CHALMERS | GÖTEBORG UNIVERSITY

Rate of Weak Convergence of the Finite Element Method for the Stochastic Heat Equation with Additive Noise

MATTHIAS GEISSERT
MIHÁLY KOVÁCS
STIG LARSSON

Rate of Weak Convergence of the Finite Element Method for the Stochastic Heat Equation with Additive Noise

Matthias Geissert, Mihály Kovács, Stig Larsson

RATE OF WEAK CONVERGENCE OF THE FINITE ELEMENT METHOD FOR THE STOCHASTIC HEAT EQUATION WITH ADDITIVE NOISE

MATTHIAS GEISSERT, MIHÁLY KOVÁCS[1], AND STIG LARSSON[1]

ABSTRACT. The stochastic heat equation driven by additive noise is discretized in the spatial variables by a standard finite element method. The weak convergence of the approximate solution is investigated and the rate of weak convergence is found to be twice that of strong convergence.

1. INTRODUCTION AND MAIN RESULT

We consider the stochastic heat equation driven by additive noise,

$$
\begin{aligned}
\mathrm{d}X - \Delta X \, \mathrm{d}t &= \mathrm{d}W, & &\text{in } D \times \mathbb{R}^+, \\
X &= 0, & &\text{on } \partial D \times \mathbb{R}^+, \\
X(\cdot, 0) &= X_0, & &\text{in } D.
\end{aligned}
\tag{1.1}
$$

Here $D \subset \mathbb{R}^d$ is a convex polygonal domain and $\Delta = \sum_{k=1}^{d} \partial^2/\partial\xi_k^2$ is the Laplace operator. Let $H = L_2(D)$ with the usual norm $\|\cdot\|$ and scalar product $\langle\cdot,\cdot\rangle$. Define $A = -\Delta$ with $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$ and write (1.1) as

$$
\mathrm{d}X + AX \, \mathrm{d}t = \mathrm{d}W, \ t > 0; \quad X(0) = X_0.
\tag{1.2}
$$

Let $(\Omega, \mathcal{F}, \mathbf{P}, \{\mathcal{F}_t\}_{t\geq 0})$ be a filtered probability space, let $Q : H \to H$ be a self-adjoint, positive semidefinite bounded linear operator, let W be a Q-Wiener process in H adapted to the filtration, and assume that X_0 is an $\mathcal{F}_0$-measurable H-valued random variable. Then (1.2) has a unique mild solution given by

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

where $E(t) = \mathrm{e}^{-tA}$ is the analytic semigroup generated by $-A$ and the stochastic convolution is well defined, see [1]. Note that A and Q need not commute.

We approximate (1.2) by a standard finite element method. Let thus D be a polygonal domain such that sufficient regularity estimates hold (see (2.7) below) and let $\{S_h\}_{h>0}$ be a family of function spaces consisting of continuous piecewise polynomials of degree $\leq r - 1$ with respect to a quasi-uniform family of triangulations of D and such that $S_h \subset H_0^1(D)$. The parameter h is the maximal mesh

---

1

size of the triangulation and r may be referred to as the order of the finite element method.

The finite element approximation of (1.2) is, see [6],

$$(1.4) \qquad \mathrm{d}X_h + A_h X_h \, \mathrm{d}t = P_h \, \mathrm{d}W, \ t > 0; \quad X(0) = P_h X_0,$$

where $P_h : H \to S_h$ denotes the orthogonal projection and $A_h : S_h \to S_h$ is the "discrete Laplacian" defined by

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

It can be verified that $P_h W$ is a Q_h-Wiener process in S_h with $Q_h = P_h Q P_h$ and the solution of (1.4) is given by

$$(1.5) \qquad X_h(t) = E_h(t) P_h X_0 + \int_0^t E_h(t-s) P_h \, \mathrm{d}W(s),$$

where $E_h(t) = \mathrm{e}^{-t A_h}$ is the analytic semigroup generated by $-A_h$.

In order to describe the spatial regularity of functions we introduce the following spaces and norms. Let

$$\dot{H}^\gamma = \mathcal{D}(A^{\frac{\gamma}{2}}), \quad \|f\|_{\dot{H}^\gamma} = \|A^{\frac{\gamma}{2}} f\|, \quad \gamma \in \mathbb{R}.$$

It is clear that $\dot{H}^0 = H$, $\dot{H}^1 = H_0^1(D)$, $\dot{H}^2 = H^2(D) \cap H_0^1(D)$ with equivalent norms. We also need the Hilbert-Schmidt norm for bounded linear operators,

$$(1.6) \qquad \|K\|_{\mathrm{HS}}^2 = \sum_{k=1}^\infty \|K\phi_k\|^2 = \mathrm{Tr}(K^* K),$$

where $\{\phi_k\}_{k=1}^\infty$ is an arbitrary ON-basis in H, and we define $L_2(\Omega, \dot{H}^\beta)$ by

$$\|f\|_{L_2(\Omega, \dot{H}^\beta)} = \mathbf{E}\big(\|f\|_{\dot{H}^\beta}^2\big)^{\frac{1}{2}} = \bigg(\int_\Omega \|f(\omega)\|_{\dot{H}^\beta}^2 \, \mathrm{d}\mathbf{P}(\omega) \bigg)^{\frac{1}{2}}.$$

In [6] the following strong convergence result was proved:

Let $r \geq 2$ and $0 \leq \beta \leq 1$ and assume that $\|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ and $X_0 \in L_2(\Omega, \dot{H}^\beta)$. Then there is C such that, for $t \geq 0$,

$$(1.7) \qquad \|X_h(t) - X(t)\|_{L_2(\Omega, H)} \leq C h^\beta \Big(\|X_0\|_{L_2(\Omega, \dot{H}^\beta)} + \|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} \Big).$$

In this work we prove weak convergence, that is, convergence of $\mathbf{E} g(X_h(t))$ for all g in a suitable class of functions. We denote by $C_{\mathrm{b}}^2(H, \mathbb{R})$ the set of all real-valued functions on H with continuous and bounded derivatives of order ≤ 2.

Our main result is the following.

Theorem 1.1. *Let X and X_h be given by (1.3) and (1.5), respectively. Assume that $g \in C_{\mathrm{b}}^2(H, \mathbb{R})$ and, in addition,*

$$(1.8) \qquad \sup_{f \in \dot{H}^\beta} \|g'(f)\|_{\dot{H}^\beta} < \infty.$$

Let $r \geq 3$, $0 < \beta \leq 1$, and assume that W is a Q-Wiener process in H with $\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ and $X_0 \in L_2(\Omega, \dot{H}^\beta)$. Then there are $C > 0$, $h_0 > 0$, independent of g, T, and h, such that for $h \leq h_0$, $T \geq 0$,

$$\left|\mathbf{E}\big(g(X_h(T)) - g(X(T))\big)\right| \leq Ch^{2\beta}|\log(h)|$$
$$\times \left(\|X_0\|_{L_2(\Omega,\dot{H}^\beta)} + \|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} + \|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2\right)$$
$$\times \sup_{f \in \dot{H}^\beta}\left(\|g'(f)\|_{\dot{H}^\beta} + \|g''(f)\|\right).$$

Note that the assumption on Q, $\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$, is the same as for the strong convergence estimate (1.7). The rate of weak convergence is thus twice the rate of strong convergence, except for the logarithmic factor.

Two special cases can be highlighted:

- If Q is of trace class, then $\beta = 1$, because $\|Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(Q) < \infty$. Thus, the rate of convergence is $O(h^2|\log(h)|)$.

- If $Q = I$, which is space-time white noise, then we assume $\|A^{-\frac{1-\beta}{2}}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(A^{-(1-\beta)}) < \infty$. Since the eigenvalues of A behave like $\lambda_k \sim k^{2/d}$ as $k \to \infty$, this is possible if and only if $d = 1$ and $\beta < \frac{1}{2}$. Then the rate of convergence is almost $O(h)$.

The theorem has two unnatural restrictions which are due only to our proof, namely $r \geq 3$ and the extra assumption (1.8) on g'. Indeed, after this work was completed we learned that in [4] the same rate of convergence (without logarithm) was proved without these assumptions. However, the estimate in [4] blows up as $T \to 0$, while ours is uniform in T (cf. (3.3) and (3.4) below). The proof is also more complicated than ours and the condition $\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ is replaced by a different, but essentially equivalent, one. Time stepping by the theta method is also included in [4].

While the literature on strong convergence of numerical approximations of parabolic stochastic partial differential equations is abundant, there is very little on weak convergence. Apart from [4] we are only aware of [5], which proves a similar result but under a much stronger restriction on the test function g, and [3] which is concerned with the Schrödinger equation.

2. Preliminaries

In this section we collect some facts and make some preparations for the proof of Theorem 1.1.

Let S, T be bounded linear operators on the Hilbert space H and assume that T is of trace class. Then,

$$(2.1) \qquad \mathrm{Tr}(ST) = \mathrm{Tr}(TS), \quad \mathrm{Tr}(T^*) = \mathrm{Tr}(T).$$

If, instead, T is Hilbert-Schmidt, then by (1.6) and (2.1),

$$(2.2) \qquad \|T\|_{\mathrm{HS}} = \|T^*\|_{\mathrm{HS}}, \quad \|TS\|_{\mathrm{HS}} = \|S^*T^*\|_{\mathrm{HS}}, \quad \|ST\|_{\mathrm{HS}} \leq \|S\|\|T\|_{\mathrm{HS}}.$$

Under the assumptions on the finite elements that we made in Section 1 we have the following inequalities. We state them only for the ranges of parameters that we need. First we have the inverse inequality,

$$(2.3) \qquad \|A_h^\gamma u_h\| \leq Ch^{-2\gamma}\|u_h\|, \quad u_h \in S_h, \ \gamma \in [0,1].$$

We also have the following equivalence of norms in S_h:

$$(2.4) \qquad \|A^\gamma u_h\| \leq C\|A_h^\gamma u_h\|, \quad u_h \in S_h, \gamma \in [0, \tfrac{1}{2}],$$

and

$$(2.5) \qquad \|A_h^\gamma u_h\| \leq C\|A^\gamma u_h\|, \quad u_h \in S_h, \ \gamma \in [0, \tfrac{1}{2}].$$

Moreover, for $0 \leq \delta \leq \tfrac{1}{2}$, $\delta \leq \rho \leq 1$, we have

$$(2.6) \qquad \|A^\delta(P_h - I)A^{-\rho}\| \leq Ch^{2\rho - 2\delta}.$$

This is a well-known approximation property for P_h. Finally, the standard error estimate for the linear elliptic finite element problem can be formulated as

$$(2.7) \qquad \|(A^{-1} - A_h^{-1}P_h)f\| \leq Ch^s\|A^{\frac{s-2}{2}}f\|, \quad s \in [2, 3].$$

(This requires some assumptions on the polygonal domain D.) The above inequalities are well-known in the special cases when the parameters γ, δ, ρ, s are at the endpoints of the parameter intervals. For intermediate parameter values we use an interpolation procedure. The error estimate (2.7) is the only place where we need $r \geq 3$.

It is also well known that there are positive numbers C, α such that

$$(2.8) \qquad \|A^\gamma E(t)v\| \leq Ct^{-(\gamma - \delta)}\mathrm{e}^{-\alpha t}\|A^\delta v\|, \quad t > 0, \ 0 \leq \delta \leq \gamma \leq 1,$$

and that

$$(2.9) \qquad \int_0^T \|A^{\frac{1}{2}}E(t)v\|^2 \,\mathrm{d}t \leq \tfrac{1}{2}\|v\|^2, \quad \int_0^T \|A_h^{\frac{1}{2}}E_h(t)P_hv\|^2 \,\mathrm{d}t \leq \tfrac{1}{2}\|v\|^2.$$

We also recall Itô's isometry, which takes the form

$$(2.10) \qquad \mathbf{E}\Big(\Big\| \int_0^t F(t-s)\,\mathrm{d}W(s)\Big\|^2\Big) = \int_0^t \|F(t-s)Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \,\mathrm{d}s,$$

when the integrand is a deterministic operator as in (1.3) and (1.5).

Let

$$(2.11) \qquad Z(t, \tau, \xi) = E(t - \tau)\xi + \int_\tau^t E(t - s)\,\mathrm{d}W(s), \quad 0 \leq \tau \leq t \leq T.$$

If ξ is $\mathcal{F}_\tau$-measurable, then $Y(t) = Z(t, \tau, \xi)$ is the unique mild solution to

$$\mathrm{d}Y(t) + AY(t)\,\mathrm{d}t = \mathrm{d}W(t), \ t > \tau; \quad Y(\tau) = \xi.$$

In particular, the solution of (1.2) is $X(t) = Z(t, 0, X_0)$. The following lemma describes the spatial regularity of Z.

Lemma 2.1. *Let* $\beta \in [0, 1]$. *If* $\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ *and* $\xi \in L_2(\Omega, \dot{H}^\beta)$, *then* $Z(T, t, \xi) \in L_2(\Omega, \dot{H}^\beta)$ *for all* $0 \leq t \leq T$.

Proof. This is proved in [6, Theorem 3.1]; see also [1, Section 5.4] and Lemma 2.2 below. □

We have the same regularity for X_h:

Lemma 2.2. *Let* $\beta \in [0, 1]$, *assume that* $\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$, $X_0 \in L^2(\Omega, \dot{H}^\beta)$, *and let* X_h *be given by* (1.5). *Then*

$$\|X_h(t)\|_{L_2(\Omega, \dot{H}^\beta)} \leq C\Big(\|X_0\|_{L_2(\Omega, \dot{H}^\beta)} + \|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}\Big).$$

Proof. By taking norms in (1.5) and using (2.4) we have

$$\|X_h(t)\|^2_{L_2(\Omega,\dot{H}^\beta)} \le C\mathbf{E}\Big(\|A_h^{\frac{\beta}{2}} X_h(t)\|^2\Big)$$

$$\le C\mathbf{E}\Big(\|A_h^{\frac{\beta}{2}} E_h(t)P_h X_0\|^2\Big) + C\mathbf{E}\Big(\Big\|\int_0^t A_h^{\frac{\beta}{2}} E_h(t-s)P_h\,\mathrm{d}W(s)\Big\|^2\Big)$$

$$= C\big(I_1 + I_2\big).$$

By Itô's isometry (2.10), (2.2), and selfadjointness, we obtain

$$I_2 = \int_0^t \|A_h^{\frac{\beta}{2}} E_h(t-s)P_h Q^{\frac{1}{2}}\|^2_{\mathrm{HS}}\,\mathrm{d}s = \int_0^t \|(Q^{\frac{1}{2}})^*(A_h^{\frac{\beta}{2}} E_h(s)P_h)^*\|^2_{\mathrm{HS}}\,\mathrm{d}s$$

$$= \int_0^t \|Q^{\frac{1}{2}} A_h^{\frac{\beta}{2}} E_h(s)P_h\|^2_{\mathrm{HS}}\,\mathrm{d}s.$$

Using (2.2) and (2.9) we calculate

$$I_2 \le \|Q^{\frac{1}{2}} A_h^{-\frac{1-\beta}{2}} P_h\|^2_{\mathrm{HS}} \int_0^t \|A_h^{\frac{1}{2}} E_h(s)P_h\|^2\,\mathrm{d}s$$

$$\le \tfrac{1}{2}\|Q^{\frac{1}{2}} A_h^{-\frac{1-\beta}{2}} P_h\|^2_{\mathrm{HS}} = \tfrac{1}{2}\|Q^{\frac{1}{2}} A^{-\frac{1-\beta}{2}} A^{\frac{1-\beta}{2}} A_h^{-\frac{1-\beta}{2}} P_h\|^2_{\mathrm{HS}}.$$

Hence, using (2.4) and (2.2), we obtain

$$I_2 \le \tfrac{1}{2}\|Q^{\frac{1}{2}} A^{-\frac{1-\beta}{2}}\|^2_{\mathrm{HS}} \|A^{\frac{1-\beta}{2}} A_h^{-\frac{1-\beta}{2}} P_h\|^2 \le C\|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|^2_{\mathrm{HS}}.$$

Finally, by (2.5),

$$I_1 \le \mathbf{E}\Big(\|A_h^{\frac{\beta}{2}} P_h X_0\|^2\Big) \le C\,\mathbf{E}\Big(\|A^{\frac{\beta}{2}} X_0\|^2\Big) = C\|X_0\|^2_{L_2(\Omega,\dot{H}^\beta)}.$$

The proof is complete. □

For $g \in C_{\mathrm{b}}^2(H,\mathbb{R})$ and with Z as in (2.11) we define

$$(2.12) \qquad u(x,t) = \mathbf{E}\Big(g(Z(T,t,x))\Big), \quad x \in H,\ 0 \le t \le T.$$

If ξ is an $\mathcal{F}_t$-measurable random variable, then by [1, Theorem 9.8],

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

Therefore, by the law of double expectation,

$$(2.13) \qquad \mathbf{E}\Big(u(\xi,t)\Big) = \mathbf{E}\Big(\mathbf{E}\Big(g(Z(T,t,\xi))\Big|\mathcal{F}_t\Big)\Big) = \mathbf{E}\Big(g(Z(T,t,\xi))\Big).$$

For a function $f \in C^2(H,\mathbb{R})$, by the Riesz representation theorem, we may identify the bounded linear functional $f'(x)$ with an element in H and the bounded bilinear functional $f''(x)$ with a symmetric bounded linear operator on H. Hence, using the explicit formula for Z in (2.11), we have

$$\langle \partial_1 u(x,t), y\rangle = \mathbf{E}\Big(\langle g'(Z(T,t,x)), E(T-t)y\rangle\Big), \quad y \in H,$$

$$\langle \partial_1^2 u(x,t)y_1, y_2\rangle = \mathbf{E}\Big(\langle g''(Z(T,t,x))E(T-t)y_1, E(T-t)y_2\rangle\Big), \quad y_1, y_2 \in H,$$

so that

$$(2.14) \qquad \partial_1 u(x,t) = \mathbf{E}\Big(E(T-t)g'(Z(T,t,x))\Big),$$

$$(2.15) \qquad \partial_1^2 u(x,t) = \mathbf{E}\Big(E(T-t)g''(Z(T,t,x))E(T-t)\Big).$$

The function u satisfies Kolmogorov's equation (see [1, Chapter 9] for details),

$$(2.16) \quad \begin{aligned} \partial_2 u(x,t) - \langle Ax, \partial_1 u(x,t) \rangle + \tfrac{1}{2} \operatorname{Tr}\left(\partial_1^2 u(x,t) Q\right) &= 0, \quad x \in \mathcal{D}(A),\ t \in (0,T), \\ u(x,T) &= g(x), \qquad\qquad\quad x \in H. \end{aligned}$$

3. Proof of Theorem 1.1

Let $u(x,t)$ be defined by (2.12) and apply (2.13) with $\xi = X(0)$,

$$\begin{aligned} \mathbf{E}\Big(u(X(0),0)\Big) &= \mathbf{E}\Big(\mathbf{E}\Big(g(Z(T,0,X(0)))\Big|\mathcal{F}_0\Big)\Big) \\ &= \mathbf{E}\Big(g(Z(T,0,X(0)))\Big) = \mathbf{E}\Big(g(X(T))\Big), \end{aligned}$$

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

$$\mathbf{E}\Big(u(X_h(T),T)\Big) = \mathbf{E}\Big(\mathbf{E}\Big(g(Z(T,T,X_h(T)))\Big|\mathcal{F}_T\Big)\Big) = \mathbf{E}\Big(g(X_h(T))\Big).$$

Hence,

$$\begin{aligned} \mathbf{E}\Big(g(X_h(T)) - g(X(T))\Big) &= \mathbf{E}\Big(u(X_h(T),T) - u(X(0),0)\Big) \\ &= \mathbf{E}\Big(u(X_h(0),0) - u(X(0),0)\Big) \\ &\quad + \mathbf{E}\Big(u(X_h(T),T) - u(X_h(0),0)\Big). \end{aligned}$$

Using Itô's formula (see, for example, [1]) for $u(X_h(t),t)$, where X_h satisfies (1.4) with covariance operator $Q_h = P_h Q_h P_h$, and using the Kolmogorov equation (2.16) to replace the term $\partial_2 u$, we have at least formally

$$(3.1)$$
$$\begin{aligned} &\mathbf{E}\Big(u(X_h(T),T) - u(X_h(0),0)\Big) \\ &= \mathbf{E}\int_0^T \Big(\partial_2 u(X_h(t),t) - \langle A_h X_h(t), \partial_1 u(X_h(t),t)\rangle + \tfrac{1}{2}\operatorname{Tr}(\partial_1^2 u(X_h(t),t)Q_h)\Big)\,\mathrm{d}t \\ &= \mathbf{E}\int_0^T \Big(-\langle X_h(t), (A_h - A)\partial_1 u(X_h(t),t)\rangle + \tfrac{1}{2}\operatorname{Tr}\big(\partial_1^2 u(X_h(t),t)(Q_h - Q)\big)\Big)\,\mathrm{d}t. \end{aligned}$$

We note here that the use of Itô's formula requires a function $u : H \times \mathbb{R}^+ \to \mathbb{R}$ with uniformly continuous partial derivatives on bounded subsets of $H \times [0,T]$. The above calculation can be made rigorous as follows. Define u_n in the same way as u in (2.12) but with A replaced by the Yosida approximant $A_n = nA(nI - A)^{-1}$. Then u_n has continuous partial derivatives ([1, Theorem 9.16]) and, since $X_h(t) \in S_h$, we may consider u_n as a function $S_h \times \mathbb{R}^+ \to \mathbb{R}$. Then the derivatives are uniformly continuous on bounded subsets of $S_h \times [0,T]$. Writing Itô's formula and Kolmogorov's equation for u_n we get (3.1) for u_n. Passing to the limit in (3.1) using standard arguments (see, for example [1, Chapter 9] and [2, Chapter 7]) we obtain (3.1) for u.

We therefore have

$$\mathbf{E}\Big(g(X_h(T)) - g(X(T))\Big) = \mathbf{E}\Big(u(X_h(0),0) - u(X(0),0)\Big)$$

$$- \mathbf{E}\int_0^T \langle (A_h P_h - P_h A)\partial_1 u(X_h(t),t), X_h(t)\rangle \, \mathrm{d}t$$

(3.2)

$$+ \tfrac{1}{2}\mathbf{E}\int_0^T \mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)(Q_h - Q)\right)\mathrm{d}t$$

$$= I_1 + I_2 + I_3.$$

We estimate the three terms separately. For the first term we have, by the chain rule and recalling $X_h(0) = P_h X_0$, $X(0) = X_0$,

$$I_1 = \mathbf{E}\Big\{\int_0^1 \langle \partial_1 u(X_0 + s(P_h X_0 - X_0),0), P_h X_0 - X_0\rangle \, \mathrm{d}s\Big\}$$

$$= \mathbf{E}\Big\{\int_0^1 \langle (P_h - I)\partial_1 u(X_0 + s(P_h X_0 - X_0),0), (P_h - I)X_0\rangle \, \mathrm{d}s\Big\},$$

so that, by (2.6) with $\rho = \beta/2$, $\delta = 0$, and since $X_0 \in \dot{H}^\beta$ almost surely,

$$|I_1| \leq \mathbf{E}\Big\{\int_0^1 Ch^\beta \|\partial_1 u(X_0 + s(P_h X_0 - X_0),0)\|_{\dot{H}^\beta} \cdot Ch^\beta \|X_0\|_{\dot{H}^\beta} \, \mathrm{d}s\Big\}$$

$$\leq Ch^{2\beta} \sup_{x \in \dot{H}^\beta} \|\partial_1 u(x,0)\|_{\dot{H}^\beta} \|X_0\|_{L_2(\Omega,\dot{H}^\beta)}.$$

Here, by (2.14) and (2.8) with $\gamma = \delta = \beta/2$,

$$\|\partial_1 u(x,0)\|_{\dot{H}^\beta} = \|\mathbf{E}\{E(T)g'(Z(T,0,x))\}\|_{\dot{H}^\beta}$$

$$\leq \mathbf{E}\{\|E(T)g'(Z(T,0,x))\|_{\dot{H}^\beta}\}$$

$$\leq C \sup_{f \in \dot{H}^\beta} \|g'(f)\|_{\dot{H}^\beta},$$

since it follows from Lemma 2.1 that $Z(T,0,x) \in \dot{H}^\beta$ almost surely. Hence,

(3.3)
$$|I_1| \leq Ch^{2\beta} \sup_{f \in \dot{H}^\beta} \|g'(f)\|_{\dot{H}^\beta} \|X_0\|_{L_2(\Omega,\dot{H}^\beta)}.$$

We remark that by using, instead, (2.8) with $\gamma = \beta/2$, $\delta = 0$, in the last step, we would get

(3.4)
$$|I_1| \leq Ch^{2\beta}T^{-\beta/2} \sup_{f \in \dot{H}^\beta} \|g'(f)\| \|X_0\|_{L_2(\Omega,\dot{H}^\beta)},$$

where the assumption (1.8) is not needed.

For the second term, we use the identity

$$A_h P_h - P_h A = A_h P_h (A^{-1} - A_h^{-1} P_h)A,$$

to get

$$I_2 = -\mathbf{E}\int_0^T \langle (A_h P_h - P_h A)\partial_1 u(X_h(t),t), X_h(t)\rangle \, \mathrm{d}t$$

$$= -\mathbf{E}\int_0^T \langle A_h P_h (A^{-1} - A_h^{-1} P_h)A\partial_1 u(X_h(t),t), X_h(t)\rangle \, \mathrm{d}t$$

$$= -\mathbf{E}\int_0^T \langle A_h^{1-\frac{\beta}{2}} P_h (A^{-1} - A_h^{-1} P_h)A\partial_1 u(X_h(t),t), A_h^{\frac{\beta}{2}} X_h(t)\rangle \, \mathrm{d}t.$$

Therefore,

$$
\begin{aligned}
|I_2| &\leq \mathbf{E} \int_0^T \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A\partial_1 u(X_h(t),t)\| \, \|X_h(t)\|_{\dot{H}^\beta} \, \mathrm{d}t \\
&\leq \int_0^T \sup_{x \in \dot{H}^\beta} \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A\partial_1 u(x,t)\| \, \mathrm{d}t \\
&\quad \times \sup_{t \in [0,T]} \|X_h(t)\|_{L_2(\Omega, \dot{H}^\beta)}.
\end{aligned}
$$

Using (2.14) we have here

$$
\begin{aligned}
&\|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A\partial_1 u(x,t)\| \\
&\quad = \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A\mathbf{E}\{E(T-t)g'(Z(T,t,x))\}\| \\
&\quad \leq \mathbf{E}\{\|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(T-t)A^{\frac{\beta}{2}} g'(Z(T,t,x))\|\} \\
&\quad \leq \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(T-t)\| \sup_{f \in \dot{H}^\beta} \|g'(f)\|_{\dot{H}^\beta}.
\end{aligned}
$$

Therefore,

(3.5)
$$
\begin{aligned}
|I_2| &\leq \int_0^T \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\| \, \mathrm{d}t \\
&\quad \times \sup_{f \in \dot{H}^\beta} \|g'(f)\|_{\dot{H}^\beta} \sup_{t \in [0,T]} \|X_h(t)\|_{L_2(\Omega, \dot{H}^\beta)}.
\end{aligned}
$$

The latter two factors are bounded according to (1.8) and Lemma 2.2. It remains to estimate the first factor. In view of (2.3) we have

$$
\begin{aligned}
\int_0^T &\|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\| \, \mathrm{d}t \\
&\qquad\qquad \leq Ch^{-2+\beta} \int_0^T \|(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\| \, \mathrm{d}t.
\end{aligned}
$$

Using (2.7) with $s = 2$ and $s = 2 + \beta$ (this is the only place where we need $r \geq 3$) and (2.8), we obtain

$$
\begin{aligned}
\int_0^T &\|(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\| \, \mathrm{d}t \\
&= \int_0^{h^2} \|(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\| \, \mathrm{d}t \\
&\quad + \int_{h^2}^T \|(A^{-1} - A_h^{-1}P_h)A^{-\frac{\beta}{2}} AE(t)\| \, \mathrm{d}t \\
&\leq Ch^2 \int_0^{h^2} t^{-1+\frac{\beta}{2}} \, \mathrm{d}t + Ch^{2+\beta} \int_{h^2}^T t^{-1} \mathrm{e}^{-\alpha t} \, \mathrm{d}t \\
&\leq \frac{C}{\beta} h^{2+\beta} + Ch^{2+\beta} |\log(h)| \leq Ch^{2+\beta} |\log(h)|,
\end{aligned}
$$

for $h \leq h_0$ (C and h_0 depend on β). Inserting this together with (1.8) and Lemma 2.2 into (3.5) shows

$$(3.6) \qquad |I_2| \leq Ch^{2\beta}|\log(h)|\left(\|X_0\|_{L_2(\Omega,\dot{H}^\beta)} + \|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}\right)\sup_{f \in \dot{H}^\beta}\|g'(f)\|_{\dot{H}^\beta}.$$

For the third term in (3.2) we use the identity

$$Q_h - Q = P_hQP_h - Q = P_hQP_h - QP_h + QP_h - Q = (P_h - I)QP_h + Q(P_h - I).$$

Therefore,

$$2I_3 = \mathbf{E}\int_0^T \mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)(Q_h - Q)\right)\mathrm{d}t$$
$$= \mathbf{E}\int_0^T \mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)\{(P_h - I)QP_h + Q(P_h - I)\}\right)\mathrm{d}t.$$

Note that, since $E(t)$ is trace class for $t > 0$, it follows from (2.15) that $\partial_1^2 u(X_h(t),t)$ is trace class as well, so that by (2.1) the above integrands are well defined. Moreover, it follows from (2.1) and selfadjointness that

$$\mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)(P_h - I)(QP_h)\right) = \mathrm{Tr}\left((QP_h)^*(P_h - I)^*\partial_1^2 u(X_h(t),t)^*\right)$$
$$= \mathrm{Tr}\left(\{P_hQ(P_h - I)\}\{\partial_1^2 u(X_h(t),t)\}\right)$$
$$= \mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)P_hQ(P_h - I)\right).$$

Therefore,

$$2I_3 = \mathbf{E}\int_0^T \mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)(P_h + I)Q(P_h - I)\right)\mathrm{d}t.$$

We bound the integrand. By (2.15) we have

$$\left|\mathrm{Tr}\left(\partial_1^2 u(X_h(t),t)(P_h + I)Q(P_h - I)\right)\right|$$
$$\leq \sup_{x \in \dot{H}^\beta}\left|\mathrm{Tr}\left(\partial_1^2 u(x,t)(P_h + I)Q(P_h - I)\right)\right|$$
$$= \sup_{x \in \dot{H}^\beta}\left|\mathrm{Tr}\left(\mathbf{E}\{E(T - t)g''(Z(T,t,x))E(T - t)\}(P_h + I)Q(P_h - I)\right)\right|$$
$$\leq \mathbf{E}\left\{\sup_{x \in \dot{H}^\beta}\left|\mathrm{Tr}\left(E(T - t)g''(Z(T,t,x))E(T - t)(P_h + I)Q(P_h - I)\right)\right|\right\}$$
$$\leq \sup_{f \in \dot{H}^\beta}\left|\mathrm{Tr}\left(E(T - t)g''(f)E(T - t)(P_h + I)Q(P_h - I)\right)\right|,$$

so that

$$(3.7) \qquad 2|I_3| \leq \mathbf{E}\int_0^T \sup_{f \in \dot{H}^\beta}\left|\mathrm{Tr}\left(E(t)g''(f)E(t)(P_h + I)Q(P_h - I)\right)\right|\mathrm{d}t.$$

Let $\varepsilon \in (0, \beta/2)$. Since $E(t)$ is of trace class for $t > 0$, we may use (2.1) to obtain

$$
\mathrm{Tr}\left(E(t)g''(f)E(t)(P_h + I)Q(P_h - I)\right)
$$
$$
= \mathrm{Tr}\left(A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\right.
$$
$$
\left. \times A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h + I)Q(P_h - I)\right)
$$
$$
= \mathrm{Tr}\left(A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h + I)Q(P_h - I)\right.
$$
$$
\left. \times A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\right)
$$
$$
= \mathrm{Tr}\left(A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h + I)Q^{\frac{1}{2}}\right.
$$
$$
\left. \times Q^{\frac{1}{2}}(P_h - I)A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\right),
$$

so that, in view of $\mathrm{Tr}(AB) \leq \|A\|_{\mathrm{HS}}\|B\|_{\mathrm{HS}}$,

$$
\left|\mathrm{Tr}\left(E(t)g''(f)E(t)(P_h + I)Q(P_h - I)\right)\right|
$$
$$
\leq \|A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h + I)Q^{\frac{1}{2}}\|_{\mathrm{HS}}\|Q^{\frac{1}{2}}(P_h - I)A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\|_{\mathrm{HS}}
$$
$$
= J_1 \times J_2.
$$

For the first factor we use (2.2) to get

$$
J_1 = \|\{A^{\frac{\beta}{2}-\varepsilon}E(t/2)\}\{g''(f)E(t)(P_h + I)Q^{\frac{1}{2}}\}\|_{\mathrm{HS}}
$$
$$
= \|\{Q^{\frac{1}{2}}(P_h + I)E(t)g''(f)\}\{A^{\frac{\beta}{2}-\varepsilon}E(t/2)\}\|_{\mathrm{HS}}
$$
$$
\leq \|Q^{\frac{1}{2}}(P_h + I)E(t)\|_{\mathrm{HS}}\|g''(f)\|\,\|A^{\frac{\beta}{2}-\varepsilon}E(t/2)\|.
$$

Using also (2.6) with $\delta = \rho = \frac{1-\beta}{2}$, we have here

$$
\|Q^{\frac{1}{2}}(P_h + I)E(t)\|_{\mathrm{HS}}
$$
$$
= \|Q^{\frac{1}{2}}A^{-\frac{1-\beta}{2}}A^{\frac{1-\beta}{2}}(P_h + I)A^{-\frac{1-\beta}{2}}A^{\frac{1-\beta}{2}}E(t)\|_{\mathrm{HS}}
$$
$$
\leq \|Q^{\frac{1}{2}}A^{-\frac{1-\beta}{2}}\|_{\mathrm{HS}}\|A^{\frac{1-\beta}{2}}(P_h + I)A^{-\frac{1-\beta}{2}}\|\,\|A^{\frac{1-\beta}{2}}E(t)\|
$$
$$
\leq C\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}\|A^{\frac{1-\beta}{2}}E(t)\|.
$$

Therefore,

$$
J_1 \leq C\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}\|g''(f)\|\,\|A^{\frac{1-\beta}{2}}E(t)\|\,\|A^{\frac{\beta}{2}-\varepsilon}E(t/2)\|.
$$

For the other factor we have, by (2.2),

$$
J_2 = \|Q^{\frac{1}{2}}A^{-\frac{1-\beta}{2}}A^{\frac{1-\beta}{2}}(P_h - I)A^{-\frac{1+\beta}{2}+\varepsilon}A^{\frac{1}{2}}E(t/2)\|_{\mathrm{HS}}
$$
$$
\leq \|Q^{\frac{1}{2}}A^{-\frac{1-\beta}{2}}\|_{\mathrm{HS}}\|A^{\frac{1-\beta}{2}}(P_h - I)A^{-\frac{1+\beta}{2}+\varepsilon}\|\,\|A^{\frac{1}{2}}E(t/2)\|.
$$

Then, by (2.6) with $\delta = \frac{1-\beta}{2} \leq \rho = \frac{1+\beta}{2} - \varepsilon$,

$$
J_2 \leq Ch^{2\beta - 2\varepsilon}\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}\|A^{\frac{1}{2}}E(t/2)\|.
$$

Inserting this into (3.7) and using (2.8), we obtain

$$|I_3| \leq Ch^{2\beta - 2\varepsilon} \|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \sup_{f \in \dot{H}^\beta} \|g''(f)\|$$

$$\times \int_0^T \|A^{\frac{1-\beta}{2}} E(t)\| \, \|A^{\frac{\beta}{2}-\varepsilon} E(t/2)\| \, \|A^{\frac{1}{2}} E(t/2)\| \, \mathrm{d}t$$

$$\leq Ch^{2\beta - 2\varepsilon} \|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \sup_{f \in \dot{H}^\beta} \|g''(f)\| \int_0^T t^{-1+\varepsilon} \mathrm{e}^{-3\alpha t} \, \mathrm{d}t.$$

By estimating the integral we conclude

$$|I_3| \leq C\varepsilon^{-1} h^{2\beta - 2\varepsilon} \|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \sup_{f \in \dot{H}^\beta} \|g''(f)\|.$$

Finally, we set $\varepsilon = 1/|\log(h)| < \beta/2$ (for $h \leq h_0$ with h_0 small enough) to get

$$\varepsilon^{-1} h^{2\beta - 2\varepsilon} = h^{2\beta} \mathrm{e}^{-2\varepsilon \log(h)} \varepsilon^{-1} = \mathrm{e}^{-2} h^{2\beta} |\log(h)|.$$

Thus, there are $C = C(\beta)$, $h_0 = h_0(\beta)$ independent of g, T and h such that

$$|I_3| \leq Ch^{2\beta} |\log(h)| \, \|A^{-\frac{1-\beta}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \sup_{f \in \dot{H}^\beta} \|g''(f)\|,$$

for $h \leq h_0$, $t \in [0, T]$. Together with (3.3) and (3.6) this completes the proof.

It is important to note that the strong assumptions (2.7), where $r \geq 3$, and (1.8) are only needed in the estimate of the second term I_2. For I_3 we only use (2.6) and

$$\sup_{f \in \dot{H}^\beta} \|g''(f)\| < \infty.$$

For I_1 we only use (2.6) and

$$\sup_{f \in \dot{H}^\beta} \|g'(f)\| < \infty$$

in order to get (3.4), while (2.6) and (1.8) are needed for (3.3).

References

[1] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.

[2] ———, *Second Order Partial Differential Equations in Hilbert Spaces*, London Mathematical Society Lecture Note Series, vol. 293, 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 and J. Printems, *Weak order for the discretization of the stochastic heat equation*, preprint 2007, `arXiv:0710.5450v1 [math.NA]`.

[5] 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.

[6] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), 829–847.

Technische Universität Darmstadt, Fachbereich Mathematik, Schlossgartenstr. 7, D–64289 Darmstadt, Germany

E-mail address: `geissert@mathematik.tu-darmstadt.de`

Department of Mathematics and Statistics, University of Otago, P.O. Box 56, Dunedin, New Zealand and Department of Mathematical Sciences, Chalmers University of Technology and Göteborg University, SE–412 96 Göteborg, Sweden

E-mail address: `mihaly@chalmers.se`

Department of Mathematical Sciences, Chalmers University of Technology and Göteborg University, SE–412 96 Göteborg, Sweden

E-mail address: `stig@chalmers.se`