WEAK CONVERGENCE OF FINITE ELEMENT APPROXIMATIONS OF LINEAR STOCHASTIC EVOLUTION EQUATIONS WITH ADDITIVE NOISE

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

Abstract. A unified approach is given for the analysis of the weak error of spatially semidiscrete finite element methods for linear stochastic partial differential equations driven by additive noise. An error representation formula is found in an abstract setting based on the semigroup formulation of stochastic evolution equations. This is then applied to the stochastic heat, linearized Cahn-Hilliard, and wave equations. In all cases it is found that the rate of weak convergence is twice the rate of strong convergence, sometimes up to a logarithmic factor, under the same or, essentially the same, regularity requirements.

1. Introduction

Let U, 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 H, $B \in \mathcal{B}(U, H)$, where $\mathcal{B}(U, H)$ denotes the space of bounded linear operators from U to H, $\{W(t)\}_{t\geq 0}$ is a U-valued Wiener process with covariance operator Q with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, P)$, and X_0 is an $\mathcal{F}_0$-measurable H-valued random variable. The covariance operator $Q \in \mathcal{B}(U, U)$ with $Q \geq 0$ (selfadjoint, positive semidefinite). Under appropriate conditions, see (3.5) below, the unique weak solution is given by

$$X(t) = E(t)X_0 + \int_0^t E(t-s)B\,\mathrm{d}W(s). \tag{1.2}$$

Let $V_h \subset H$ be a family finite dimensional subspaces and let $B_h \in \mathcal{B}(U, H)$ be a family of operators with $B_h : U \to V_h$, $0 < h \leq 1$. We consider approximating stochastic Cauchy problems of the form

$$\mathrm{d}X_h(t) + A_hX_h(t)\,\mathrm{d}t = B_h\,\mathrm{d}W(t),\ t > 0; \quad X_h(0) = X_{h0}, \tag{1.3}$$

Date: September 6, 2010.

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.

where $-A_h$ is the generator of a strongly continuous semigroup $\{E_h(t)\}_{t\geq 0}$ on V_h. We take X_{h0} to be an $\mathcal{F}_0$-measurable random variable. As above the unique weak solution is given by

$$X_h(t) = E_h(t)X_{h0} + \int_0^t E_h(t-s)B_h \,\mathrm{d}W(s). \tag{1.4}$$

This framework is designed to accomodate standard spatial finite element discretizations of various linear stochastic evolution problems including the heat equation, the linearized Cahn-Hilliard equation, and the wave equation. For further details on stochastic integration and the semigroup approach to stochastic partial differential equations we refer to [1].

Let $G : H \to \mathbb{R}$ be a function with globally bounded, continuous Fréchet derivatives of order 1 and 2, that is, $G \in C_{\mathrm{b}}^2(H, \mathbb{R})$. We consider the weak error $e_h(T)$ at $T > 0$ defined as

$$e_h(T) := \mathbf{E}\big(G(X_h(T))\big) - \mathbf{E}\big(G(X(T))\big). \tag{1.5}$$

While the literature on strong convergence of numerical approximations of stochastic partial differential equations is abundant, especially for parabolic problems (see [5] for an exhaustive list of references), there is very little on weak convergence. In particular, there are no results on the weak error of the finite element method for the linear stochastic Cahn-Hilliard and wave equations. The papers [4, 5, 9] consider the stochastic heat equation and so does [7], which proves similar results but under a stronger restriction on the test function G. The results in [3] are concerned with the Schrödinger equation and [8] proves weak convergence of the leap frog scheme for the stochastic wave equation. In all cases it is observed that the rate of weak convergence is twice that of strong convergence.

We now present a brief outline of this paper. Precise definitions and statements are given in the following sections. In Section 2 we recall basic facts about trace class and Hilbert-Schmidt operators. In Section 3 we work in the abstract setting (1.2), (1.4) and derive a formula for the weak error in Theorem 3.1. This is then applied to semidiscretizations of parabolic equations in Section 4 and a hyperbolic equation in Section 5. An important difference is that the semigroup $E(t) = \mathrm{e}^{-tA}$ is analytic in Section 4 but only strongly continuous in Section 5.

Let $\mathcal{D} \subset \mathbb{R}^d$ be a spatial domain and consider the Laplace operator $\Lambda = -\Delta$ as an unbounded operator on $L_2(\mathcal{D})$ with domain of definition $D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$. In Subsection 4.1 we study the stochastic heat equation,

$$\mathrm{d}X + \Lambda X \,\mathrm{d}t = \mathrm{d}W,\ t > 0; \quad X(0) = X_0. \tag{1.6}$$

This is of the form (1.1) with $H = U = L_2(\mathcal{D})$, $A = \Lambda$, $B = I$.

Let $S_h \subset H_0^1(\mathcal{D})$ be a family of standard finite element spaces consisting of continuous piecewise polynomials of degree $\leq r-1$ parametrized by meshsize h. Thus $r \geq 2$ is the formal convergence order of the finite element method. The spatially discrete approximation of (1.6) is

$$\mathrm{d}X_h + \Lambda_h X_h \,\mathrm{d}t = P_h \,\mathrm{d}W,\ t > 0; \quad X_h(0) = P_h X_0. \tag{1.7}$$

Here Λ_h denotes the discrete Laplacian and $P_h : L_2(\mathcal{D}) \to S_h$ is the orthogonal projection. This is clearly of the form (1.3) with $V_h = S_h$, $A_h = \Lambda_h$, $B_h = P_h$, $X_{h0} = P_h X_0$.

In [14] it was assumed that

$$\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty, \quad \text{for some } \beta \geq 0, \tag{1.8}$$

where $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm of bounded linear operators, see (2.1). Under appropriate smoothness of the initial value it was then shown that the solution has regularity of order β in the mean square,

$$\Big(\mathbf{E}\big(\|\Lambda^{\frac{\beta}{2}}X(t)\|^2\big)\Big)^{1/2} < \infty, \tag{1.9}$$

and that the finite element approximation has strong convergence of order β,

$$\Big(\mathbf{E}\big(\|X_h(t)-X(t)\|\big)^2\Big)^{1/2} = O(h^\beta), \quad 0 \leq \beta \leq r. \tag{1.10}$$

Here $\|\cdot\|$ denotes the norm in $H = L_2(\mathcal{D})$.

In the present work we first show in Theorem 4.1 that under the condition (1.8) we have weak convergence of essentially order 2β,

$$e_h(T) = O(h^{2\beta}|\log(h)|), \quad 0 < \beta \leq 1. \tag{1.11}$$

For larger β we assume in Theorem 4.2 that

$$\|\Lambda^{\beta-1}Q\|_{\mathrm{Tr}} < \infty, \tag{1.12}$$

where the trace norm (2.3) is used, and we show weak order $O(h^{2\beta}|\log(h)|)$ for $1 \leq \beta \leq \frac{r}{2}$.

In order to compare (1.12) and (1.8), we show in Theorem 2.1 that

$$\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} \leq \|\Lambda^{\beta-1}Q\|_{\mathrm{Tr}} \leq \|\Lambda^{\beta-1+\alpha}Q\|_{\mathcal{B}(H)}\|\Lambda^{-\alpha}\|_{\mathrm{Tr}}, \quad \beta \geq 0,\ \alpha > 0. \tag{1.13}$$

It is clear that (1.12) implies (1.8) and that they coincide in two important cases: (i) if Λ and Q commute, in particular, if $Q = I$; and (ii) if $\beta = 1$, that is, if $\mathrm{Tr}(Q) < \infty$. Thus, the rate of weak convergence is essentially twice the rate of strong convergence under essentially the same regularity assumption.

A result similar to (1.11) was first proved in [5]. More precisely, there it was assumed that

$$\|\Lambda^{\delta}Q\|_{\mathcal{B}(H)} < \infty,\ \|\Lambda^{-\alpha}\|_{\mathrm{Tr}} < \infty, \quad \text{for some } \alpha > 0,\ \alpha - 1 \leq \delta \leq \alpha. \tag{1.14}$$

In view of (1.13) is is clear that (1.14) implies (1.8) with $\beta = 1 - \alpha + \delta$. Under this assumption was shown in [5] that we have weak convergence of order $O(h^{2\gamma})$ for $0 < \gamma < \beta \leq 1$, which is almost (1.11). Weak convergence of the form (1.11) was also proved in [9] under assumption (1.8) but with stronger restrictions on the test function G and on r.

Hence, for the stochastic heat equation, we slightly sharpen and simplify the results of [5] and [9] and we extend them to higher order.

In Subsection 4.2 study the linearized stochastic Cahn-Hilliard equation (linearized Cahn-Hilliard-Cook equation),

$$\mathrm{d}X + \Lambda^2 X\,\mathrm{d}t = \mathrm{d}W,\ t > 0; \quad X(0) = X_0, \tag{1.15}$$

with finite element approximation

$$\mathrm{d}X_h + \Lambda_h^2 X_h\,\mathrm{d}t = P_h\mathrm{d}W,\ t > 0; \quad X_h(0) = P_h X_0. \tag{1.16}$$

Under the assumption $\|\Lambda^{\frac{\beta-2}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$, it was shown in [11] that we have regularity of order β for $\beta \geq 0$ and strong convergence of order $O(h^\beta|\log(h)|)$ for $1 \leq \beta \leq r$. Here, in Theorem 4.4, if we assume, for example, that S_h is based on a quasi-uniform mesh family and that for some $\alpha > 0$ we have $0 < \beta \le \min(2, \frac{r}{2})$, $0 \le \beta - 2 + \alpha \le 1$ and

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

then the weak convergence is of order $O(h^{2\beta}|\log(h)|)$.

Our most novel result concerns the stochastic wave equation in Section 5,

$$\begin{aligned} &\mathrm{d}X_1 - X_2\,\mathrm{d}t = 0,\ t > 0; && X_1(0) = X_{0,1}, \\ &\mathrm{d}X_2 + \Lambda X_1\,\mathrm{d}t = \mathrm{d}W,\ t > 0; && X_2(0) = X_{0,2}, \end{aligned} \tag{1.17}$$

with its straight-forward finite element approximation based on S_h and Λ_h. This is of the form (1.1) with $H = L_2(\mathcal{D}) \times (H_0^1(\mathcal{D}))^*$, $U = L_2(\mathcal{D})$, and

$$A = \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad X_0 = \begin{bmatrix} X_{0,1} \\ X_{0,2} \end{bmatrix}.$$

Under assumption (1.8) it was shown in [10] for the first component X_1 (the displacement) that we have regularity of order β for $\beta \ge 0$ and strong convergence of order $O(h^{\frac{r}{r+1}\beta})$ for $0 \le \beta \le r+1$. Here, in Theorem 5.1, we assume

$$\|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}} < \infty \tag{1.18}$$

and show weak convergence of order $O(h^{\frac{r}{r+1}2\beta})$ for $0 \le \beta \le \frac{r+1}{2}$. Again, we show in Theorem 2.1 that the new condition (1.18) implies (1.8) and that they coincide if Λ and Q commute.

2. Preliminaries

Let H be a separable real Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and norm $\|\cdot\|$ and let $\mathcal{B}(H)$ denote the space of bounded linear operators on H with the usual norm $\|\cdot\|_{\mathcal{B}(H)}$. An operator $T \in \mathcal{B}(H)$ is called Hilbert-Schmidt, if for some orthonormal basis $\{e_k\}_{k=1}^\infty$ the sum

$$\|T\|_{\mathrm{HS}}^2 := \sum_{k=1}^\infty \|Te_k\|^2 \tag{2.1}$$

is finite. In this case the sum is independent of the choice of the orthonormal basis and the quantity $\|T\|_{\mathrm{HS}}$ is called the Hilbert-Schmidt norm of T. The set of Hilbert-Schmidt operators is denoted by $\mathcal{L}_2(H)$. If $S \in \mathcal{B}(H)$ and $T \in \mathcal{L}_2(H)$, then T^*, TS, and ST belong to $\mathcal{L}_2(H)$ and

$$\|T^*\|_{\mathrm{HS}} = \|T\|_{\mathrm{HS}},\ \|TS\|_{\mathrm{HS}} \le \|T\|_{\mathrm{HS}} \|S\|_{\mathcal{B}(H)},\ \|ST\|_{\mathrm{HS}} \le \|T\|_{\mathrm{HS}} \|S\|_{\mathcal{B}(H)}. \tag{2.2}$$

Let $\mathcal{L}_1(H)$ denote the set of nuclear operators from H to H, that is, $T \in \mathcal{L}_1(H)$ if $T \in \mathcal{B}(H)$ and there are sequences $\{a_j\}, \{b_j\} \subset H$ with $\sum_{j=1}^\infty \|a_j\| \|b_j\| < \infty$ and such that

$$Tx = \sum_{j=1}^\infty \langle x, b_j \rangle a_j, \quad x \in H.$$

Sometimes these operators are referred to as trace class operators. It is well known that $\mathcal{L}_1(H)$ becomes a Banach space under the norm

$$\|T\|_{\mathrm{Tr}} = \inf\Big\{ \sum_{j=1}^\infty \|a_j\| \|b_j\| : Tx = \sum_{j=1}^\infty \langle x, b_j \rangle a_j \Big\}. \tag{2.3}$$

If $T \in \mathcal{L}_1(H)$, then for any orthonormal basis $\{e_k\}_{k=1}^\infty \subset H$ the trace of T, defined as

$$\operatorname{Tr}(T) = \sum_{k=1}^{\infty} \langle Te_k, e_k \rangle, \tag{2.4}$$

is finite and the sum is independent of the choice of the orthonormal basis. If $T \geq 0$ (selfadjoint, positive semidefinite) and the sum in (2.4) converges for one particular orthonormal basis, then $T \in \mathcal{L}_1(H)$. We recall the following well known properties of the trace and trace norm which we frequently use, see [1, App. C], [12, Chapt. 30] and [15, Chapt. 7]. If $S \in \mathcal{B}(H)$ and $T \in \mathcal{L}_1(H)$, then both TS and ST belong to $\mathcal{L}_1(H)$ and

$$\operatorname{Tr}(TS) = \operatorname{Tr}(ST), \tag{2.5}$$

$$|\operatorname{Tr}(TS)| = |\operatorname{Tr}(ST)| \leq \|T\|_{\mathrm{Tr}} \|S\|_{\mathcal{B}(H)}, \tag{2.6}$$

$$\|TS\|_{\mathrm{Tr}} \leq \|T\|_{\mathrm{Tr}} \|S\|_{\mathcal{B}(H)}, \quad \|ST\|_{\mathrm{Tr}} \leq \|T\|_{\mathrm{Tr}} \|S\|_{\mathcal{B}(H)}. \tag{2.7}$$

Furthermore, if $T \in \mathcal{L}_1(H)$, then its adjoint $T^* \in \mathcal{L}_1(H)$ and

$$\operatorname{Tr}(T) = \operatorname{Tr}(T^*), \quad \|T\|_{\mathrm{Tr}} = \|T^*\|_{\mathrm{Tr}}. \tag{2.8}$$

If both $T, S \in \mathcal{L}_2(H)$, then $TS \in \mathcal{L}_1(H)$ and

$$\|TS\|_{\mathrm{Tr}} \leq \|T\|_{\mathrm{HS}} \, \|S\|_{\mathrm{HS}}. \tag{2.9}$$

The following theorem compares the conditions (1.8), (1.12), (1.14), and (1.18) on the covariance operator Q. Since $\|T\|_{\mathrm{HS}}^2 = \operatorname{Tr}(T^*T) = \|T^*T\|_{\mathrm{Tr}}$, we note that (1.8) is expressed as the trace of a symmetric, positive semidefinite operator:

$$\|\Lambda^{\frac{\beta-1}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 = \operatorname{Tr}([\Lambda^{\frac{\beta-1}{2}} Q^{\frac12}]^* \Lambda^{\frac{\beta-1}{2}} Q^{\frac12}) = \|[\Lambda^{\frac{\beta-1}{2}} Q^{\frac12}]^* \Lambda^{\frac{\beta-1}{2}} Q^{\frac12}\|_{\mathrm{Tr}},$$

while (1.12) and (1.18) involve the trace norm of a nonsymmetric operator.

Theorem 2.1. *Assume that $Q \in \mathcal{B}(H)$ 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 $s \in \mathbb{R}$, $\alpha > 0$,*

$$\|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 \leq \|A^s Q\|_{\mathrm{Tr}} \leq \|A^{s+\alpha} Q\|_{\mathcal{B}(H)} \| \|A^{-\alpha}\|_{\mathrm{Tr}}, \tag{2.10}$$

$$\|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 \leq \|A^{s+\frac12} Q A^{-\frac12}\|_{\mathrm{Tr}}, \tag{2.11}$$

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

$$\|A^{\frac{s}{2}} Q^{\frac12}\|_{\mathrm{HS}}^2 = \|A^s Q\|_{\mathrm{Tr}} = \|A^{s+\frac12} Q A^{-\frac12}\|_{\mathrm{Tr}}. \tag{2.12}$$

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 $[A^{\frac{s}{2}}Q^{\frac{1}{2}}]^*$ is not equal to $Q^{\frac{1}{2}}A^{\frac{s}{2}}$ in general, we do have $[A^{\frac{s}{2}}Q^{\frac{1}{2}}]^*\phi_k = Q^{\frac{1}{2}}A^{\frac{s}{2}}\phi_k$, and we compute using (2.2), (2.1), (2.4), (2.6), and (2.7),

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

This is (2.10). Similarly, (2.11) 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 (2.12) 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 (2.12) in view of (2.10). The second equality in (2.12) can be shown in a similar fashion. □

Finally, we define $C_{\mathrm{b}}^2(H,\mathbb{R})$ to be the set of all real-valued, twice Fréchet differentiable functions G, whose first and second derivatives are continuous and bounded. By the Riesz representation theorem, we may identify the first derivative $DG(x)$ at $x \in H$ with an element $G'(x) \in H$ such that

$$DG(x)y = \langle G'(x), y\rangle, \quad y \in H,$$

and the second derivative $D^2G(x)$ with a selfadjoint linear operator $G''(x) \in \mathcal{B}(H)$ such that

$$D^2G(x)(y,z) = \langle G''(x)y, z\rangle, \quad y, z \in H.$$

We say that $G \in C^2(H,\mathbb{R})$ if G, G', and G'' are continuous, that is, $G \in C(H,\mathbb{R})$, $G' \in C(H,H)$, and $G'' \in C(H,\mathcal{B}(H))$. Thus, we define

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

with the seminorm

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

Note that we do not assume that the function G itself is bounded.

3. Error Representation

In this section we derive a representation of the weak error in the general framework. In the following sections we use this to obtain the weak convergence order for finite element approximations of various equations.

If condition (3.5) below holds, then there is a unique weak solution of

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

which is given by

$$Y(t) = E(T)X_0 + \int_0^t E(T-s)B\,\mathrm{d}W(s), \quad t \in [0,T].$$

Notice that $X(T) = Y(T)$, where X is given by (1.2). Similarly, we define

$$Y_h(t) = E_h(T)X_{h0} + \int_0^t E_h(T-s)B_h\,\mathrm{d}W(s), \quad t \in [0,T],$$

and note that $X_h(T) = Y_h(T)$, where X_h is given by (1.4). We also consider the auxiliary problem

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

where ξ is an $\mathcal{F}_\tau$-measurable random variable. Its unique weak solution is given by

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

For $G \in C_{\mathrm{b}}^2(H,\mathbb{R})$, we define a function $u : H \times [0,T] \to \mathbb{R}$ by

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

It follows from (3.1) that its partial derivatives are given by

$$u_x(x,t) = \mathbf{E}\big(G'(Z(T,t,x))\big), \tag{3.2}$$

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

It is known (see, for example, [2, Chapters 3 and 6]) that u is a solution to Kolmogorov's equation

(3.4)
$$\begin{aligned} &u_t(x,t) + \frac{1}{2}\operatorname{Tr}\big(u_{xx}(x,t)E(T-t)BQB^*E(T-t)^*\big) = 0, \quad (x,t) \in H \times [0,T),\\ &u(x,T) = G(x), \quad x \in H. \end{aligned}$$

We are now ready to prove a representation formula for the weak error.

Theorem 3.1. *If*

$$\operatorname{Tr}\Big(\int_0^T E(t)BQB^*E(t)^*\,\mathrm{d}t\Big) < \infty \tag{3.5}$$

and $G \in C_{\mathrm{b}}^2(H,\mathbb{R})$, then the weak error $e_h(T)$ in (1.5) has the representation

$$
\begin{aligned}
e_h(T) = {} & \mathbf{E}\big(u(Y_h(0),0) - u(Y(0),0)\big) \\
& + \frac{1}{2}\mathbf{E}\int_0^T \operatorname{Tr}\Big(u_{xx}(Y_h(t),t) \\
& \quad \times \big[E_h(T-t)B_h + E(T-t)B\big]Q\big[E_h(T-t)B_h - E(T-t)B\big]^*\Big)\,\mathrm{d}t \qquad (3.6)\\
= {} & \mathbf{E}\big(u(Y_h(0),0) - u(Y(0),0)\big) \\
& + \frac{1}{2}\mathbf{E}\int_0^T \operatorname{Tr}\Big(u_{xx}(Y_h(t),t) \\
& \quad \times \big[E_h(T-t)B_h - E(T-t)B\big]Q\big[E_h(T-t)B_h + E(T-t)B\big]^*\Big)\,\mathrm{d}t. \qquad (3.7)
\end{aligned}
$$

Proof. Condition (3.5) guarantees that the stochastic convolution in (1.2) exists. Since $E_h(t)B_h$ acts in a finite-dimensional space, a condition analogous to (3.5) holds and hence (1.4) exists. As in [1, Theorem 9.8], if ξ is $\mathcal{F}_t$-measurable, then

$$
u(\xi,t) = \mathbf{E}\Big(G(Z(T,t,\xi))\Big|\mathcal{F}_t\Big). \qquad (3.8)
$$

Therefore, by the law of double expectation,

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

Thus, with $\xi = Y(0) = E(T)X_0$, and since $Y(T) = X(T)$,

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

and, with $\xi = Y_h(T)$,

$$
\begin{aligned}
\mathbf{E}\Big(u(Y_h(T),T)\Big) &= \mathbf{E}\Big(G(Z(T,T,Y_h(T)))\Big) \\
&= \mathbf{E}\Big(G(Y_h(T))\Big) = \mathbf{E}\Big(G(X_h(T))\Big).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
e_h(T) &= \mathbf{E}\Big(G(X_h(T)) - G(X(T))\Big) = \mathbf{E}\Big(u(Y_h(T),T) - u(Y(0),0)\Big) \\
&= \mathbf{E}\Big(u(Y_h(0),0) - u(Y(0),0)\Big) + \mathbf{E}\Big(u(Y_h(T),T) - u(Y_h(0),0)\Big).
\end{aligned} \qquad (3.9)
$$

Using Itô's formula for $u(Y_h(t),t)$ and Kolmogorov's equation (3.4) we get

$$
\begin{aligned}
&\mathbf{E}\Big(u(Y_h(T),T) - u(Y_h(0),0)\Big) \\
&\quad = \mathbf{E}\int_0^T \Big\{u_t(Y_h(t),t) \\
&\qquad + \frac{1}{2}\operatorname{Tr}\Big(u_{xx}(Y_h(t),t)[E_h(T-t)B_h]Q[E_h(T-t)B_h]^*\Big)\Big\}\,\mathrm{d}t \\
&\quad = \frac{1}{2}\mathbf{E}\int_0^T \operatorname{Tr}\Big(u_{xx}(Y_h(t),t) \\
&\qquad \times \big\{[E_h(T-t)B_h]Q[E_h(T-t)B_h]^* - E(T-t)BQB^*E(T-t)^*\big\}\Big)\,\mathrm{d}t.
\end{aligned} \qquad (3.10)
$$

(Note that $B_h \in \mathcal{B}(U,H)$ with $B_h : U \to V_h$, $E_h(s) : V_h \to V_h$, and that we consider $E_h(s)B_h$ as an operator in $\mathcal{B}(U,H)$. Since $E_h(s)$ acts only on V_h the corresponding adjoint $[E_h(s)B_h]^*$ is not equal to $B_h^* E_h(s)^*$.) Now consider the identity

$$\begin{aligned} &u_{xx}(\xi,r)\big\{[E_h(s)B_h]Q[E_h(s)B_h]^* - E(s)BQB^*E(s)^*\big\} \\ &\quad = u_{xx}(\xi,r)[E_h(s)B_h - E(s)B]Q[E_h(s)B_h]^* \\ &\qquad + u_{xx}(\xi,r)E(s)BQ[E_h(s)B_h - E(s)B]^* =: S_1 + S_2. \end{aligned}$$

The first term has finite trace since, for example, $E_h(s)B_h$ has finite trace and so has the second term, since $E(s)BQB^*E(s)^*$ has finite trace for almost every s by (3.5). Therefore, using (2.8), (2.5), and that Q, $u_{xx}(\xi,r)$ are selfadjoint, we get

$$\begin{aligned} \mathrm{Tr}(S_1+S_2) &= \mathrm{Tr}(S_1) + \mathrm{Tr}(S_2) = \mathrm{Tr}(S_1) + \mathrm{Tr}(S_2^*) \\ &= \mathrm{Tr}(S_1) + \mathrm{Tr}([E_h(s)B_h - E(s)B]QB^*E(s)^*u_{xx}(\xi,r)) \\ &= \mathrm{Tr}(S_1) + \mathrm{Tr}(u_{xx}(\xi,r)[E_h(s)B_h - E(s)B]QB^*E(s)^*) \\ &= \mathrm{Tr}\Big(u_{xx}(\xi,r)[E_h(s)B_h - E(s)B]Q[E_h(s)B_h + E(s)B]^*\Big) \qquad (3.11) \\ &= \mathrm{Tr}\Big([E_h(s)B_h + E(s)B]Q[E_h(s)B_h - E(s)B]^*u_{xx}(\xi,r)\Big) \\ &= \mathrm{Tr}\Big(u_{xx}(\xi,r)[E_h(s)B_h + E(s)B]Q[E_h(s)B_h - E(s)B]^*\Big). \qquad (3.12) \end{aligned}$$

The proof is completed by inserting (3.11) or (3.12) into (3.10) and using (3.9). □

4. Application to Parabolic Equations

In this section we apply the error representation in Section 3 to finite element approximations of the linear stochastic heat and Cahn-Hilliard equations.

4.1. **The stochastic heat equation.** Let $\mathcal{D} \subset \mathbb{R}^d$ be a bounded domain, let $\Lambda := -\Delta$, where $\Delta = \sum_{k=1}^d \partial^2/\partial\xi_k^2$ is the Laplace operator, and set $D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$. Let $U = H := L_2(\mathcal{D})$ with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$, $B := I$ and $A := \Lambda$. Then (1.1) takes the form of the stochastic heat equation (1.6). In order to quantify spatial regularity we introduce the following spaces and norms. Let

$$\dot H^\alpha := D(\Lambda^{\alpha/2}), \quad |v|_\alpha := \|\Lambda^{\alpha/2}v\| = \Big(\sum_{j=1}^\infty \lambda_j^\alpha \langle v,\phi_j\rangle^2\Big)^{1/2}, \quad \alpha\in\mathbb{R},$$

where $\{(\lambda_j,\phi_j)\}_{j=1}^\infty$ are the eigenpairs of Λ with orthonormal eigenvectors. Then $\dot H^\alpha \subset \dot H^\beta$ for $\alpha \ge \beta$. It is known that $\dot H^0 = L_2(\mathcal{D})$, $\dot H^1 = H_0^1(\mathcal{D})$, $\dot H^2 = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ with equivalent norms and that $\dot H^{-\beta}$ can be identified with the dual space $(\dot H^\beta)^*$ for $\beta > 0$, see [13, Chapt. 3].

Let $\{S_h\}_{h>0}$ be a family of function spaces consisting of continuous piecewise polynomials of degree $\le r-1$ with respect to a family of triangulations of $\mathcal{D}$ and such that $S_h \subset H_0^1(D)$. The parameter h is the maximal mesh size of the triangulation and r may be referred to as the order of the finite element method. Let $P_h : H \to S_h$ denote the orthogonal projection and let $\Lambda_h : S_h \to S_h$ be the "discrete Laplacian" defined by

$$\langle \Lambda_h\psi,\chi\rangle = \langle\nabla\psi,\nabla\chi\rangle, \quad \forall\psi,\chi\in S_h.$$

Our basic assumption on the finite element method is that the Ritz projection $R_h : \dot{H}^1 \to S_h$ defined as

$$\langle \nabla R_h v, \nabla \chi \rangle = \langle \nabla v, \nabla \chi \rangle, \quad \forall v \in \dot{H}^1,\ \chi \in S_h, \tag{4.1}$$

satisfies the error bound

$$\|R_h v - v\| \le C h^\beta |v|_\beta, \quad v \in \dot{H}^\beta,\ 1 \le \beta \le r. \tag{4.2}$$

This holds, for example, with $r = 2$ if $\mathcal{D}$ is a convex polygonal domain and S_h consists of piecewise linear functions. See [13] for further details.

If we set $V_h := S_h$, $B_h := P_h$, $A_h := \Lambda_h$, and $X_{h0} := P_h X_0$, then (1.4) takes the form of the semidiscrete finite element approximation (1.7). We have the following result for the weak error.

Theorem 4.1. *Let X and X_h be the solutions of* (1.6) *and* (1.7)*, respectively. Let $g \in C^2_{\mathrm{b}}(H, \mathbb{R})$ and assume that $\|A^{\frac{\beta-1}{2}} Q^{\frac12}\|_{\mathrm{HS}} = \|\Lambda^{\frac{\beta-1}{2}} Q^{\frac12}\|_{\mathrm{HS}} < \infty$ for some $\beta \in (0,1]$. Then there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, β, and T but not on h, such that for $h \le h_0$,*

$$\big|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\big| \le C h^{2\beta} |\log(h)|.$$

If, in addition $X_0 \in L_1(\Omega, \dot{H}^{2\beta})$, then C is independent of T as well.

Proof. If $\|A^{\frac{\beta-1}{2}} Q^{\frac12}\|_{\mathrm{HS}} < \infty$ for some $\beta \in (0,1]$, then (3.5) holds, see [14]. This guarantees that $X(t)$ and $X_h(t)$ are defined. Let $F_h(t) := E_h(t)P_h - E(t)$ be the deterministic error operator with $h \le h_0$ small enough. We recall the error estimates, see [13, Chapt. 3],

$$\|F_h(t) v\| \le C h^s t^{-\frac{s-\gamma}{2}} |v|_\gamma, \quad 0 \le \gamma \le s \le r. \tag{4.3}$$

We use Theorem 3.1 to estimate the weak error with $G := g$. First, by the chain rule, and $Y_h(0) - Y(0) = E_h(T)P_h X_0 - E(T)X_0 = F_h(T)X_0$,

$$\begin{aligned}
&\mathbf{E}\left(u(Y_h(0),0) - u(Y(0),0)\right) \\
&\quad = \mathbf{E}\int_0^1 \langle u_x(Y(0) + s(Y_h(0) - Y(0)), 0), Y_h(0) - Y(0)\rangle \,\mathrm{d}s \\
&\quad = \mathbf{E}\int_0^1 \langle u_x(E(T)X_0 + sF_h(T)X_0, 0), F_h(T)X_0\rangle \,\mathrm{d}s.
\end{aligned}$$

Thus, using (3.2) and (4.3), we obtain

$$\begin{aligned}
|\mathbf{E}\left(u(Y_h(0),0) - u(Y(0),0)\right)| &\le \sup_{x\in H} \|u_x(x,0)\|\, \mathbf{E}\big(\|F_h(T)X_0\|\big) \\
&\le C h^{2\beta} T^{-\frac{2\beta-\gamma}{2}} \mathbf{E}\big(|X_0|_\gamma\big) \sup_{x\in H} \|g'(x)\|, \quad 0 \le \gamma \le 2\beta.
\end{aligned}$$

If $\gamma = 2\beta$ there is no dependence on T. Next, we estimate the second term (3.6) in the error representation in Theorem 3.1. Since $E(t)$, $E_h(t)P_h$, and hence also

$F_h(t)$, are selfadjoint, we obtain by means of (2.6), (2.7), (2.8), and (2.9),

$$\begin{aligned}
&\Big|\mathbf{E}\int_0^T \mathrm{Tr}\,\Big(u_{xx}(Y_h(t),t) \\
&\qquad \times [E_h(T-t)B_h + E(T-t)B]Q[E_h(T-t)B_h - E(T-t)B]^*\Big)\,\mathrm{d}t\Big| \\
&= \Big|\mathbf{E}\int_0^T \mathrm{Tr}\,\Big(u_{xx}(Y_h(t),t)[E_h(T-t)P_h + E(T-t)]^* \\
&\qquad \times A^{\frac{1-\beta}{2}}A^{\frac{\beta-1}{2}}Q^{\frac12}Q^{\frac12}A^{\frac{\beta-1}{2}}A^{\frac{1-\beta}{2}}F_h(T-t)\Big)\,\mathrm{d}t\Big| \\
&= \Big|\mathbf{E}\int_0^T \mathrm{Tr}\,\Big(u_{xx}(Y_h(t),t)(A^{\frac{1-\beta}{2}}[E_h(T-t)P_h + E(T-t)])^* \\
&\qquad \times A^{\frac{\beta-1}{2}}Q^{\frac12}Q^{\frac12}A^{\frac{\beta-1}{2}}A^{\frac{1-\beta}{2}}F_h(T-t)\Big)\,\mathrm{d}t\Big| \\
&\le \mathbf{E}\int_0^T \|u_{xx}(Y_h(t),t)(A^{\frac{1-\beta}{2}}[E_h(T-t)P_h + E(T-t)])^*A^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}} \\
&\qquad \times \|Q^{\frac12}A^{\frac{\beta-1}{2}}A^{\frac{1-\beta}{2}}F_h(T-t)\|_{\mathrm{HS}}\,\mathrm{d}t
\end{aligned}$$

$$\le \sup_{(x,t)\in H\times[0,T]}\|u_{xx}(x,t)\|_{\mathcal{B}(H)}\|A^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 \tag{4.4}$$

$$\times \int_0^T \|A^{\frac{1-\beta}{2}}(E_h(t)P_h + E(t))\|_{\mathcal{B}(H)}\|A^{\frac{1-\beta}{2}}F_h(t)\|_{\mathcal{B}(H)}\,\mathrm{d}t. \tag{4.5}$$

Since $\|A^{\frac12}v_h\| = \|\nabla v_h\| = \|A_h^{\frac12}v_h\|$ for $v_h \in S_h$, we conclude $\|A^\delta v_h\| \le \|A_h^\delta v_h\|$ for $v_h \in S_h$, $\delta \in [0,\frac12]$, and using also the analyticity of the semigroups we have

$$\|A^\delta(E_h(t)P_h + E(t))\|_{\mathcal{B}(H)} \le Ce^{-\omega t}t^{-\delta}, \quad \delta\in[0,\tfrac12]. \tag{4.6}$$

To estimate $\|A^{\frac{1-\beta}{2}}F_h(t)\|_{\mathcal{B}(H)}$ we use interpolation. By analyticity, as above,

$$\|A^\delta F_h(t)\|_{\mathcal{B}(H)} \le Ct^{-\delta}, \quad \delta\in[0,\tfrac12]. \tag{4.7}$$

Interpolation between (4.7) with $\delta = \frac12$ and (4.3) with $s = 2$ and $\gamma = 0$ yields

$$\|A^{\frac{1-\beta}{2}}F_h(t)\|_{\mathcal{B}(H)} \le \|F_h(t)\|_{\mathcal{B}(H)}^{\beta}\|A^{\frac12}F_h(t)\|_{\mathcal{B}(H)}^{1-\beta} \le Ch^{2\beta}t^{-\frac{1+\beta}{2}}, \quad \beta\in[0,1].$$

Therefore, for $\beta\in(0,1]$ one may estimate the integral in (4.5) as follows

$$\begin{aligned}
&\int_0^T \|A^{\frac{1-\beta}{2}}(E_h(t)P_h + E(t))\|_{\mathcal{B}(H)}\|A^{\frac{1-\beta}{2}}F_h(t)\|_{\mathcal{B}(H)}\,\mathrm{d}t \\
&= \left(\int_0^{h^2} + \int_{h^2}^T\right)\|A^{\frac{1-\beta}{2}}(E_h(t)P_h + E(t))\|_{\mathcal{B}(H)}\|A^{\frac{1-\beta}{2}}F_h(t)\|_{\mathcal{B}(H)}\,\mathrm{d}t \\
&\le C\int_0^{h^2} t^{-\frac{1-\beta}{2}}t^{-\frac{1-\beta}{2}}\,\mathrm{d}t + C\int_{h^2}^T e^{-\omega t}t^{-\frac{1-\beta}{2}}h^{2\beta}t^{-\frac{1+\beta}{2}}\,\mathrm{d}t \le Ch^{2\beta}|\log(h)|.
\end{aligned}$$

Finally, using (3.3), we obtain

$$\sup_{(x,t)\in H\times[0,T]}\|u_{xx}(x,t)\|_{\mathcal{B}(H)} \le \sup_{x\in H}\|g''(x)\|_{\mathcal{B}(H)}, \tag{4.8}$$

and the proof is complete in view of (4.4). □

By inspection of the above proof we see that the error estimate is

$$\begin{aligned}\big|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\big| &\le Ch^{2\beta}T^{-\frac{2\beta-\gamma}{2}}\mathbf{E}\big(|X_0|_\gamma\big)\sup_{x\in H}\|g'(x)\| \\ &\quad + Ch^{2\beta}|\log(h)|\beta^{-1}\sup_{x\in H}\|g''(x)\|\|A^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2.\end{aligned}$$

Similar remarks can be made about the theorems to follow.

Theorem 4.1 does not allow $\beta > 1$. This is satisfactory if $r = 2$, but for higher order elements, that is, $r > 2$, it is insufficient. Under a slightly stronger condition on A and Q we now extend the result to the case $\beta > 1$.

Theorem 4.2. *Let X and X_h be the solutions of* (1.6) *and* (1.7)*, respectively. Let $g \in C_{\mathrm{b}}^2(H,\mathbb{R})$ and assume that $\|A^{\beta-1}Q\|_{\mathrm{Tr}} = \|\Lambda^{\beta-1}Q\|_{\mathrm{Tr}} < \infty$ for some $\beta \in [1, \frac{r}{2}]$. Then there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, β, and T but not on h, such that for $h \le h_0$,*

$$\big|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\big| \le Ch^{2\beta}|\log(h)|.$$

If, in addition $X_0 \in L_1(\Omega, \dot H^{2\beta})$, then C is independent of T as well.

Proof. If $\|A^{\beta-1}Q\|_{\mathrm{Tr}} < \infty$ for some $\beta \in [1, \frac{r}{2}]$, then (3.5) holds by Theorem 2.1 and [14], so that $X(t)$ and $X_h(t)$ are defined. From (4.3) with $\gamma = 2\beta - 2 \le s = 2\beta$ it follows that

$$\|F_h(t)A^{1-\beta}\|_{\mathcal{B}(H)} \le Ch^{2\beta}t^{-1}, \quad 1 \le \beta \le \frac{r}{2}, \tag{4.9}$$

and, by (4.7) and since $\beta \ge 1$,

$$\|F_h(t)A^{1-\beta}\|_{\mathcal{B}(H)} \le \|F_h(t)\|_{\mathcal{B}(H)}\|A^{1-\beta}\|_{\mathcal{B}(H)} \le C. \tag{4.10}$$

The first term in the error representation in Theorem 3.1 can be estimated the same way as in Theorem 4.1. To bound the second term (3.7) we use (4.6) with $\delta = 0$ and (2.5) to obtain

$$\begin{aligned}&\Big|\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t) \\ &\qquad \times [E_h(T-t)B_h - E(T-t)B]Q[E_h(T-t)B_h + E(T-t)B]^*\Big)\,\mathrm{d}t\Big| \\ &= \Big|\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t) \\ &\qquad \times F_h(t)A^{1-\beta}A^{\beta-1}Q[E_h(T-t)P_h + E(T-t)]^*\Big)\,\mathrm{d}t\Big| \\ &\le C\sup_{(x,t)\in H\times[0,T]}\|u_{xx}(x,t)\|_{\mathcal{B}(H)}\|A^{\beta-1}Q\|_{\mathrm{Tr}}\int_0^T\|F_h(t)A^{\beta-1}\|_{\mathcal{B}(H)}e^{-\omega t}\,\mathrm{d}t.\end{aligned}$$

Using (4.9) and (4.10) we now have

$$\begin{aligned}\int_0^T\|F_h(t)A^{1-\beta}\|_{\mathcal{B}(H)}e^{-\omega t}\,\mathrm{d}t &= \Big(\int_0^{h^{2\beta}} + \int_{h^{2\beta}}^T\Big)\|F_h(t)A^{1-\beta}\|_{\mathcal{B}(H)}e^{-\omega t}\,\mathrm{d}t \\ &\le C\int_0^{h^{2\beta}}\mathrm{d}t + Ch^{2\beta}\int_{h^{2\beta}}^T t^{-1}e^{-\omega t}\,\mathrm{d}t \le Ch^{2\beta}|\log(h)|,\end{aligned}$$

and the proof is complete in view of (4.8). □

In [14] the strong rate of convergence is found to be $O(h^{\beta})$ under the condition $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}<\infty$. Theorem 2.1 shows that $\|\Lambda^{\beta-1}Q\|_{\mathrm{Tr}}<\infty$ provides a sufficient condition for this and the conditions coincide if $\beta=1$ or if Λ and Q commute.

In the special case $Q=I$ a simple calculation using the asymptotics $\lambda_j\sim j^{2/d}$, $j\to\infty$, of the eigenvalues of Λ shows that the spatial dimension d has to be 1 and $\beta<\frac{1}{2}$, which gives a weak order of almost h. If $\mathrm{Tr}(Q)<\infty$, then we may take $\beta=1$ and hence the rate of weak convergence is at least $O(h^2|\log(h)|)$.

4.2. **The linear Cahn-Hilliard-Cook equation.** Let $\mathcal{D}$ be a bounded domain in $\mathbb{R}^d$ for $d\le 3$. Let $H=U$ be the subspace of $L_2(\mathcal{D})$, which is orthogonal to constants with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$, i.e., $H=U=\{v\in L_2:\langle v,1\rangle=0\}$, and let $B=I$. Let $H^s=H^s(\mathcal{D})$ be the usual Sobolev space. We define the linear operator $\Lambda:=-\Delta$ with domain of definition

$$D(\Lambda)=\Big\{v\in H^2\cap H:\frac{\partial v}{\partial n}=0\text{ on }\partial\mathcal{D}\Big\}.$$

Then Λ is a selfadjoint, positive definite, densely defined operator on H. If we set $A:=\Lambda^2$, then $-A$ generates an analytic semigroup on H. We also define $\dot H^s=\mathcal{D}(\Lambda^{\frac{s}{2}})$ with norms $|v|_s=\|\Lambda^{\frac{s}{2}}v\|$ for real s. It is well known that, for integer $s\ge 0$, $\dot H^s$ is a subspace of $H^s\cap H$ characterized by certain boundary conditions and that the norm $|\cdot|_s$ is equivalent to the standard norm $\|\cdot\|_{H^s}$ on $\dot H^s$. In particular, we have $\dot H^1=H^1\cap H$ and the norm $|v|_1=\|\Lambda^{\frac12}v\|=\|\nabla v\|$ is equivalent to $\|v\|_{H^1}$ on $\dot H^1$. With these definitions (1.1) takes the form of the linear Cahn-Hilliard-Cook equation (1.15).

With $S_h\subset H^1$ being a family of finite dimensional subspaces we set $V_h:=\{\chi\in S_h:\langle\chi,1\rangle=0\}$ and define $\Lambda_h:V_h\to V_h$ by

$$\langle\Lambda_h\chi,\eta\rangle=\langle\nabla\chi,\nabla\eta\rangle,\quad \chi,\,\eta\in V_h.$$

Finally, we set $A_h:=\Lambda_h^2$, $B_h:=P_h:H\to V_h$ the orthogonal projection, and set $X_{h0}:=P_hX_0$. Then (1.3) takes the form (1.16).

As for the heat equation, we assume that the Ritz projection $R_h:\dot H^1\to V_h$ defined as in (4.1) satisfies an error estimate of the form (4.2). This holds, for example, with $r=2$ if $\mathcal{D}$ is a convex polygonal domain and with S_h being the standard family of finite element spaces consisting of continuous piecewise linear functions on a regular family of triangulations of $\mathcal{D}$ with maximum mesh size h.

Theorem 4.3. *Let X and X_h be the solutions of* (1.15) *and* (1.16)*, respectively. Let $g\in C_{\mathrm{b}}^2(H,\mathbb{R})$, assume that $0<\beta\le\min(2,\frac{r}{2})$, and, for some K,*

$$\|A^{\frac{\beta-2}{2}}Q\|_{\mathrm{Tr}}=\|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}}\le K,\tag{4.11}$$

$$\|A_h^{\frac{\beta-2}{2}}P_hQ\|_{\mathrm{Tr}}=\|\Lambda_h^{\beta-2}P_hQ\|_{\mathrm{Tr}}\le K,\quad 0<h\le 1.\tag{4.12}$$

Then there are $C>0$, $h_0>0$, depending on g, X_0, K, β, and T but not on h, such that for $h\le h_0$,

$$\big|\mathbf{E}\big(g(X_h(T))-g(X(T))\big)\big|\le Ch^{2\beta}|\log(h)|.$$

If, in addition $X_0\in L_1(\Omega,\dot H^{2\beta})$, then C is independent of T as well.

Proof. If $\|A^{\frac{\beta-2}{2}}Q\|_{\mathrm{Tr}}<\infty$ for some $\beta\ge 0$, then (3.5) holds, see [11], and X and X_h exist. Let $F_h(t):=E_h(t)P_h-E(t)$ be the deterministic error operator with $h\le h_0$ small enough. We recall from [6] the error estimate

$$\|F_h(t)v\|\le Ch^s t^{-\frac{s-\gamma}{4}}|v|_\gamma,\quad 0\le\gamma\le s\le r. \tag{4.13}$$

We use Theorem 3.1 to estimate the weak error. As in the proof of Theorem 4.1 we get, for $0\le\gamma\le 2\beta\le r$,

$$\big|\mathbf{E}\big(u(Y_h(0),0)-u(Y(0),0)\big)\big|\le Ch^{2\beta}T^{-\frac{2\beta-\gamma}{4}}\mathbf{E}\big(|X_0|_\gamma\big)\sup_{x\in H}\|g'(x)\|.$$

To estimate the second term (3.6) in the error representation we proceed as in (4.4), (4.5). By inserting both $A^{\pm\frac{\beta-2}{2}}=\Lambda^{\pm(\beta-2)}$ and $A_h^{\pm\frac{\beta-2}{2}}=\Lambda_h^{\pm(\beta-2)}$, we obtain this time

$$\begin{aligned}
&\Big|\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t)\\
&\qquad\times[E_h(T-t)B_h+E(T-t)B]Q[E_h(T-t)B_h-E(T-t)B]^*\Big)\,\mathrm{d}t\Big|\\
&=\Big|\mathbf{E}\int_0^T \mathrm{Tr}\Big(u_{xx}(Y_h(t),t)[A_h^{\frac{2-\beta}{2}}E_h(T-t)A_h^{\frac{\beta-2}{2}}P_h\\
&\qquad+A^{\frac{2-\beta}{2}}E(T-t)A^{\frac{\beta-2}{2}}]QF_h(T-t)^*\Big)\,\mathrm{d}t\Big|\\
&\le\sup_{(x,t)\in H\times[0,T]}\|u_{xx}(x,t)\|_{\mathcal{B}(H)}\Big(\|A_h^{\frac{\beta-2}{2}}P_hQ\|_{\mathrm{Tr}}+\|A^{\frac{\beta-2}{2}}Q\|_{\mathrm{Tr}}\Big)\\
&\qquad\times\int_0^T\Big(\|A_h^{\frac{2-\beta}{2}}E_h(t)P_h\|_{\mathcal{B}(H)}+\|A^{\frac{2-\beta}{2}}E(t)\|_{\mathcal{B}(H)}\Big)\|F_h(t)\|_{\mathcal{B}(H)}\,\mathrm{d}t.
\end{aligned}$$

In view of (4.8) and (4.11), (4.12) it remains to bound the integral. By analyticity of the semigroups we have

$$\|A_h^\delta(E_h(t)P_h\|_{\mathcal{B}(H)}+\|A^\delta E(t))\|_{\mathcal{B}(H)}\le Ce^{-\omega t}t^{-\delta},\quad\delta\ge 0. \tag{4.14}$$

Therefore, by using (4.14) with $\delta=\frac{2-\beta}{2}\in[0,1)$, that is, $\beta\in(0,2]$, and (4.13) with $s=0$ and $s=2\beta\le r$, $\gamma=0$,

$$\begin{aligned}
&\int_0^T\Big(\|A_h^{\frac{2-\beta}{2}}E_h(t)P_h\|_{\mathcal{B}(H)}+\|A^{\frac{2-\beta}{2}}E(t)\|_{\mathcal{B}(H)}\Big)\|F_h(t)\|_{\mathcal{B}(H)}\,\mathrm{d}t\\
&\quad=\Big(\int_0^{h^4}+\int_{h^4}^T\Big)\|A^{\frac{2-\beta}{2}}[E_h(t)P_h+E(t)]\|_{\mathcal{B}(H)}\|F_h(t)\|_{\mathcal{B}(H)}\,\mathrm{d}t\\
&\quad\le C\int_0^{h^4}t^{-\frac{2-\beta}{2}}\,\mathrm{d}t+C\int_{h^4}^T e^{-\omega t}t^{-\frac{2-\beta}{2}}h^{2\beta}t^{-\frac{2\beta}{4}}\,\mathrm{d}t\le Ch^{2\beta}|\log(h)|.
\end{aligned}$$

This completes the proof. □

In [11] the strong rate of convergence is found to be $O(h^\beta|\log(h)|)$ under the condition $\|A^{\frac{\beta-2}{4}}Q^{\frac12}\|_{\mathrm{HS}}<\infty$. Theorem 2.1 shows that (4.11) provides a sufficient condition for this and that the conditions coincide if A and Q commute or if $\beta=2$, that is, $\mathrm{Tr}(Q)<\infty$.

It remains to identify conditions under which we have (4.12) together with (4.11). This is addressed in the next theorem.

Theorem 4.4. *Let X and X_h be the solutions of* (1.15) *and* (1.16)*, respectively. Let $g \in C_{\mathrm{b}}^2(H,\mathbb{R})$ and make one of the following assumptions.*

(i) *Assume that $Q = I$, $0 < \beta \le \min(2, \frac{r}{2})$, and $\|\Lambda^{\beta-2}\|_{\mathrm{Tr}} < \infty$.*
(ii) *Assume that $\mathrm{Tr}(Q) < \infty$, $r = 4$, and $\beta = 2$.*
(iii) *Assume that $r \ge 3$, $\frac{3}{2} \le \beta \le \min(2, \frac{r}{2})$, and $\|\Lambda^{\beta-2}Q\|_{\mathrm{Tr}} < \infty$.*
(iv) *Assume that S_h is based on a quasi-uniform mesh family and that, for some $\alpha > 0$, we have $0 < \beta \le \min(2, \frac{r}{2})$, $0 \le \beta - 2 + \alpha \le 1$ and*

$$\|\Lambda^{\beta-2+\alpha}Q\|_{\mathcal{B}(H)} < \infty, \quad \|\Lambda^{-\alpha}\|_{\mathrm{Tr}} < \infty. \tag{4.15}$$

Then there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, β, and T but not on h, such that for $h \le h_0$,

$$\big|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\big| \le Ch^{2\beta}|\log(h)|.$$

If, in addition $X_0 \in L_1(\Omega, \dot{H}^{2\beta})$, then C is independent of T as well.

Proof. We must show that (4.11) and (4.12) hold in each of the four cases.

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

$$\|\Lambda_h^{-\alpha}P_h\|_{\mathrm{Tr}} \le \|\Lambda_h^{-\alpha}\|_{\mathrm{Tr}} = \sum_{j=1}^{N_h} \lambda_{h,j}^{-\alpha} \le \sum_{j=1}^{\infty} \lambda_j^{-\alpha} = \|\Lambda^{-\alpha}\|_{\mathrm{Tr}}, \quad \alpha \ge 0. \tag{4.16}$$

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

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

In case (ii) we have

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

For case (iii) we use the fact that

$$\|\Lambda_h^{-\delta}P_h\Lambda^{\delta}\| \le C, \quad 0 \le \delta \le \frac{1}{2}.$$

For $\delta = \frac{1}{2}$ this follows by using $\|\Lambda_h^{\frac{1}{2}}w_h\| = \|\Lambda^{\frac{1}{2}}w_h\|$ for $w_h \in S_h$ in the calculation

$$\begin{aligned}\|\Lambda_h^{-\frac{1}{2}}P_hf\| &= \sup_{v_h\in S_h} \frac{|\langle \Lambda_h^{-\frac{1}{2}}P_hf, v_h\rangle|}{\|v_h\|} = \sup_{v_h\in S_h} \frac{|\langle f, \Lambda_h^{-\frac{1}{2}}v_h\rangle|}{\|v_h\|} = \sup_{w_h\in S_h} \frac{|\langle f, w_h\rangle|}{\|\Lambda_h^{\frac{1}{2}}w_h\|} \\ &= \sup_{w_h\in S_h} \frac{|\langle f, w_h\rangle|}{\|\Lambda^{\frac{1}{2}}w_h\|} \le \sup_{v\in \dot{H}^0} \frac{|\langle f, \Lambda^{-\frac{1}{2}}v\rangle|}{\|v\|} = \|\Lambda^{-\frac{1}{2}}f\|.\end{aligned}$$

The case $\delta = 0$ is obvious and the general case follows by interpolation. Hence, with $\delta = 2 - \beta \in [0, \frac{1}{2}]$, that is, $\frac{3}{2} \le \beta \le 2$, we have

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

Finally, for case (iv) we first note that Theorem 2.1 shows that (4.15) implies (4.11). For quasi-uniform mesh families we have the inverse inequality $\|\nabla v_h\| \le Ch^{-1}\|v_h\|$, $v_h \in S_h$, so that

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

Hence, using also $R_h = \Lambda_h^{-1} P_h \Lambda$ and (4.2), we get

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

We conclude

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

With $\delta = \beta - 2 + \alpha \in [0,1]$ and (4.16) we obtain

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

□

Finally, we comment on two of the cases of the previous theorem.

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

(iv) 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.

5. Application to a hyperbolic equation

In this section we apply the general theory to the stochastic wave equation. As for the heat equation in Subsection 4.1 we use the space $L_2(\mathcal{D})$ with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$ and the Laplace operator $\Lambda = -\Delta$ with $D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$. We introduce, using the notation from Subsection 4.1,

$$H^\alpha := \dot H^\alpha \times \dot H^{\alpha-1}, \quad |||v|||_\alpha^2 := |v_1|_\alpha^2 + |v_2|_{\alpha-1}^2, \quad \alpha \in \mathbb{R},$$

and set $H := H^0 = \dot H^0 \times \dot H^{-1}$ with corresponding norm $|||\cdot||| = |||\cdot|||_0$ and inner product $(\cdot,\cdot)$. We define $U := \dot H^0 = L_2(\mathcal{D})$ and

$$A := \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, \quad B := \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad X := \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad X_0 := \begin{bmatrix} X_{0,1} \\ X_{0,2} \end{bmatrix},$$

with

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

Here Λ is regarded as an operator $\dot H^1 \to \dot H^{-1}$. The operator $-A$ is the generator of a strongly continuous semigroup $E(t) = \mathrm{e}^{-tA}$ on H and

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

where $C(t) = \cos(t\Lambda^{1/2})$ and $S(t) = \sin(t\Lambda^{1/2})$ are the so-called cosine and sine operators. For example, using $\{(\lambda_j, \phi_j)\}_{j=1}^{\infty}$, orthonormal eigenpairs of Λ, we have

$$\Lambda^{-1/2}S(t)v = \Lambda^{-1/2}\sin(t\Lambda^{1/2})v = \sum_{j=1}^{\infty} \lambda_j^{-1/2}\sin(t\lambda_j^{1/2})\langle v, \phi_j\rangle\phi_j.$$

With the above definition, the stochastic wave equation (1.17) can be written in the form of (1.1).

Let $S_h \subset \dot{H}^1$, Λ_h, P_h be as in Subsection 4.1 with the error estimate (4.2). The semidiscrete approximation of (1.17) is

$$\begin{aligned} &\mathrm{d}X_{h,1} - X_{h,2}\,\mathrm{d}t = 0,\ t > 0; && X_{h,1}(0) = P_hX_{0,1}, \\ &\mathrm{d}X_{h,2} + \Lambda_hX_{h,1}\,\mathrm{d}t = P_h\mathrm{d}W,\ t > 0; && X_{h,2}(0) = P_hX_{0,2}, \end{aligned} \tag{5.1}$$

We put this is the form (1.3) by defining $V_h := S_h \times S_h$ and

$$A_h := \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix}, \quad B_h := \begin{bmatrix} 0 \\ P_h \end{bmatrix}, \quad X_{h0} = P_hX_0.$$

It can be shown that $-A_h$ generates a C_0-semigroup $E_h(t)$ given by

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

with $C_h(t) = \cos(t\Lambda_h^{1/2})$, $S_h(t) = \sin(t\Lambda_h^{1/2})$, which can be expressed in terms of the eigenpairs $\{(\lambda_{h,j}, \phi_{h,j})\}_{j=1}^{N_h}$ of Λ_h.

Our weak convergence result follows.

Theorem 5.1. *Let X and X_h be the solutions of* (1.17) *and* (5.1)*, respectively. Let $g \in C_{\mathrm{b}}^2(\dot{H}^0, \mathbb{R})$ and assume that $\|\Lambda^{\beta-\frac{1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}} < \infty$ and that $X_0 \in L_1(\Omega, H^{2\beta})$ for some $\beta \in [0, \frac{r+1}{2}]$. Then, there are $C > 0$, $h_0 > 0$, depending on g, X_0, Q, and T but not on h, such that for $h \le h_0$,*

$$\big|\mathbf{E}\big(g(X_{h,1}(T)) - g(X_1(T))\big)\big| \le Ch^{\frac{r}{r+1}2\beta}.$$

Proof. If $\|\Lambda^{\beta-\frac{1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathrm{Tr}} < \infty$ for some $\beta \in (0, \frac{r+1}{2}]$, then (3.5) holds by Theorem 2.1 and [10]. Let P_1 denote the canonical projection $H \to \dot{H}^0$. Define the function $G : H \to \mathbb{R}$ by $G(x) := g(P_1x) = g(x_1)$, for $x = [x_1, x_2]^\top \in H$. Then, by (3.8), for $y, z \in H$,

$$(u_x(Y(t), t), y) = \mathbf{E}\big(\langle g'(P_1Z(Y(t), t, T)), P_1y\rangle\big|\mathcal{F}_t\big) \tag{5.2}$$

and

$$(u_{xx}(Y(t), t)y, z) = \mathbf{E}\big(\langle g''(P_1Z(Y(t), t, T))P_1y, P_1z\rangle\big|\mathcal{F}_t\big). \tag{5.3}$$

Let us introduce the error operators

$$\begin{aligned} K_h(t) &:= \Lambda_h^{-\frac{1}{2}}S_h(t)P_h - \Lambda^{-\frac{1}{2}}S(t), \\ G_h(t) &:= C_h(t)P_h - C(t). \end{aligned}$$

From [10] we quote an error estimate for the finite element approximation of the deterministic wave equation

$$\ddot{u} + \Lambda u = 0,\ t > 0; \quad u(0) = w_1,\ \dot{u}(0) = w_2,$$

with solution $u(t) = C(t)w_1 + \Lambda^{-\frac{1}{2}}S(t)w_2$. With $w = [w_1, w_2]^\top$ we have

$$\|G_h(t)w_1 + K_h(t)w_2\| \le C(T)h^{\frac{r}{r+1}s}|||w|||_s, \quad t \in [0,T], s \in [0, r+1]. \tag{5.4}$$

In particular, with $w_1 = 0$,

$$\|K_h(t)w_2\| \le C(T)h^{\frac{r}{r+1}s}|w_2|_{s-1}, \quad w_2 \in \dot{H}^{s-1},$$

or

$$\|K_h(t)\Lambda^{\frac{1-s}{2}}v\| \le C(T)h^{\frac{r}{r+1}s}\|v\|, \quad v \in \dot{H}^{1-s}.$$

The operator $K_h(t)\Lambda^{\frac{1-s}{2}}$ is bounded on $\dot{H}^0$ for $s \ge 1$. For $0 \le s \le 1$ the latter estimate shows that $K_h(t)\Lambda^{\frac{1-s}{2}}$ extends uniquely to a bounded linear operator on $\dot{H}^0$ and we use the same notation for the extended operator. Hence, we may write with the operator norm and $s = 2\beta$,

$$\|K_h(t)\Lambda^{\frac{1}{2}-\beta}\|_{\mathcal{B}(\dot{H}^0)} \le C(T)h^{\frac{r}{r+1}2\beta}, \quad t \in [0,T], \ 0 \le 2\beta \le r+1. \tag{5.5}$$

We use Theorem 3.1 with $G(\cdot) = g(P_1\cdot)$. Since $X_{h,1}(0) = P_hX_1(0)$ and $X_{h,2}(0) = P_hX_2(0)$, we have from (5.2) and (5.4) with $s = 2\beta \le r+1$

$$\begin{aligned}
&|\mathbf{E}\left(u(Y_h(0),0) - u(Y(0),0)\right)| \\
&\quad = \Big|\mathbf{E}\Big(\int_0^1 (u_x(Y(0) + s(Y_h(0) - Y(0)), 0), Y_h(0) - Y(0))\,\mathrm{d}s\Big)\Big| \\
&\quad = \Big|\int_0^1 \mathbf{E}\big(\langle g'(P_1Z(Y(0) + s(Y_h(0) - Y(0)))), P_1(Y_h(0) - Y(0))\rangle\big|\mathcal{F}_0\big)\,\mathrm{d}s\Big| \\
&\quad \le \sup_{x \in \dot{H}^0} \|g'(x)\|\,\mathbf{E}\big(\|P_1(Y_h(0) - Y(0))\|\big) \\
&\quad = \sup_{x \in \dot{H}^0} \|g'(x)\|\,\mathbf{E}\big(\|G_h(T)X_1(0) + K_h(T)X_2(0)\|\big) \\
&\quad \le \sup_{x \in \dot{H}^0} \|g'(x)\|C(T)h^{\frac{r}{r+1}2\beta}\mathbf{E}\big(|||X_0|||_{2\beta}\big).
\end{aligned}$$

To bound the second term (3.6) in the error representation in Theorem 3.1 we can simplify the integrand due to the special choice of G. With $y = [y_1, y_2]^\top$ and the abbreviation $s = T - t$ we calculate, using (5.3),

$$\begin{aligned}
&[E_h(s)B_h - E(s)B]Q[E_h(s)B_h + E(s)B]^*u_{xx}(Y_h(t),t)^*y \\
&\quad = \mathbf{E}\big([E_h(s)B_h - E(s)B]Q[E_h(s)B_h + E(s)B]^*P_1^*g''(P_1Z(Y(t),t,T))P_1y\big|\mathcal{F}_t\big).
\end{aligned}$$

We have, using selfadjointness, that

$$\begin{aligned}
&[E_h(s)B_h + E(s)B]^*P_1^*g''(P_1Z(Y(t),t,T))^*P_1y \\
&\quad = (P_1[E_h(s)B_h + E(s)B])^*g''(P_1Z(Y(t),t,T))y_1 \\
&\quad = [\Lambda_h^{-\frac{1}{2}}S_h(s)P_h + \Lambda^{-\frac{1}{2}}S(s)]g''(P_1Z(Y(t),t,T))y_1.
\end{aligned}$$

Therefore it follows that

$$\begin{aligned}
&[E_h(s)B_h - E(s)B]Q[E_h(s)B_h + E(s)B]^*u_{xx}(Y_h(t),t)^*y \\
&\quad = \begin{bmatrix} K_h(s)Q[\Lambda_h^{-\frac{1}{2}}S_h(s)P_h + \Lambda^{-\frac{1}{2}}S(s)]g''(P_1Z(Y(t),t,T))y_1 \\ G_h(s)Q[\Lambda_h^{-\frac{1}{2}}S_h(s)P_h + \Lambda^{-\frac{1}{2}}S(s)]g''(P_1Z(Y(t),t,T))y_1 \end{bmatrix}.
\end{aligned}$$

Note that the above operator acts only on y_1. Therefore, when we compute its trace as in (2.4) by using an orthonormal basis for H of the form $\{(e_k, 0), (0, f_l)\}_{k,l=1}^\infty$, where $\{e_k\}$ is an orthonormal basis of $\dot H^0$ and $\{f_k\}$ is an orthonormal basis of $\dot H^{-1}$, only terms involving e_k remain. Hence,

$$\begin{aligned}
&\Big|\mathbf{E}\Big(\operatorname{Tr}\big(u_{xx}(Y_h(t),t)[E_h(s)B_h+E(s)B]Q[E_h(s)B_h-E(s)B]^*\big)\Big)\Big|\\
&\quad=\Big|\mathbf{E}\Big(\operatorname{Tr}\big([E_h(s)B_h-E(s)B]Q[E_h(s)B_h+E(s)B]^*u_{xx}(Y_h(t),t)^*\big)\Big)\Big|\\
&\quad=\Big|\mathbf{E}\Big(\sum_{k=1}^{\infty}\mathbf{E}\big(\langle K_h(s)Q[\Lambda_h^{-\frac12}S_h(s)P_h+\Lambda^{-\frac12}S(s)]g''(P_1Z(Y(t),t,T))e_k,e_k\rangle\big|\mathcal{F}_t\big)\Big)\Big|\\
&\quad=\Big|\mathbf{E}\Big(\operatorname{Tr}\big(K_h(s)Q[\Lambda_h^{-\frac12}S_h(s)P_h+\Lambda^{-\frac12}S(s)]g''(P_1Z(Y(t),t,T))\big)\Big)\Big|\\
&\quad\le\|K_h(s)\Lambda^{\frac12-\beta}\|_{\mathcal{B}(\dot H^0)}\|\Lambda^{\beta-\frac12}Q\Lambda^{-\frac12}\|_{\mathrm{Tr}}\\
&\qquad\times\|\Lambda^{\frac12}[\Lambda_h^{-\frac12}S_h(s)P_h+\Lambda^{-\frac12}S(s)]\|_{\mathcal{B}(\dot H^0)}\sup_{x\in\dot H^0}\|g''(x)\|_{\mathcal{B}(\dot H^0)}.
\end{aligned}$$

Noting that $\|\Lambda^{\frac12}v_h\|=\|\nabla v_h\|=\|\Lambda_h^{\frac12}v_h\|$ for $v_h\in S_h$, and using (5.5), we conclude

$$\begin{aligned}
&\Big|\mathbf{E}\Big(\operatorname{Tr}\big(u_{xx}(Y_h(t),t)[E_h(s)B_h+E(s)B]Q[E_h(s)B_h-E(s)B]^*\big)\Big)\Big|\\
&\quad\le C(T)h^{\frac{r}{r+1}2\beta}\|\Lambda^{\beta-\frac12}Q\Lambda^{-\frac12}\|_{\mathrm{Tr}}\sup_{x\in\dot H^0}\|g''(x)\|_{\mathcal{B}(\dot H^0)}.
\end{aligned}$$

Now, we may estimate the second term (3.6),

$$\begin{aligned}
&\Big|\mathbf{E}\Big(\int_0^T\operatorname{Tr}\big(u_{xx}(Y_h(t),t)\\
&\qquad\times[E_h(T-t)B_h+E(T-t)B]Q[E_h(T-t)B_h-E(T-t)B]^*\big)\,\mathrm{d}t\Big)\Big|\\
&\quad\le C(T)h^{\frac{r}{r+1}2\beta}\|\Lambda^{\beta-\frac12}Q\Lambda^{-\frac12}\|_{\mathrm{Tr}}\sup_{x\in\dot H^0}\|g''(x)\|_{\mathcal{B}(\dot H^0)}.
\end{aligned}$$

□

In [10] the strong rate of convergence is found to be $O(h^{\frac{r}{r+1}\beta})$ for $\beta\in[0,r+1]$ under the condition $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}<\infty$. Theorem 2.1 shows that $\|\Lambda^{\beta-\frac12}Q\Lambda^{-\frac12}\|_{\mathrm{Tr}}<\infty$ provides a sufficient condition for this and the conditions coincide if A and Q commute, in particular, if $Q=I$.

As a special case, if $Q=I$, then $d=1$ and we may take $\beta<\frac12$. Hence the order of weak convergence is almost $O(h^{\frac{r}{r+1}})$.

References

[1] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, 1992.

[2] ———, *Second Order Partial Differential Equations in Hilbert Spaces*, Cambridge University Press, Cambridge, 2002.

[3] 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.

[4] A. Debussche, *Weak approximation of stochastic partial differential equations: the non linear case*, preprint 2008, `arXiv:0804.1304v1 [math.NA]`.

[5] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comp. **78** (2009), 845–863.

[6] 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.

[7] E. Hausenblas, *Weak approximation for semilinear stochastic evolution equations*, Stochastic Analysis and Related Topics VIII, Progr. Probab., vol. 53, Birkhäuser, Basel, 2003, pp. 111–128.
[8] E. Hausenblas, *Weak approximation of the stochastic wave equation*, preprint.
[9] 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.
[10] 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.
[11] S. Larsson and A. Mesforush, *Finite element approximation of the linearized Cahn-Hilliard-Cook equation*, IMA J. Numer. Anal, to appear.
[12] P. D. Lax, *Functional Analysis*, Wiley, 2002.
[13] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer, 2006.
[14] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), 829–847.
[15] J. Weidmann, *Linear Operators in Hilbert Spaces*, 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