3.2.2.

We will prove Theorem 3.2.6 in the next subsection. Let us briefly record an alternative representation of $e(T)$ which follows from Taylor's formula. It will be the starting point for our error estimates in Sections 3.3 and 3.5. For $t\in[0,T]$, $\theta\in[0,1]$ and $y\in U_1$ set

$$\begin{aligned}
&F(t) := \tilde{E}(t)B - E(t)B,\\
&\Psi_1(t,\theta,y)\\
&\quad := (1-\theta)\Big\langle u_{xx}\big(t,\tilde{Y}(t)+E(T-t)By+\theta F(T-t)y\big)F(T-t)y\,,\,F(T-t)y\Big\rangle_H,\\
&\Psi_2(t,\theta,y) := \Big\langle u_{xx}\big(t,\tilde{Y}(t)+\theta E(T-t)By\big)E(T-t)By\,,\,F(T-t)y\Big\rangle_H.
\end{aligned}$$

Corollary 3.2.8. *In the setting of Theorem 3.2.6 we have*

$$\mathbb{E}\int_0^T\int_{U_1}\int_0^1\big\{|\Psi_1(t,\theta,y)|+|\Psi_2(t,\theta,y)|\big\}\,\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t<\infty, \tag{3.2.10}$$

and the following alternative error representation holds:

$$\begin{aligned}
e(T) &= \mathbb{E}\big\{u(0,\tilde{E}(T)X_0)-u(0,E(T)X_0)\big\}\\
&\quad + \mathbb{E}\int_0^T\int_{U_1}\int_0^1\big\{\Psi_1(t,\theta,y)+\Psi_2(t,\theta,y)\big\}\,\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t.
\end{aligned}\tag{3.2.11}$$

Proof. The integrand of the iterated integral in (3.2.9) can be rewritten as

$$\begin{aligned}
u\big(t,\tilde{Y}(t)&+\tilde{E}(T-t)By\big)-u\big(t,\tilde{Y}(t)+E(T-t)By\big)-\big\langle u_x(t,\tilde{Y}(t)),F(T-t)y\big\rangle_H\\
&=\Big\{u\big(t,\tilde{Y}(t)+\tilde{E}(T-t)By\big)-u\big(t,\tilde{Y}(t)+E(T-t)By\big)\\
&\quad-\big\langle u_x\big(t,\tilde{Y}(t)+E(T-t)By\big),F(T-t)y\big\rangle_H\Big\}\\
&\quad+\big\langle u_x\big(t,\tilde{Y}(t)+E(T-t)By\big)-u_x(t,\tilde{Y}(t)),F(T-t)y\big\rangle_H\\
&=\int_0^1\big\{\Psi_1(t,\theta,y)+\Psi_2(t,\theta,y)\big\}\,\mathrm{d}\theta,
\end{aligned}$$

where the last step is due to an application of Taylor's theorem to the function $\theta\mapsto u\big(t,\tilde{Y}(t)+E(T-t)By+\theta\big(\tilde{E}(T-t)By-E(T-t)By\big)\big)$ (using the integral form of the remainder for the second order expansion about 0) and an application of the fundamental theorem of calculus to the function $\theta\mapsto\Psi_2(t,\theta,y)$. By (3.2.8) we have

$$\mathbb{E}\int_0^T\int_{U_1}\Big|\int_0^1\big\{\Psi_1(t,\theta,y)+\Psi_2(t,\theta,y)\big\}\,\mathrm{d}\theta\Big|\,\nu(\mathrm{d}y)\,\mathrm{d}t<\infty.$$

The stronger assertion (3.2.10) follows from the boundedness of $G'':H\to\mathcal{L}(H)$, Lemma 3.2.1, (3.1.17) and (3.1.18). □

Remark 3.2.9. Let us note that Theorem 3.2.6 can be generalized with an analogous proof to stochastic convolution processes of the form $X(t)=E(t)X_0+\int_0^t E(t-s)B\,\mathrm{d}(W+L)(s)$, $t\in[0,T]$, and respective approximations $\tilde{X}(t)=\tilde{E}(t)X_0+\int_0^t\tilde{E}(t-s)B\,\mathrm{d}(W+L)(s)$, $t\in[0,T]$, where L is as described in Subsection 3.1.1 and

dc_1541_18

W is a (cylindrical) Wiener process on U whose reproducing kernel Hilbert space is continuously embedded in U_0, say. Replacing L by $W+L$ in the definitions (3.2.1), (3.2.2), (3.2.3) and (3.2.4), the corresponding error $e(T) = \mathbb{E}\big(G(\tilde{X}(T)) - G(X(T))\big)$ can be represented by adding to the right hand side of (3.2.9) the term

$$\frac{1}{2}\,\mathbb{E}\int_0^T \mathrm{Tr}\Big\{u_{xx}\big(\tilde{Y}(t),t\big)\Big(\tilde{E}(T-t)BQ_W B^*\tilde{E}(T-t)^* \\ - E(T-t)BQ_W B^* E(T-t)^*\Big)\Big\}\,\mathrm{d}t,$$

where $Q_W \in \mathcal{L}(U)$ is the covariance operator of W, defined by

$$\mathbb{E}\langle W(1),x\rangle_U\langle W(1),y\rangle_U = \langle Q_W x, y\rangle_U,\ x,y\in U,$$

see Theorem 1.2.1, in particular, equation (1.2.6) in its proof. Similarly, the considerations in Sections 3.3–3.5 below can be generalized by using arguments from Chapter 1.

3.2.3. Proof of the error representation formula. In this subsection, we give the proof of Theorem 3.2.6.

For $\xi \in L^2(\Omega,\mathcal{F}_t,\mathbb{P};H)$ we have

$$\begin{aligned}\mathbb{E}\big(G(Z(T,t,\xi))\big) &= \int_H\int_H G(x+y)\,\mathbb{P}\left(\int_t^T E(T-s)B\,\mathrm{d}L(s)\in\mathrm{d}y\right)\mathbb{P}(\xi\in\mathrm{d}x)\\ &= \mathbb{E}\big(u(t,\xi)\big)\end{aligned}$$

by (3.2.3), (3.2.4), the independence of $\int_t^T E(T-s)B\,\mathrm{d}L(s)$ and $\mathcal{F}_t$, and Fubini's theorem. Since $X(T)=Y(T)$ and $\tilde{X}(T)=\tilde{Y}(T)$ it follows that

$$\begin{aligned}e(T) &= \mathbb{E}\big(G(\tilde{Y}(T)) - G(Y(T))\big)\\ &= \mathbb{E}\big(G(Z(T,T,\tilde{Y}(T))) - G(Z(T,0,Y(0)))\big)\\ &= \mathbb{E}\big(u(T,\tilde{Y}(T)) - u(0,Y(0))\big)\\ &= \mathbb{E}\big(u(0,\tilde{Y}(0)) - u(0,Y(0))\big) + \mathbb{E}\big(u(T,\tilde{Y}(T)) - u(0,\tilde{Y}(0))\big).\end{aligned}\tag{3.2.12}$$

By (3.2.1) and (3.2.2), the first term in the last line equals $\mathbb{E}(u(0,\tilde{E}(T)X_0) - u(0,E(T)X_0))$.

To handle the second term in the last line of (3.2.12), we first assume that $G: H\to\mathbb{R}$ and $G': H\to H$ are bounded, so that $G\in C_{\mathrm{b}}^2(H,\mathbb{R})$. We will remove this restriction later on. We now want to apply Itô's formula to the function $(t,x)\mapsto u(t,x)$ and the martingale $\tilde{Y} = (\tilde{Y}(t))_{t\in[0,T]}$. For this we need the following properties of u.

Proposition 3.2.10. *Let Assumption 3.1.7 hold and $G\in C_{\mathrm{b}}^2(H,\mathbb{R})$. The function $u:[0,T]\times H\to\mathbb{R}$, $(t,x)\mapsto u(t,x)$ defined in (3.2.4) and its Fréchet partial derivatives u_x, u_{xx} are continuous and bounded on $[0,T]\times H$. The time derivative u_t of u exists on $[0,T)\times H$ and is continuous. Moreover, for every $\varepsilon>0$ there exists some $C_\varepsilon<\infty$ such that*

$$\int_{U_1}\big|u\big(t,x+E(T-t)By\big) - u(t,x) - \big\langle u_x(t,x),E(T-t)By\big\rangle_H\big|\,\nu(\mathrm{d}y) < C_\varepsilon \tag{3.2.13}$$

dc_1541_18

for all $t \in [0, T-\varepsilon]$, *and* u *satisfies the backward Kolmogorov equation*

$$\left.\begin{aligned} &u_t(t,x) \\ &\quad = -\int_{U_1} \Big\{ u\big(t, x + E(T-t)By\big) - u(t,x) - \big\langle u_x(t,x), E(T-t)By \big\rangle_H \Big\} \,\nu(\mathrm{d}y), \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (t,x) \in [0,T) \times H, \\ &u(T,x) = G(x), \quad x \in H. \end{aligned}\right\} \tag{3.2.14}$$

Proof. We begin with the continuity and boundedness of u, u_x and u_{xx}. The boundedness is obvious by the definition (3.2.4) of u and by (3.2.6). Pick $0 \leqslant s \leqslant t \leqslant T$, $x, y \in H$. Using (3.2.4), Jensen's inequality, the mean value theorem, (3.2.3) and Itô's isometry, we have

$$\begin{aligned} |u(t,x) - u(s,y)|^2 &\leqslant \mathbb{E}\big(|G(Z(T,t,x)) - G(Z(T,s,y))|^2 \big) \\ &\leqslant \sup_{x \in H} \|G'(x)\|_H^2 \, \mathbb{E}\Big(\Big\| x - y - \int_s^t E(T-r)B \,\mathrm{d}L(r) \Big\|_H^2 \Big) \\ &\leqslant 2 \sup_{x \in H} \|G'(x)\|_H^2 \Big(\|x-y\|_H^2 + \int_s^t \|E(T-r)B\|_{\mathcal{L}_2(U_0,H)}^2 \,\mathrm{d}r \Big). \end{aligned}$$

Thus, the continuity of u follows from (3.1.17) and the boundedness of G'. Since $u_x(t,x) = \mathbb{E}\, G'(Z(T,t,x))$, the continuity of $u_x : [0,T] \times H \to H$ follows analogously from the boundedness of G''. To show the continuity of $u_{xx} : [0,T] \times H \to \mathcal{L}(H)$, define $F \in C_{\mathrm{b}}(H \times H; \mathbb{R})$ by

$$F(x,y) := \|G''(x) - G''(y)\|_{\mathcal{L}(H)}, \quad x, y \in H,$$

and fix $(t,x) \in [0,T] \times H$, $((t_k, x_k))_{k \in \mathbb{N}} \subset [0,T] \times H$ with $(t_k, x_k) \to (t,x)$ as $k \to \infty$. Note that $Z(T, t_k, x_k) \to Z(T,t,x)$ in $L^2(\Omega; H)$ by Itô's isometry. As a consequence, $(Z(T,t,x), Z(T,t_k,x_k)) \to (Z(T,t,x), Z(T,t,x))$ in distribution (on $H \times H$) and we obtain

$$\begin{aligned} \|u_{xx}(t,x) - u_{xx}(t_k,x_k)\|_{\mathcal{L}(H)} &\leqslant \mathbb{E}\, F\big(Z(T,t,x), Z(T,t_k,x_k) \big) \\ &\xrightarrow{k \to \infty} \mathbb{E}\, F\big(Z(T,t,x), Z(T,t,x) \big) = 0, \end{aligned}$$

yielding the continuity of u_{xx}.

By Taylor's formula and Lemma 3.2.1,

$$\begin{aligned} &\int_{U_1} \big| u\big(t, x + E(T-t)By\big) - u(t,x) - \big\langle u_x(t,x), E(T-t)By \big\rangle_H \big| \,\nu(\mathrm{d}y) \\ &\qquad\qquad \leqslant \frac{1}{2} \sup_{x \in H} \|G''(x)\|_{\mathcal{L}(H)} \int_{U_1} \|E(T-t)By\|_H^2 \nu(\mathrm{d}y) \\ &\qquad\qquad = \frac{1}{2} \sup_{x \in H} \|G''(x)\|_{\mathcal{L}(H)} \|E(T-t)B\|_{\mathcal{L}_2(U_0,H)}^2. \end{aligned}$$

Using Assumption 3.1.7(v), condition (3.2.13) follows with

$$C_\varepsilon = 1/2 \sup_{x \in H} \|G''(x)\|_{\mathcal{L}(H)} \|\Phi_\varepsilon\|_{\mathcal{L}_2(U_0,H)}^2.$$

In order to verify the Kolmogorov equation (3.2.14), we first note that for fixed $t \in [0,T]$ the H-valued random variables $\int_0^t E(s)B \,\mathrm{d}L(s)$ and $\int_{T-t}^T E(T-s)B \,\mathrm{d}L(s)$

have the same distribution, so that

$$v(t,x) := \mathbb{E}\, G\big(x + \int_0^t E(s)B\,\mathrm{d}L(s)\big) = u(T-t,x), \quad (t,x) \in [0,T] \times H. \tag{3.2.15}$$

Next, we fix $x \in H$ and apply Itô's formula as stated in Theorem 3.2.5 to the function $H \ni y \mapsto G(x+y) \in \mathbb{R}$ and the martingale $M = (M(t))_{t\in[0,T]} := (\int_0^t E(s)B\,\mathrm{d}L(s))_{t\in[0,T]} \in \mathcal{M}_T^2(H)$. We obtain

$$\begin{aligned} &G(x+M(t)) \\ &= G(x) + \int_0^t \int_{U_1} \big\{G\big(x+M(s-)+E(s)By\big) - G\big(x+M(s-)\big)\big\}\, q(\mathrm{d}s,\mathrm{d}y) \\ &\quad + \int_0^t \int_{U_1} \big\{G\big(x+M(s)+E(s)By\big) - G\big(x+M(s)\big) \\ &\qquad\qquad - \big\langle G'\big(x+M(s)\big), E(s)By\big\rangle_H\big\}\, \nu(\mathrm{d}y)\,\mathrm{d}s, \end{aligned} \tag{3.2.16}$$

where q is the compensated Poisson random measure on $[0,\infty)\times U_1$ associated with L and where the term $E(s)By$ is defined for $\lambda\otimes\nu$-almost all $(s,y)\in[0,T]\times U_1$ according to (3.1.17), Lemma 3.2.1 and Remark 3.2.2. The integrand appearing in the integral w.r.t. q belongs to $L^2(\Omega_T\times U_1, \mathbb{P}_T\otimes\nu;\mathbb{R})$ as a consequence of Taylor's formula, the boundedness of G', Lemma 3.2.1 and (3.1.17). Similarly, the second integral in (3.2.16) exists for all $\omega\in\Omega$ and belongs to $L^1(\Omega;\mathbb{R})$ since

$$\begin{aligned} &\int_0^t \int_{U_1} \big|G\big(x+M(s)+E(s)By\big) - G\big(x+M(s)\big) \\ &\qquad\qquad - \big\langle G'\big(x+M(s)\big), E(s)By\big\rangle_H\big|\, \nu(\mathrm{d}y)\,\mathrm{d}s \\ &\leqslant \frac{1}{2}\sup_{x\in H}\|G''(x)\|_{\mathcal{L}(H)} \int_0^t \|E(s)B\|^2_{\mathcal{L}_2(U_0,H)}\mathrm{d}s. \end{aligned}$$

Taking expectations in (3.2.16) and using the martingale property of the integral w.r.t. q yields

$$v(t,x) = G(x) + \int_0^t \int_{U_1} \big\{v(s,x+E(s)By) - v(s,x) - \langle v_x(s,x), E(s)By\rangle_H\big\}\, \nu(\mathrm{d}y)\,\mathrm{d}s. \tag{3.2.17}$$

By the fundamental theorem of calculus, (3.2.14) follows from (3.2.15), (3.2.17) if the mapping

$$(0,T] \ni s \to \int_{U_1} \big\{v(s,x+E(s)By) - v(s,x) - \langle v_x(s,x), E(s)By\rangle_H\big\}\, \nu(\mathrm{d}y) \in \mathbb{R}. \tag{3.2.18}$$

is continuous.

Note that we cannot apply directly the continuity theorem for parameter-dependent integrals to show the continuity of the mapping (3.2.18). The reason is that the term $E(s)By$ in the integral in (3.2.18) is defined only in an $L^2([0,T]\times U_1, \lambda\otimes\nu; H)$-sense, cf. Lemma 3.2.1 and Remark 3.2.2, so that we have no information about the continuity of $(0,T] \in s \mapsto E(s)By \in H$ for fixed $y \in U_1$. Therefore, we use an approximation argument: For $s\in(0,T]$, $x\in H$,

$y \in U_1$ and $k \in \mathbb{N}$ set

$$\begin{aligned} f(s,x,y) &:= v(s,x+E(s)By) - v(s,x) - \langle v_x(s,x), E(s)By\rangle_H, \\ f_k(s,x,y) &:= f(s,x,\Pi_k y), \end{aligned}$$

where Π_k is the orthogonal projection of U_1 onto $\operatorname{span}\{f_1,\dots,f_k\}$, $(f_k)_{k\in\mathbb{N}} \subset U_0$ being an orthonormal basis of U_1 as in Lemma 3.2.1. For fixed $x \in H$, $f(s,x,y)$ is defined in an $L^2([0,T]\times U_1, \lambda\otimes\nu;\mathbb{R})$-sense whereas $f_k(s,x,y)$ is defined pointwise. The continuity theorem for parameter-dependent integrals and the strong continuity of $(E(t))_{t\geqslant 0}$ yield the continuity of in $\int_{U_1} f_k(s,x,y)\nu(\mathrm{d}y)$ in $(s,x) \in [0,T]\times H$. Moreover, we have $f_k(s,x,\cdot) \xrightarrow{k\to\infty} f(s,x,\cdot)$ in $L^1(U_1,\nu;\mathbb{R})$, uniformly in $(s,x) \in [\varepsilon,T]\times H$ for all $\varepsilon > 0$. Indeed, setting $\Pi^k y := y - \Pi_k y$ and using Taylor's theorem, Lemma 3.2.1 and Assumption 3.1.7(v), we obtain

$$\begin{aligned} &\int_{U_1} |f(s,x,y) - f_k(s,x,y)|\,\nu(\mathrm{d}y) \\ &\leqslant \int_{U_1}\int_0^1 \big|\big\langle v_{xx}\big(s,x+E(s)B(\Pi_k y+\theta\Pi^k y)\big)E(s)B\Pi^k y, E(s)B\Pi^k y\big\rangle_H\big| \\ &\qquad \times (1-\theta)\,\mathrm{d}\theta\,\nu(\mathrm{d}y) \\ &\quad + \int_{U_1}\int_0^1 \big|\big\langle v_{xx}\big(s,x+\theta E(s)B\Pi_k y\big)E(s)B\Pi_k y, E(s)B\Pi^k y\big\rangle_H\big|\,\mathrm{d}\theta\,\nu(\mathrm{d}y) \\ &\leqslant \sup_{x\in H}\|G''(x)\|_{\mathcal{L}(H)}\big(\|E(s)B\Pi^k\|^2_{\mathcal{L}_2(U_0,H)} \\ &\quad + \|E(s)B\Pi_k\|_{\mathcal{L}_2(U_0,H)}\|E(s)B\Pi^k\|_{\mathcal{L}_2(U_0,H)}\big) \\ &\leqslant \sup_{x\in H}\|G''(x)\|_{\mathcal{L}(H)}\big(\|\Phi_\varepsilon\Pi^k\|^2_{\mathcal{L}_2(U_0,H)} + \|\Phi_\varepsilon\Pi_k\|_{\mathcal{L}_2(U_0,H)}\|\Phi_\varepsilon\Pi^k\|_{\mathcal{L}_2(U_0,H)}\big) \end{aligned}$$

for all $s \in [\varepsilon,T]$ and some $\Phi_\varepsilon \in \mathcal{L}_2(U_0,H)$. The expression in the last line tends to zero as $k\to\infty$. As a consequence, $\int_{U_1} f_k(s,x,y)\nu(\mathrm{d}y) \xrightarrow{k\to\infty} \int_{U_1} f(s,x,y)\nu(\mathrm{d}y)$ uniformly in $(s,x) \in [\varepsilon,T]\times H$. Thus, the continuity of $\int_{U_1} f_k(s,x,y)\nu(\mathrm{d}y)$ in $(s,x)\in[0,T]\times H$ implies the continuity of $\int_{U_1} f(s,x,y)\nu(\mathrm{d}y)$ in $(s,x) \in (0,T]\times H$. In particular, we obtain the continuity of the mapping (3.2.18) as well as the continuity of u_t on $[0,T)\times H$. □

For $G \in C^2_{\mathrm{b}}(H,\mathbb{R})$, the regularity assertions in Propostition 3.2.10 allow us to apply Itô's formula as stated in Theorem 3.2.5 to the function $(t,x)\mapsto u(t,x)$ and the H-valued martingale $\tilde Y = (\tilde Y(t))_{t\in[0,T]}$ defined in (3.2.2). For $T' \in (0,T)$ we obtain

(3.2.19)

$$\begin{aligned} u(T',\tilde Y(T')) \;=\; & u(0,\tilde Y(0)) + \int_0^{T'} u_t(t,\tilde Y(t))\,\mathrm{d}t \\ & + \int_0^{T'}\int_{U_1} \big\{u\big(t,\tilde Y(t-) + \tilde E(T-t)By\big) - u(t,\tilde Y(t-))\big\}\,q(\mathrm{d}t,\mathrm{d}y) \\ & + \int_0^{T'}\int_{U_1} \big\{u\big(t,\tilde Y(t) + \tilde E(T-t)By\big) - u(t,\tilde Y(t)) \\ & \qquad - \big\langle u_x(t,\tilde Y(t)), \tilde E(T-t)By\big\rangle_H\big\}\,\nu(\mathrm{d}y)\,\mathrm{d}t; \end{aligned}$$

as before, q is the compensated Poisson random measure on $[0,\infty)\times U_1$ associated with L as described previously and the term $\tilde E(T-t)By$ is defined for $\lambda\otimes\nu$-almost all $(t,y)\in[0,T]\times U_1$ according to (3.1.18), Lemma 3.2.1 and Remark 3.2.2. Using the boundedness of u, u_x and u_{xx}, (3.2.14), (3.1.17) and applying similar arguments as in the proof of Proposition 3.2.10, one sees that all terms in (3.2.19) are well-defined and integrable w.r.t. $\mathbb{P}$. Thus, we can take expectations and use the martingale property of the integral w.r.t. q and the backward Kolmogorov equation (3.2.14) to obtain

(3.2.20)
$$
\begin{aligned}
&\mathbb{E}\big(u(T',\tilde Y(T'))-u(0,\tilde Y(0))\big)=\\
&\quad\mathbb{E}\int_0^{T'}\int_{U_1}\Big\{u\big(t,\tilde Y(t)+\tilde E(T-t)By\big)-u\big(t,\tilde Y(t)+E(T-t)By\big)\\
&\qquad\qquad-\big\langle u_x(t,\tilde Y(t)),\big(\tilde E(T-t)B-E(T-t)B\big)y\big\rangle_H\Big\}\,\nu(\mathrm{d}y)\,\mathrm{d}t
\end{aligned}
$$

for all $T'\in(0,T)$. Taking the limit $T'\to T$ on both sides of (3.2.20), we can replace T' by T. Here we used the stochastic continuity of $\tilde Y$ and the continuity of u for the limit on the left hand side. For the limit on the right hand side we used (3.2.8), which is again a consequence of Taylor's formula, the boundedness of G'', (3.1.17), (3.1.18) and Lemma 3.2.1, using similar arguments as in the proof of Proposition 3.2.10. The combination of (3.2.12) and (3.2.20) yields the error representation formula (3.2.9) for the case $G\in C^2_{\mathrm b}(H,\mathbb{R})$.

Finally, we consider the general case of a test function $G\in C^2(H,\mathbb{R})$ such that $\sup_{x\in H}\|G''(x)\|_{\mathcal{L}(H)}<\infty$. For $\varepsilon>0$, define $G_\varepsilon\in C^2_{\mathrm b}(H,\mathbb{R})$ by

$$G_\varepsilon(x):=e^{-\varepsilon\|x\|_H^2}G(x),\quad x\in H.$$

The Fréchet derivatives $G_\varepsilon'(x)\in H$ and $G_\varepsilon''(x)\in\mathcal{L}(H)$ are given by

$$
\begin{aligned}
G_\varepsilon'(x)&=e^{-\varepsilon\|x\|_H^2}\big(G'(x)-2\varepsilon G(x)x\big),\\
G_\varepsilon''(x)&=e^{-\varepsilon\|x\|_H^2}\big[G''(x)+2\varepsilon\big(\langle G'(x),\cdot\rangle_H x-\langle x,\cdot\rangle_H G'(x)\big)-4\varepsilon^2G(x)\langle x,\cdot\rangle_H x\big],
\end{aligned}
$$

so that the boundedness of $G_\varepsilon:H\to\mathbb{R}$, $G_\varepsilon':H\to H$ and $G_\varepsilon'':H\to\mathcal{L}(H)$ follows from (3.2.5). Moreover, note that $G_\varepsilon(x)\xrightarrow{\varepsilon\to0}G(x)$ in $\mathbb{R}$ and $G_\varepsilon'(x)\xrightarrow{\varepsilon\to0}G'(x)$ in H for all $x\in H$, as well as $\sup_{\varepsilon\in(0,1]}\sup_{x\in H}\|G_\varepsilon''(x)\|_{\mathcal{L}(H)}<\infty$. The latter is a consequence of (3.2.5) and the standard estimate $\sup_{s>0}s^ne^{-s}<n!$, $n\in\mathbb{N}$. We set

$$u_\varepsilon(t,x):=\mathbb{E}\,G_\varepsilon(Z(T,t,x)),\quad (t,x)\in[0,T]\times H.$$

By what has been shown above, the assertion of Theorem 3.2.6 holds with u replaced by u_ε. By a dominated convergence argument using (3.2.5), we obtain that $u_\varepsilon(t,x)\xrightarrow{\varepsilon\to0}u(t,x)$ in $\mathbb{R}$ and $(u_\varepsilon)_x(t,x)\xrightarrow{\varepsilon\to0}u_x(t,x)$ in H for all $(t,x)\in[0,T]\times H$. For all $t\in[0,T]$ and $y\in U_1$, we have

$$
\begin{aligned}
&\Big|u_\varepsilon\big(t,\tilde Y(t)+\tilde E(T-t)By\big)-u_\varepsilon\big(t,\tilde Y(t)+E(T-t)By\big)\\
&\qquad-\big\langle(u_\varepsilon)_x(t,\tilde Y(t)),\big(\tilde E(T-t)B-E(T-t)B\big)y\big\rangle_H\Big| \qquad (3.2.21)\\
&\leqslant\sup_{\varepsilon\in(0,1]}\sup_{x\in H}\|G_\varepsilon''(x)\|_{\mathcal{L}(H)}\Big(\|\tilde E(T-t)By\|_H^2+\|E(T-t)B\|_H^2\Big)
\end{aligned}
$$

due to Taylor's theorem and the fact that $(u_\varepsilon)_{xx}(t,x) = \mathbb{E}G''_\varepsilon(Z(T,t,x))$. Note that the term right hand side of (3.2.21) is integrable w.r.t. $\mathrm{d}t\,\nu(\mathrm{d}y)$ as a consequence of (3.1.17), (3.1.18) and Lemma 3.2.1. Thus, considering (3.2.8) and (3.2.9) with u replaced by u_ε, we can use dominated convergence as $\varepsilon \to 0$ to finish the proof.

3.3. Stochastic equations of parabolic type

3.3.1. The stochastic heat equation. Here we give a detailed error analysis of a space-time discretization of the linear stochastic heat equation with additive Lévy noise. We will use some notation introduced in Subsection 1.3.1, in particular the smoothness spaces $\dot{H}^\alpha$. Let $\mathcal{D} \subset \mathbb{R}^d$ be a bounded convex domain and let $\Lambda := -\Delta = -\sum_{j=1}^d \partial^2/\partial\xi_j^2$ be the Laplace operator on $L^2(\mathcal{D})$ with zero-Dirichlet boundary condition, i.e., with domain $D(\Lambda) := \{v \in H_0^1(\mathcal{D}) : \ \Lambda u \in L^2(\mathcal{D})\} = \dot{H}^2$ as described in Subsection 1.3.1. Then, setting

$$H := U := L^2(\mathcal{D}), \quad (A, D(A)) := (\Lambda, D(\Lambda)), \quad B := \mathrm{id}_{L^2(\mathcal{D})},$$

the abstract equation (3.1.2) becomes the stochastic heat equation (3.1.3). It is not difficult to see that the condition $\|A^{-1/2}Q^{1/2}\|_{\mathcal{L}_2(H)} = \|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ implies (3.1.17), where

$$(E(t))_{t\geqslant 0} := (e^{-tA})_{t\geqslant 0} \subset \mathcal{L}(H) \tag{3.3.1}$$

is the semigroup generated by $-A$. Hence, there exists a unique weak solution $X = (X(t))_{t\geqslant 0}$ to Eq. (3.1.3), given by the variation-of-constants formula (3.1.1). Furthermore, for some $C > 0$ independent of T,

$$\|X(T)\|_{L^2(\Omega;H)} \leq C(\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}} + \|X_0\|_{L^2(\Omega;H)}). \tag{3.3.2}$$

As before, for the spatial discretization of (3.1.3), we consider a family of finite-dimensional spaces $(V_h)_{h>0} \subset H_0^1(\mathcal{D})$. As before, by $P_h : H \to V_h$ and $R_h : \dot{H}^1 \to V_h$ we denote the orthogonal projections with respect to the inner products in H and $\dot{H}^1$, respectively. The discrete Laplacian $A_h : V_h \to V_h$ is defined by (1.5.52). Our assumption on the spatial approximation, as in Chapter 1, is that R_h satisfies (1.5.51).

The time discretization of (3.1.3) on a finite interval $[0,T]$ is done via the backward Euler scheme with time step $\Delta t = T/N$, $N \in \mathbb{N}$, and grid points $t_n = n\Delta t$, $n = 0, \dots, N$. For $h > 0$ and $N \in \mathbb{N}$, the discretization $(X^n_{h,\Delta t})_{n=1,\dots,N}$ of $(X(t))_{t\in[0,T]}$ in space and time is given as the solution to

$$X^n_{h,\Delta t} - X^{n-1}_{h,\Delta t} + \Delta t A_h X^n_{h,\Delta t} = P_h(L(t_n) - L(t_{n-1})),\ n = 1, \dots, N;\ X^0_{h,\Delta t} = P_h X_0. \tag{3.3.3}$$

Remark 3.3.1 (strong error). If the covariance operator $Q \in \mathcal{L}(H)$ of L is such that

$$\|A^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty \tag{3.3.4}$$

for some $\beta \geqslant 0$, then the solution $X(t)$ takes values in $\dot{H}^\beta$ for all $t > 0$. For the Gaussian case, i.e., the case where L in (3.1.3) is a Q-Wiener process, it has been shown in [**110**, Theorem 1.2] (see also Theorem 2.3.2) that, if (3.3.4) holds and $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \dot{H}^\beta)$ for some $\beta \in (0,1]$, then the scheme (3.3.3) has strong convergence of order β in space and $\beta/2$ in time:

$$\|X^n_{h,\Delta t} - X(t_n)\|_{L^2(\Omega;H)} \leqslant C(h^\beta + \Delta t^{\frac{\beta}{2}}), \quad n = 0, \dots, N.$$

Unlike weak error estimates, strong L^2-error estimates are the same in the Gaussian case and in our setting, since the only stochastic tool that is needed is Itô's isometry (3.1.19) which looks the same if the driving noise is a Lévy process which is an L^2-martingale. Thus the strong error result in [**110**, Theorem 1.2] carries over to our setting.

Remark 3.3.2. The V_h-valued random variables $P_h(L(t_n)-L(t_{n-1}))$ in (3.3.3) can be defined in two ways. On the one hand, we may set

$$P_h(L(t_n)-L(t_{n-1})) := L^2(\Omega;V_h)\text{-}\lim_{K\to\infty}\sum_{k=1}^{K}(L_k(t_n)-L_k(t_{n-1}))P_h e_k,$$

with an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of U_0 and real-valued uncorrelated Lévy processes $L_k=(L_k(t))_{t\geqslant 0}$, $k\in\mathbb{N}$, as in Remark 3.1.3. The limit exists since, by the finite-dimensionality of V_h, one has $P_h\in\mathcal{L}_2(H,V_h)=\mathcal{L}_2(U,V_h)\subset\mathcal{L}_2(U_0,V_h)$. On the other hand, we can extend the orthogonal projection $P_h: H\to V_h$ to a generalized L^2-projection $P_h:\dot H^{-1}\to V_h$ defined by

$$\langle P_h v,w\rangle_H=\langle v,w\rangle_{\dot H^{-1}\times\dot H^{1}},\quad v\in\dot H^{-1},\ w\in V_h.$$

Then, the assumption $\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}}<\infty$ implies that we can take

$$U_1:=D(A^{-1/2})=\dot H^{-1}$$

as the state space of L, so that the expression $P_h(L(t_n)-L(t_{n-1}))$ makes sense ω-wise. Obviously, both definitions are compatible. In practice, one has to find a suitable way to sample (an approximation of) the discretized noise increment $P_h(L(t_n)-L(t_{n-1}))$. We do not treat this problem here but refer to [**18, 37**] and [**75**, Remark 4] for related considerations.

With $R(\lambda):=1/(1+\lambda)$ and $E_{h,\Delta t}:=R(\Delta tA_h):=(I+\Delta tA_h)^{-1}$ as well as $E^n_{h,\Delta t}:=R^n(\Delta tA_h):=((I+\Delta tA_h)^{-1})^n$, the scheme (3.3.3) can be rewritten as

$$X^n_{h,\Delta t}=E^n_{h,\Delta t}P_hX_0+\sum_{j=1}^{n}E^{n-j+1}_{h,\Delta t}P_h(L(t_j)-L(t_{j-1})),\quad n=0,\dots,N.$$

For $t\in[0,T]$, let $\tilde E(t)=\tilde E_{h,\Delta t}(t)\in\mathcal{L}(H)$ be defined by

$$\tilde E(t)=\tilde E_{h,\Delta t}(t):=\mathbb{1}_{\{0\}}(t)P_h+\sum_{n=1}^{N}\mathbb{1}_{(t_{n-1},t_n]}(t)E^n_{h,\Delta t}P_h \tag{3.3.5}$$

and set

$$\tilde X(t)=\tilde X_{h,\Delta t}(t):=\tilde E_{h,\Delta t}(t)X_0+\int_0^t\tilde E_{h,\Delta t}(t-s)\,\mathrm{d}L(s). \tag{3.3.6}$$

Then $X^n_{h,\Delta t}=\tilde X_{h,\Delta t}(t_n)$, $\mathbb{P}$-almost surely. This follows from the construction of the stochastic integral, using an approximation argument and Itô's isometry (3.1.19).

The following deterministic estimates will be used in the proof of our weak error result stated in Theorem 3.3.4 below.

Lemma 3.3.3. *The operators $E(t)$ and $\tilde E(t)=\tilde E_{h,\Delta t}(t)$ defined in* (3.3.1) *and* (3.3.5) *satisfy the error estimates*

$$\|\tilde E(s)-E(s)\|_{\mathcal{L}(H)}\leqslant C(h^2+\Delta t)s^{-1}, \tag{3.3.7}$$

$$\|A^\alpha E(s)\|_{\mathcal{L}(H)}+\|A^\alpha\tilde E(s)\|_{\mathcal{L}(H)}\leqslant Cs^{-\alpha},\quad 0\leqslant\alpha\leqslant 1/2, \tag{3.3.8}$$

$s\in(0,T]$, where $C>0$ does not depend on h, Δt and s.

Proof. Estimate (3.3.7) follows from

$$\|E^n_{h,\Delta t}P_h - E(t_n)\|_{\mathcal{L}(H)} \leqslant C(h^2+\Delta t)t_n^{-1}, \tag{3.3.9}$$

see, for example, [**103**, Theorem 7.7]. We note here that while the latter result is proved under the assumption that $\mathcal{D}$ has smooth boundary, the proof remains valid under (1.5.51) on R_h, which is our basic assumption, and hence (3.3.9) is valid in our setting. For $s\in(t_{n-1},t_n]$ we have

$$\begin{aligned}\|(E(t_n)-E(s))v\|_H &= \|AE(s)(E(t_n-s)-\mathrm{id}_H)A^{-1}v\|_H\\ &\leqslant \|AE(s)\|_{\mathcal{L}(H)}\|(E(t_n-s)-\mathrm{id}_H)A^{-1}v\|_H\\ &\leqslant Cs^{-1}\Delta t\|v\|_H,\end{aligned}$$

where we used [**87**, Chapter 2, Theorem 6.13(c),(d)] on analytic semigroups. Estimate (3.3.8) is due to [**87**, Chapter 2, Theorem 6.13(c)], [**103**, Lemma 7.3], interpolation, and the fact that $\|A^\alpha v_h\|\leqslant\|A_h^\alpha v_h\|$ for $v_h\in V_h$, $0\leqslant\alpha\leqslant 1/2$. The latter follows from the basic identity $\|A^{1/2}v_h\|=\|A_h^{1/2}v_h\|$ and interpolation. □

Next we state our result concerning the weak error of the discretization of the stochastic heat equation. Note that the rate of convergence is the same as in the Gaussian case under identical regularity assumptions, c.f. Remark 1.4.4.

Theorem 3.3.4. *Assume that $X_0\in L^2(\Omega,\mathcal{F}_0,\mathbb{P};H)$ and $\|A^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}}<\infty$ for some $\beta\in(0,1]$. Let $(X(t))_{t\geqslant0}$ be the weak solution* (3.1.1) *of* (3.1.2) *and let $(X^n_{h,\Delta t})_{n=0,\dots,N}$ be given by* (3.3.3)*. If $g\in C^2(H,\mathbb{R})$ with $\sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}<\infty$, then there exists a constant $C=C(g,T)>0$, that does not depend on h and Δt, such that*

$$\big|\mathbb{E}\big(g(X^N_{h,\Delta t})-g(X(T))\big)\big|\leqslant C(h^{2\beta}+\Delta t^\beta)|\log(h^2+\Delta t)|$$

for $h^2+\Delta t\leqslant 1/e$.

Proof. We are in the setting of Section 3.1 with $H=U=L^2(\mathcal{D})$, $B=\mathrm{id}_H$, and $(E(t))_{t\geqslant0}$, $(\tilde E(t))_{t\in[0,T]}=(\tilde E_{h,\Delta t}(t))_{t\in[0,T]}$, $(\tilde X(t))_{t\in[0,T]}=(\tilde X_{h,\Delta t}(t))_{t\in[0,T]}$ being given by (3.3.1), (3.3.5), (3.3.6) respectively. In particular, Assumption 3.1.7 is fulfilled. Since $X^N_{h,\Delta t}=\tilde X(T)$, we can use Corollary 3.2.8 with $G:=g$ to estimate the weak error. Let $F(t):=\tilde E(t)-E(t)$ be the deterministic error operator.

We begin with the first term on the right hand side of (3.2.11) in Corollary 3.2.8. The stability estimate (3.3.2) and the deterministic estimate (3.3.7)

dc_1541_18

yield, for $\max(h^2, \Delta t) \leqslant 1$,

(3.3.10)

$$\begin{aligned}
&\big|\mathbb{E}\big\{u(0,\tilde{E}(T)X_0) - u(0,E(T)X_0)\big\}\big| \\
&= \big|\mathbb{E}\big\{u(0,\tilde{Y}(0)) - u(0,Y(0))\big\}\big| \\
&= \Big|\mathbb{E}\int_0^1 \big\langle u_x\big(0, Y(0)+\theta(\tilde{Y}(0)-Y(0))\big), \tilde{Y}(0)-Y(0)\big\rangle_H \mathrm{d}\theta\Big| \\
&= \Big|\mathbb{E}\int_0^1 \big\langle \mathbb{E}\big(g'(Z(T,0,x))\big)\big|_{x=Y(0)+\theta(\tilde{Y}(0)-Y(0))}, \tilde{Y}(0)-Y(0)\big\rangle_H \mathrm{d}\theta\Big| \\
&\leqslant \int_0^1 \big\|g'\big(Z\big(T,0,Y(0)+\theta(\tilde{Y}(0)-Y(0))\big)\big)\big\|_{L^2(\Omega,H)} \mathrm{d}\theta\, \|(\tilde{E}(T)-E(T))X_0\|_{L^2(\Omega,H)} \\
&\leqslant C\big(1+\int_0^1 \big\|Z\big(T,0,Y(0)+\theta(\tilde{Y}(0)-Y(0))\big)\big\|_{L^2(\Omega,H)}\mathrm{d}\theta\big) \\
&\qquad \times (h^2+\Delta t)\, T^{-1}\, \|X_0\|_{L^2(\Omega;H)} \\
&\leqslant C\big(1+\|A^{-1/2}Q^{1/2}\|_{\mathrm{HS}} + \|X_0\|_{L^2(\Omega;H)}\big)\, \|X_0\|_{L^2(\Omega;H)}\, T^{-1}\,(h^{2\beta}+\Delta t^{\beta}).
\end{aligned}$$

Next, we consider the second term on the right hand side of (3.2.11). We estimate the integrals of the functions Ψ_1 and Ψ_2 separately. Using Lemma 3.2.1 and Remark 3.2.2, we obtain

$$\begin{aligned}
&\big|\mathbb{E}\int_0^T\int_{U_1}\int_0^1 \Psi_1(t,\theta,y)\,\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\big| \\
&\quad\leqslant \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\int_0^T\int_{U_1}\|F(T-t)y\|_H^2\,\nu(\mathrm{d}y)\,\mathrm{d}t \\
&\quad= \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\int_0^T\|F(T-t)\|^2_{\mathcal{L}_2(U_0,H)}\,\mathrm{d}t \\
&\quad\leqslant C\sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}(h^{2\beta}+\Delta t^{\beta}).
\end{aligned} \tag{3.3.11}$$

The last step is due to the fact that, by Itô's isometry (3.1.19), the integral in the penultimate line is the square of the strong error $\|X^N_{h,\Delta t} - X(T)\|_{L^2(\Omega;H)}$ for zero initial condition $X_0 = 0$, which can be estimated as in the Gaussian case [**110**, Theorem 1.2], compare with Remark 3.3.1. Further, by the Cauchy-Schwarz inequality, Lemma 3.2.1, and the fact that $U_0 = Q^{1/2}(U)$,

(3.3.12)

$$\begin{aligned}
&\big|\mathbb{E}\int_0^T\int_{U_1}\int_0^1 \Psi_2(t,\theta,y)\,\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\big| \\
&\quad\leqslant \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\int_0^T\int_{U_1}\|E(T-t)y\|_H\|F(T-t)y\|_H\,\nu(\mathrm{d}y)\,\mathrm{d}t \\
&\quad\leqslant \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\int_0^T\|E(T-t)\|_{\mathcal{L}_2(U_0,H)}\|F(T-t)\|_{\mathcal{L}_2(U_0,H)}\,\mathrm{d}t \\
&\quad\leqslant \sup_{x\in H}\|g''(x)\|_{\mathcal{L}(H)}\|A^{\frac{\beta-1}{2}}Q^{1/2}\|^2_{\mathcal{L}_2(H)}\int_0^T\|E(t)A^{\frac{1-\beta}{2}}\|_{\mathcal{L}(H)}\|F(t)A^{\frac{1-\beta}{2}}\|_{\mathcal{L}(H)}\,\mathrm{d}t.
\end{aligned}$$

By (3.3.8) we have

$$\|E(t)A^{\frac{1-\beta}{2}}\|_{\mathcal{L}(H)} = \|A^{\frac{1-\beta}{2}}E(t)\|_{\mathcal{L}(H)} \leqslant Ct^{-\frac{1-\beta}{2}} \tag{3.3.13}$$

and

$$\|A^{\alpha}F(t)\|_{\mathcal{L}(H)} \leqslant Ct^{-\alpha}, \quad 0 \leqslant \alpha \leqslant 1/2. \tag{3.3.14}$$

Interpolation between (3.3.7) and (3.3.14) with $\alpha = 1/2$ gives

$$\|A^{\frac{1-\beta}{2}}F(t)\|_{\mathcal{L}(H)} \leqslant C\|F(t)\|^{\beta}_{\mathcal{L}(H)}\|A^{\frac{1}{2}}F(t)\|^{1-\beta}_{\mathcal{L}(H)} \leqslant C(h^2+\Delta t)^{\beta}t^{-\frac{1+\beta}{2}}. \tag{3.3.15}$$

Note that $\|F(t)A^{\alpha}\|_{\mathcal{L}(H)} = \|A^{\alpha}F(t)\|_{\mathcal{L}(H)}$ due to the self adjointness of $\tilde{E}(t)$, $E(t)$ and A^{α}. Altogether, using (3.3.13), (3.3.14) and (3.3.15), the integral in the last line of (3.3.12) can be estimated by

$$\begin{aligned}
&\int_0^T \|E(t)A^{\frac{1-\beta}{2}}\|_{\mathcal{L}(H)}\|F(t)A^{\frac{1-\beta}{2}}\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\quad= \Big(\int_0^{h^2+\Delta t} + \int_{h^2+\Delta t}^T\Big)\|A^{\frac{1-\beta}{2}}E(t)\|_{\mathcal{L}(H)}\|A^{\frac{1-\beta}{2}}F(t)\|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\quad\leqslant C\int_0^{h^2+\Delta t} t^{-\frac{1-\beta}{2}}t^{-\frac{1-\beta}{2}}\,\mathrm{d}t + C\int_{h^2+\Delta t}^T t^{-\frac{1-\beta}{2}}(h^2+\Delta t)^{\beta}t^{-\frac{1+\beta}{2}}\,\mathrm{d}t \\
&\quad= C(h^2+\Delta t)^{\beta}(1+|\log(h^2+\Delta t)|) \\
&\quad\leqslant C(h^{2\beta}+\Delta t^{\beta})|\log(h^2+\Delta t)|
\end{aligned} \tag{3.3.16}$$

for $h^2 + \Delta t \leqslant 1/e$, where $C > 0$ depends on T. The combination of (3.3.10), (3.3.11), (3.3.12) and (3.3.16) finishes the proof. □

3.4. Stochastic Volterra integro-differential equations

Here we consider a stochastic integro-differential equation of Volterra-type where the deterministic equation exhibits a parabolic character. The error analysis is basically analogous to the heat equation and therefore we skip some computational details. The weak solution of the Volterra-type stochastic evolution equation, a simple model of viscoelastic materials in the presence of noise,

$$\mathrm{d}X(t) + \left(\int_0^t b(t-s)AX(s)\,\mathrm{d}s\right)\mathrm{d}t = \mathrm{d}L(t),\ t\in(0,T];\ X(0)=X_0\in H, \tag{3.4.1}$$

is also given by (3.1.1), where $E(t) := S(t)$ is the solution operator of the linear, homogeneous deterministic problem as described in Subsection 1.5.1 and $(A, D(A)) := (\Lambda, D(\Lambda))$ as in the Section 3.3. Using a finite element approximation in space as for the heat equation and a convolution quadrature in time, in exact analogy with the Gaussian case (1.5.61), we consider the following recurrence,

$$X^n_{h,\Delta t} - X^{n-1}_{h,\Delta t} + \Delta t\left(\sum_{k=1}^n \omega_{n-k}\,A_hX^k_{h,\Delta t}\right) = P_h(L(t_n)-L(t_{n-1})), \quad n\geqslant 1, \tag{3.4.2}$$

with $X^0_{h,\Delta t} = P_hX_0$ and convolution weights $(\omega_k)_{k\geq 0}$ chosen according to

$$\left(\frac{1-z}{\Delta t}\right)^{1-\rho} = \sum_{k\geqslant 0}\omega_k z^k, \quad |z|<1.$$

Let $\tilde{E}(t) = \tilde{B}_{h,\Delta t}(t) = \tilde{E}_{h,\Delta t}(t)$, where $\tilde{B}_{h,\Delta t}(t)$ is defined in (1.5.66) and let $(\tilde{X}(t))_{t\in[0,T]}$ be given according to (3.3.6). Suppose that that b satisfies Assumptions 1.5.1 and 1.5.30. Then, according to Theorem 1.5.32 and Proposition 1.5.6,

$$\|\tilde{E}(s) - E(s)\|_{\mathcal{L}(H)} \leqslant C(h^{2/\rho} + \Delta t)s^{-1}, \tag{3.4.3}$$

$$\|A^{\alpha} E(s)\|_{\mathcal{L}(H)} + \|A^{\alpha}\tilde{E}(s)\|_{\mathcal{L}(H)} \leqslant Cs^{-\rho\alpha}, \quad 0 \leqslant \alpha \leqslant 1/(2\rho), \tag{3.4.4}$$

$s \in (0, T]$, where $C > 0$ does not depend on h, Δt and s. Note further, that if $\|A^{-1/(2\rho)}Q^{1/2}\|_{\mathcal{L}_2(H)} < \infty$, then using Itô's isometry, we get the stability estimate as in the Gaussian case (see the proof of Proposition 1.5.10 with $\beta = 0$),

$$\|X(T)\|_{L^2(\Omega,H)} \leq C(\|A^{-1/(2\rho)}Q^{1/2}\|_{\mathrm{HS}} + \|X_0\|_{L^2(\Omega,H)}),$$

and, in particular, Assumption 3.1.7 (iv) holds. Furthermore, for $t \in [\varepsilon, T]$ and $x \in H$, we have that

$$\|E(t)x\|_H = \|A^{1/(2\rho)}E(t)A^{-1/(2\rho)}x\|_H$$
$$\leq \|A^{1/(2\rho)}E(t)\|_{\mathcal{L}(H)}\| \, \|A^{-1/(2\rho)}x\|_H \leq \|\Phi_\varepsilon x\|_H.$$

Therefore, Assumption 3.1.7 (v) holds with

$$\Phi_\varepsilon := \sup_{t\in[\varepsilon,T]} \|A^{1/(2\rho)}E(t)\|_{\mathcal{L}(H)}A^{-1/(2\rho)},$$

where the supremum is finite because of (3.4.4). Hence, via an analogous calculation as for the heat equation above using now (3.4.3) and (3.4.4), setting $H = U = L^2(\mathcal{D})$, $B = \mathrm{id}_H$ we arrive at the following result.

THEOREM 3.4.1. *Suppose that b satisfies Assumptions 1.5.1 and 1.5.30. Assume further that $\|A^{\frac{\beta-1/\rho}{2}}Q^{\frac{1}{2}}\|^2_{\mathrm{HS}} < \infty$, $\beta \in (0, 1/\rho)$ and $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$. Let $(X(t))_{t\geqslant 0}$ be the weak solution of (3.4.1) and let $(X^n_{h,\Delta t})_{n=0,\ldots,N}$ be defined by (3.4.2). If $g \in C^2(H, \mathbb{R})$ with $\sup_{x\in H} \|g''(x)\|_{\mathcal{L}(H)} < \infty$, then there exists a constant $C = C(g, T) > 0$, that does not depend on h and Δt, such that*

$$\left|\mathbb{E}\big(g(X^N_{h,\Delta t}) - g(X(T))\big)\right| \leqslant C(h^{2\beta} + \Delta t^{\rho\beta})|\log(h^{2/\rho} + \Delta t)|,$$

for $h^{2/\rho} + \Delta t \leqslant 1/e$.

This is essentially twice the strong rate where the latter is the same as in the Gaussian case found Theorem 1.5.28, as the strong error analysis carries over to our setting, cf. Remark 3.3.1. Note also that the weak rate agrees with the one in the Gaussian case from Theorem 1.5.33 under the same regualrity assumptions on the data.

3.5. The stochastic wave equation

Here, we apply the general error representation from Section 3.2 to a discretization of the stochastic wave equation (3.1.4).

Let $\mathcal{D} \subset \mathbb{R}^d$ be a convex bounded domain and let the spaces $\mathcal{H}^\alpha$ be defined by (1.3.3). We set

$$H := \mathcal{H}^0 = \dot{H}^0 \times \dot{H}^{-1} = L^2(\mathcal{D}) \times H^{-1}(\mathcal{D}), \quad U := \dot{H}^0 = L^2(\mathcal{D})$$

and let $A : D(A) \subset H \to H$ and $B \in \mathcal{L}(U, H)$ defined by (1.3.4) and let the semigroup $(E(t))_{t\geqslant 0}$ generated by $-A$ be given by (1.3.5). Then abstract equation (3.1.2) becomes the stochastic wave equation (3.1.4) with H-valued solution

dc_1541_18

$(X(t))_{t\geqslant 0} = ((X_1(t), X_2(t))^\top)_{t\geqslant 0}$. As in the Gaussian case, cf. Theorem 1.3.3, one sees that the condition $\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ implies (3.1.17) and hence the existence of a unique weak solution $X = (X(t))_{t\geqslant 0}$ given by (3.1.1), provided that the initial condition $X_0 = (X_{0,1}, X_{0,2})^\top$ is H-valued and $\mathcal{F}_0$-measurable. Furthermore,

$$\|X(T)\|_{L^2(\Omega,H)} \leq C(\|X_0\|_{L^2(\Omega,H)} + T^{\frac{1}{2}}\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}}). \tag{3.5.1}$$

The discretization of (3.1.4) is done via finite elements in space and an I-stable rational single step scheme of order in time as in Subsection 1.3.3. Indeed, let the discretization $A_h : V^r_{h,0} \times V^r_{h,0} \to V^r_{h,0} \times V^r_{h,0}$ of the operator $A : D(A) \subset H \to H$ be defined by (1.3.25) and let $(E_h(t))_{t\geqslant 0} \subset \mathcal{L}(V^r_{h,0} \times V^r_{h,0})$ denote the strongly continuous semigroup generated by $-A_h$ given by (1.3.27). We consider for $N \in \mathbb{N}$ a uniform grid $t_n = n\Delta t = n(T/N)$, $n = 0, \ldots, N$, on a finite time interval $[0, T]$. We approximate the operators $E^n_{h,\Delta t} \in \mathcal{L}(V^r_{h,0} \times V^r_{h,0})$ by

$$E^n_{h,\Delta t} := (R(\Delta t A_h))^n,$$

where R is a rational function that fullfills (1.3.16) for some positive integer p.

The numerical scheme for the stochastic wave equation (3.1.2) can now be formulated as follows: For $h > 0$ and $N \in \mathbb{N}$, the discretization $(X^n_{h,\Delta t})_{n=0,\ldots,N}$ of $(X(t))_{t\in[0,T]}$ in space and time is given as the solution to

(3.5.2)

$$X^n_{h,\Delta t} = E_{h,\Delta t}\big(X^{n-1}_{h,\Delta t} + P_h B(L(t_n) - L(t_{n-1}))\big), \quad n = 1, \ldots, N; \quad X^0_{h,\Delta t} = P_h X_0,$$

where $P_h = [P_{h,1}, P_{h,2}]^T$ with $P_{h,1} \colon \dot{H}^0 \to V^r_{h,0}$ and $P_{h,2} \colon \dot{H}^{-1} \to V^r_{h,0}$ being the orthogonal projectors defined by $\langle P_{h,1} f, \chi\rangle = \langle f, \chi\rangle$, $\forall \chi \in V^r_{h,0}$, for $f \in \dot{H}^0$ and $\langle P_{h,2} f, \chi\rangle = \langle f, \chi\rangle_{\dot{H}^{-1}\times\dot{H}^1}$, $\forall \chi \in V^r_{h,0}$, for $f \in \dot{H}^{-1}$. As in the Gaussian case, we assume that the Ritz projection R_h satisfies (1.3.22).

Remark 3.5.1 (strong error). As observed for the discretization of the heat equation in Remark 3.3.1, strong L^2-error estimates for the scheme (3.5.2) carry over from the Gaussian case in the Lévy L^2-martingale case since they only use Itô's isometry (3.1.19). Arguing as in the proof of Theorem 1.3.14, we obtain that, if

$$\|\Lambda^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty \quad \text{and} \quad X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathcal{H}^\beta) \tag{3.5.3}$$

for some $\beta > 0$, then the scheme (3.5.2) approximates the first component $X_1 = P^1 X$ of the solution X to (3.1.2) with strong order $\min(\beta r/(r+1), r)$ in space and $\min(\beta p/(p+1), 1)$ in time, where $P^1 \colon H \to \dot{H}^0$ is defined as $P^1 x = x_1$ for $x = [x_1, x_2]^T \in H$:

$$\|X^n_{h,\Delta t,1} - X_1(t_n)\|_{L^2(\Omega;\dot{H}^0)} \leqslant C\big(h^{\min(\beta\frac{r}{r+1}, r)} + \Delta t^{\min(\beta\frac{p}{p+1}, 1)}\big), \quad n = 0, \ldots, N.$$

Here we have set $X^n_{h,\Delta t,1} := P^1 X^n_{h,\Delta t}$. The condition (3.5.3) implies that the solution $X = (X(t))_{t\geqslant 0}$ takes values in $\mathcal{H}^\beta$, cf. Theorem 1.3.3.

The solution to the scheme (3.5.2) is given by

$$X^n_{h,\Delta t} = E^n_{h,\Delta t} P_h X_0 + \sum_{j=1}^{n} E^{n-j+1}_{h,\Delta t} P_h B(L(t_n) - L(t_{n-1})), \quad n = 0, \ldots, N.$$

For $t \in [0,T]$, define operators $\tilde{E}(t) = \tilde{E}_{h,\Delta t}(t) \in \mathcal{L}(H)$ by

$$\tilde{E}(t) = \tilde{E}_{h,\Delta t}(t) := \mathbb{1}_{\{0\}}(t) P_h + \sum_{j=1}^{N} \mathbb{1}_{(t_{n-1},t_n]}(t) E^n_{h,\Delta t} P_h. \tag{3.5.4}$$

Then, analogously to the corresponding argument in Section 3.3, one sees that the $V^r_{h,0} \times V^r_{h,0}$-valued process $(\tilde{X}(t))_{t\in[0,T]} = (\tilde{X}_{h,\Delta t}(t))_{t\in[0,T]}$ defined by

$$\tilde{X}(t) = \tilde{X}_{h,\Delta t}(t) := \tilde{E}_{h,\Delta t}(t) X_0 + \int_0^t \tilde{E}_{h,\Delta t}(t-s) B \, \mathrm{d}L(s)$$

satisfies $X^n_{h,\Delta t} = \tilde{X}(t_n)$ $\mathbb{P}$-almost surely. We need the following deterministic error estimate.

Proposition 3.5.2. *Let $\alpha \geqslant 0$. The operators $E(t)$ and $\tilde{E}(t) = \tilde{E}_{h,\Delta t}(t)$ defined by* (1.3.5) *and* (3.5.4)*, respectively, satisfy the error estimate*

$$\sup_{t\in[0,T]} \big(\|P^1(\tilde{E}(t) - E(t))\|_{\mathcal{L}(\mathcal{H}^\alpha, \dot{H}^0)} + \|P^1(\tilde{E}(t) - E(t))B\|_{\mathcal{L}(\dot{H}^{(\alpha/2)-1}, \dot{H}^{-\alpha/2})} \big) \leqslant C \big(h^{\min(\alpha\frac{r}{r+1}, r)} + \Delta t^{\min(\alpha\frac{p}{p+1}, 1)} \big), \tag{3.5.5}$$

for $\Delta t \leqslant 1$, where $C = C(T) > 0$ does not depend on h and Δt.

Proof. We use the estimates

$$\sup_{n\in\{0,\dots,N\}} \|P^1(E^n_{h,\Delta t} P_h - E(t_n))\|_{\mathcal{L}(\mathcal{H}^\alpha, \dot{H}^0)} \leqslant C(T) \big(h^{\min(\alpha\frac{r}{r+1}, r)} + \Delta t^{\min(\alpha\frac{p}{p+1}, p)} \big) \tag{3.5.6}$$

and

$$\|E(t) - E(s)\|_{\mathcal{L}(\mathcal{H}^\delta, H)} \leqslant C|t-s|^\delta, \quad t,\, s \geqslant 0,\ \delta \in [0,1]. \tag{3.5.7}$$

from Corollary 1.3.12 and Lemma 1.3.5. Because of the 'piecewise' definition of $\tilde{E}(t)$ in (3.5.4), the combination of (3.5.6) and (3.5.7) gives

$$\begin{aligned} &\sup_{t\in[0,T]} \|P^1(\tilde{E}(t) - E(t))\|_{\mathcal{L}(\mathcal{H}^\alpha, \dot{H}^0)} \\ &\leqslant \sup_{n\in\{0,\dots,N\}} \|P^1(\tilde{E}(t_n) - E(t_n))\|_{\mathcal{L}(\mathcal{H}^\alpha, \dot{H}^0)} \\ &\qquad + \sup_{n\in\{1,\dots,N\}} \sup_{t\in(t_{n-1},t_n)} \|E(t_n) - E(t)\|_{\mathcal{L}(\mathcal{H}^\alpha, H)} \\ &\leqslant C(T) \big(h^{\min(\alpha\frac{r}{r+1}, r)} + \Delta t^{\min(\alpha\frac{p}{p+1}, p)} + \Delta t^{\min(\alpha, 1)} \big) \\ &= C(T) \big(h^{\min(\alpha\frac{r}{r+1}, r)} + \Delta t^{\min(\alpha\frac{p}{p+1}, 1)} \big) \end{aligned} \tag{3.5.8}$$

for $\Delta t \leqslant 1$. It remains to show that

$$\sup_{t\in[0,T]} \|P^1(\tilde{E}(t) - E(t))B\|_{\mathcal{L}(\dot{H}^{(\alpha/2)-1}, \dot{H}^{-\alpha/2})} \leqslant C(T) \big(h^{\min(\alpha\frac{r}{r+1}, r)} + \Delta t^{\min(\alpha\frac{p}{p+1}, 1)} \big). \tag{3.5.9}$$

To this end, we will prove the estimate

$$\sup_{n\in\{0,\dots,N\}} \|P^1(\tilde{E}(t_n) - E(t_n))B\|_{\mathcal{L}(\dot{H}^{(\alpha/2)-1}, \dot{H}^{-\alpha/2})} \leqslant C(T) \big(h^{\min(\alpha\frac{r}{r+1}, r)} + \Delta t^{\min(\alpha\frac{p}{p+1}, p)} \big). \tag{3.5.10}$$

dc_1541_18

Then, (3.5.9) follows from (3.5.10) and (3.5.7) arguing similarly as in (3.5.8) and using the fact that

$$\begin{aligned}\|P^1(E(t_n)-E(t))B\|_{\mathcal{L}(\dot H^{(\alpha/2)-1},\dot H^{-\alpha/2})} &= \|\Lambda^{-\frac{\alpha}{4}}P^1(E(t_n)-E(t))B\Lambda^{\frac12-\frac{\alpha}{4}}\|_{\mathcal{L}(\dot H^0)}\\ &= \|P^1(E(t_n)-E(t))B\Lambda^{\frac{1-\alpha}{2}}\|_{\mathcal{L}(\dot H^0)}\\ &\leqslant \|P^1(E(t_n)-E(t))\|_{\mathcal{L}(\mathcal{H}^\alpha,\dot H^0)}\|B\Lambda^{\frac{1-\alpha}{2}}\|_{\mathcal{L}(\dot H^0,\mathcal{H}^\alpha)},\end{aligned}$$

where $\|B\Lambda^{\frac{1-\alpha}{2}}\|_{\mathcal{L}(\dot H^0,\mathcal{H}^\alpha)} = \|B\|_{\mathcal{L}(\dot H^{\alpha-1},\mathcal{H}^\alpha)} = 1$.

In order to show (3.5.10), we distinguish the cases $\alpha > 2$ and $0 \leqslant \alpha \leqslant 2$. For $\alpha > 2$ we have by (3.5.6)

$$\begin{aligned}&\sup_{n\in\{0,\dots,N\}}\|P^1(\tilde E(t_n)-E(t_n))B\|_{\mathcal{L}(\dot H^{\alpha-1},\dot H^0)}\\ &\leqslant \sup_{n\in\{0,\dots,N\}}\|P^1(\tilde E(t_n)-E(t_n))\|_{\mathcal{L}(\mathcal{H}^\alpha,\dot H^0)}\|B\|_{\mathcal{L}(\dot H^{\alpha-1},\mathcal{H}^\alpha)}\\ &\leqslant C(T)\big(h^{\min(\alpha\frac{r}{r+1},r)}+\Delta t^{\min(\alpha\frac{p}{p+1},p)}\big).\end{aligned} \tag{3.5.11}$$

As the operator $P^1(\tilde E(t)-E(t))B \in \mathcal{L}(\dot H^0)$ is symmetric in $\dot H^0$ and since $\dot H^{-\alpha+1}$ can be identified with the dual space of $\dot H^{\alpha-1}$, we have

$$\|P^1(\tilde E(t)-E(t))B\|_{\mathcal{L}(\dot H^{\alpha-1},\dot H^0)} = \|P^1(\tilde E(t)-E(t))B\|_{\mathcal{L}(\dot H^0,\dot H^{-\alpha+1})}$$

and therefore also

$$\begin{aligned}\sup_{n\in\{0,\dots,N\}}&\|P^1(\tilde E(t_n)-E(t_n))B\|_{\mathcal{L}(\dot H^0,\dot H^{-\alpha+1})}\\ &\leqslant C(T)\big(h^{\min(\alpha\frac{r}{r+1},r)}+\Delta t^{\min(\alpha\frac{p}{p+1},p)}\big).\end{aligned} \tag{3.5.12}$$

Next, we use the fact that $\dot H^{(\alpha/2)-1}$ and $\dot H^{-\alpha/2}$ can be represented as the real interpolation spaces $(\dot H^0,\dot H^{\alpha-1})_{\theta,2}$ and $(\dot H^{-\alpha+1},\dot H^0)_{\theta,2}$, respectively, where $\theta = ((\alpha/2)-1)/(\alpha-1) \in (0,1)$, cf. Remark 1.3.1. Thus, interpolation between (3.5.11) and (3.5.12) yields

$$\begin{aligned}&\sup_{n\in\{0,\dots,N\}}\|P^1(\tilde E(t_n)-E(t_n))B\|_{\mathcal{L}(\dot H^{(\alpha/2)-1},\dot H^{-\alpha/2})}\\ &\leqslant \sup_{n\in\{0,\dots,N\}} C(\alpha)\|P^1(\tilde E(t_n)-E(t_n))B\|^{1-\theta}_{\mathcal{L}(\dot H^0,\dot H^{-\alpha+1})}\\ &\qquad\times\|P^1(\tilde E(t_n)-E(t_n))B\|^{\theta}_{\mathcal{L}(\dot H^{\alpha-1},\dot H^0)}\\ &\leqslant C(T,\alpha)\big(h^{\min(\alpha\frac{r}{r+1},r)}+\Delta t^{\min(\alpha\frac{p}{p+1},p)}\big),\end{aligned}$$

see, e.g., Definition 1.2.2/2 and Theorem 1.3.3(a) in [**104**].

For $0 \leqslant \alpha \leqslant 2$, we note that

$$\begin{aligned}\|P^1(\tilde E(t_n)-E(t_n))B\|_{\mathcal{L}(\dot H^0,\dot H^{-1})} &= \|P^1(\tilde E(t_n)-E(t_n))B\|_{\mathcal{L}(\dot H^1,\dot H^0)}\\ &\leqslant \|P^1(\tilde E(t_n)-E(t_n))\|_{\mathcal{L}(\mathcal{H}^2,\dot H^0)}\|B\|_{\mathcal{L}(\dot H^1,\mathcal{H}^2)},\end{aligned}$$

where we used again the symmetry of $P^1(\tilde E(t)-E(t))B \in \mathcal{L}(\dot H^0)$. By (3.5.6) we obtain

$$\sup_{n\in\{0,\dots,N\}}\|P^1(\tilde E(t_n)-E(t_n))B\|_{\mathcal{L}(\dot H^0,\dot H^{-1})} \leqslant C(T)\big(h^{\min(2\frac{r}{r+1},r)}+\Delta t^{\min(2\frac{p}{p+1},p)}\big), \tag{3.5.13}$$

dc_1541_18

which is (3.5.10) for $\alpha = 0$. Moreover, also by (3.5.6),

$$\begin{aligned} &\sup_{n\in\{0,\dots,N\}} \|P^1(\tilde{E}(t_n) - E(t_n))B\|_{\mathcal{L}(\dot{H}^{-1},\dot{H}^0)} \\ &\qquad\leqslant \sup_{n\in\{0,\dots,N\}} \|P^1(\tilde{E}(t_n) - E(t_n))\|_{\mathcal{L}(H,\dot{H}^0)} \|B\|_{\mathcal{L}(\dot{H}^{-1},H)} \leqslant C(T), \end{aligned} \tag{3.5.14}$$

i.e., we have (3.5.10) for $\alpha = 2$. Finally, if $\alpha \in (0,2)$, interpolation with $\theta = (\alpha/2) - 1 \in (0,1)$ between (3.5.13) and (3.5.14) gives

$$\begin{aligned} &\sup_{n\in\{0,\dots,N\}} \|P^1(\tilde{E}(t_n) - E(t_n))B\|_{\mathcal{L}(\dot{H}^{(\alpha/2)-1},\dot{H}^{-\alpha/2})} \\ &\quad\leqslant \sup_{n\in\{0,\dots,N\}} C(\alpha) \|P^1(\tilde{E}(t_n) - E(t_n))B\|^{1-\theta}_{\mathcal{L}(\dot{H}^0,\dot{H}^{-1})} \\ &\quad\quad \times \|P^1(\tilde{E}(t_n) - E(t_n))B\|^{\theta}_{\mathcal{L}(\dot{H}^{-1},\dot{H}^0)} \\ &\quad\leqslant C(T,\alpha)\big(h^{\min(2\frac{r}{r+1},r)} + \Delta t^{\min(2\frac{p}{p+1},p)}\big)^{\frac{\alpha}{2}} \\ &\quad= C(T,\alpha)\big(h^{2\frac{r}{r+1}} + \Delta t^{2\frac{p}{p+1}}\big)^{\frac{\alpha}{2}} \\ &\quad\leqslant C(T,\alpha)\big(h^{\min(\alpha\frac{r}{r+1},r)} + \Delta t^{\min(\alpha\frac{p}{p+1},p)}\big). \end{aligned}$$

□

We are now in the position to prove the following result concerning the weak error of the approximation $X^N_{h,\Delta t,1} := P^1 X^N_{h,\Delta t}$ of the first component $X_1(T) = P^1X(T)$ of the solution to the stochastic wave eqation (3.1.2) at time T.

THEOREM 3.5.3. *Let $X_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathcal{H}^{2\beta})$ for some $\beta > 0$ and $g \in C^2(\dot{H}^0, \mathbb{R})$ with $\sup_{x\in\dot{H}^0} \|g''(x)\|_{\mathcal{L}(\dot{H}^0)} < \infty$. Suppose that either of the following conditions holds.*

(i) $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$ *and*

$$\sup_{x\in\dot{H}^0} \|\Lambda^{\frac{\beta}{2}} g''(x)\Lambda^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot{H}^0)} < \infty. \tag{3.5.15}$$

(ii)

$$\lim_{m\to\infty} \int_{U_1} \|\Lambda^{-\frac{1}{2}} p_m y\|_{\dot{H}^0} \|\Lambda^{\beta-\frac{1}{2}} p_m y\|_{\dot{H}^0}\, \nu(\mathrm{d}y) < \infty, \tag{3.5.16}$$

where p_m denotes the orthogonal projection from $\dot{H}^0$ to $\operatorname{span}\{\varphi_1, \dots, \varphi_m\}$, $(\varphi_k)_{k\in\mathbb{N}}$ *being an orthonormal basis of $\dot{H}^0$ consisting of eigenvectors of Λ.*

Then, there is a unique weak solution $(X(t))_{t\geqslant 0}$ to Eq. (3.1.2) *given by* (3.1.1).

Let $(X^n_{h,\Delta t})_{n=0,\dots,N}$ be given by (3.5.2). *Then, there exists a constant $C = C(g,T) > 0$, that does not depend on h and Δt, such that for $\Delta t \leqslant 1$,*

$$\big|\mathbb{E}\big(g(X^N_{h,\Delta t,1}) - g(X_1(T))\big)\big| \leqslant C\big(h^{\min(2\beta\frac{r}{r+1},r)} + \Delta t^{\min(2\beta\frac{p}{p+1},1)}\big).$$

PROOF. First suppose that (i) holds. We apply Theorem 3.2.6 and Corollary 3.2.8 with $G = g \circ P^1$. Note that $G'(x) = (P^1)^* g'(P^1x) \in H$ and

$$G''(x) = (P^1)^* g''(P^1x)P^1 \in \mathcal{L}(H)$$

for all $x \in H$, where $(P^1)^* \in \mathcal{L}(\dot{H}^0, H)$ is the Hilbert space adjoint of $P^1 \in \mathcal{L}(H, \dot{H}^0)$. Using (3.2.6) one obtains that

$$\begin{aligned} u_x(t,\xi) &= \mathbb{E}\big((P^1)^* g'(P^1 Z(T,t,x))\big)\big|_{x=\xi}, \\ u_{xx}(t,\xi) &= \mathbb{E}\big((P^1)^* g''(P^1 Z(T,t,x))P^1\big)\big|_{x=\xi} \end{aligned} \tag{3.5.17}$$

dc_1541_18

for all H-valued random variables ξ and $t \in [0,T]$.

We combine (3.5.1), (3.5.17) and the deterministic error estimate (3.5.5) with $\alpha = 2\beta$ in order to estimate the first term on the right hand side of the error representation formula (3.2.11) in Corollary 3.2.8. We have, where the first inequality follows similarly as for the stochastic heat equation, that

$$
\begin{aligned}
&\big|\mathbb{E}\big\{u(0,\tilde{E}(T)X_0) - u(0,E(T)X_0)\big\}\big| \\
&\quad\leqslant \int_0^1 \big\|g'\big(P^1 Z\big(T,0,Y(0)+\theta(\tilde{Y}(0)-Y(0))\big)\big)\big\|_{L^2(\Omega,\dot{H}^0)}\,\mathrm{d}\theta \\
&\qquad\times \|P^1(\tilde{E}(T)-E(T))X_0\|_{L^2(\Omega,\dot{H}^0)} \\
&\quad\leqslant C\big(1+\int_0^1 \big\|Z\big(T,0,Y(0)+\theta(\tilde{Y}(0)-Y(0))\big)\big\|_{L^2(\Omega,H)}\,\mathrm{d}\theta\big) \\
&\qquad\times \|P^1(\tilde{E}(T)-E(T))\|_{\mathcal{L}(\mathcal{H}^{2\beta},\dot{H}^0)}\,\|X_0\|_{L^2(\Omega;\mathcal{H}^{2\beta})} \\
&\quad\leqslant C\big(1+\|X_0\|_{L^2(\Omega;H)} + T^{\frac12}\|\Lambda^{-1/2}Q^{1/2}\|_{\mathrm{HS}}\big)\,\|X_0\|_{L^2(\Omega;\mathcal{H}^{2\beta})} \\
&\qquad\times \big(h^{\min(2\beta\frac{r}{r+1},r)} + \Delta t^{\min(2\beta\frac{p}{p+1},1)}\big).
\end{aligned}
\tag{3.5.18}
$$

Using (3.5.17), Lemma 3.2.1 and Remark 3.2.2, the integral of the function Ψ_1 in the second term on the right hand side of the formula (3.2.11) can be estimated as follows:

$$
\begin{aligned}
&\Big|\mathbb{E}\int_0^T\int_{U_1}\int_0^1 \Psi_1(t,\theta,y)\,\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&= \Big|\mathbb{E}\int_0^T\int_{U_1}\int_0^1 (1-\theta)\Big\langle \mathbb{E}\big(g''\big(P^1Z(T,t,x+E(T-t)By+\theta F(T-t)y)\big)\big)\big|_{x=\tilde{Y}(t)} \\
&\quad\times P^1F(T-t)y,\ P^1F(T-t)y\Big\rangle_{\dot{H}^0}\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&\leqslant \sup_{x\in\dot{H}^0}\|g''(x)\|_{\mathcal{L}(\dot{H}^0)}\int_0^T \|P^1F(T-t)\|^2_{\mathcal{L}_2(U_0,\dot{H}^0)}\mathrm{d}t \\
&\leqslant \sup_{x\in\dot{H}^0}\|g''(x)\|_{\mathcal{L}(\dot{H}^0)} C\big(h^{\min(\beta\frac{r}{r+1},r)} + \Delta t^{\min(\beta\frac{p}{p+1},1)}\big)^2.
\end{aligned}
\tag{3.5.19}
$$

The last step in (3.5.19) is due to the fact that, by Itô's isometry (3.1.19), the integral in the penultimate line is the square of the strong error $\|X^N_{h,k,1} - X_1(T)\|_{L^2(\Omega;\dot{H}^0)}$ for zero initial condition $X_0 = 0$; it can be estimated as in the Gaussian case, see (1.3.32).

Concerning the integral of the function Ψ_2 in the second term on the right hand side of (3.2.11), we have by (3.5.17), Lemma 3.2.1, (1.1.4) and since $U_0 =$

dc_1541_18

$Q^{1/2}(U) = Q^{1/2}(\dot H^0)$, that

(3.5.20)

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^T \int_{U_1}\int_0^1 \Psi_2(t,\theta,y)\,\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&\quad = \Big|\mathbb{E}\int_0^T \int_{U_1}\int_0^1 \Big\langle \mathbb{E}\big(g''\big(P^1 Z(T,t,x+\theta E(T-t)By)\big)\big)\big|_{x=\tilde Y(t)} \\
&\qquad \times P^1 E(T-t)By, P^1 F(T-t)y\Big\rangle_{\dot H^0}\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&\quad = \Big|\mathbb{E}\int_0^T \int_{U_1}\int_0^1 \Big\langle \mathbb{E}\big(\Lambda^{\frac{\beta}{2}} g''\big(P^1 Z(T,t,x+\theta E(T-t)By)\big)\Lambda^{-\frac{\beta}{2}}\big)\big|_{x=\tilde Y(t)} \\
&\qquad \times \Lambda^{\frac{\beta}{2}} P^1 E(T-t)B\Lambda^{\frac{1-\beta}{2}}\Lambda^{\frac{\beta-1}{2}}y,\ \Lambda^{-\frac{\beta}{2}}P^1F(T-t)\Lambda^{\frac{1-\beta}{2}}\Lambda^{\frac{\beta-1}{2}}y\Big\rangle_{\dot H^0}\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&\quad \leqslant \sup_{x\in\dot H^0}\|\Lambda^{\frac{\beta}{2}}g''(x)\Lambda^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 \\
&\qquad \times \int_0^T \|\Lambda^{\frac{\beta}{2}}P^1E(T-t)B\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\|\Lambda^{-\frac{\beta}{2}}P^1F(T-t)\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\,\mathrm{d}t.
\end{aligned}$$

Note that

(3.5.21)

$$\|\Lambda^{\frac{\beta}{2}}P^1E(T-t)B\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot H^0)} = \|\Lambda^{\frac{\beta-1}{2}}S(T-t)\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot H^0)} = \|S(T-t)\|_{\mathcal{L}(\dot H^0)} \leqslant 1,$$

and thus it remains to estimate the integral

$$\begin{aligned}
\int_0^T \|\Lambda^{-\frac{\beta}{2}}P^1F(T-t)\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot H^0)}\mathrm{d}t &= \int_0^T \|P^1F(t)\|_{\mathcal{L}(\dot H^{\beta-1},\dot H^{-\beta})}\mathrm{d}t \\
&= \int_0^T \|P^1(\tilde E(t)-E(t))B\|_{\mathcal{L}(\dot H^{\beta-1},\dot H^{-\beta})}\mathrm{d}t.
\end{aligned}$$

To this end, it suffices to apply the deterministic error estimate (3.5.5) with $\alpha = 2\beta$. The combination of (3.5.18), (3.5.19) and (3.5.20) finishes the proof.

Next, suppose that (ii) holds. By Lemma 3.2.1 we have that

$$\begin{aligned}
\|p_m\Lambda^{(\beta-1)/2}Q^{1/2}\|^2_{\mathcal{L}_2(\dot H^0)} = \|\Lambda^{(\beta-1)/2}p_mQ^{1/2}\|^2_{\mathcal{L}_2(\dot H^0)} &= \int_{U_1}\|\Lambda^{(\beta-1)/2}p_my\|^2_{\dot H^0}\,\nu(\mathrm{d}y) \\
&= \int_{U_1}\langle\Lambda^{-1/2}p_my,\Lambda^{\beta-1/2}p_my\rangle_{\dot H^0}\,\nu(\mathrm{d}y) \\
&\leqslant \int_{U_1}\|\Lambda^{-1/2}p_my\|_{\dot H^0}\|\Lambda^{\beta-1/2}p_my\|_{\dot H^0}\,\nu(\mathrm{d}y),
\end{aligned}$$

whence,

$$\|\Lambda^{(\beta-1)/2}Q^{1/2}\|^2_{\mathrm{HS}} = \|\Lambda^{(\beta-1)/2}Q^{1/2}\|^2_{\mathcal{L}_2(\dot H^0)} = \lim_{m\to\infty}\|p_m\Lambda^{(\beta-1)/2}Q^{1/2}\|^2_{\mathcal{L}_2(\dot H^0)} < \infty,$$

proving that there is a unique weak solution $(X(t))_{t\geqslant 0}$ to Eq. (3.1.2) given by (3.1.1). To estimate the weak error, we apply an approximation procedure and consider for $m\in\mathbb{N}$ the H-valued random variables $X^{[m]}(T) := E(T)X_0 + \int_0^T E(T-s)Bp_m\,\mathrm{d}L(s)$ and $\tilde X^{[m]}(T) = \tilde X^{[m]}_{h,\Delta t}(T) := \tilde E(T)X_0 + \int_0^T \tilde E(T-s)Bp_m\,\mathrm{d}L(s)$. Using Itô's isometry and the fact that $\|(I-p_m)\Lambda^{-1/2}Q^{1/2}\|_{\mathcal{L}_2(\dot H^0)} \to 0$ as $m\to\infty$, we get that $X^{[m]}(T) \xrightarrow{m\to\infty} X(T)$ and $\tilde X^{[m]}(T) \xrightarrow{m\to\infty} \tilde X(T)$ in $L^2(\Omega;H)$. As a consequence of this and (3.2.5), we obtain

$$e^{[m]}(T) := \mathbb{E}\big(G(\tilde{X}^{[m]}(T)) - G(X^{[m]}(T))\big) \xrightarrow{m\to\infty} e(T).$$

Thus, it suffices to show the desired rate for the error $e^{[m]}(T)$ with a constant that does not depend on m. To this end, we observe that the estimate (3.5.18) can be used without any changes, and that the analogue to the estimate (3.5.19) gives indeed the desired rate if we use that $\|p_m\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} \leqslant \|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$. Finally, the estimate corresponding to (3.5.20) reads as

$$\begin{aligned}
&\Big|\mathbb{E}\int_0^T\int_{U_1}\int_0^1 \Big\langle \mathbb{E}\big(g''\big(P^1Z^{[m]}(T,t,x+\theta E(T-t)Bp_my)\big)\big)\big|_{x=\tilde{Y}^{[m]}(t)} \\
&\qquad\times P^1E(T-t)Bp_my, P^1F(T-t)p_my\Big\rangle_{\dot{H}^0}\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&\quad= \Big|\mathbb{E}\int_0^T\int_{U_1}\int_0^1 \Big\langle \mathbb{E}\big(g''\big(P^1Z^{[m]}(T,t,x+\theta E(T-t)Bp_my)\big)\big)\big|_{x=\tilde{Y}^{[m]}(t)} \\
&\qquad\times P^1E(T-t)B\Lambda^{1/2}\Lambda^{-1/2}p_my, P^1F(T-t)\Lambda^{1/2-\beta}\Lambda^{\beta-1/2}p_my\Big\rangle_{\dot{H}^0}\mathrm{d}\theta\,\nu(\mathrm{d}y)\,\mathrm{d}t\Big| \\
&\quad\leqslant \sup_{x\in\dot{H}^0}\|g''(x)\|_{\mathcal{L}(\dot{H}^0)}\int_0^T \|P^1E(T-t)B\Lambda^{1/2}\|_{\mathcal{L}(\dot{H}^0)}\|P^1F(T-t)\Lambda^{1/2-\beta}\|_{\mathcal{L}(\dot{H}^0)}\,\mathrm{d}t \\
&\qquad\times\int_{U_1}\|\Lambda^{-1/2}p_my\|_{\dot{H}^0}\|\Lambda^{\beta-1/2}p_my\|_{\dot{H}^0}\,\nu(\mathrm{d}y),
\end{aligned}$$

where $Z^{[m]}$ and $\tilde{Y}^{[m]}$ are defined by replacing B by Bp_m in the definitions of Z and $\tilde{Y}$. By (3.5.5) with $\alpha = 2\beta$ and the fact that $\|B\|_{\mathcal{L}(\dot{H}^{\alpha-1},\mathcal{H}^\alpha)} = 1$ we have that

$$\begin{aligned}
\sup_{t\in[0,T]}\|P^1F(T-t)\Lambda^{1/2-\beta}\|_{\mathcal{L}(\dot{H}^0)} &= \sup_{t\in[0,T]}\|P^1(\tilde{E}(t)-E(t))B\Lambda^{1/2-\beta}\|_{\mathcal{L}(\dot{H}^0)} \\
&= \sup_{t\in[0,T]}\|P^1(\tilde{E}(t)-E(t))B\|_{\mathcal{L}(\dot{H}^{2\beta-1},\dot{H}^0)} \leqslant C\big(h^{\min(2\beta\frac{r}{r+1},r)} + \Delta t^{\min(2\beta\frac{p}{p+1},1)}\big).
\end{aligned}$$

Finally, by (3.5.16) and (3.5.21) with $\beta = 0$, the proof is complete. □

Finally, we state two remarks and discuss some examples where the conditions of Theorem 3.5.3, in particular (3.5.15) and (3.5.16), are satisfied.

REMARK 3.5.4. It follows from Lemma 3.2.1 and the fact that $\Lambda^\alpha p_m \in \mathcal{L}_2(\dot{H}^0)$ for all $\alpha\in\mathbb{R}$ and $m\in\mathbb{N}$, that the terms $\Lambda^{-1/2}p_my$ and $\Lambda^{\beta-1/2}p_my$ in (3.5.16) are defined in an $L^2(U_1,\nu(\mathrm{d}y);\dot{H}^0)$-sense. Note that the sequence

$$\big(\|\Lambda^{-1/2}p_my\|_{\dot{H}^0}\|\Lambda^{\beta-1/2}p_my\|_{\dot{H}^0}\big)_{m\in\mathbb{N}}$$

is monotonically increasing for ν-almost all $y\in U_1$, so that the limit in (3.5.16) is in fact a supremum. Moreover, if we explicitly choose $U_1 = \dot{H}^{\beta-1}$ as the state space of L, then the condition (ii) is equivalent to assuming that $\operatorname{supp}\nu\subset\dot{H}^{2\beta-1}$ and

$$\int_{\dot{H}^{2\beta-1}}\|\Lambda^{-\frac{1}{2}}y\|_{\dot{H}^0}\|\Lambda^{\beta-\frac{1}{2}}y\|_{\dot{H}^0}\,\nu(\mathrm{d}y) < \infty.$$

This choice of U_1 is possible w.l.o.g. whenever $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$, since then the natural embedding of $U_0 = Q^{1/2}U = Q^{1/2}\dot{H}^0$ into $\dot{H}^{\beta-1}$ is Hilbert-Schmidt and we can re-expand L in the form (3.1.12) as an $\dot{H}^{\beta-1}$-valued martingale, compare Remark 3.1.3. However, in the spirit of, e.g., [**4, 94, 95**], we prefer a formulation of our results that is independent of the specific choice of the state space U_1.

REMARK 3.5.5. Instead of the symmetric condition $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$, the sufficient asymmetric condition $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ is imposed in Theorem 1.3.13 in the Wiener case in order to double the rate of strong convergence for the wave equation. The asymmetric condition (3.5.16) appearing in (ii) above, which is again sufficient for $\|\Lambda^{(\beta-1)/2}Q^{1/2}\|_{\mathrm{HS}} < \infty$, resembles the same situation in the present case.

EXAMPLE 3.5.6. An important example which satisfies (3.5.15) is $g(x) = \|x\|^2_{\dot H^0}$.

EXAMPLE 3.5.7. As another example for a function g satisfying (3.5.15) consider $g(x) := f(\langle\varphi_1, x\rangle_{\dot H^0}, \dots, \langle\varphi_n, x\rangle_{\dot H^0})$, $x \in \dot H^0$, where $f \in C^2(\mathbb{R}^n, \mathbb{R})$ has bounded second order derivatives and $(\varphi_k)_{k\in\mathbb{N}} \subset D(\Lambda)$ is an orthonormal basis of $\dot H^0 = L^2(\mathcal{D})$ consisting of eigenfunctions of Λ with corresponding eigenvalues $(\lambda_k)_{k\in\mathbb{N}} \subset (0,\infty)$. Then, for $x, y \in \dot H^0$,

$$\Lambda^{\beta/2} g''(x)\Lambda^{-\beta/2} y = \sum_{j,k=1}^{n} \lambda_j^{-\beta/2}\lambda_k^{\beta/2}(\partial_j\partial_k f)\big(\langle\varphi_1, x\rangle_{\dot H^0}, \dots, \langle\varphi_n, x\rangle_{\dot H^0}\big)\langle\varphi_j, y\rangle_{\dot H^0}\varphi_k$$

and (3.5.15) holds. More generally, (3.5.15) is satisfied by all $g \in C^2(\dot H^0, \mathbb{R})$ of the form $g = \tilde g \circ \Lambda^{-\beta/2}$ with $\tilde g \in C^2(\dot H^0, \mathbb{R})$ satisfying $\sup_{x\in\dot H^0} \|\tilde g''(x)\|_{\mathcal{L}(\dot H^0)} < \infty$. For such g we have $g''(x) = \Lambda^{-\beta/2}\tilde g''(\Lambda^{-\beta/2}x)\Lambda^{-\beta/2}$.

EXAMPLE 3.5.8. Consider the situation of Example 3.1.5; that is when

$$\nu = \sum_{k\in\mathbb{N}} \nu_k \circ \pi_k^{-1},$$

where ν_k is the Lévy measure of L_k and $\pi_k : \mathbb{R} \to U_1$ is defined by $\pi_k(\xi) := \xi e_k$. Let us assume that $e_k = \sqrt{q_k}\,\varphi_k$, where $(q_k)_{k\in\mathbb{N}}$ is a bounded sequence of positive numbers and $(\varphi_k)_{k\in\mathbb{N}}$ is an orthonormal basis of $\dot H^0$ consisting of eigenfunctions of Λ with corresponding eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$. Then, with p_m as in (ii) in Theorem 3.5.3, we have that

$$\begin{aligned}
&\lim_{m\to\infty} \int_{U_1} \|\Lambda^{-\frac12} p_m y\|_{\dot H^0} \|\Lambda^{\beta-\frac12} p_m y\|_{\dot H^0}\,\nu(\mathrm{d}y)\\
&= \sum_{k\in\mathbb{N}} \int_{\mathbb{R}} \xi^2 q_k \|\Lambda^{-\frac12}\varphi_k\|_{\dot H^0} \|\Lambda^{\beta-\frac12}\varphi_k\|_{\dot H^0}\,\nu_k(\mathrm{d}\xi)\\
&= \sum_{k\in\mathbb{N}} \int_{\mathbb{R}} \xi^2 q_k \lambda_k^{-\frac12} \|\varphi_k\|_{\dot H^0} \lambda_k^{\beta-\frac12} \|\varphi_k\|_{\dot H^0}\,\nu_k(\mathrm{d}\xi)\\
&= \sum_{k\in\mathbb{N}} \int_{\mathbb{R}} \xi^2 q_k \|\Lambda^{\frac{\beta-1}{2}}\varphi_k\|_{\dot H^0} \|\Lambda^{\frac{\beta-1}{2}}\varphi_k\|_{\dot H^0}\,\nu_k(\mathrm{d}\xi)\\
&= \lim_{m\to\infty} \int_{U_1} \|\Lambda^{\frac{\beta-1}{2}} p_m y\|_{\dot H^0} \|\Lambda^{\frac{\beta-1}{2}} p_m y\|_{\dot H^0}\,\nu(\mathrm{d}y) = \|\Lambda^{\frac{\beta-1}{2}} Q^{\frac12}\|^2_{\mathrm{HS}},
\end{aligned}$$

where we used Lemma 3.2.1 in the last step. That is, when ν is concentrated on the eigenspaces $\{r\varphi_k : r \in \mathbb{R}\}$, $k \in \mathbb{N}$, of Λ, then the abstract asymmetric condition (3.5.16) coincides with the familiar symmetric Hilbert-Schmidt condition. The situation is similar in the Wiener case when Λ and Q commute, c.f. (1.1.10).

Summary

In this dissertation we develop techniques to study the weak rate of convergence of finite element approximations of certain class of linear and semilinear SPDEs driven by additive Wiener noise or Lévy noise of pure jump type. The main challenge is to prove optimal rates of weak convergence that corresponds to twice the rate of strong convergence, where the latter usually coincides with the root-mean-squared regularity of the solution of the SPDE. Below we summarise the main results of the dissertation, all of them being own results of the author.

The main weak convergence results of Chapter 1 are based on a general abstract error representation formula, stated in Theorem 1.2.1, which allows one to treat several different kind of linear stochastic equations driven by additive Wiener noise, parabolic, hyperbolic and Volterra type, from a common structural point of view. Important results from this chapter are Theorem 1.3.13 and Theorem 1.3.14 which consider the weak, respectively, the strong rate of convergence of a finite element approximation of the linear stochastic wave equation with additive noise. Furthermore, we establish the strong and weak rate of convergence, in Theorem 1.5.28, respectively, in Theorem 1.5.33, for a finite element approximation scheme for a class of stochastic Volterra integro-differential equations, with convolution kernels typical is the theory of linear viscoelasticity.

In Chapter 2 we consider mild solutions of semilinear parabolic type stochastic equations driven by additive Wiener noise, in particular, mild solutions of the semilinear heat equation and a class of semilinear Volterra integro-differential equations. We treat these equations in a common framework as their mild solutions satisfy the same integral equation with the only difference being the deterministic evolution operator appearing in the equation with various degree of smoothing property. We use tools from Malliavin calculus and a duality argument to obtain the main weak convergence result of the chapter, Theorem 2.3.7, which we apply to a finite element discretization of a semilinear heat equation, in Subsection 2.4.1, and then to a class of semilinear stochastic Volterra integro-differential equations in Subsection 2.4.2. We would like to point out that we allow the test functions to depend on the paths of the solution in a special way and therefore we obtain convergence rates of approximations of covariances and higher order statistics of the solution in Corollary 2.3.8. The rate of strong convergence for these equations, which is half that of the weak convergence rate, is stated in Theorem 2.3.2.

In Chapter 3 we consider linear equations driven by additive square integrable Lévy noise of pure jump type. As in Chapter 1, but requiring significantly more effort, we develop a general abstract error representation formula, stated in Theorem

3.2.6, which we then use to study the rate of weak convergence of finite element approximations of the linear stochastic heat equation, in Theorem 3.3.4, a class of of stochastic Volterra integro-differential equations, in Theorem 3.4.1, and the stochastic wave equation in Theorem 3.5.3. The other main contributions in this chapter, which are applicable, but not directly related to the numerical analysis of SPDEs, are Lemma 3.2.4, where we establish the equivalence of two different integration theories for infinite dimensional Lévy processes and Proposition 3.2.10, where we develop a backward Kolmogorov equation associated with an infinite dimensional drift-free Lévy-Itô integral process.

There are some new unpublished results in the dissertation in Sections 1.5.4 and 1.5.6. In [**69**] the strong convergence results result on semidiscretization in time ([**69**, Theorem 4.6], Theorem 1.5.18 here) and also the full discretization ([**69**, Theorem 5.1], Theorem 1.5.28 here) is stated and proved assuming that the convolution kernel satisfies Assumptions 1.5.1 and 1.5.30. Note that the latter assumption implies that the kernel can be extended to an analytic function in a sectorial region around the positive half-axis. Instead, here, we assume that kernel satisfies Assumption 1.5.2 which is just a little stronger than Assumption 1.5.1 but it allows for non-analytic kernels and hence avoids the heavy restriction posed by Assumption 1.5.30. One price to pay is that one has to assume smooth initial data in Theorem 1.5.28 compared to [**69**, Theorem 5.1] as there are no non-smooth data estimates available for the finite element method for such general kernels. The key deterministic result is stated in Theorem 1.5.13 which in turn can be used to prove estimates (1.5.48) and (1.5.49). The latter estimates correspond to the result in [**69**, Proposition 4.4] (which is [**78**, Theorem 3.2]) and [**69**, Corollary 4.5], respectively. The smoothing property (1.5.42) is also new which is then used, together with Theorem 1.5.13, to prove the smooth data estimate (1.5.47) and the error estimate in Lemma 1.5.27.

Finally, we mention that, in order to make the dissertation concise and focused, results from the author's other main research area, fractional order PDEs and related topics, contained in [**7, 8, 9, 10, 11, 12, 13, 14, 15, 58**] and further results concerning strong convergence of numerical approximations of SPDEs from [**42, 59, 60, 63, 64, 65**] are not presented in the dissertation.

Bibliography

[1] A. Andersson, M. Kovács, and S. Larsson, Weak error analysis for semilinear stochastic Volterra equations with additive noise, *J. Math. Anal. Appl.* **437**(3) (2016) 1283–1304.

[2] A. Andersson, R. Kruse, and S. Larsson, Duality in refined Sobolev-Malliavin spaces and weak approximation of SPDE, *Stochastic Partial Differential Equations: Analysis and Computations* **4** (2016) 113–149.

[3] A. Andersson and S. Larsson, Weak convergence for a spatial approximation of the nonlinear stochastic heat equation, *Math. Comp.* **85** (2016) 1335–1358.

[4] D. Applebaum and M. Riedle, Cylindrical Lévy processes in Banach spaces, *Proc. London Math. Soc.* **101**(3) (2010) 697–726.

[5] W. Arendt, C.J.K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, Birkhäuser, Basel, 2011.

[6] B. Baeumer, M. Geissert, and M. Kovács, Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise, *J. Diff. Eq.* **258**(2) (2015) 535–554.

[7] B. Baeumer, M. Haase, and M. Kovács, Unbounded functional calculus for bounded groups with applications, *J. Evol. Equ.* **9**(1) (2009) 171–195.

[8] B. Baeumer and M. Kovács, Approximating multivariate tempered stable processes, *J. Appl. Probab.* **49**(1) (2012) 167–183.

[9] B. Baeumer, M. Kovács, and M. M. Meerschaert, Fractional reproduction-dispersal equations and heavy tail dispersal kernels, *Bull. Math. Biol.* **69**(7) (2007) 2281–2297.

[10] B. Baeumer, M. Kovács, and M. M. Meerschaert, Numerical solutions for fractional reaction-diffusion equations, *Comput. Math. Appl.* **55**(10) (2008) 2212–2226.

[11] B. Baeumer, M. Kovács, and M. M. Meerschaert, Subordinated multiparameter groups of linear operators: properties via the transference principle, *Functional analysis and evolution equations*, 35–50, Birkhäuser, Basel, 2008.

[12] B. Baeumer, M. Kovács, M. M. Meerschaert, and H. Sankaranarayanan, Boundary conditions for fractional diffusion, *J. Comput. and Appl. Math.* **336** (2018) 408–424.

[13] B. Baeumer, M. Kovács, M. M. Meerschaert, R. L. Schilling, and P. Straka, Reflected spectrally negative stable processes and their governing equations, *Trans. Amer. Math. Soc.* **368** (1) (2016) 227–248.

[14] B. Baeumer, M. Kovács, and H. Sankaranarayanan, Higher order Grünwald approximations of fractional derivatives and fractional powers of operators, *Trans. Amer. Math. Soc.* **367**(2) (2015), 813–834.

[15] B. Baeumer, M. Kovács, and H. Sankaranarayanan, Fractional partial differential equations with boundary conditions, *J. Differential Equations* **264**(2) (2018) 1377–1410.

[16] G.A. Baker and J.H. Bramble, Semidiscrete and single step fully discrete approximations for second order hyperbolic equations, *RAIRO Numer. Anal.* **13** (1979) 76–100.

[17] A. Barth, A finite element method for martingale-driven stochastic partial differential equations, *Commun. Stoch. Anal.* **4**(3) (2010) 355–375.

[18] A. Barth and A. Lang, Simulation of stochastic partial differential equations using finite element methods, *Stochastics* **84**(2-3) (2012) 217–231.

[19] A. Barth and A. Lang, Milstein approximation for advection-diffusion equations driven by multiplicative noncontinuous martingale noises, *Appl. Math. Optim.* **66** (2012) 387–413.

[20] A. Barth and T. Stüwe, Weak convergence of Galerkin approximations of stochastic partial differential equations driven by additive Lévy noise, *Math. Comput. Simulation* **143** (2018) 215–225.

dc_1541_18

[21] V. Bentkus, Asymptotic expansions and convergence rates for Euler's approximations of semigroups, *Lith. Math. J.* **49**(2) (2009) 140–157.

[22] C. E. Bréhier, M. Hairer, and A. Stewart, Weak error estimates for trajectories of SPDEs under Spectral Galerkin discretization, J. Comput. Math. **36**(2) (2018) 159–182.

[23] P. Brenner and V. Thomeé, On rational approximations of groups of operators, *SIAM J. Numer. Anal.* **17**(1) (1980) 119–125.

[24] Z. Brzeźniak, Some remarks on Itô and Stratonovich integration in 2-smooth Banach spaces, Probabilistic methods in fluids, World Sci. Publ., River Edge, NJ, 2003, pp. 48–69.

[25] Z. Brzeźniak, J.M.A.M. van Neerven, M.C. Veraar, and L. Weis, Itô's formula in UMD Banach spaces and regularity of solutions of the Zakai equation, *J. Differential Equations* **245** (2008) 30–58.

[26] Z. Brzeźniak and J. Zabczyk, Regularity of Ornstein-Uhlenbeck processes driven by a Lévy white noise, *Potential Anal.* **32** (2010) 153–188.

[27] U. J. Choi and R. C. Maccamy, Fractional order Volterra equations, Volterra integrodifferential equations in Banach spaces and applications (Trento, 1987), Pitman Res. Notes Math. Ser., 190, Longman Sci. Tech., Harlow, 1989, 231–245.

[28] P. Clément, G. Da Prato, and J. Prüss, White noise perturbation of the equations of linear parabolic viscoelasticity, Rend. Istit. Mat. Univ. Trieste **XXIX** (1997) 207–220.

[29] D. Conus, A. Jentzen, and R. Kurniawan, Weak convergence rates of spectral Galerkin approximations for SPDEs with nonlinear diffusion coefficients, *Preprint*, 2014, arXiv:1408.1108v1.

[30] G. Da Prato, A. Jentzen, and M. Röckner, A mild Itô formula for SPDEs, *Preprint*, 2012, arXiv:1009.3526v4.

[31] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and Its Applications 44, Cambridge University Press, Cambridge, 1992.

[32] A. de Bouard and A. Debussche, Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation, *Appl. Math. Optim.* **54**(3) (2006) 369–399.

[33] A. Debussche, Weak approximation of stochastic partial differential equations: the nonlinear case, *Math. Comp.* **80** (2011) 89–117.

[34] A. Debussche and J. Printems, Weak order for the discretization of the stochastic heat equation, *Math. Comp.* **78** (2009) 845–863.

[35] L. J. de Naurois, A. Jentzen, and T. Welti, Weak convergence rates for spatial spectral Galerkin approximations of semilinear stochastic wave equations with multiplicative noise, *Preprint*, 2015, arXiv:1508.05168v1.

[36] J. Dixon and S. McKee, Weakly singular discrete Gronwall inequalities, *Z. angew. Math. Mech.* **66** (1986) 535–544.

[37] T. Dunst, E. Hausenblas, and A. Prohl, Approximate Euler method for parabolic stochastic partial differential equations driven by space-time Lévy noise, *SIAM J. Numer. Anal.* **50**(6) (2012) 2873–2896.

[38] P. P. B. Eggermont, On the quadrature error in operational quadrature methods for convolutions, *Numer. Math.* **62** (1992) 35–48.

[39] C. M. Elliott and S. Larsson, Error estimates with smooth and nonsmooth data for the finite element method for the Cahn-Hilliard equation, *Math. Comp.* **58** (1992) 603–630.

[40] S. J. Fromm, Potential space estimates for Green potentials in convex domains, *Proc. Amer. Math. Soc.* **119**(1) (1993) 225–233.

[41] M. Fuhrman and G. Tessitore, Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control, *Ann. Probab.* **30** (2002) 1397–1465.

[42] D. Furihata, M. Kovács, S. Larsson, and F. Lindgren, Strong convergence of a fully discrete finite element approximation of the stochastic Cahn–Hilliard equation, *SIAM. J. Numer. Anal.* **56**(2) (2018) 708–731.

[43] M. Geissert, M. Kovács, and S. Larsson, Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise, *BIT* **49**(2) (2009) 343–356.

[44] J.B. Gravereaux and J. Pellaumail, Formule de Ito pour des processus non continus à valeurs dans des espaces de Banach, *Ann. Inst. Henri Poincaré Probab. Statist.* **10**(4) (1974) 399–422.

[45] I. Gyöngy and N. Krylov, On the rate of convergence of splitting-up approximations for SPDEs, Stochastic inequalities and applications, Progr. Probab., vol. 56, Birkhäuser, Basel, 2003, 301–321.

[46] I. Gyöngy and N. Krylov, On the splitting-up method and stochastic partial differential equations, *Ann. Probab.* **31** (2003) 564–591.

[47] I. Gyöngy and A. Millet, On discretization schemes for stochastic evolution equations, *Potential Anal.* **23** (2005) 99–134.

[48] I. Gyöngy and A. Millet, Rate of convergence of implicit approximations for stochastic evolution equations, in: Stochastic Differential Equations: Theory and Applications, Interdiscip. Math. Sci., Vol. 2 (World Sci. Publ., Hackensack, NJ, 2007), 281–310.

[49] E. Hausenblas, Weak approximation for semilinear stochastic evolution equations, Stochastic analysis and related topics VIII, 111–128, Progr. Probab., 53, Birkhäuser, Basel, 2003.

[50] E. Hausenblas, Finite element approximation of stochastic partial differential equations driven by Poisson random measures of jump type, *SIAM J. Numer. Anal.* **46** (2008) 437–471.

[51] E. Hausenblas and I. Marchis, A numerical approximation of parabolic stochastic partial differential equations driven by a Poisson random measure, *BIT* **46** (2006) 773–811.

[52] E. Hausenblas, Weak approximation of the stochastic wave equation, *J. Comput. Appl. Math.* **235**(1) (2010) 33–58.

[53] E. Hille and J. Tamarkin, On the absolute integrability of Fourier transforms, *Fundamenta Mathematicae* **25**(1) (1935) 329–352.

[54] J. Jacod, T.G. Kurtz, S. Méléard, and P. Protter, The approximate Euler method for Lévy driven stochastic differential equations, *Ann. Inst. Henri Poincaré Probab. Stat.* **41**(3) (2005) 523–558.

[55] A. Jentzen and P. E. Kloeden, The numerical approximation of stochastic partial differential equations, *Milan J. Math.* **77** (2009) 205–244.

[56] A. Jentzen and M. Röckner, Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise, *J. Diff. Eq.* **252**(1) (2012) 114–136.

[57] A. Karczewska and P. Rozmej, On Numerical Solutions to stochastic Volterra equations, arXiv:math/0409026.

[58] M. Kovács and M. M. Meerschaert, Ultrafast subordinators and their hitting times, *Publ. Inst. Math.* (Beograd) (N.S.) **80**(94) (2006) 193–206.

[59] M. Kovács, S. Larsson, and F. Lindgren, Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise, *Numer. Algorithms* **53**(2-3) (2010) 309–320.

[60] M. Kovács, S. Larsson, and F. Lindgren, Spatial approximation of stochastic convolutions, *J. Comput. Appl. Math.* **235**(12) (2011), 3554–3570.

[61] M. Kovács, S. Larsson, and F. Lindgren, Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise, *BIT* **52** (2012) 85–108.

[62] M. Kovács, S. Larsson, and F. Lindgren, Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II: Fully discrete schemes, *BIT* **53** (2013) 497–525.

[63] M. Kovács, S. Larsson, F. Lindgren, On the backward Euler approximation of the stochastic Allen–Cahn equation, *J. Appl. Probab.* **52**(2) (2015) 323–338.

[64] M. Kovács, S. Larsson, and F. Lindgren, On the discretisation in time of the stochastic Allen–Cahn equation, *Mathematische Nachrichten* **291**(5-6) (2018) 966–995.

[65] M. Kovács, S. Larsson, and A. Mesforush, Finite element approximation of the Cahn–Hilliard–Cook equation, *SIAM J. Numer. Anal.* **49**(6) (2011) 2407–2429.

[66] M. Kovács, S. Larsson, and F. Saedpanah, Finite element approximation of the linear stochastic wave equation with additive noise, *SIAM J. Numer. Anal.* **48** (2010) 408–427.

[67] M. Kovács, S. Larsson, and K. Urban, On wavelet-Galerkin methods for semilinear parabolic equations with additive noise, Monte Carlo and quasi-Monte Carlo methods 2012, 481–499, Springer Proc. Math. Stat., 65, Springer, Heidelberg, 2013.

[68] M. Kovács, F. Lindner, and R. L. Schilling, Weak convergence of finite element approximations of linear stochastic evolution equations with additive Lévy noise, *SIAM/ASA J. Uncertain. Quantif.* **3**(1) (2015) 1159–1199.

[69] M. Kovács and J. Printems, Strong order of convergence of a fully discrete approximation of a linear stochastic Volterra type evolution equation, *Math. Comp.* **83**(298) (2014) 2325– 2346.

[70] M. Kovács and J. Printems, Weak convergence of a fully discrete approximation of a linear stochastic evolution equation with a positive-type memory term, *J. Math. Anal. Appl.* **413** (2014) 939–952.

[71] A. Lang, Mean square convergence of a semidiscrete scheme for SPDEs of Zakai type driven by square integrable martingales, *Procedia Computer Science* **1** (2012) 1615–1623.

[72] S. Larsson and A. Mesforush, Finite element approximation of the linear stochastic Cahn-Hilliard equation, *IMA J. Numer. Anal.* **31** (2011) 1315–1333.

[73] P. D. Lax, *Functional Analysis*, Wiley-Interscience, New York, 2002.

[74] M. N. Le Roux, Semidiscretization in time for parabolic problems, *Math. Comp* **33** (1979) 919–931.

[75] F. Lindner and R.L. Schilling, Weak order for the discretization of the stochastic heat equation driven by impulsive noise, *Potential Anal.* **38**(2) (2013) 345–379.

[76] C. Lubich, Convolution quadrature and discretized operational calculus. I., *Numer. Math.* **52** (1988) 129–145.

[77] C. Lubich, Convolution quadrature and discretized operational calculus. II., *Numer. Math.* **52** (1988) 413–425.

[78] C. Lubich, I. Sloan, and V. Thomée, Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term, *Math. Comp.* **65** (1996) 1–17.

[79] C. Lubich, Convolution quadrature revisited, *B.I.T.* **44**(3) (2004) 503–514.

[80] V. Mandrekar, B. Rüdiger, and S. Tappe, Itô's formula for Banach space-valued jump processes driven by Poisson random measures, In: C. Dalang, M. Dozzi, F. Russo (eds.): *Seminar on stochastic analysis, random fields and applications VII. Centro Stefano Franscini, Ascona, May 2011.* Progress in Probability **67**. Birkhäuser, Basel 2013, 171–186.

[81] W. McLean and V. Thomée, Numerical solution of an evolution equation with a positive-type memory term, *J. Austral. Math. Soc. Ser. B* **35** (1993) 23–70.

[82] R. Mikulevičius and C. Zhang, On the rate of convergence of weak Euler approximation for nondegenerate SDEs driven by Lévy processes, *Stochastic Process. Appl.* **121** (2011) 1720–1748.

[83] M. Métivier, *Semimartingales. A course on stochastic processes*, de Gruyter Studies in Mathematics 2, de Gruyter, Berlin, 1982.

[84] M. Métivier and J. Pellaumail, *Stochastic integration*, Probability and Mathematical Statistics, Academic Press, New York, 1980.

[85] S. Monniaux and J. Prüss, A theorem of the Dore-Venni type for noncommuting operators, *Trans. Amer. Math. Soc.* **349** (1997) 4787–4814.

[86] S.P. Norsett and G. Wanner, The real pole sandwich for rational approximations and oscillation equations, *BIT* **19** (1979) 79–94.

[87] A. Pazy, *Semigroups of linear operators and applications to partial differential equations,* Springer, New York, 1983.

[88] S. Peszat and J. Zabczyk, *Stochastic partial differential equations with Lévy noise. An evolution equation approach,* Encyclopedia of Mathematics and Its Applications 113, Cambridge University Press, Cambridge, 2007.

[89] E. Platen and N. Bruti-Liberati, *Numerical solutions of stochastic differential equations with jumps in finance,* Springer, Berlin, 2010.

[90] C. Prévôt, Existence, uniqueness and regularity w.r.t. the initial condition of mild solutions of SPDEs driven by Poisson noise, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **13**(1) (2010) 133–163.

[91] C. Prévôt and M. Röckner, *A concise course on stochastic partial differential equations*, Springer, Lecture Notes in Mathematics 1905, Berlin, 2007.

[92] P. Protter and D. Talay, The Euler scheme for Lévy driven stochastic differential equations, *Ann. Probab.* **25**(1) (1997) 393–423.

[93] J. Prüss, *Evolutionary Integral Equations and Applications,* Birkhäuser, Basel, 1993.

[94] M. Riedle, Stochastic integration with respect to cylindrical Lévy processes in Hilbert spaces: an L^2 approach, *Infin. Dimens. Anal. Quantum. Probab. Relat. Top.* **17**(1) (2014) 1450008.

[95] M. Riedle, Ornstein-Uhlenbeck processes driven by cylindrical Lévy processes, *Potential Anal.* **42**(4) (2015) 809–838.

[96] B. Rüdiger and S. Tappe, Isomorphisms for spaces of predictable processes and an extension of the Itô integral, *Stoch. Anal. Appl.* **30**(3) (2012) 529–537.

[97] L. Quer-Sardanyons and M. Sanz-Solé, Space semi-discretisations for a stochastic wave equation, *Potential Anal.* **24** (2006) 303–332.

[98] K. Sato, *Lévy processes and infinitely divisible distributions*, Cambridge studies in advanced mathematics 68, Cambridge University Press, Cambridge, 2013.

[99] R.L. Schilling, R. Song, and Z. Vondraček, *Bernstein functions. Theory and applications,* de Gruyter Studies in Mathematics, de Gruyter, Berlin, 2010.

[100] T. Shardlow, Weak convergence of a numerical method for a stochastic heat equation, *BIT Numerical Mathematics* **43** (2003) 179–193.

[101] S. Sperlich, On parabolic Volterra equations disturbed by fractional Brownian motions, *Stoch. Anal. Appl.* **27**(1) (2009) 74–94.

[102] D. Talay, Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE and applications, Filtering and control of random processes (Paris, 1983), 294–313, Lect. Notes Control Inf. Sci., 61, Springer, Berlin, 1984.

[103] V. Thomée, *Galerkin finite element methods for parabolic problems,* (2nd edn), Springer Series in Computational Mathematics 25, Springer, Berlin, 2006.

[104] H. Triebel, *Interpolation theory, function spaces, differential operators,* (2nd edn), Johann Ambrosius Barth, Heidelberg, 1995.

[105] J. B. Walsh, On numerical solutions of the stochastic wave equation, *Illinois J. Math.* **50** (2006) 991–1018.

[106] X. Wang, Weak error estimates of the exponential Euler scheme for semi-linear SPDEs without Malliavin calculus, *Discrete Contin. Dyn. Syst.* **36** (2016) 481–497.

[107] X. Wang, An exponential integrator scheme for time discretization of nonlinear stochastic wave equation, *J. Sci. Comput.* **64**(1) (2015) 234–263.

[108] X. Wang and S. Gan, Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise, *J. Math. Anal. Appl.* **398**(1) (2013)151–169.

[109] J. Weidmann, *Linear operators in Hilbert spaces,* Graduate texts in mathematics 68, Springer, New York, 1980.

[110] Y. Yan, Galerkin finite element methods for stochastic parabolic partial differential equations, *SIAM J. Numer. Anal.* **43**(4) (2005) 1363–1384.