BIT Numer Math (2009) 49: 343–356
DOI 10.1007/s10543-009-0227-y

BIT

Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise

Matthias Geissert · Mihály Kovács · Stig Larsson

Received: 22 January 2008 / Accepted: 15 March 2009 / Published online: 22 April 2009

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.

Keywords Finite element · Parabolic equation · Stochastic · Additive noise · Wiener process · Error estimate · Weak convergence

Mathematics Subject Classification (2000) 65M60 · 60H15 · 60H35 · 65C30

Communicated by Anders Szepessy.

M. Kovács and S. Larsson supported by the Swedish Research Council (VR). Part of this work was done at Institut Mittag-Leffler.
S. Larsson supported by the Swedish Foundation for Strategic Research (SSF) through GMMC, the Gothenburg Mathematical Modelling Centre.

M. Geissert
Technische Universität Darmstadt, Fachbereich Mathematik, Schlossgartenstr. 7, 64289 Darmstadt, Germany
e-mail: geissert@mathematik.tu-darmstadt.de

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

S. Larsson (✉)
Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Göteborg, Sweden
e-mail: stig@chalmers.se

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

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

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 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 [7],

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

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

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

where $E_h(t) = \mathrm{e}^{-tA_h}$ 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,

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

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

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

In [7] the following strong convergence result was proved:

Let $r \geq 2$ and $0 \leq \beta \leq r$ 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 $t \geq 0$,

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

(In [7] there is an unnecessary restriction $0 \leq \beta \leq 1$.)

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, twice Fréchet differentiable functions g whose first and second derivatives are continuous and bounded. To be more precise, 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 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 $g'' \in C(H, \mathcal{B}(H))$. Then 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. One may construct $g \in C_{\mathrm{b}}^2(H)$ by the following procedure: Take $v \in H$ and $G \in C^2(\mathbb{R}, \mathbb{R})$ such that

G' and G'' are bounded. Then g, defined by $g(u) := G(\langle v, u\rangle)$, is in $C_{\mathrm{b}}^2(H)$. Since $g'(u) = G'(\langle v, u\rangle)v$, (1.8) below is fulfilled if $v \in \dot{H}^\beta$. A more interesting example is provided by $g(u) = \langle G(u), v\rangle = \int_D G(u(\xi))v(\xi)\,\mathrm{d}\xi$, where $g'(u) = G'(u)v$ satisfies (1.8) if $v \in \dot{H}^\beta \cap L_\infty(D)$. A particular case is the linear functional $g(u) = \langle u, v\rangle$.

Our main result is the following.

Theorem 1.1 *Let X and X_h be given by* (1.3) *and* (1.5), *respectively. 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 $X_0 \in L_2(\Omega, \dot{H}^\beta)$. Assume that $g \in C_{\mathrm{b}}^2(H, \mathbb{R})$ and, in addition,*

$$\|g'(u)\|_{\dot{H}^\beta} \leq K(1 + \|u\|_{\dot{H}^\beta}), \quad u \in \dot{H}^\beta. \tag{1.8}$$

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

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

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 high-lighted:

- 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' (see also Remark 3.1). Indeed, after this work was completed we learned that in [5] the same rate of convergence (without logarithm) was proved without these assumptions. However, the estimate in [5] blows up as $T \to 0$, while ours is uniform in T (cf. (3.4) and (3.5) 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 [5]. In a more recent work, [4], a similar result is proved for a semilinear equation in 1-D.

In this light we may say that condition (1.8) means that our result has the flavor of superconvergence. For example, when applied with a linear functional $g(u) = \langle u, v\rangle$, $v \in \dot{H}^1$, and $\beta = 1$, it essentially coincides with the following result from [7]:

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

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, 5], we are only aware of [6], 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,

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

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

$$\|T\|_{\mathrm{HS}} = \|T^*\|_{\mathrm{HS}}, \qquad \|TS\|_{\mathrm{HS}} = \|S^*T^*\|_{\mathrm{HS}}, \qquad \|ST\|_{\mathrm{HS}} \le \|S\|\,\|T\|_{\mathrm{HS}}. \tag{2.2}$$

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,

$$\|A_h^{\gamma} u_h\| \le C h^{-2\gamma} \|u_h\|, \quad u_h \in S_h,\ \gamma \in [0,1]. \tag{2.3}$$

We also have the following equivalence of norms in S_h:

$$\|A^{\gamma} u_h\| \le C \|A_h^{\gamma} u_h\|, \quad u_h \in S_h,\ \gamma \in \left[0, \frac{1}{2}\right], \tag{2.4}$$

and (by the H^1-boundedness of P_h)

$$\|A_h^{\gamma} P_h u\| \le C \|A^{\gamma} u\|, \quad u \in \dot{H}^{\gamma},\ \gamma \in \left[0, \frac{1}{2}\right]. \tag{2.5}$$

Moreover, for $0 \le \delta \le \frac{1}{2}$, $\delta \le \rho \le 1$, we have

$$\|A^{\delta}(P_h - I)A^{-\rho}\| \le C h^{2\rho - 2\delta}. \tag{2.6}$$

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

$$\|(A^{-1} - A_h^{-1} P_h) f\| \le C h^s \|A^{\frac{s-2}{2}} f\|, \quad s \in [2,3]. \tag{2.7}$$

(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 $r \ge 3$.

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

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

and that

$$\int_0^T \|A^{\frac{1}{2}} E(t)v\|^2 \,\mathrm{d}t \le \frac{1}{2}\|v\|^2, \qquad \int_0^T \|A_h^{\frac{1}{2}} E_h(t) P_h v\|^2 \,\mathrm{d}t \le \frac{1}{2}\|v\|^2. \tag{2.9}$$

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

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

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

Let

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

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

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

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

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

Proof This is proved in [7, Theorem 3.1]; see also [1, Sect. 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*

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

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

$$\begin{aligned}
\|X_h(t)\|_{L_2(\Omega,\dot{H}^\beta)}^2 &\le C\mathbf{E}\big(\|A_h^{\frac{\beta}{2}} X_h(t)\|^2\big) \\
&\le C\mathbf{E}\big(\|A_h^{\frac{\beta}{2}} E_h(t) P_h X_0\|^2\big) + C\mathbf{E}\Big(\Big\| \int_0^t A_h^{\frac{\beta}{2}} E_h(t-s) P_h \,\mathrm{d}W(s)\Big\|^2\Big) \\
&= C\big(I_1 + I_2\big).
\end{aligned}$$

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

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

Using (2.2) and (2.9) we calculate

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

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

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

Finally, by (2.5),

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

The proof is complete. □

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

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

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

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

Therefore, by the law of double expectation,

$$\mathbf{E}\Big(u(\xi,t)\Big) = \mathbf{E}\Big(\mathbf{E}\Big(g(Z(T,t,\xi))\Big|\mathcal{F}_t\Big)\Big) = \mathbf{E}\Big(g(Z(T,t,\xi))\Big). \tag{2.14}$$

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

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

$$\partial_1^2 u(x,t) = \mathbf{E}\Big(E(T-t)g''(Z(T,t,x))E(T-t)\Big). \tag{2.16}$$

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

$$\begin{aligned} &\partial_2 u(x,t) - \langle Ax, \partial_1 u(x,t)\rangle + \frac{1}{2}\operatorname{Tr}\big(\partial_1^2 u(x,t)Q\big) = 0, \quad x \in \mathcal{D}(A),\ t \in (0,T), \\ &u(x,T) = g(x), \quad x \in H. \end{aligned} \tag{2.17}$$

3 Proof of Theorem 1.1

Let $u(x,t)$ be defined by (2.12) and apply (2.14) 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 (2.17) to replace the term $\partial_2 u$, we have at least formally

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

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, Chap. 9] and [2, Chap. 7]) we obtain (3.1) for u.

We therefore have

$$
\begin{aligned}
\mathbf{E}\Big(g(X_h(T)) - g(X(T))\Big) &= \mathbf{E}\Big(u(X_h(0), 0) - u(X(0), 0)\Big) \\
&\quad - \mathbf{E}\int_0^T \langle (A_h P_h - P_h A)\partial_1 u(X_h(t), t), X_h(t)\rangle \, \mathrm{d}t \\
&\quad + \frac{1}{2}\mathbf{E}\int_0^T \mathrm{Tr}\Big(\partial_1^2 u(X_h(t), t)(Q_h - Q)\Big)\, \mathrm{d}t \\
&= I_1 + I_2 + I_3.
\end{aligned} \tag{3.2}
$$

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

$$
\begin{aligned}
I_1 &= \mathbf{E}\bigg\{\int_0^1 \langle \partial_1 u(X_0 + s(P_h X_0 - X_0), 0), P_h X_0 - X_0\rangle \, \mathrm{d}s\bigg\} \\
&= \mathbf{E}\bigg\{\int_0^1 \langle (P_h - I)\partial_1 u(X_0 + s(P_h X_0 - X_0), 0), (P_h - I)X_0\rangle \, \mathrm{d}s\bigg\},
\end{aligned}
$$

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

$$
\begin{aligned}
|I_1| &\le \mathbf{E}\bigg\{\int_0^1 Ch^\beta \|\partial_1 u(X_0 + s(P_h X_0 - X_0), 0)\|_{\dot{H}^\beta} \cdot Ch^\beta \|X_0\|_{\dot{H}^\beta} \, \mathrm{d}s\bigg\} \\
&\le Ch^{2\beta} \int_0^1 \|\partial_1 u(X_0 + s(P_h X_0 - X_0), 0)\|_{L_2(\Omega,\dot{H}^\beta)} \|X_0\|_{L_2(\Omega,\dot{H}^\beta)} \, \mathrm{d}s \\
&\le Ch^{2\beta} \|X_0\|_{L_2(\Omega,\dot{H}^\beta)} \sup_{s\in[0,1]} \|\partial_1 u(X_0 + s(P_h X_0 - X_0), 0)\|_{L_2(\Omega,\dot{H}^\beta)}.
\end{aligned}
$$

Now, by (1.8), (2.13), and (2.15),

$$
\begin{aligned}
&\|\partial_1 u(X_0 + s(P_h X_0 - X_0), 0)\|^2_{L_2(\Omega,\dot{H}^\beta)} \\
&\quad = \mathbf{E}\|\partial_1 u(X_0 + s(P_h X_0 - X_0), 0)\|^2_{\dot{H}^\beta} \\
&\quad = \mathbf{E}\{\|\mathbf{E}\left(E(T) g'(Z(T, 0, X_0 + s(P_h X_0 - X_0)))|\mathcal{F}_0\right)\|^2_{\dot{H}^\beta}\} \\
&\quad \le \mathbf{E}\{\|E(T) g'(Z(T, 0, X_0 + s(P_h X_0 - X_0)))\|^2_{\dot{H}^\beta}\}
\end{aligned}
$$

$$\leq C(1+\mathbf{E}\{\|Z(T,0,X_0+s(P_hX_0-X_0))\|^2_{\dot{H}^\beta}\})$$

$$\leq C(1+\|X_0+s(P_hX_0-X_0)\|^2_{L_2(\Omega,\dot{H}^\beta)}+\|A^{-\frac{1-\beta}{2}}Q^{\frac{1}{2}}\|^2_{\mathrm{HS}}), \tag{3.3}$$

where we used Lemma 2.1 for the last inequality. Thus,

$$\sup_{s\in[0,1]}\|\partial_1 u(X_0+s(P_hX_0-X_0),0)\|_{L_2(\Omega,\dot{H}^\beta)}$$

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

and therefore,

$$|I_1|\leq Ch^{2\beta}\|X_0\|_{L_2(\Omega,\dot{H}^\beta)}. \tag{3.4}$$

We remark that by using, instead, (2.8) with $\gamma=\beta/2$, $\delta=0$, in (3.3), we would get

$$|I_1|\leq Ch^{2\beta}T^{-\beta/2}\sup_{f\in\dot{H}^\beta}\|g'(f)\|\,\|X_0\|_{L_2(\Omega,\dot{H}^\beta)}, \tag{3.5}$$

where the assumption (1.8) is not needed.

For the second term, we use the identity

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

to get

$$\begin{aligned}
I_2&=-\mathbf{E}\int_0^T\langle(A_hP_h-P_hA)\partial_1 u(X_h(t),t),X_h(t)\rangle\,\mathrm{d}t\\
&=-\mathbf{E}\int_0^T\langle A_hP_h(A^{-1}-A_h^{-1}P_h)A\partial_1 u(X_h(t),t),X_h(t)\rangle\,\mathrm{d}t\\
&=-\mathbf{E}\int_0^T\langle A_h^{1-\frac{\beta}{2}}P_h(A^{-1}-A_h^{-1}P_h)A\partial_1 u(X_h(t),t),A_h^{\frac{\beta}{2}}X_h(t)\rangle\,\mathrm{d}t.
\end{aligned}$$

Therefore,

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

Using (2.15) (see also (2.13)) and (1.8) we have here

$$\begin{aligned}
&\|A_h^{1-\frac{\beta}{2}}P_h(A^{-1}-A_h^{-1}P_h)A\partial_1 u(X_h(t),t)\|_{L_2(\Omega,H)}\\
&\quad=\|A_h^{1-\frac{\beta}{2}}P_h(A^{-1}-A_h^{-1}P_h)A\mathbf{E}\{E(T-t)g'(Z(T,t,X_h(t)))|\mathcal{F}_t\}\|_{L_2(\Omega,H)}
\end{aligned}$$

$$\leq \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(T-t)\mathbf{E}\{A^{\frac{\beta}{2}} g'(Z(T,t,X_h(t)))|\mathcal{F}_t\}\|_{L_2(\Omega,H)}$$
$$\leq \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(T-t)\|\, K(1 + \|Z(T,t,X_h(t))\|_{L_2(\Omega,\dot{H}^\beta)}).$$

Lemma 2.2 provides a bound $\|X_h(t)\|_{L_2(\Omega,\dot{H}^\beta)} \leq C$, which inserted into Lemma 2.1 leads to a bound for $\|Z(T,t,X_h(t))\|_{L_2(\Omega,\dot{H}^\beta)}$. Therefore,

$$|I_2| \leq C \int_0^T \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\|\, \mathrm{d}t. \tag{3.6}$$

In view of (2.3) we have

$$\begin{aligned}
&\int_0^T \|A_h^{1-\frac{\beta}{2}} P_h(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\|\, \mathrm{d}t \\
&\quad \leq Ch^{-2+\beta} \int_0^T \|(A^{-1} - A_h^{-1}P_h)A^{1-\frac{\beta}{2}} E(t)\|\, \mathrm{d}t.
\end{aligned} \tag{3.7}$$

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

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

for $h \leq h_0$ (C and h_0 depend on β). Inserting this into (3.7), (3.6) shows that

$$|I_2| \leq Ch^{2\beta}|\log(h)|. \tag{3.8}$$

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

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

Therefore,

$$\begin{aligned}
2I_3 &= \mathbf{E} \int_0^T \mathrm{Tr}\Big(\partial_1^2 u(X_h(t),t)(Q_h - Q)\Big)\, \mathrm{d}t \\
&= \mathbf{E} \int_0^T \mathrm{Tr}\Big(\partial_1^2 u(X_h(t),t)\{(P_h - I) Q P_h + Q(P_h - I)\}\Big)\, \mathrm{d}t.
\end{aligned}$$

Note that, since $E(t)$ is trace class for $t>0$, it follows from (2.16) 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

$$\begin{aligned}\operatorname{Tr}\Big(\partial_1^2 u(X_h(t),t)(P_h-I)\,(QP_h)\Big) &= \operatorname{Tr}\Big(\,(QP_h)^*\,(P_h-I)^*\,\partial_1^2 u(X_h(t),t)^*\Big)\\ &= \operatorname{Tr}\Big(\{P_hQ(P_h-I)\}\{\partial_1^2 u(X_h(t),t)\}\Big)\\ &= \operatorname{Tr}\Big(\partial_1^2 u(X_h(t),t)P_hQ(P_h-I)\Big).\end{aligned}$$

Therefore,

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

We bound the integrand. By (2.16) we have

$$\begin{aligned}&\Big|\operatorname{Tr}\Big(\partial_1^2 u(X_h(t),t)(P_h+I)Q(P_h-I)\Big)\Big|\\ &\quad\le \sup_{x\in\dot H^\beta}\Big|\operatorname{Tr}\Big(\partial_1^2 u(x,t)(P_h+I)Q(P_h-I)\Big)\Big|\\ &\quad= \sup_{x\in\dot H^\beta}\Big|\operatorname{Tr}\Big(\mathbf{E}\big\{E(T-t)g''(Z(T,t,x))E(T-t)\big\}(P_h+I)Q(P_h-I)\Big)\Big|\\ &\quad\le \mathbf{E}\Big\{\sup_{x\in\dot H^\beta}\Big|\operatorname{Tr}\Big(E(T-t)g''(Z(T,t,x))E(T-t)(P_h+I)Q(P_h-I)\Big)\Big|\Big\}\\ &\quad\le \sup_{f\in\dot H^\beta}\Big|\operatorname{Tr}\Big(E(T-t)g''(f)E(T-t)(P_h+I)Q(P_h-I)\Big)\Big|,\end{aligned}$$

so that

$$2|I_3|\le \mathbf{E}\int_0^T \sup_{f\in\dot H^\beta}\Big|\operatorname{Tr}\Big(E(t)g''(f)E(t)(P_h+I)Q(P_h-I)\Big)\Big|\,\mathrm{d}t. \tag{3.9}$$

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

$$\begin{aligned}&\operatorname{Tr}\Big(E(t)g''(f)E(t)(P_h+I)Q(P_h-I)\Big)\\ &\quad=\operatorname{Tr}\Big(A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\\ &\qquad\times A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h+I)Q(P_h-I)\Big)\\ &\quad=\operatorname{Tr}\Big(A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h+I)Q(P_h-I)\\ &\qquad\times A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\Big)\end{aligned}$$

$$= \mathrm{Tr}\Big(A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h+I)Q^{\frac12} \times Q^{\frac12}(P_h-I)A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\Big),$$

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

$$\begin{aligned}
&\Big|\mathrm{Tr}\Big(E(t)g''(f)E(t)(P_h+I)Q(P_h-I)\Big)\Big| \\
&\quad\le \|A^{\frac{\beta}{2}-\varepsilon}E(t/2)g''(f)E(t)(P_h+I)Q^{\frac12}\|_{\mathrm{HS}}\,\|Q^{\frac12}(P_h-I)A^{-\frac{\beta}{2}+\varepsilon}E(t/2)\|_{\mathrm{HS}} \\
&\quad= J_1 \times J_2.
\end{aligned}$$

For the first factor we use (2.2) to get

$$\begin{aligned}
J_1 &= \|\{A^{\frac{\beta}{2}-\varepsilon}E(t/2)\}\{g''(f)E(t)(P_h+I)Q^{\frac12}\}\|_{\mathrm{HS}} \\
&= \|\{Q^{\frac12}(P_h+I)E(t)g''(f)\}\{A^{\frac{\beta}{2}-\varepsilon}E(t/2)\}\|_{\mathrm{HS}} \\
&\le \|Q^{\frac12}(P_h+I)E(t)\|_{\mathrm{HS}}\,\|g''(f)\|\,\|A^{\frac{\beta}{2}-\varepsilon}E(t/2)\|.
\end{aligned}$$

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

$$\begin{aligned}
&\|Q^{\frac12}(P_h+I)E(t)\|_{\mathrm{HS}} \\
&\quad= \|Q^{\frac12}A^{-\frac{1-\beta}{2}}A^{\frac{1-\beta}{2}}(P_h+I)A^{-\frac{1-\beta}{2}}A^{\frac{1-\beta}{2}}E(t)\|_{\mathrm{HS}} \\
&\quad\le \|Q^{\frac12}A^{-\frac{1-\beta}{2}}\|_{\mathrm{HS}}\,\|A^{\frac{1-\beta}{2}}(P_h+I)A^{-\frac{1-\beta}{2}}\|\,\|A^{\frac{1-\beta}{2}}E(t)\| \\
&\quad\le C\|A^{-\frac{1-\beta}{2}}Q^{\frac12}\|_{\mathrm{HS}}\,\|A^{\frac{1-\beta}{2}}E(t)\|.
\end{aligned}$$

Therefore,

$$J_1 \le C\|A^{-\frac{1-\beta}{2}}Q^{\frac12}\|_{\mathrm{HS}}\,\|g''(f)\|\,\|A^{\frac{1-\beta}{2}}E(t)\|\,\|A^{\frac{\beta}{2}-\varepsilon}E(t/2)\|.$$

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

$$\begin{aligned}
J_2 &= \|Q^{\frac12}A^{-\frac{1-\beta}{2}}A^{\frac{1-\beta}{2}}(P_h-I)A^{-\frac{1+\beta}{2}+\varepsilon}A^{\frac12}E(t/2)\|_{\mathrm{HS}} \\
&\le \|Q^{\frac12}A^{-\frac{1-\beta}{2}}\|_{\mathrm{HS}}\,\|A^{\frac{1-\beta}{2}}(P_h-I)A^{-\frac{1+\beta}{2}+\varepsilon}\|\,\|A^{\frac12}E(t/2)\|.
\end{aligned}$$

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

$$J_2 \le Ch^{2\beta-2\varepsilon}\|A^{-\frac{1-\beta}{2}}Q^{\frac12}\|_{\mathrm{HS}}\,\|A^{\frac12}E(t/2)\|.$$

Inserting this into (3.9) and using (2.8), we obtain

$$|I_3| \le Ch^{2\beta-2\varepsilon}\|A^{-\frac{1-\beta}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 \sup_{f\in\dot H^\beta}\|g''(f)\|$$

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

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

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

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

for $h \leq h_0$, $t \in [0, T]$. Together with (3.4) and (3.8) this completes the proof.

Remark 3.1 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 (3.5), while (2.6) and (1.8) are needed for (3.4).

References

1. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Encyclopedia of Mathematics and Its Applications, vol. 44. Cambridge University Press, Cambridge (1992)
2. Da Prato, G., Zabczyk, J.: Second Order Partial Differential Equations in Hilbert Spaces. London Mathematical Society Lecture Note Series, vol. 293. Cambridge University Press, Cambridge (2002)
3. de Bouard, A., Debussche, A.: Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation. Appl. Math. Optim. **54**, 369–399 (2006)
4. Debussche, A.: Weak approximation of stochastic partial differential equations: the non linear case. Preprint 2008, arXiv:0804.1304v1 [math.NA]
5. Debussche, A., Printems, J.: Weak order for the discretization of the stochastic heat equation. Math. Comput. **78**, 845–863 (2009)
6. Hausenblas, E.: Weak Approximation for Semilinear Stochastic Evolution Equations. Stochastic Analysis and Related Topics VIII, Progr. Probab., vol. 53, pp. 111–128. Birkhäuser, Basel (2003)
7. Yan, Y.: Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise. BIT **44**, 829–847 (2004)