WEAK CONVERGENCE OF FINITE ELEMENT APPROXIMATIONS OF LINEAR STOCHASTIC EVOLUTION EQUATIONS WITH ADDITIVE NOISE II. FULLY DISCRETE SCHEMES.

MIHÁLY KOVÁCS, STIG LARSSON[1,2], AND FREDRIK LINDGREN[1]

Abstract. We present an abstract framework for analyzing the weak error of fully discrete approximation schemes for linear evolution equations driven by additive Gaussian noise. First, an abstract representation formula is derived for sufficiently smooth test functions. The formula is then applied to the wave equation, where the spatial approximation is done via the standard continuous finite element method and the time discretization via an I-stable rational approximation to the exponential function. It is found that the rate of weak convergence is twice that of strong convergence. Furthermore, in contrast to the parabolic case, higher order schemes in time, such as the Crank-Nicolson scheme, are worthwhile to use if the solution is not very regular. Finally we apply the theory to parabolic equations and detail a weak error estimate for the linearized Cahn-Hilliard-Cook equation as well as comment on the stochastic heat equation.

1. Introduction

Let $\mathcal{U}, \mathcal{H}$ be real separable Hilbert spaces and consider the following abstract stochastic Cauchy problem

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = B\,\mathrm{d}W(t),\ t > 0; \quad X(0) = X_0, \tag{1.1}$$

where $-A$ is the generator of a strongly continuous semigroup $\{E(t)\}_{t\geq 0}$ on $\mathcal{H}$, $B \in \mathcal{B}(\mathcal{U}, \mathcal{H})$, where $\mathcal{B}(\mathcal{U}, \mathcal{H})$ denotes the space of bounded linear operators from $\mathcal{U}$ to $\mathcal{H}$, $\{W(t)\}_{t\geq 0}$ is a $\mathcal{U}$-valued Wiener process with covariance operator $Q \geq 0$ (self-adjoint, positive semidefinite) with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We take X_0 to be an $\mathcal{F}_0$-measurable $\mathcal{H}$-valued random variable with finite mean. If

$$\mathrm{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) < \infty, \tag{1.2}$$

then the unique weak solution is given by (see [5, Chapter 5])

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}W(s). \tag{1.3}$$

2000 *Mathematics Subject Classification.* 65M60, 60H15, 60H35, 65C30.

Key words and phrases. finite element, parabolic equation, hyperbolic equation, stochastic, heat equation, Cahn-Hilliard-Cook equation, wave equation, additive noise, Wiener process, error estimate, weak convergence.

[1]Supported by the Swedish Research Council (VR).

[2]Supported by the Swedish Foundation for Strategic Research (SSF) through GMMC, the Gothenburg Mathematical Modelling Centre.

With $G\colon \mathcal{H} \to \mathbb{R}$ being a twice Fréchet differentiable function with bounded and continuous first and second derivatives, we study the weak error

$$e(T) = \mathbf{E}\big(G(\tilde{X}(T)) - G(X(T))\big), \tag{1.4}$$

where $\tilde{X}(T)$ is some approximation of the process X at time T.

This paper is a sequel to [14]. It consists of three parts. In Section 2 we define some central concepts and state important background results used throughout the paper. In Section 3 we show that the error formula in [14, Theorem 3.1] holds for a much wider class of approximations of the solution to (1.1) than stated in that paper, which is concerned with spatial semidiscretization by finite elements. Finally, in Sections 4 and 5, the usefulness of the general error formula in Theorem 3.1 is demonstrated through the fact that it can be applied to analyze fully discrete schemes for a wide class of stochastic evolution equations: hyperbolic and parabolic alike. The basic line of proof of the general formula is adapted from [9] which is concerned with the stochastic heat equation.

The statement of Theorem 3.1 deserves a motivation. Consider, for example, the case when $\tilde{X}(T)$ is the value of a semidiscretization in space with finite elements, so that

$$\tilde{X}(T) = \tilde{E}(T)\tilde{X}_0 + \int_0^T \tilde{E}(T-s)\tilde{B}\,\mathrm{d}W(s)$$

has the same form as the solution (1.3) of the original problem (1.1). Then an error formula may be derived with the aid of Kolmogorov's backward equation and Itô's formula as in [14]. It turns out that the error analysis is substantially simplified if new processes are constructed by multiplying $X(t)$ and $\tilde{X}(t)$ by suitable integrating factors. That is, define new, drift-free processes

$$Y(t) = E(T-t)X(t) = E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}W(s) \tag{1.5}$$

and

$$\tilde{Y}(t) = \tilde{E}(T-t)\tilde{X}(t) = \tilde{E}(T)\tilde{X}_0 + \int_0^t \tilde{E}(T-s)\tilde{B}\,\mathrm{d}W(s), \tag{1.6}$$

where

$$X(T) = Y(T), \quad \tilde{X}(T) = \tilde{Y}(T), \tag{1.7}$$

with $\{Y(t)\}_{t\geq 0}$ being the solution of the equation

$$\mathrm{d}Y(t) = E(T-t)B\,\mathrm{d}W(t),\ t>0; \quad Y(0) = E(T)X_0 \tag{1.8}$$

and $\{\tilde{Y}(t)\}_{t\geq 0}$ solves

$$\mathrm{d}\tilde{Y}(t) = \tilde{E}(T-t)\tilde{B}\,\mathrm{d}W(t),\ t>0; \quad \tilde{Y}(0) = \tilde{E}(T)\tilde{X}_0. \tag{1.9}$$

However, if $\tilde{X}(T)$ is the result of a time-stepping scheme, then a process of the form (1.3) is not immediately available, nor is the rewriting in (1.6) possible to perform. On the other hand, the construction of a continuous drift-free process as in the right hand side of (1.6) such that (1.7) holds is still possible. In this case $\{\tilde{E}(t)\}_{t\geq 0}$ will not be continuous in t and does not have the semigroup property but it turns out that these are not necessary. All we need to obtain the fundamental formulas (3.6)–(3.8) for the error is to assume that there exists a well-defined process $\{\tilde{Y}(t)\}_{t\in[0,T]}$

on the form

$$\tilde{Y}(t) = \tilde{E}(T)\tilde{X}_0 + \int_0^t \tilde{E}(T-s)\tilde{B}\,\mathrm{d}W(s) \tag{1.10}$$

such that (1.7) holds. Here $\{\tilde{E}(t)\}_{t\in[0,T]} \subset \mathcal{B}(\mathcal{S},\mathcal{S})$ and $\tilde{B} \in \mathcal{B}(\mathcal{U},\mathcal{S})$, where $\mathcal{S}$ is a Hilbert subspace of $\mathcal{H}$ with the same norm (typically $\mathcal{S} = \mathcal{H}$ or $\mathcal{S}$ is a finite-dimensional subspace of $\mathcal{H}$). The process $\tilde{Y}$ is then well defined if

$$\mathrm{Tr}\Big(\int_0^T \tilde{E}(t)\tilde{B}Q[\tilde{E}(t)\tilde{B}]^*\,\mathrm{d}t\Big) < \infty, \tag{1.11}$$

and in this case it will be the unique weak solution of (1.9). Here the adjoint $[\tilde{E}(t)\tilde{B}]^* = \tilde{B}^*\tilde{E}(t)^* \colon \mathcal{S} \to \mathcal{U}$ is taken with respect to the scalar product in $\mathcal{H}$.

In Section 4, we apply the general formula from Theorem 3.1 to the error analysis of semi- and fully discrete numerical schemes for the stochastic wave equation

$$\mathrm{d}\dot{U}(t) - \Delta U(t)\,\mathrm{d}t = \mathrm{d}W(t),\ U(t)|_{\partial\mathcal{D}} = 0,\ t>0; \quad U(0) = U_0,\ \dot{U}(0) = V_0, \tag{1.12}$$

where the solution process $\{U(t)\}_{t\geq 0}$ and the Wiener process $\{W(t)\}_{t\geq 0}$ take values in $\mathcal{U} = L_2(\mathcal{D})$, where $\mathcal{D}$ is a sufficiently nice open bounded domain in $\mathbb{R}^d$. Writing $X(t) = \big[X_1(t), X_2(t)\big]^T := \big[U(t), \dot{U}(t)\big]^T$, $X_0 = \big[X_{0,1}, X_{0,2}\big]^T := \big[U_0, V_0\big]^T$ and $\Lambda := -\Delta$, the wave equation (1.12) can be written in the form (1.1) with

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

It is well known that $-A$ is the generator of a unitary, strongly continuous semigroup (and thus a group) on the space $\mathcal{H} = L_2(\mathcal{D}) \times (H_0^1(\mathcal{D}))^*$.

The first result in Subsection 4.1 is a bound of the error $e(T)$ in the more specific context of the wave equation but keeping the approximating process general. The bound is expressed in terms of the operators and initial data in (1.10) and (1.3). In Subsection 4.2 we apply this to single step rational approximations of (1.1), that is, to solutions of the scheme

$$X^j = R(kA)\big(X^{j-1} + W(t_j) - W(t_{j-1})\big),$$

where k is the step size and where the rational function R fulfills the approximation and stability properties

$$\begin{aligned} |R(iy) - e^{-iy}| &\leq C|y|^{p+1}, \quad && |y| \leq b, \\ |R(iy)| &\leq 1, && y \in \mathbb{R}, \end{aligned} \tag{1.13}$$

for some positive integer p and some $b > 0$. If the initial value is smooth enough and

$$\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty, \tag{1.14}$$

where $\|\cdot\|_{\mathrm{Tr}}$ denotes the trace norm, then for the weak error in (1.4) we have

$$|e(T)| = O(k^{\max(\frac{p}{p+1}2\beta,1)}) \quad \text{as } k \to 0.$$

It is important to note that, in contrast to the stochastic heat equation, the convergence rate is improved with higher order schemes if the noise is sufficiently irregular because $\frac{p}{p+1}$ increases with the order, p, of the method. This feature appears for fully discrete schemes investigated in Subsection 4.3 as well, where the spatial approximation is performed with standard, continuous, piecewise polynomial finite elements of order r and the temporal approximation again with rational functions as in (1.13). If h denotes the size of the finite element mesh and if (1.14) holds, then for the first component $X_1 = U$ we have

$$\big|\mathbf{E}\big(G(\tilde{X}_1(T)) - G(X_1(T))\big)\big| = O(k^{\max(\frac{p}{p+1}2\beta,1)}) + O(h^{\max(\frac{r}{r+1}2\beta,r)}) \text{ as } h,k\to 0.$$

It is a general phenomenon that for non-smooth noise, the rate of weak convergence is twice that of the strong (mean-square) convergence. To be able to compare the weak rate with the strong rate of fully discrete schemes for the wave equation, and also because strong results for such schemes are absent in the literature, we prove a strong convergence result in Section 4.4 for the same algorithm. Indeed, we find that the strong rate is half the weak rate also in the case of the wave equation.

Our main motivation in applying the results of Section 3 to the stochastic wave equation is the fact that this is one of the canonical SPDE's that is less understood from the numerical point of view. Results on weak convergence can be found in [12] but are confined to a leap-frog scheme in one spatial variable. Also, the test-functions in that paper differ from ours, making the results difficult to compare. In [14] the stochastic wave equation is discretized with finite elements in space only but in several spatial dimensions. The findings are in accordance with the results of the present paper.

For the stochastic wave equation on the one-dimensional real line with white noise the strong rate 1/2 for the leap-frog scheme is proved in [24]. This rate is also proved to be optimal. For white noise in one dimension the algorithm discussed in the present paper only gives a strong rate of $p/2(p+1)$ and hence it is not optimal. The reason for this is explained in [15] with the fact that the Green's function of the leap-frog scheme coincides at the meshpoints with the Green's function for the wave equation, which is not true for the present scheme. The latter paper, [15], studies spatial discretization with finite elements in several dimensions with findings in agreement (in one dimension) with the finite difference spatial approximation studied in [22].

In connection to hyperbolic equations [7] should be mentioned, where the authors are concerned with weak as well as strong convergence of a time discretization scheme for a nonlinear stochastic Schrödinger equation.

Finally, in Section 5, we apply Theorem 3.1 to the backward Euler in time and finite element in space approximation of the linearized Cahn-Hilliard-Cook equation. The reason for doing this is twofold. First, we demonstrate that the general error representation formula is useful in the parabolic setting as well. Second, the stochastic heat equation, which is the canonical parabolic equation, has been studied in [9]. While our general approach would certainly be applicable to the heat equation as well, it would just reprove a known result, maybe with a more transparent proof, see Remark 5.3. Similarly to the wave equation, also here we find that the rate of weak convergence is twice that of the strong convergence [16] under essentially the same assumptions.

The literature on weak convergence for parabolic equations is richer. We have already mentioned [9] that proves results for fully discrete schemes of the linear stochastic heat equation. In [13] spatial, finite element schemes are considered for the same equation. Such schemes are also studied for the linear Cahn-Hilliard-Cook equation as well as the linear heat equation in [14]. Semidiscrete temporal schemes are investigated in [11] for the linear heat equation and in [8] for the nonlinear heat equation in one dimension. The techniques of the latter paper are extended to spatially semidiscrete schemes in multiple dimensions in [1]. A recent paper, [19], successfully uses the methods of [9] to study a linear parabolic SPDE driven by impulsive noise instead of a Wiener noise. This indicates the possibility of extending the results of the present paper to larger classes of noise.

2. Preliminaries

Here we collect some background material from infinite-dimensional stochastic analysis and stochastic PDEs. We use the semigroup approach of DaPrato and Zabczyk and we refer to the monograph [5] for details and proofs.

Let $\mathcal{U}$ and $\mathcal{H}$ 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 $\mathcal{U}$ to $\mathcal{H}$ by $\mathcal{B}(\mathcal{U},\mathcal{H})$ and the p:th Schatten class of operators from $\mathcal{U}$ to $\mathcal{H}$ by $\mathcal{L}_p(\mathcal{U},\mathcal{H})$. They are Banach spaces for all integers $p \geq 1$ and we will denote their norms by $\|\cdot\|_{\mathcal{L}_p(\mathcal{U},\mathcal{H})}$. The operators in $\mathcal{L}_1(\mathcal{U},\mathcal{H})$ are also refered to as trace class operators and operators in $\mathcal{L}_2(\mathcal{U},\mathcal{H})$ as Hilbert-Schmidt operators. The space $\mathcal{L}_2(\mathcal{U},\mathcal{H})$ is a Hilbert space with inner product denoted $\langle\cdot,\cdot\rangle_{\mathcal{L}_2(\mathcal{U},\mathcal{H})}$. When the underlying Hilbert spaces are understood from the context we will write $\|\cdot\|_{\mathrm{Tr}} = \|\cdot\|_{\mathcal{L}_1(\mathcal{U},\mathcal{H})}$, $\|\cdot\|_{\mathrm{HS}} = \|\cdot\|_{\mathcal{L}_2(\mathcal{U},\mathcal{H})}$ and $\langle\cdot,\cdot\rangle_{\mathrm{HS}} = \langle\cdot,\cdot\rangle_{\mathcal{L}_2(\mathcal{U},\mathcal{H})}$ in order to – we hope – increase the readability of the paper.

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

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

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, [5, Appendix C], [17] and [25].

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

$$\|T\|_{\mathcal{L}_p(\mathcal{U},\mathcal{H})} = \|T^*\|_{\mathcal{L}_p(\mathcal{H},\mathcal{U})}. \tag{2.1}$$

If $\mathcal{U} = \mathcal{H}$ and $p = 1$, then also

$$\mathrm{Tr}(T) = \mathrm{Tr}(T^*) \tag{2.2}$$

and

$$|\,\mathrm{Tr}(T)| \leq \|T\|_{\mathrm{Tr}}. \tag{2.3}$$

Further, if T is selfadjoint and positive semidefinite, then $\mathrm{Tr}(T) \geq 0$ and (2.3) holds with equality.

If $\mathcal{U}_1$, $\mathcal{U}_2$, and $\mathcal{H}$ are separable Hilbert spaces and $T \in \mathcal{L}_p(\mathcal{U}_2,\mathcal{H})$ and if $S_1 \in \mathcal{B}(\mathcal{U}_1,\mathcal{U}_2)$ and $S_2 \in \mathcal{B}(\mathcal{H},\mathcal{U}_1)$, then

$$\begin{aligned} \|TS_1\|_{\mathcal{L}_p(\mathcal{U}_1,\mathcal{H})} &\leq \|T\|_{\mathcal{L}_p(\mathcal{U}_2,\mathcal{H})}\|S_1\|_{\mathcal{B}(\mathcal{U}_1,\mathcal{U}_2)}, \\ \|S_2T\|_{\mathcal{L}_p(\mathcal{U}_2,\mathcal{U}_1)} &\leq \|T\|_{\mathcal{L}_p(\mathcal{U}_2,\mathcal{H})}\|S_2\|_{\mathcal{B}(\mathcal{H},\mathcal{U}_1)}. \end{aligned} \tag{2.4}$$

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

$$\mathrm{Tr}(TS) = \mathrm{Tr}(ST). \tag{2.5}$$

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

$$\begin{aligned}\|T^*T\|_{\mathrm{Tr}} &= \mathrm{Tr}(T^*T)=\|T\|_{\mathrm{HS}}^2=\|T^*\|_{\mathrm{HS}}^2\\ &= \mathrm{Tr}(TT^*)=\|TT^*\|_{\mathrm{Tr}}.\end{aligned} \tag{2.6}$$

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

$$\|TS\|_{\mathrm{Tr}}\le\|T\|_{\mathrm{HS}}\|S\|_{\mathrm{HS}}=(\mathrm{Tr}(TT^*)\,\mathrm{Tr}(SS^*))^{1/2}. \tag{2.7}$$

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 cite Theorem 2.1 in [14].

Theorem 2.1. *Assume that $Q\in\mathcal{B}(\mathcal{H})$ is selfadjoint, positive semidefinite and that A is a densely defined, unbounded, selfadjoint, positive definite, linear operator in $\mathcal{H}$ with an orthonormal basis of eigenvectors. Then, for $s\in\mathbb{R}$, $\alpha>0$, we have*

$$\begin{aligned}\|A^{\frac{s}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 &\le \|A^sQ\|_{\mathrm{Tr}}\le\|A^{s+\alpha}Q\|_{\mathcal{B}(\mathcal{H})}\|A^{-\alpha}\|_{\mathrm{Tr}},\\ \|A^{\frac{s}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 &\le \|A^{s+\frac12}QA^{-\frac12}\|_{\mathrm{Tr}}.\end{aligned} \tag{2.8}$$

Furthermore, if A and Q have a common basis of eigenvectors, in particular, if $Q=I$, then

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

Let $(\Omega,\mathcal{F},\mathbf{P})$ be a probability space and $L_p(\Omega,\mathcal{H})$ denote the space of random variables $X\colon(\Omega,\mathcal{F})\to(\mathcal{H},\mathrm{Bor}(\mathcal{H}))$, where $\mathrm{Bor}(\mathcal{H})$ denotes the Borel σ-algebra of the separable Hilbert space $\mathcal{H}$, such that

$$\|X\|_{L_p(\Omega,\mathcal{H})}^p=\mathbf{E}\big(\|X\|_{\mathcal{H}}^p\big)=\int_\Omega\|X(\omega)\|_{\mathcal{H}}^p\,\mathrm{d}\mathbf{P}(\omega)<\infty.$$

By "strong convergence" we mean norm convergence in $L_2(\Omega,\mathcal{H})$.

If $\{F(t)\}_{t\in[0,T]}$ is a family of (deterministic) bounded linear operators from a Hilbert space $\mathcal{U}$ to another $\mathcal{H}$, then the Itô integral (also called the Wiener integral as the integrand is deterministic)

$$\int_0^T F(t)\,\mathrm{d}W(t)$$

with respect to a $\mathcal{U}$-valued Q-Wiener process is well defined if

$$\int_0^T\|F(t)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t=\int_0^T\mathrm{Tr}(F(t)QF^*(t))\,\mathrm{d}t<\infty. \tag{2.9}$$

If (2.9) holds then we have Itô's isometry

$$\Big\|\int_0^T F(t)\,\mathrm{d}W(t)\Big\|_{L_2(\Omega,\mathcal{H})}^2=\int_0^T\|F(t)Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}t. \tag{2.10}$$

The functionals G in (1.4) are called test functions and throughout this paper we will assume that they are mappings from $\mathcal{H}$ to $\mathbb{R}$ with bounded and continuous first and second Fréchet derivatives. That is, they belong to the space

$$C_{\mathrm{b}}^2=C_{\mathrm{b}}^2(\mathcal{H},\mathbb{R})=\Big\{G\in C^2(\mathcal{H},\mathbb{R}):\|G\|_{C_{\mathrm{b}}^2(\mathcal{H},\mathbb{R})}<\infty\Big\},$$

where

$$\|G\|_{C_{\mathrm{b}}^2(\mathcal{H},\mathbb{R})}:=\sup_{x\in\mathcal{H}}\|G'(x)\|_{\mathcal{H}}+\sup_{x\in H}\|G''(x)\|_{\mathcal{B}(\mathcal{H})}.$$

The derivatives $G'(x)$ and $G''(x)$ are identified with elements in $\mathcal{H}$ and $\mathcal{B}(\mathcal{H})$, respectively, by the Riesz representation theorem. Note that $\|\cdot\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})}$ is only a seminorm and that we do not assume that G itself is bounded.

3. An error representation formula

Following [9] and [14], we start by developing a representation of the weak error $e(T)$ in (1.4) in the abstract setting. To do this we use the process $\{Y(t)\}_{t\geq 0}$ in (1.5), being the unique weak solution of the drift-free differential equation (1.8) with the important property that $Y(T) = X(T)$. We also introduce the auxiliary problem

$$\mathrm{d}Z(t) = E(T-t)B\,\mathrm{d}W(t),\ t \in (\tau, T]; \quad Z(\tau) = \xi,$$

where $\xi \in L_1(\Omega, \mathcal{H})$ is an $\mathcal{F}_\tau$-measurable random variable. Its unique weak solution is given by

$$Z(t,\tau,\xi) = \xi + \int_\tau^t E(T-s)B\,\mathrm{d}W(s), \quad t \in [\tau, T]. \tag{3.1}$$

We note that $Z(t, 0, E(T)X_0) = Y(t)$ and that $Z(t,t,\xi) = \xi$. For $G \in C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})$, we define the continuous function $u\colon \mathcal{H}\times[0,T] \to \mathbb{R}$ by

$$u(x,t) = \mathbf{E}\big(G(Z(T,t,x))\big). \tag{3.2}$$

It follows from (3.1) and (3.2) that the partial derivatives of u are given by

$$\begin{aligned} u_x(x,t) &= \mathbf{E}\big(G'(Z(T,t,x))\big),\\ u_{xx}(x,t) &= \mathbf{E}\big(G''(Z(T,t,x))\big). \end{aligned} \tag{3.3}$$

Hence,

$$\begin{aligned} \sup_{(x,t)\in\mathcal{H}\times[0,T]} \|u_x(x,t)\| &\leq \|G\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})},\\ \sup_{(x,t)\in\mathcal{H}\times[0,T]} \|u_{xx}(x,t)\|_{\mathcal{B}(\mathcal{H})} &\leq \|G\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})}. \end{aligned} \tag{3.4}$$

It is known that u is a solution to Kolmogorov's equation

(3.5)
$$\begin{aligned} &u_t(x,t) + \tfrac{1}{2}\,\mathrm{Tr}\left(u_{xx}(x,t)E(T-t)BQB^*E(T-t)^*\right) = 0, && (x,t)\in\mathcal{H}\times[0,T),\\ &u(x,T) = G(x), && x \in \mathcal{H}, \end{aligned}$$

see, for example, [6, Lemma 6.1.1], where we note that $G \in C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})$ is enough for existence and $G \in UC^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})$ is only needed for uniqueness as the proofs of [6, Theorems 3.2.3 and 3.2.7] show and the global boundedness of G is not needed by [6, Remark 3.2.1]. Finally, the partial derivatives of u in (3.5) are continuous on $[0,T)\times\mathcal{H}$.

Our key result is the following representation formula for the weak error.

Theorem 3.1. *Assume that* (1.2) *and* (1.11) *hold and let* $\{X(t)\}_{t\in[0,T]}$ *be the unique mild solution* (1.3) *of* (1.1) *and that* $\tilde{X}(T)$ *can be represented as* $\tilde{X}(T) = \tilde{Y}(T)$*, where* $\tilde{Y}$ *is given by* (1.10).

If $G \in C_{\mathrm{b}}^2(\mathcal{H}, \mathbb{R})$, then the weak error $e(T)$ in (1.4) *has the representation*

$$
\begin{aligned}
e(T) &= \mathbf{E} \int_0^1 \Big\langle u_x\big(Y(0) + s(\tilde{Y}(0) - Y(0)), 0\big), \tilde{Y}(0) - Y(0) \Big\rangle \,\mathrm{d}s \\
&\quad + \tfrac{1}{2} \mathbf{E} \int_0^T \mathrm{Tr}\, \Big(u_{xx}(\tilde{Y}(t), t) \mathcal{O}(t) \Big) \,\mathrm{d}t,
\end{aligned} \tag{3.6}
$$

where

$$
\mathcal{O}(t) = \big(\tilde{E}(T-t)\tilde{B} + E(T-t)B\big) Q \big(\tilde{E}(T-t)\tilde{B} - E(T-t)\tilde{B}\big)^*, \tag{3.7}
$$

or

$$
\mathcal{O}(t) = \big(\tilde{E}(T-t)\tilde{B} - E(T-t)B\big) Q \big(\tilde{E}(T-t)\tilde{B} + E(T-t)B\big)^*. \tag{3.8}
$$

Proof. As in [5, Theorem 9.8], since ξ is $\mathcal{F}_t$-measurable, we have that

$$
u(\xi, t) = \mathbf{E}\Big(G(Z(T, t, \xi)) \Big| \mathcal{F}_t \Big). \tag{3.9}
$$

Thus, by the law of double expectation,

$$
\mathbf{E}\Big(u(\xi, t) \Big) = \mathbf{E}\Big(\mathbf{E}\Big(G(Z(T, t, \xi)) \Big| \mathcal{F}_t \Big) \Big) = \mathbf{E}\Big(G(Z(T, t, \xi)) \Big). \tag{3.10}
$$

Therefore, taking also into account that $X(T) = Y(T)$, it follows that

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

and, since $\tilde{X}(T) = \tilde{Y}(T)$, we also have that

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

Hence,

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

For the first term we note that due to the differentiability of u we can write

$$
\begin{aligned}
&\mathbf{E}\Big(u(\tilde{Y}(0), 0) - u(Y(0), 0) \Big) \\
&\qquad = \mathbf{E} \int_0^1 \Big\langle u_x\big(Y(0) + s(\tilde{Y}(0) - Y(0))\big), \tilde{Y}(0) - Y(0) \Big\rangle \,\mathrm{d}s.
\end{aligned}
$$

For the second term, we use Itô's formula (see [4, Theorem 2.1], where, in contrast to [5, Theorem 4.17], uniform continuity of the appearing derivatives on bounded subsets of $\mathcal{H} \times [0, T)$ is not assumed) for $u(\tilde{Y}(t), t)$ on $[0, T - \epsilon]$ and passing to the limit $\epsilon \to 0+$ using the continuity of u on $\mathcal{H} \times [0, T]$ and the continuity of the paths of $\tilde{Y}(t)$ on $[0, T]$. Thus, taking also Kolmogorov's equation (3.5) into account, we get

(3.11)

$$
\begin{aligned}
&\mathbf{E}\Big(u(\tilde{Y}(T),T)-u(\tilde{Y}(0),0)\Big)\\
&\quad=\mathbf{E}\int_0^T\Big\{u_t(\tilde{Y}(t),t)+\tfrac12\operatorname{Tr}\Big(u_{xx}(\tilde{Y}(t),t)[\tilde{E}(T-t)\tilde{B}]Q[\tilde{E}(T-t)\tilde{B}]^*\Big)\Big\}\,\mathrm{d}t\\
&\quad=\tfrac12\mathbf{E}\int_0^T\operatorname{Tr}\Big(u_{xx}(\tilde{Y}(t),t)\big\{[\tilde{E}(T-t)\tilde{B}]Q[\tilde{E}(T-t)\tilde{B}]^*\\
&\qquad-[E(T-t)B]Q[E(T-t)B]^*\big\}\Big)\,\mathrm{d}t.
\end{aligned}
$$

The operator $u_{xx}(\xi,r)$ 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 assumptions (1.2) and (1.11). Hence, the trace above is well defined for almost every t since by (2.4) with $p=1$,

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

where the last step is (2.3) 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 $u_{xx}(\xi,r)[E(s)B]Q[\tilde{E}(s)\tilde{B}]^*$ is also of trace class for almost every s, since, by (2.1), (2.4), and (2.7),

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

Therefore we may rewrite the operator in the trace in (3.11) by adding and subtracting $u_{xx}(\xi,r)[E(s)B]Q[\tilde{E}(s)\tilde{B}]^*$ to get

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

Further, using (2.2), (2.5), and that Q and $u_{xx}(\xi,r)$ are selfadjoint, we obtain

$$\begin{aligned}
\mathrm{Tr}(O_1+O_2) &= \mathrm{Tr}(O_1)+\mathrm{Tr}(O_2) = \mathrm{Tr}(O_1)+\mathrm{Tr}(O_2^*)\\
&= \mathrm{Tr}(O_1)+\mathrm{Tr}([\tilde{E}(s)\tilde{B}-E(s)B]Q[E(s)B]^*u_{xx}(\xi,r))\\
&= \mathrm{Tr}(O_1)+\mathrm{Tr}(u_{xx}(\xi,r)[\tilde{E}(s)\tilde{B}-E(s)B]Q[E(s)B]^*)\\
&= \mathrm{Tr}\left(u_{xx}(\xi,r)[\tilde{E}(s)\tilde{B}-E(s)B]Q[\tilde{E}(s)\tilde{B}+E(s)B]^*\right) \qquad (3.12)\\
&= \mathrm{Tr}\left([\tilde{E}(s)\tilde{B}+E(s)B]Q[\tilde{E}(s)\tilde{B}-E(s)B]^*u_{xx}(\xi,r)\right)\\
&= \mathrm{Tr}\left(u_{xx}(\xi,r)[\tilde{E}(s)\tilde{B}+E(s)B]Q[\tilde{E}(s)\tilde{B}-E(s)B]^*\right). \qquad (3.13)
\end{aligned}$$

Finally, by inserting (3.12) or (3.13) into (3.11) we finish the proof. □

4. Application to the wave equation

4.1. **A general error formula.** In this section we apply the general result from Section 3 to the stochastic wave equation (1.12). As mentioned in the Introduction the wave equation may be re-written in the form (1.1). We start out by making this statement precise. At the same time we introduce a framework for measuring the regularity of the solutions and to perform a careful error analysis. To this aim we let $\mathcal{D}$ denote a convex open bounded domain in $\mathbb{R}^d$ with polygonal boundary $\partial\mathcal{D}$, and equip $L_2(\mathcal{D})$ with the usual norm $\|\cdot\|_{L_2(\mathcal{D})}$ and inner product $\langle\cdot,\cdot\rangle_{L_2(\mathcal{D})}$ and let $\Lambda=-\Delta$ be the Laplace operator with $D(\Lambda)=H^2(\mathcal{D})\cap H_0^1(\mathcal{D})$. To measure regularity we introduce a scale of Hilbert spaces. Let $\{(\lambda_j,\phi_j)\}_{j=1}^\infty$ be eigenpairs of Λ with a nondecreasing sequence of positive eigenvalues λ_j and corresponding orthonormal eigenvectors ϕ_j. For $\alpha\in\mathbb{R}$ we endow $D(\Lambda^{\alpha/2})$ with inner product

$$\langle v,w\rangle_{\dot{H}^\alpha} := \langle\Lambda^{\alpha/2}v,\Lambda^{\alpha/2}w\rangle_{L_2(\mathcal{D})} = \sum_{j=1}^\infty \lambda_j^\alpha\langle v,\phi_j\rangle_{L_2(\mathcal{D})}\langle w,\phi_j\rangle_{L_2(\mathcal{D})},$$

and the corresponding norm $\|v\|_{\dot{H}^\alpha}^2=\langle v,v\rangle_{\dot{H}^\alpha}$. For $\alpha\geq 0$ the space $\dot{H}^\alpha$ now may be defined through

$$\dot{H}^\alpha := \{v\in L_2(\mathcal{D}) : \|v\|_{\dot{H}^\alpha}<\infty\}. \qquad (4.1)$$

If $\alpha<0$, then $\dot{H}^\alpha$ is taken to be the closure of $L_2(\mathcal{D})$ with respect to $\|\cdot\|_{\dot{H}^\alpha}$. It is notable that with $\alpha>0$ the space $\dot{H}^{-\alpha}$ may be identified with the dual of $\dot{H}^\alpha$ and that $\dot{H}^\alpha\subset\dot{H}^\beta$ if $\alpha\geq\beta$. Further, it is known that $\dot{H}^0=L_2(\mathcal{D})$, $\dot{H}^1=H_0^1(\mathcal{D})$, $\dot{H}^2=H^2(\mathcal{D})\cap H_0^1(\mathcal{D})$, see [23, Chapt. 3]. In addition we define the product space

$$\mathcal{H}^\alpha := \dot{H}^\alpha\times\dot{H}^{\alpha-1}, \quad \alpha\in\mathbb{R}, \qquad (4.2)$$

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

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

We take $\mathcal{H}$ to be the special case of (4.2) 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 $\dot{H}^{-1}$ as $(\Lambda x)(y) = \langle \nabla x, \nabla y \rangle_{L_2(D)}$ and let A and B be defined by

$$A := \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ I \end{bmatrix}. \tag{4.3}$$

Here B is considered as an operator from $\dot{H}^{-1}$ to $\mathcal{H}$ and the domain of A is

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

We have some freedom in defining $\mathcal{U}$ but a natural choice is $\mathcal{U} = \dot{H}^0 = L_2(\mathcal{D})$. Thus, we consider the process $\{W(t)\}_{t\geq 0}$ to be a Q-Wiener process on $L_2(\mathcal{D})$ and thus Q to be bounded, selfadjoint and positive semidefinite on $\mathcal{U} = L_2(\mathcal{D})$. Note that $\dot{H}^0 \subset \dot{H}^{-1}$, and therefore $B \in \mathcal{B}(\mathcal{U}, \mathcal{H})$. It is well known that the operator $-A$ is the generator of a strongly continuous semigroup $E(t) = \mathrm{e}^{-tA}$ on $\mathcal{H}$, in fact, a unitary group, that can be written as

$$E(t) = \mathrm{e}^{-tA} = \begin{bmatrix} C(t) & \Lambda^{-1/2} S(t) \\ -\Lambda^{1/2} S(t) & C(t) \end{bmatrix}, \tag{4.4}$$

where $C(t) = \cos(t\Lambda^{1/2})$ and $S(t) = \sin(t\Lambda^{1/2})$ are the so-called cosine and sine operators.

With these definitions (1.1) becomes the stochastic wave equation (1.12) with solution $\big[X_1(t), X_2(t)\big]^T \in \mathcal{H}$ if also the initial value $X_0 = \big[X_{0,1}, X_{0,2}\big]^T \in \mathcal{H}$ and (1.2) holds.

We are know ready to use the framework set forth in Section 3, starting by assuming as little as possible about the type of perturbations of $X(t)$. To get a result that enables error analysis of the full solution X, as well as either of the coordinates X_1 or X_2, we make three additional, rather weak assumption on the data of the problem. First, we take G to be the composition of a function $g \in C_{\mathrm{b}}^2(\mathcal{V}, \mathbb{R})$, where $\mathcal{V}$ is a real separable Hilbert space, with an operator $L \in \mathcal{B}(\mathcal{H}, \mathcal{V})$, i.e.,

$$G(v) = g(Lv), \quad v = [v_1, v_2]^T \in \mathcal{H}.$$

It is clear that $G \in C_{\mathrm{b}}^2(\mathcal{H}, \mathbb{R})$. Typically we have $\mathcal{V} = \mathcal{H}$ and $L = I$, or $\mathcal{V} = \dot{H}^0$ and $Lv = P^1 v = v_1$ (projection to the first coordinate), or possibly $\mathcal{V} = \dot{H}^{-1}$ and $Lv = P^2 v = v_2$ (projection to the second coordinate). Second, since in all interesting cases the function $\tilde{X}_0$ will be related to X_0 we will assume that this relation can be described by an operator $\tilde{P} \in \mathcal{B}(\mathcal{H})$ through

$$\tilde{X}_0 = \tilde{P} X_0.$$

In the cases we consider below, $\tilde{P}$ will be the orthogonal projection onto a finite element subspace of $\mathcal{H}$ or the identity operator but it could also be an interpolation operator. The third assumption we make is that $\tilde{B} = \tilde{P} B$. This is unnecessarily restrictive in general but it suffices for the purposes of this paper and it also makes the presentation clearer. We begin with a technical result.

Lemma 4.1. *If $E(t)$ is given by (4.4), 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

$$\mathrm{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) = T\,\mathrm{Tr}(\Lambda^{-1/2}Q\Lambda^{-1/2}). \tag{4.5}$$

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,

$$\mathrm{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) = \int_0^T \mathrm{Tr}\Big(E(t)BQB^*E(t)^*\Big)\,\mathrm{d}t \tag{4.6}$$

and by (2.6) and since $E(t)^* = E(t)^{-1}$ we have

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

Thus, the integrand in the right-hand side of (4.6) is constant and is equal to $\mathrm{Tr}_{\dot H^0}(\Lambda^{-1/2}Q\Lambda^{-1/2})$, which implies (4.5). Therefore (i) must be true. This argument is reversible so (i) $\Leftrightarrow$ (iv). But from this computation it is also evident that $\|\Lambda^{-1/2}Q^{1/2}\|^2_{\mathrm{HS}} = \mathrm{Tr}(\Lambda^{-1/2}Q\Lambda^{-1/2})$, i.e., that (i) $\Leftrightarrow$ (ii). ☐

Corollary 4.2. *If* $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ *for some* $\beta \ge 0$, *then* (iv) *holds and* (1.1) *has a unique weak solution* $\{X(t)\}_{t\in[0,T]}$.

Proof. By (2.3) and (2.4) we have that

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

The result now follows by Lemma 4.1. ☐

Since $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|^2_{\mathrm{HS}} \le \|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}}$ by (2.8), we may conclude that under the assumption in the previous corollary, we actually have mean-square regularity of order β. This follows from the following theorem, which is quoted from [15, Theorem 3.1].

Theorem 4.3. *If* $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|^2_{\mathrm{HS}} < \infty$, $X_0 \in L_2(\Omega,\mathcal{H}^\beta)$ *for some* $\beta \ge 0$ *and* A, B *as in* (4.3), *then the mild solution* (1.3) *satisfies*

$$\|X(t)\|_{L_2(\Omega,\mathcal{H}^\beta)} \le C\big(\|X_0\|_{L_2(\Omega,\mathcal{H}^\beta)} + T^{1/2}\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}\big).$$

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 4.4. *If* $\{E(t)\}_{t\ge0}$ *is the semigroup in* (4.4), *then*

$$\|(E(t)-E(s))x\| \le C|t-s|^\alpha\|x\|_{\mathcal{H}^\alpha}, \quad x\in\mathcal{H}^\alpha,\ t,s\ge0,\ \alpha\in[0,1].$$

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

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

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

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

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 $\mathcal{B}(\dot H^1, \dot H^0)$. To this aim we note that for real y the inequality

$$|\sin(ty) - \sin(sy)| \le |t-s||y| \tag{4.7}$$

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

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

The inequality (4.7) 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.

Theorem 4.5. *Assume that $\{X(t)\}_{t\in[0,T]}$ is the mild solution* (1.3) *of the stochastic wave equation* (1.1) *with $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ and $X_0 \in L_1(\Omega, \mathcal{H}^{2\beta})$ for some $\beta \ge 0$. Assume further that $\tilde X(T)$ can be represented as $\tilde X(T) = \tilde Y(T)$, where $\tilde Y(t)$ is given by* (1.10) *with $\tilde X_0 = \tilde P X_0$, $\tilde P \in \mathcal{B}(\mathcal{H})$ and $\tilde B = \tilde P B$ and such that* (1.11) *holds. Let $g \in C_{\mathrm{b}}^2(\mathcal{V}, \mathbb{R})$ and $L \in \mathcal{B}(\mathcal{H}, \mathcal{V})$. Define*

$$K_1 := \sup_{t\in[0,T]} \|\tilde E(t)\tilde P\|_{\mathcal{H}}, \tag{4.8}$$

$$K_2 := \|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}}. \tag{4.9}$$

Then there is $C = C(T, \|L\|_{\mathcal{B}(\mathcal{H},\mathcal{V})}, \|X_0\|_{L_1(\Omega,\mathcal{H}^{2\beta})}, \|g\|_{C_{\mathrm{b}}^2}, K_1, K_2)$ such that

$$\left|\mathbf{E}\left(g(L\tilde X(T)) - g(LX(T))\right)\right| \le C \sup_{t\in[0,T]} \|L(\tilde E(t)\tilde P - E(t))\|_{\mathcal{B}(\mathcal{H}^{2\beta},\mathcal{V})}. \tag{4.10}$$

We want to emphasize 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

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

Proof. First note that (4.9) implies (1.2) by Corollary 4.2 and hence we may use Theorem 3.1 with $G(X)=g(LX)$. To this aim we note that $G'(X)=L^*g'(LX)$ and $G''(X)=L^*g''(LX)L$. The terms in (3.6) will be estimated in order of appearance with $\mathcal{O}$ as in (3.8). For the first term, by (3.3), (3.9) and the fact that both $Y(0)$ and $\tilde{Y}(0)$ are $\mathcal{F}_0$-measurable, we have

$$\begin{aligned}
&\Big|\mathbf{E}\Big(\int_0^1 \Big\langle u_x\big(Y(0)+s(\tilde{Y}(0)-Y(0)),0\big),\tilde{Y}(0)-Y(0)\Big\rangle\,\mathrm{d}s\Big)\Big|\\
&\quad=\Big|\mathbf{E}\Big(\int_0^1 \Big\langle \mathbf{E}\Big(g'\big(LZ(T,0,Y(0)+s(\tilde{Y}(0)-Y(0)))\big)\Big|\mathcal{F}_0\Big),L(\tilde{Y}(0)-Y(0))\Big\rangle\,\mathrm{d}s\Big)\Big|\\
&\quad\le \sup_{x\in\mathcal{V}}\|g'(x)\|\,\mathbf{E}\big(\|L(\tilde{Y}(0)-Y(0))\|\big)\\
&\quad= \sup_{x\in\mathcal{V}}\|g'(x)\|\,\mathbf{E}\big(\|L(\tilde{E}(T)\tilde{X}_0-E(T)X_0)\|\big)\\
&\quad\le \|g\|_{C_{\mathrm{b}}^2}\,\|L(\tilde{E}(T)\tilde{P}-E(T))\|_{\mathcal{B}(\mathcal{H}^{2\beta},\mathcal{V})}\|X_0\|_{L_1(\Omega,\mathcal{H}^{2\beta})}.
\end{aligned}$$

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 (3.3), (3.9), we may write

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

Using (2.5), (2.3) and (2.4) we bound the integrand above as follows:

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

Here we have $\sup_{x\in\mathcal{V}}\|g''(x)\|_{\mathcal{B}(\mathcal{V})}\le\|g\|_{C_{\mathrm{b}}^2}$ and by (2.1) and (2.4),

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

The first factor can be estimated as

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

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

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

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

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

and the proof is complete. □

4.2. **Weak convergence of temporally semidiscrete schemes.** We begin by applying Theorem 4.5 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 [3]. An I-stable rational approximation of order p is a rational function R such that (1.13) holds. A class of such functions are constructed in [2] and analyzed further in connection to oscillation equations in [20]. 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 $E_k = R(kA)$, $k>0$, are well defined on $\mathcal{H}$ and they are contractions, and hence stable, as $-A$ generates a unitary group. Here $k = T/N$, $N\in\mathbb{N}$, is the time step. We can approximate the solution of (1.1) on the uniform grid $t_j = jk$, $j=0,\dots,N$, by the solution of the difference equation

$$X_k^j = E_k(X_k^{j-1} + B\Delta W^j),\ j=1,\dots,N;\quad X_k^0 = X_0, \tag{4.11}$$

given by

$$X_k^n = E_k^n X_0 + \sum_{j=1}^n E_k^{n-j+1}B\Delta W^j,\quad n\ge 1,$$

where $\Delta W^j = W(t_j) - W(t_{j-1})$. We want to define a process $\{\tilde Y_k(t)\}_{t\in[0,T]}$ of the form (1.10) that is as close as possible to (1.5) and such that $X_k^N = \tilde Y_k(T)$. To this aim we first define a new discrete process

$$Y_k^n = E_k^{N-n}X_k^n = E_k^N X_0 + \sum_{j=1}^n E_k^{N-j+1}B\Delta W^j. \tag{4.12}$$

Clearly $Y_k^N = X_k^N$. In order to make a piecewise constant time interpolation of (4.12) we introduce the time intervals $I_j = [t_{j-1}, t_j)$ for $j=1,\dots,N$ and $I_{N+1} =$

$\{t_N\} = \{T\}$. With χ being the indicator function we then write

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

It may easily be checked that this corresponds to writing

$$\tilde{E}_k(t) = \sum_{j=0}^{N} E_k^j \chi_{\widehat{I}_j}(t)$$

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

$$\tilde{Y}_k(t) := \tilde{E}_k(T)X_0 + \int_0^t \tilde{E}_k(T-s)B \,\mathrm{d}W(s). \tag{4.13}$$

The process $\tilde{Y}_k$ has the desired properties and, in addition, $\tilde{Y}_k(t_n) = Y_k^n$. To apply Theorem 4.5 it remains to show that (1.11) holds with $\tilde{E}(t) = \tilde{E}_k(t)$ and $\tilde{P} = I$ (hence $\tilde{B} = B$). This is indeed the case as soon as we are guaranteed a weak solution of (1.1), as stated in the following lemma.

Lemma 4.6. *If $\tilde{E}(t) = \tilde{E}_k(t)$ and $\tilde{B} = B$, then* (1.2) *implies* (1.11).

Proof. If (1.2) holds then the trace of $\Lambda^{-1/2}Q\Lambda^{-1/2}$ is finite by Lemma 4.1. Thus, for all $t \geq 0$, using (2.3) and (2.4), it follows that

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

where the last equality is shown in the proof of Lemma 4.1 as $\mathrm{Tr}(BQB^*) = \mathrm{Tr}(Q^{1/2}B^*BQ^{1/2})$ by (2.6). The statement of the lemma follows by the monotone convergence theorem again as in the proof of Lemma 4.1. □

In order to make use of Theorem 4.5 we need a bound on $E(t) - \tilde{E}_k(t)$. The results in [3] are concerned with the difference at the grid points. With our notation their conclusion reads that, for $x \in \mathcal{H}^{p+1}$,

$$\|(E(t_n) - E_k^n)x\| \leq C t_n k^p \|x\|_{\mathcal{H}^{p+1}}. \tag{4.14}$$

As already mentioned, the conditions in (1.13) ensures that the operator E_k^n is a contraction on $\mathcal{H}$ for any $n \geq 0$, so (4.14) can be extended to fractional order by interpolation, i.e.,

$$\|(E(t_n) - E_k^n)x\| \leq C t_n k^{\alpha\frac{p}{p+1}} \|x\|_{\mathcal{H}^\alpha}, \quad \alpha \in [0, p+1]. \tag{4.15}$$

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

Lemma 4.7. *For the operators $\tilde{E}_k(t)$ and $E(t)$ defined above, we have that*

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

where p is a nonnegative integer as in (1.13) *and $\alpha \geq 0$.*

Proof. The statement of the lemma follows from (4.15) and Lemma 4.4. Indeed, for $t \in \widehat{I}_j$, we have

$$\tilde{E}_k(t) - E(t) = \big(\tilde{E}_k(t_j) - E(t_j)\big) + \big(E(t_j) - E(t)\big) = \big(E_k^j - E(t_j)\big) + \big(E(t_j) - E(t)\big).$$

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

$$\begin{aligned}\sup_{t\in I} &\|\tilde{E}_k(t) - E(t)\|_{\mathcal{B}(\mathcal{H}^\alpha,\mathcal{H})} \\ &\le \sup_{j\in\mathcal{I}} \big(\|E_k^j - E(t_j)\|_{\mathcal{B}(\mathcal{H}^\alpha,\mathcal{H})} + \sup_{t\in\widehat{I}_j} \|E(t_j) - E(t)\|_{\mathcal{B}(\mathcal{H}^\alpha,\mathcal{H})}\big) \\ &\le C(T)(k^{\min(\alpha\frac{p}{p+1},p)} + k^{\min(\alpha,1)}) \le C(T)k^{\min(\alpha\frac{p}{p+1},1)},\ k \le 1.\end{aligned}$$

Finally, for $k > 1$, the statement follows by stability. □

We are now ready to prove a bound for the weak error of the pure time-discretization via (4.11) of the stochastic wave equation.

Theorem 4.8. *Assume that $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ and $X_0 \in L_1(\Omega, \mathcal{H}^{2\beta})$ for some $\beta \ge 0$ and $G \in C_{\mathrm{b}}^2(\mathcal{H}, \mathbb{R})$. Then the weak error of the rational approximation algorithm* (4.11) *of the stochastic wave equation described above is bounded by*

$$\big|\mathbf{E}\left(G(X_k^N) - G(X(T))\right)\big| \le Ck^{\min(2\beta\frac{p}{p+1},1)},\ k > 0. \tag{4.16}$$

Proof. We may use Theorem 4.5 with $\tilde{E}(t) = \tilde{E}_k(t)$, $\tilde{P} = I$, $\mathcal{V} = \mathcal{H}$ and $L = I$, because $\|\Lambda^{\beta-1/2}Q\Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ implies (1.11) by Corollary 4.2 and Lemma 4.6. Thus, our claim follows by applying Lemma 4.7 with $\alpha = 2\beta$ to (4.10). □

4.3. **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 [2], where estimates for the deterministic wave equation are proved. The spatial discretization is performed by a standard continuous 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 [2].

We assume that $\mathcal{D}$ is a convex polygonal domain and we let $\{S_h^r\}_{0<h\le1}$, $r = 2, 3$, be a standard family of finite element function spaces consisting of continuous piecewise polynomials of degree $r-1$ with respect to a regular family of triangulations of $\mathcal{D}$. Moreover, we define $S_{h,0}^r = \{v \in S_h^r : v|_{\partial\mathcal{D}} = 0\}$, so that $S_{h,0}^r \subset \dot{H}^1$. With H^β denoting the standard Sobolev space we then have the error estimate

$$\|R_h v - v\| \le Ch^\beta \|v\|_{H^\beta}, \quad v \in \dot{H}^1 \cap H^\beta,\ \beta \in [1, r], \tag{4.17}$$

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

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

Further, we define the discrete Laplacian $\Lambda_h\colon S_{h,0}^r \to S_{h,0}^r$ by

$$(\Lambda_h\eta, \chi) = (\nabla\eta, \nabla\chi), \quad \forall \eta, \chi \in S_{h,0}^r.$$

The homogeneous spatially semidiscrete wave equation is to find

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

such that

$$\begin{bmatrix}\dot{u}_{h,1}\\ \dot{u}_{h,2}\end{bmatrix} + \begin{bmatrix}0 & -I\\ \Lambda_h & 0\end{bmatrix}\begin{bmatrix}u_{h,1}(t)\\ u_{h,2}(t)\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix},\ t > 0; \quad \begin{bmatrix}u_{h,1}(0)\\ u_{h,2}(0)\end{bmatrix} = \begin{bmatrix}P_{h,1}u_{0,1}\\ P_{h,2}u_{0,2}\end{bmatrix} \tag{4.18}$$

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

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

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

and if $P_h = [P_{h,1}, P_{h,2}]^T$ and $u_0 := [u_{0,1}, v_{0,2}]^T$, then (4.18) may be written

$$\dot u_h + A_h u_h = 0,\ t > 0; \quad u_h(0) = P_h u_0. \tag{4.19}$$

The operator $-A_h$ is the infinitesimal generator of a strongly continuous semigroup $E_h(t)$ and the solution of (4.19) is given by

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

Similarly to (4.4) the operator $E_h(t)$ has a representation in terms of sine and cosine operators; i.e.,

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

with $S_h(t) = \sin(t\Lambda_h^{1/2})$ and $C_h(t) = \cos(t\Lambda_h^{1/2})$.

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 fulfills (1.13) for some positive integer p. The fully discrete problem on the same uniform grid as in Subsection 4.2 then reads

$$v^n_{h,k} = R(kA_h) v^{n-1}_{h,k},\ n = 1, \dots, N; \quad v^0_{h,k} = P_h u_0. \tag{4.20}$$

We will henceforth write $E_{h,k} = R(kA_h)$ and the solution of (4.20) may then be written as

$$v^n_{h,k} = E^n_{h,k} P_h u_0. \tag{4.21}$$

The error estimate proved in [2] 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 4.9. *If $P^1\colon \mathcal{H} \to \dot H^0$ is defined as $P^1 x = x_1$ for $x = [x_1, x_2]^T \in \mathcal{H}$, then*

$$\|P^1(E^n_{h,k} P_h - E(t_n)) u_0\|_{\dot H^0} \le C(t_n)\big(h^r \|u_0\|_{\mathcal{H}^{r+1}} + k^p \|u_0\|_{\mathcal{H}^{p+1}}\big),\ t_n = nk \ge 0.$$

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

Corollary 4.10. *Under the assumptions of Theorem 4.9 we have, for $\beta \ge 0$,*

$$\|P^1(E^n_{h,k} P_h - E(t_n))\|_{\mathcal{B}(\mathcal{H}^\beta, \dot H^0)} \le C(t_n)\big(h^{\min(\beta\frac{r}{r+1}, r)} + k^{\min(\beta\frac{p}{p+1}, p)}\big),\ t_n = nk \ge 0.$$

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

$$X^j_{h,k} = E_{h,k}(X^{j-1}_{h,k} + B_h \Delta W^j),\ j = 1, \dots, N; \quad X^0_{h,k} = P_h X_0. \tag{4.22}$$

The solution is given by

$$X^n_{h,k} = E^n_{h,k} P_h X_0 + \sum_{j=1}^n E^{n-j+1}_{h,k} B_h \Delta W^j. \tag{4.23}$$

As in the previous section we multiply by $E_{h,k}^{N-n}$ and arrive at the drift free version

$$Y_{h,k}^n = E_{h,k}^N P_h X_0 + \sum_{j=1}^n E_{h,k}^{N-j+1} B_h W^j$$

and with piecewice constant interpolation

$$\tilde{Y}_{h,k}(t) = \tilde{E}_{h,k}(T) P_h X_0 + \int_0^t \tilde{E}_{h,k}(T-s) B_h \,\mathrm{d}W(s) \tag{4.24}$$

in exact analogy with the temporally semidiscrete case in (4.13).

Next we bound the weak error for fully discrete schemes given by (4.22). We only prove a result for the first component in X.

Theorem 4.11. *Assume that* $\|\Lambda^{\beta-1/2} Q \Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ *and* $X_0 \in L_1(\Omega, \mathcal{H}^{2\beta})$ *for some* $\beta \geq 0$. *If* $X_{h,k}^n$ *is given by* (4.20) *and* $X(t)$ *is the weak solution* (1.3) *of* (1.1) *with* A, B *as in* (4.3), *then for* $g \in C_{\mathrm{b}}^2(\dot{H}^0, \mathbb{R})$, *we have*

$$\big|\mathbf{E}\big(g(X_{h,k,1}^N) - g(X_1(T))\big)\big| \leq C(T)\big(h^{\min(2\beta\frac{r}{r+1}, r)} + k^{\min(2\beta\frac{p}{p+1}, 1)}\big).$$

Proof. The function in (4.24) is clearly of the form (1.10) with $\tilde{Y}_{h,k}(T) = X_{h,k}^N$ and we have already seen that $\|\Lambda^{\beta-1/2} Q \Lambda^{-1/2}\|_{\mathrm{Tr}} < \infty$ implies (1.2). Furthermore,

$$\mathrm{Tr}(\tilde{E}_{h,k}(t) B_h Q B_h^* \tilde{E}_{h,k}(t)^*) \leq \mathrm{Tr}(B_h Q B_h) < \infty,$$

as $B_h Q B_h$ is a bounded operator with finite-dimensional range and hence it is of trace class. Therefore also (1.11) holds and Theorem 4.5 can be applied with $\mathcal{V} = \dot{H}^0$, $L = P^1$ (as defined in Theorem 4.9), $\tilde{E}(t) = \tilde{E}_{h,k}(t)$, $\tilde{B} = B_h$ and $\tilde{P} = P_h$. From Corollary 4.10 and Lemma 4.4, as in the proof of Lemma 4.7, it follows that

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

Finally the statement of the theorem follows from inserting this into (4.10). □

4.4. **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 of 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. As we are not aware of any results on strong convergence of a fully discrete approximation of the stochastic wave equation using finite elements in the spatial domain, we give a short derivation of the strong order. We remark that the strong convergence is studied for a spatially semidiscrete finite element method in [15], for a fully discrete leap-frog scheme in one spatial dimension in [24], and for a spatially semidiscrete scheme in one dimension in [22].

First we form the strong error by taking the difference of (4.23) and (1.3), projecting onto the first component, and taking norms:

$$\begin{aligned}
\mathbf{E}\Big(\|P^1(X_{h,k}^N - X(T))\|_{\dot{H}^0}^2\Big) &\leq C\mathbf{E}\Big(\|P^1(E_{h,k}^N P_h - E(T))X_0\|_{\dot{H}^0}^2\Big) \\
&\quad + C\mathbf{E}\Big(\Big\|P^1 \int_0^T (\tilde{E}_{h,k}(T-s)P_h - E(T-s))B \,\mathrm{d}W(s)\Big\|_{\dot{H}^0}^2\Big) =: I_1 + I_2.
\end{aligned}$$

If $X_0 \in L_2(\Omega, \mathcal{H}^\beta)$, then

$$\begin{aligned} I_1 &\le C\|P^1(E_{h,k}P_h - E(T))\|^2_{\mathcal{B}(\mathcal{H}^\beta,\dot H^0)}\mathbf{E}\big(\|X_0\|^2_{\mathcal{H}^\beta}\big) \\ &\le C(T)\big(h^{\min(\beta\frac{r}{r+1},r)} + k^{\min(\beta\frac{p}{p+1},p)}\big)^2\|X_0\|^2_{L_2(\Omega,\mathcal{H}^\beta)} \end{aligned}$$

by Corollary 4.10. For I_2 we use Itô's isometry (2.10) to get

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

where the first inequality follows from the fact that $\|B\Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{B}(\dot H^0,\mathcal{H}^\beta)} = 1$ combined with (2.4), and the last inequality from Corollary 4.10 and Lemma 4.4 as in the proof of Lemma 4.7. Combining the bounds for I_1 and I_2 and taking square roots, we have shown the following result.

Theorem 4.12. *Let $\|\Lambda^{\frac{\beta-1}{2}}Q^{1/2}\|^2_{\mathrm{HS}} < \infty$ and $X_0 \in L_2(\Omega, \mathcal{H}^\beta)$ for some $\beta \ge 0$. Then the strong error of the approximation $X^N_{h,k,1} = P^1X^N_{h,k}$ of the displacement $X_1(T) = P^1X(T)$ in the stochastic wave equation is bounded by*

$$\|X^N_{h,k,1} - X_1(T)\|_{L_2(\Omega,\dot H^0)} \le C(T)\big(h^{\min(\beta\frac{r}{r+1},r)} + k^{\min(\beta\frac{p}{p+1},1)}\big).$$

The regularity assumption on Q in Theorem 4.11 implies the assumption in Theorem 4.12, see Theorem 2.1 with $s = \beta - 1$ in (2.8). Thus the claim that the weak rate is essentially twice the strong rate is justified (if β is not too large) by comparing Theorems 4.11 and 4.12. Note also that the mean-square regularity is of order β according to Theorem 4.3.

5. Application to parabolic equations

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 [16], is

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

where now $\Lambda = -\Delta$ is the Laplacian together with homogeneous Neumann boundary conditions. To write the CHC equation in the form (1.1) we therefore set $\mathcal{H} = \{f \in L_2(\mathcal{D}) : \langle f, 1\rangle = 0\}$, $A = \Lambda^2$ with $D(\Lambda) = \{f \in H^2(\mathcal{D}) \cap \mathcal{H} : \frac{\partial f}{\partial n} = 0\}$, where $\mathcal{D}$ is a convex polygonal domain, and we take $\mathcal{U} = \mathcal{H}$ and $B = I$. We further define the the spaces $\dot H^\alpha = D(\Lambda^\alpha)$ in analogy with Section 4. Thus, $\mathcal{H} = \dot H^0$,

$D(A) = \dot{H}^4$ and $-A$ is known to be the infinitesimal generator of the analytic semigroup $E(t) = \mathrm{e}^{-tA} = \mathrm{e}^{-t\Lambda^2}$ on $\mathcal{H}$.

We recall the finite element spaces S_h^r (without boundary conditions) of order $r = 2, 3$ from Subsection 4.3 and set $\dot{S}_h^r = \{v \in S_h^r : \langle v, 1\rangle = 0\}$. We now define the discrete Laplacian $\Lambda_h \colon \dot{S}_h^r \to \dot{S}_h^r$ and the Ritz projector $R_h \colon \dot{H}^1 \to \dot{S}_h^r$ in the analogous way and we have an error bound of the same form as in (4.17). We set $A_h = \Lambda_h^2$ and note that $-A_h$ is the generator of an analytic semigroup $E_h(t)$ on $\dot{S}_h^r$. We consider only the backward Euler time-stepping and therefore introduce $E_{h,k} = (1 + kA_h)^{-1}$ and define $\tilde{E}_{h,k}(t)$ in an analogous fashion to the case of the wave equation, see (4.13).

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

$$\|(E_{h,k}^n P_h - E(t_n))v\| \le C(h^\alpha + k^{\alpha/4}) t_n^{-\alpha/4} \|v\|, \quad t_n = kn,\ \alpha \in [0, r], \tag{5.1}$$

where $P_h \colon \mathcal{H} \to \dot{S}_h^r$ denotes the $L_2(\mathcal{D})$-orthogonal projection to $\dot{S}_h^r$. To see this we write

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

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

$$\|E_{h,k}^n P_h v - E_h(t_n) P_h v\| \le C k^\gamma t_n^{-\gamma} \|v\|, \quad \gamma \in [0, 1], \tag{5.2}$$

holds for the backward Euler method [18]. It follows from the stability of the finite element approximation and [10, Corollary 5.3] that

$$\|E_h(t) P_h v - E(t) v\| \le C h^\gamma t^{-\gamma/4} \|v\|, \quad \gamma \in [0, r]. \tag{5.3}$$

Thus, with $\gamma = \alpha/4 \le r/4 \le 1$ in (5.2) and $\gamma = \alpha$ in (5.3), the estimate (5.1) follows.

It is also well known (see, for example, [21, Theorem 6.13]) that

$$\|(E(t) - E(s)) A^{-\gamma} v\| \le |t - s|^\gamma \|v\|, \quad \gamma \in [0, 1], \tag{5.4}$$

and therefore, taking also (5.1) into account, it follows that

$$\|(\tilde{E}_{h,k}(t) P_h - E(t))v\| \le C(h^\alpha + k^{\alpha/4}) t^{-\alpha/4} \|v\|, \quad \alpha \in [0, r]. \tag{5.5}$$

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

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

For the first term (5.1) applies and for the second term we use (5.4):

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

Finally, we recall the smoothing property of the backward Euler scheme. It follows from [23, Lemma 7.3] by stability and interpolation that for $t \in (t_{j-1}, t_j]$,

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

Therefore,

$$\|A_h^\alpha \tilde{E}_{h,k}(t) P_h v\| \le C t^{-\alpha} \|v\|, \quad \alpha \ge 0,\ t > 0. \tag{5.6}$$

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 [16] (up to a logarithmic factor) under essentially the same regularity requirements on A and Q.

Theorem 5.1. *Let X be the solution of* (1.1) *and $X^n_{h,k}$ be given by* (4.23) *with spaces and operators described above and $B_h = P_h$. Assume $G \in C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})$, $X_0 \in L_1(\Omega,\mathcal{H})$ and*

$$\|A^{(\beta-2)/2}Q\|_{\mathrm{Tr}} \le K, \quad \|A_h^{(\beta-2)/2}P_hQ\|_{\mathrm{Tr}} \le K, \tag{5.7}$$

for some $\beta \in (0,\frac{r}{2}]$ and $K>0$. Then there is C depending on T, K, $\|X_0\|_{L_1(\Omega,\mathcal{H})}$, and $\|G\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})}$ such that, for $Nk=T$, $h^4+k<T$,

$$\big|\mathbf{E}(G(X^N_{h,k}) - G(X(T)))\big| \le C\big(h^{2\beta}+k^{\beta/2}\big)\log(\tfrac{T}{h^4+k}). \tag{5.8}$$

Proof. Assumption (5.7) guarantees that $\|A^{\frac{\beta-2}{4}}Q^{\frac12}\|_{\mathrm{HS}} < \infty$ in view of Theorem 2.1. This in its turn implies that X exists, as shown in [16]. Will use Theorem 3.1 with $\mathcal{O}$ as in (3.7). Furthermore, let $\tilde B = P_h$, $\tilde X_0 = P_hX_0$, $\tilde E(t) = \tilde E_{h,k}(t)$ and $\tilde Y_{h,k}(t)$ be defined as in (4.24), whence $\tilde Y_{h,k}(T) = X^N_{h,k}$. The use of Theorem (3.1) is justified since $\tilde E_{h,k}(t)P_hQ[\tilde E_{h,k}(t)P_h]^*$ is bounded and of finite rank so that (1.11) holds.

We write $\tilde F_{h,k}(t) = \tilde E_{h,k}(t)P_h - E(t)$ and recall (5.5). For the first term in (3.6) we use that $\tilde Y_{h,k}(0) - Y(0) = \tilde F_{h,k}(T)$, (5.5), and the bound for u_x in (3.4) to get

$$\begin{aligned}
&\Big|\mathbf{E}\int_0^1 \Big\langle u_x\big(Y(0)+s(\tilde Y_{h,k}(0)-Y(0)),0\big), \tilde Y_{h,k}(0)-Y(0)\Big\rangle\,\mathrm{d}s\Big| \\
&\quad\le \sup_{x\in\mathcal{H}}\|u_x(x,0)\|\,\mathbf{E}\big(\|\tilde F_{h,k}(T)X_0\|\big) \\
&\quad\le \|G\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})}C(h^{2\beta}+k^{\beta/2})T^{-\beta/2}\mathbf{E}\big(\|X_0\|_{\mathcal{H}}\big).
\end{aligned}\tag{5.9}$$

For the second term of (3.6) we have by (2.3) and repeated use of (2.4) that

$$\begin{aligned}
&\Big|\mathbf{E}\int_0^T \mathrm{Tr}\,\Big(u_{xx}\big(\tilde Y_{h,k}(t),t\big)(\tilde E_{h,k}(T-t)P_h + E(T-t))Q\tilde F_{h,k}(T-t)^*\Big)\,\mathrm{d}t\Big| \\
&\le \mathbf{E}\Big(\int_0^T \|u_{xx}(\tilde Y_{h,k}(t),t)(\tilde E_{h,k}(T-t)P_h + E(T-t))Q\|_{\mathrm{Tr}}\|\tilde F_{h,k}(T-t)\|_{\mathcal{B}(\mathcal{H})}\,\mathrm{d}t\Big) \\
&\le \sup_{(x,t)\in\mathcal{H}\times[0,T]}\|u_{xx}(x,t)\|_{\mathcal{B}(\mathcal{H})} \\
&\quad\times\int_0^T \|(A_h^{-(\beta-2)/2}\tilde E_{h,k}(t)A_h^{(\beta-2)/2}P_h + A^{-(\beta-2)/2}E(t)A^{(\beta-2)/2})Q\|_{\mathrm{Tr}} \\
&\qquad\times\|\tilde F_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})}\,\mathrm{d}t \\
&\le \|G\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})}\int_0^T \Big(\|A_h^{-(\beta-2)/2}\tilde E_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})}\,\|A_h^{(\beta-2)/2}P_hQ\|_{\mathrm{Tr}} \\
&\quad+\|A^{-(\beta-2)/2}E(t)\|_{\mathcal{B}(\mathcal{H})}\,\|A^{(\beta-2)/2}Q\|_{\mathrm{Tr}}\Big)\,\|\tilde F_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})}\,\mathrm{d}t \\
&\le \|G\|_{C^2_{\mathrm{b}}(\mathcal{H},\mathbb{R})}\Big(\|A_h^{(\beta-2)/2}P_hQ\|_{\mathrm{Tr}} + \|A^{(\beta-2)/2}Q\|_{\mathrm{Tr}}\Big) \\
&\quad\times\int_0^T \Big(\|A_h^{1-\beta/2}\tilde E_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})} + \|A^{1-\beta/2}E(t)\|_{\mathcal{B}(\mathcal{H})}\Big)\|\tilde F_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})}\,\mathrm{d}t.
\end{aligned}$$

By (5.7) the factors in front of the integral are bounded by $2K\|G\|_{C_{\mathrm{b}}^2(\mathcal{H},\mathbb{R})}$.

We proceed by splitting the integral in two as $\int_0^T = \int_0^{h^4+k} + \int_{h^4+k}^T$. 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 (5.6) with $\alpha = -(\beta-2)/2$,

$$\begin{aligned}\int_0^{h^4+k} \Big(\|A_h^{1-\beta/2}\tilde{E}_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})} + \|A^{1-\beta/2}E(t)\|_{\mathcal{B}(\mathcal{H})}\Big)\|\tilde{F}_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})}\,\mathrm{d}t \\ \le C\int_0^{h^4+k} t^{-1+\beta/2}\,\mathrm{d}t = C(h^4+k)^{\beta/2} \le C(h^{2\beta}+k^{\beta/2}).\end{aligned}$$

For the second part we use again the analyticity of $E(t)$, (5.5) with $\alpha = 2\beta$ and (5.6) to get

$$\begin{aligned}&\int_{h^4+k}^T \Big(\|A_h^{1-\beta/2}\tilde{E}_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})} + \|A^{1-\beta/2}E(t)\|_{\mathcal{B}(\mathcal{H})}\Big)\|\tilde{F}_{h,k}(t)\|_{\mathcal{B}(\mathcal{H})}\,\mathrm{d}t \\ &\quad\le C\int_{h^4+k}^T \big(t^{-1+\beta/2}+t^{-1+\beta/2}\big)\big(h^{2\beta}+k^{\beta/2}\big)t^{-\beta/2}\,\mathrm{d}t \\ &\quad= C(h^{2\beta}+k^{\beta/2})\int_{h^4+k}^T t^{-1}\,\mathrm{d}t = C\log\big(\tfrac{T}{h^4+k}\big)(h^{2\beta}+k^{\beta/2}).\end{aligned}$$

□

Remark 5.2. We refer to [14, Theorem 4.4] for h-independent conditions guaranteeing (5.7). Furthermore, the dependence on T of C in (5.8) can be removed if we assume that $X_0 \in L_1(\Omega, \dot{H}^{2\beta})$ by using the deterministic error estimate for smooth initial data from [16] in (5.9).

Remark 5.3. The weak convergence of the finite element space discretization and backward Euler time discretization of stochastic heat equation with additive noise was considered in [9]. The results there can be recovered using the fully discrete deterministic estimates

$$\|E_{h,k}^n P_h - E(t_n)\|_{\mathcal{B}(\mathcal{H})} \le C(h^2+k)t_n^{-1}$$

and

$$\|\Lambda^\alpha E_{h,k}^n P_h\|_{\mathcal{B}(\mathcal{H})} + \|\Lambda^\alpha E(t_n)\|_{\mathcal{B}(\mathcal{H})} \le Ct_n^{-\alpha}, \quad \alpha \in [0, \tfrac{1}{2}],$$

together with Theorem 3.1. The technicalities are the same as in the spatially semidiscrete case [14, Theorem 4.1] under the same symmetric condition

$$\|\Lambda^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty.$$

We do not detail this here any further as it recovers a known result, only with perhaps a more transparent proof.

References

[1] A. Andersson and S. Larsson, *Weak error for the approximation in space of the non-linear stochastic heat equation*, preprint 2012.

[2] 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.

[3] P. Brenner and V. Thomée, *On rational approximations of groups of operators*, SIAM J. Numer. Anal. **17** (1980), 119–125.

[4] Z. Brzeźniak, *Some remarks on Itô and Stratonovich integration in 2-smooth Banach spaces*, Probabilistic methods in fluids, 48–69, World Sci. Publ., River Edge, NJ, 2003.

[5] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press, 1992.
[6] ———, Second Order Partial Differential Equations in Hilbert Spaces, London Mathematical Society Lecture Note Series **293**, Cambridge University Press, Cambridge, 2002.
[7] 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** (2006), 369–399.
[8] A. Debussche, *Weak approximation of stochastic partial differential equations: the non linear case*, Math. Comp. **80** (2011), 89–117.
[9] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comp. **78** (2009), 845–863.
[10] 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.
[11] E. Hausenblas, *Weak approximation for semilinear stochastic evolution equations*, Stochastic Analysis and Related Topics VIII, Progr. Probab. **53**, 111–128, Birkhäuser, Basel, 2003.
[12] E. Hausenblas, *Weak approximation of the stochastic wave equation*, J. Comput. Appl. Math. **235** (2010), 33–58.
[13] 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** (2009), 343–356.
[14] M. Kovács, S. Larsson and F. Lindgren, *Weak convergence of finite element approximations of stochastic evolution equations with additive noise*, BIT **52** (2012), 85–108.
[15] 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.
[16] S. Larsson and A. Mesforush, *Finite element approximation of the linear stochastic Cahn-Hilliard equation*, IMA J. Numer. Anal. **31** (2011), 1315–1333.
[17] P. D. Lax, Functional Analysis, Wiley, 2002.
[18] M. N. Le Roux, *Semidiscretization in time for parabolic problems*, Math. Comp. **33** (1979), 919–931.
[19] F. Lindner and R. L. Schilling, *Weak order for the discretization of the stochastic heat equation driven by impulsive noise*, Preprint 2010, `arXiv:0911.4681v2`.
[20] S. P. Nørsett and G. Wanner, *The real-pole sandwich for rational approximations and oscillation equations*, BIT **19** (1979), 79–94.
[21] A. Pazy, Semigroups of Linear Operators and Applications to Partial Differential Equations, Applied Mathematical Sciences **44**, Springer, 1983.
[22] L. Quer-Sardanyons and M. Sanz-Solé, *Space semi-discretisations for a stochastic wave equation*, Potential Anal. **24** (2006), 303–332.
[23] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, 2nd ed., Springer Series in Computational Mathematics **25**, Springer, 2006.
[24] J. B. Walsh, *On numerical solutions of the stochastic wave equation*, Illinois J. Math. **50** (2006), 991–1018
[25] J. Weidmann, Linear Operators in Hilbert Spaces, Graduate Texts in Mathematics, Springer, 1980.

Department of Mathematics and Statistics, University of Otago, P.O. Box 56, Dunedin, New Zealand
E-mail address: `mkovacs@maths.otago.ac.nz`

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `stig@chalmers.se`

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `fredrik.lindgren@chalmers.se`