J. Math. Anal. Appl. 413 (2014) 939–952

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

Weak convergence of a fully discrete approximation of a linear stochastic evolution equation with a positive-type memory term

Mihály Kovács [a,*], Jacques Printems [b]

[a] Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin, 9054, New Zealand
[b] Laboratoire d'Analyse et de Mathématiques Appliquées, CNRS UMR 8050, Université Paris-Est, 61, avenue du Général de Gaulle, 94 010 Créteil, France

ARTICLE INFO

Article history:
Received 5 July 2013
Available online 17 December 2013
Submitted by X. Zhang

Keywords:
Volterra equations
Euler scheme
Stochastic partial differential equations
Finite element method
Weak convergence
Convolution quadrature
Fractional differential equations

ABSTRACT

In this paper we are interested in the numerical approximation of the marginal distributions of the Hilbert space valued solution of a stochastic Volterra equation driven by an additive Gaussian noise. This equation can be written in the abstract Itô form as

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

where W^Q is a Q-Wiener process on the Hilbert space H and where the time kernel b is the locally integrable potential $t^{\rho-2}$, $\rho \in (1,2)$, or slightly more general. The operator A is unbounded, linear, self-adjoint, and positive on H. Our main assumption concerning the noise term is that $A^{(\nu-1/\rho)/2}Q^{1/2}$ is a Hilbert–Schmidt operator on H for some $\nu \in [0,1/\rho]$. The numerical approximation is achieved via a standard continuous finite element method in space (parameter h) and an implicit Euler scheme and a Laplace convolution quadrature in time (parameter $\Delta t = T/N$). We show that for $\varphi : H \to \mathbb{R}$ twice continuously differentiable test function with bounded second derivative,

$$\left|\mathbf{E}\varphi(X_h^N) - \mathbf{E}\varphi(X(T))\right| \leqslant C \ln\left(\frac{T}{h^{2/\rho} + \Delta t}\right)\left(\Delta t^{\rho\nu} + h^{2\nu}\right),$$

for any $0 \leqslant \nu \leqslant 1/\rho$. This is essentially twice the rate of strong convergence under the same regularity assumption on the noise.

* Corresponding author.
E-mail addresses: mkovacs@maths.otago.ac.nz (M. Kovács), printems@u-pec.fr (J. Printems).

0022-247X/$ – see front matter
http://dx.doi.org/10.1016/j.jmaa.2013.12.034

1. Introduction

Let $H = L^2(\mathcal{O})$ be the real separable Hilbert space of square integrable functions on some convex polygonal domain $\mathcal{O}$ of $\mathbb{R}^d$, $d \geqslant 1$, equipped with the usual inner product denoted by $(\cdot,\cdot)$ and induced norm $\|\cdot\|$. For $T > 0$, we consider the following stochastic Volterra type equation written in the abstract Itô form as

$$\mathrm{d}X(t) + \left(\int_0^t b(t-s)AX(s)\,\mathrm{d}s\right)\mathrm{d}t = \mathrm{d}W^Q(t), \quad t \in (0,T]; \qquad X(0) = X_0, \tag{1.1}$$

where $-A = \Delta$ is the Dirichlet Laplacian with domain $D(A) = H^2(\mathcal{O}) \cap H_0^1(\mathcal{O})$ and W^Q is an H-valued Wiener process on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ endowed with the normal filtration $\{\mathcal{F}_t\}_{t\geqslant 0}$ generated by W^Q with possibly unbounded covariance operator Q. We note that, strictly speaking, the process W^Q is H-valued if and only if Q is a trace class operator. The initial condition $X(0) = X_0$ is H-valued and $\mathcal{F}_0$-measurable. The convolution kernel b is given by

$$b(t) = t^{\rho-2}/\Gamma(\rho-1), \quad 1 < \rho < 2, \tag{1.2}$$

or could be somewhat more general which is made precise later in Section 2. Such equations are called stochastic Volterra equations and can be used in the modeling of diffusion of heat in materials with memory or in viscoelasticity (see [3,15] and references therein). Because of the weak singularity of the kernel b at 0, the deterministic equation exhibit certain smoothing characteristics similar to that of parabolic type evolution equations.

We study the numerical approximation of $\{X(t)\}_{t\in[0,T]}$ by an implicit Euler scheme and a Laplace transform convolution quadrature in time together with a finite element method in space. Let $N \geqslant 1$ and $\Delta t = T/N$. We set $t_n = n\Delta t$, $n = 0,\dots,N$. Let $\{\mathcal{T}_h\}_{0<h<1}$ denote a family of triangulations of $\mathcal{O}$, with mesh size $h > 0$ and consider finite element spaces $\{V_h\}_{0<h<1}$, where $V_h \subset H_0^1(\mathcal{O})$ consists of continuous piecewise linear functions vanishing at the boundary of $\mathcal{O}$. Let $X_h^n \in V_h$ be the numerical approximation of $X(t_n)$ defined via the difference equations

$$\left(X_h^n, v_h\right) - \left(X_h^{n-1}, v_h\right) + \Delta t \sum_{k=1}^{n} \omega_{n-k}\left(\nabla X_h^k, \nabla v_h\right) = \left(w^n, v_h\right), \tag{1.3}$$

for any $n \geqslant 1$, with the initial condition

$$\left(X_h^0, v_h\right) = (X_0, v_h),$$

for any $v_h \in V_h$, where we have set $w^n = W^Q(t_n) - W^Q(t_{n-1})$.

Our specific choice of the weights $\{\omega_k\}_{k\geqslant 0}$ is stemming from the deterministic framework of [12,13]. Indeed, it can be easily seen that most of the qualitative properties of the solution of (1.1) with $Q = 0$ depend heavily on the way the frequencies of the time kernel b are distributed. For example, in the case where b is a Dirac mass at 0, we formally recover the heat equation and if b is regular enough, we recover the wave equation. For that reason, the weights $\{\omega_k\}_{k\geqslant 0}$ in (1.3) have been chosen such as to mimic, at the level of the backward Euler scheme, the spectral properties of the time kernel b. Using the Laplace transform $\widehat{b}$ of b, these weights can be obtained via the relation

$$\widehat{b}\left(\frac{1-z}{\Delta t}\right) = \sum_{k\geqslant 0} \omega_k z^k, \quad |z| < 1. \tag{1.4}$$

Introducing the “discrete Laplacian”

$$A_h : V_h \to V_h, \qquad (A_h\psi, \chi) = (\nabla\psi, \nabla\chi), \quad \psi, \chi \in V_h, \tag{1.5}$$

and the orthogonal projector

$$P_h : H \to V_h, \qquad (P_h f, \chi) = (f, \chi), \quad \chi \in V_h,$$

we rewrite (1.3) in the operator form as

$$X_h^n - X_h^{n-1} + \Delta t\left(\sum_{k=1}^{n} \omega_{n-k}\, A_h X_h^k\right) = P_h w^n, \quad n \geqslant 1, \tag{1.6}$$

with $X_h^0 = P_h X_0$.

If φ is a twice differentiable real functional on $L^2(\mathcal{O})$, not necessarily bounded and with not necessarily bounded first derivative but with bounded second derivative and $A^{(\nu-\frac{1}{\rho})/2}Q^{1/2}$ is a Hilbert–Schmidt operator on H, then our main result can be stated as follows. Denoting the expectation by $\mathbf{E}$, the so-called weak error can be estimated as

$$\left|\mathbf{E}\varphi\left(X_h^N\right) - \mathbf{E}\varphi\left(X(T)\right)\right| \leqslant C \ln\left(\frac{T}{h^{2/\rho} + \Delta t}\right)\left(\Delta t^{\rho\nu} + h^{2\nu}\right), \tag{1.7}$$

where C may depend on T, φ and the initial condition $X(0)$ but not on h and N. Hence even in the presence of memory, similarly to parabolic and hyperbolic stochastic equations [1,5,6,8,9,11,19], the weak order is essentially twice the strong order, where the latter was studied in [10]. Note that in (1.7) we may even allow the test function $\varphi(x) = \|x\|^2$.

In particular, when $Q = I$ (white noise), ν has to be chosen in such a way that $A^{(\nu-1/\rho)/2}$ is a Hilbert–Schmidt operator on H. Taking the asymptotics of the eigenvalues of the Laplacian into account we must have $\nu < 1/\rho - d/2$. This yields $d = 1$ and a rate of convergence in time of $(1 - \rho/2)_-$ and in space of $(2/\rho - 1)_-$.

A popular method used in the study the weak convergence of approximations of stochastic equations relies on the associated Kolmogorov equation (see [7,16,18]). The global weak error $\mathbf{E}\varphi(X_h^N) - \mathbf{E}\varphi(X(T))$ is usually decomposed in a sum of local weak errors which are then expressed using the solution of the associated Kolmogorov equation.

Unfortunately, the presence of a non-local term in time in (1.1) prevents us to use the same method directly because the process $\{X(t)\}_{t\in[0,T]}$ is not Markovian and hence no Kolmogorov equation is associated with (1.1). However, since the equation is linear and the non-local term is in the drift part, we use the same kind of method as in [5,8,9,11] to remove the drift and obtain an equation which has a Markovian solution. Hence there is an associated Kolmogorov's equation with no drift but with a time-dependent covariance operator.

The outline of the paper is as follows. In Section 2 we introduce basic notations and the main assumptions on b, together with the existence, uniqueness and regularity properties of $\{X(t)\}_{t\in[0,T]}$. In Section 3 we recall and prove some deterministic estimates concerning the solutions of (1.1) and (1.6) with $Q = 0$. In particular, we recall the discrete mild formulation of (1.6) and establish its regularizing properties (Theorem 3.1). In Section 4 we introduce the results which are needed in order to accommodate random initial data and unbounded test functions φ. This section ends with a representation formula of the weak error (Theorem 4.3) which symmetrize the role played by the discrete and the continuous solutions. Finally, in Section 5, we state and prove the main convergence result (Theorem 5.1).

2. Preliminaries

In this section we introduce notation and collect some preliminary results. We also state the hypothesis on the convolution kernel b. We denote the set of bounded linear operators on H by $\mathcal{B}(H)$ endowed with the usual norm $\|\cdot\|_{\mathcal{B}(H)}$, where we drop the subscript from the norm if it is clear from the context. Let HS denotes the *Hilbert–Schmidt* operators on H; that is, $T \in \mathrm{HS}$ if T is linear and for an orthonormal basis (ONB) $\{e_k\}_{k\in\mathbb{N}}$ of H

$$\|T\|_{\mathrm{HS}}^2 = \sum_{k\in\mathbb{N}} \|Te_k\|^2 < \infty.$$

In this case the sum is independent of the ONB. If a linear operator T on H can be written as

$$Tx = \sum_{k\in\mathbb{N}} (x, a_k)b_k, \quad x \in H,$$

with $\sum_{k\in\mathbb{N}} \|a_k\|\|b_k\| < \infty$, then T is called a *trace-class* operator. The trace-norm of T is then defined as

$$\|T\|_{\mathrm{Tr}} = \inf\left\{\sum_{k\in\mathbb{N}} \|a_k\|\|b_k\|\colon\ Tx = \sum_{k\in\mathbb{N}} (x, a_k)b_k,\ x \in H\right\}.$$

If T is a trace class operator then for any ONB, the trace of T is defined as

$$\mathrm{Tr}(T) = \sum (Te_k, e_k)$$

is finite and the sum is independent of the ONB. Both Hilbert–Schmidt and trace class operators are bounded. If $T \geqslant 0$ is a symmetric trace class operator then, $\mathrm{Tr}(T) = \|T\|_{\mathrm{Tr}}$. It is well-known that if $T_1, T_2 \in \mathrm{HS}$, then T_1T_2 is trace class and

$$\bigl|\mathrm{Tr}(T_1T_2)\bigr| \leqslant \|T_1\|_{\mathrm{HS}}\|T_2\|_{\mathrm{HS}}. \tag{2.1}$$

Furthermore, if $T \in \mathrm{HS}$ and $S \in \mathcal{B}(H)$, then TS and ST are in HS and

$$\max\bigl\{\|TS\|_{\mathrm{HS}}, \|ST\|_{\mathrm{HS}}\bigr\} \leqslant \|T\|_{\mathrm{HS}}\|S\|_{\mathcal{B}(H)}. \tag{2.2}$$

For $1 \leqslant p < \infty$ we denote by $L^p(\Omega, H)$ the space of H-valued random variables X such that

$$\|X\|_{L^p(\Omega,H)} = \bigl(\mathbf{E}\|X\|^p\bigr)^{1/p} < \infty.$$

It is well known that our assumptions on A and on the spatial domain $\mathcal{O}$ implies the existence of a sequence of nondecreasing positive real numbers $\{\lambda_k\}_{k\geqslant 1}$ and an orthonormal basis $\{e_k\}_{k\geqslant 1}$ of H such that

$$Ae_k = \lambda_k e_k, \qquad \lim_{k\to+\infty} \lambda_k = +\infty. \tag{2.3}$$

Next, we define, by means of the spectral decomposition of A, the fractional powers A^s of A for $s \in \mathbb{R}$. That is, for $s > 0$ we set

$$A^s x = \sum_{k\geqslant 1} \lambda_k^s (x, e_k)e_k, \tag{2.4}$$

with domain $D(A^s)$ being all $x \in H$ for which the sum converges in H. In particular, $D(A^0) = H$. For $s < 0$ we define $A^s x$ as in (2.4) for all $x \in H$ and $D(A^s)$ to be the completion of H with respect to the norm of $\|x\|_s = \|A^s x\|$.

Next we state the main hypothesis on the convolution kernel b.

Assumption 1. The kernel $0 \neq b \in L^1_{loc}(\mathbb{R}_+)$, is 3-monotone; that is, b, $-\dot{b}$ are nonnegative, nonincreasing, convex, and $\lim_{t\to\infty} b(t) = 0$. Furthermore,

$$\rho := 1 + \frac{2}{\pi} \sup\bigl\{\bigl|\arg \widehat{b}(\lambda)\bigr|,\ \operatorname{Re}\lambda > 0\bigr\} \in (1,2). \tag{2.5}$$

In the special case of the Riesz kernel given in (1.2) one can easily show that ρ in the exponent coincides with the one defined in (2.5). In order to obtain non-smooth data estimates for the deterministic equation we need the following additional hypothesis.

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

In the sequel we discuss properties of the solution of (1.1). The weak solution of (1.1) is a mean-square continuous H-valued process satisfying

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

for all $\eta \in D(A^*)$ almost surely for all $t \in [0, T]$. Under Assumption 1, if $W^Q \equiv 0$; that is, the deterministic case, then there exists a resolvent family $\{S(t)\}_{t\geqslant 0} \subset \mathcal{B}(H)$ which is strongly continuous for $t \geqslant 0$, differentiable for $t > 0$ and uniformly bounded by 1, see [17, Corollary 1.2 and Corollary 3.3]. The unique weak solution in the deterministic case is given by $X(t) = S(t)X_0$, $t \in [0, T]$.

The next result, which can be found in [3] and [10] summarizes the existence, uniqueness and regularity of weak solutions of (1.1).

Proposition 2.1. *Let b satisfy Assumption 1 and let $\|A^{(\nu-\frac{1}{\rho})/2} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ and $A^{\frac{\nu}{2}} X_0 \in L^2(\Omega, H)$ for some $\nu \geqslant 0$. Then (1.1) has a unique weak solution given by the variation of constants formula*

$$X(t) = S(t)X_0 + \int_0^t S(t-s)\,\mathrm{d}W^Q(s), \quad t \geqslant 0, \tag{2.6}$$

with $\|A^{\frac{\nu}{2}} X(t)\|_{L^2(\Omega,H)} \leqslant C$, for some $C > 0$ and for all $t \geqslant 0$. Furthermore, X has a version which is Hölder continuous of order less than $\min(\frac{1}{2}, \frac{\rho\nu}{2})$.

Remark 2.1. In particular, the stochastic integral in (2.6) makes sense since $SQ^{1/2} \in L^2((0,T), \mathrm{HS})$ (see the proof of [10, Theorem 3.6]).

3. Deterministic estimates

In the following proposition we collect some of the smoothing properties of $\{S(t)\}_{t\geqslant 0}$ and $\{S_h(t)\}_{t\geqslant 0}$, where $\{S_h(t)\}_{t\geqslant 0} \subset V_h$ is the resolvent family of the deterministic equation

$$\dot{u}_h(t) + \int_0^t b(t-s)A_h u_h(s)\,\mathrm{d}s = 0, \quad t > 0, \qquad u_h(0) = P_h u_0.$$

We would like to note that in this paper the constant C denotes a generic nonnegative constant that does not depend on the parameters $h, k, t, \Delta t$ and is not necessarily the same at every occurrence.

Proposition 3.1. *If b satisfies Assumption 1, then the following estimates hold for the resolvent families $\{S(t)\}_{t\geqslant 0}$ and $\{S_h(t)\}_{t\geqslant 0}$ for some $C>0$.*

(i) $\max\{\|A^{\nu/2}S(t)\|, \|A_h^{\nu/2}S_h(t)P_h\|\} \leqslant Ct^{-\nu\rho/2}$, $0\leqslant \nu \leqslant 2/\rho$, $t>0$, $h>0$;
(ii) $\|\dot{S}(t)\| \leqslant Ct^{-1}$, $t>0$;
(iii) $\int_0^t \|A_h^{1/(2\rho)}S_h(s)P_h x\|^2\,\mathrm{d}s \leqslant C\|x\|^2$, $t>0$, $h>0$.

Proof. The statements for $\{S(t)\}_{t\geqslant 0}$ are shown in [10, Proposition 2.5]. The estimate for $\{S_h(t)\}_{t\geqslant 0}$ in (i) can be proved exactly the same way while (iii) is shown in the proof of [10, Lemma 3.1.]. □

In the sequel we derive the relevant deterministic error estimates. Using the z-transform, it is shown [10] that the solution X_h^n of (1.6) can be written using a discrete constant variation formula as

$$X_h^n = B_{h,n}P_h x + \sum_{k=0}^{n-1} B_{h,n-k}P_h w_{k+1}, \tag{3.1}$$

where, $B_{h,0} = I$ and

$$B_{h,k}P_h x = \int_0^\infty S_h(\Delta t s)P_h x \frac{e^{-s}s^{k-1}}{(k-1)!}\,\mathrm{d}s \quad \text{for } k \geqslant 1. \tag{3.2}$$

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

$$\tilde{B}_{h,N}(t) := B_{h,\sigma(t)}P_h, \quad 0 \leqslant t \leqslant T.$$

Theorem 3.1. *If b satisfies Assumptions 1 and 2, then the following estimates hold for some $C>0$ where $E_{h,N}(t) = \tilde{B}_{h,N}(t) - S(t)$, $N\Delta t = T$ and $h>0$.*

$$\big\|S(t) - S(s)\big\|_{\mathcal{B}(H)} \leqslant Cs^{-\alpha}|t-s|^\alpha, \quad 0\leqslant \alpha \leqslant 1,\ 0<s\leqslant t; \tag{3.3}$$

$$\big\|A^{\frac{\nu}{2}}\tilde{B}_{h,N}(t)\big\| \leqslant Ct^{-\rho\nu/2}, \quad 0 \leqslant \nu \leqslant \frac{1}{\rho},\ 0<t\leqslant T; \tag{3.4}$$

$$\big\|E_{h,N}(t)\big\| \leqslant Ct^{-\rho\nu}\big(\Delta t^{\rho\nu} + h^{2\nu}\big), \quad 0\leqslant \nu \leqslant \frac{1}{\rho},\ 0<t\leqslant T; \tag{3.5}$$

$$\big\|A^{\frac{1/\rho-\nu}{2}}E_{h,N}(t)\big\| \leqslant Ct^{-\frac{1}{2}-\frac{\rho\nu}{2}}\big(\Delta t^{\rho\nu} + h^{2\nu}\big), \quad 0\leqslant\nu\leqslant\frac{1}{\rho},\ 0<t\leqslant T. \tag{3.6}$$

Proof. It follows from [17, Corollary 3.3] that $\|\dot{S}(t)x\| \leqslant Ct^{-1}\|x\|$ for all $x \in H$ and $t>0$. Thus, for $0<s\leqslant t$, we have

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

Since we also have that $\|S(t)x - S(s)x\| \leqslant 2\|x\|$, the inequality in (3.3) follows.

To show (3.4), first note that it follows from (3.2) and Proposition 3.1 (i) with $\nu = 0$ that

$$\|B_{h,k}\| \leqslant C, \quad \text{for all } k \geqslant 1,\ h > 0. \tag{3.7}$$

From (3.2) and Proposition 3.1(i) with $\nu = 2/\rho$ we conclude that, for $k \geqslant 2$ and $h > 0$,

$$\begin{aligned}
\left\|A_h^{1/\rho} B_{h,k} P_h x\right\| &\leqslant C\|x\|(\Delta t)^{-1} \int_0^\infty \frac{e^{-s} s^{k-2}}{(k-1)!}\,\mathrm{d}s \\
&= C\|x\|\big((k-1)\Delta t\big)^{-1} \int_0^\infty \frac{e^{-s} s^{k-2}}{(k-2)!}\,\mathrm{d}s \\
&= C\|x\| t_{k-1}^{-1} = C\|x\| \frac{k}{k-1} t_k^{-1} \leqslant C\|x\| t_k^{-1}.
\end{aligned}$$

For $k = 1$, by Hölder's inequality and Proposition 3.1(iii), we have

$$\begin{aligned}
\left\|A_h^{1/(2\rho)} B_{h,1} P_h x\right\| &\leqslant \int_0^\infty \left\|A_h^{1/(2\rho)} S_h(\Delta t s) x\right\| e^{-s}\,\mathrm{d}s \\
&\leqslant C\left(\int_0^\infty \left\|A_h^{1/(2\rho)} S_h(\Delta t s) P_h x\right\|^2 \mathrm{d}s\right)^{1/2} \\
&\leqslant C(\Delta t)^{-1/2}\|x\|.
\end{aligned} \tag{3.8}$$

By interpolation, using (3.7), (3.8) and (3.8) we conclude that

$$\left\|A_h^{\frac{\nu}{2}} B_{h,k} P_h\right\| \leqslant C t_k^{-\frac{\rho\nu}{2}}, \quad 0 \leqslant \nu \leqslant \frac{1}{\rho},\ k \geqslant 1,\ h > 0.$$

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

$$\left\|A^{\frac{\nu}{2}} B_{h,k} P_h\right\| \leqslant C t_k^{-\frac{\rho\nu}{2}}, \quad 0 \leqslant \nu \leqslant \frac{1}{\rho},\ k \geqslant 1,\ h > 0.$$

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

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

and the proof of (3.4) is complete.

Next we prove (3.5). First, we write

$$\left\|B_{h,k} P_h - S(t_k)\right\| \leqslant \left\|B_{h,k} - S_h(t_k)\right\| + \left\|S_h(t_k) - S(t_k)\right\| := e_1 + e_2.$$

It is show in [14] that if b satisfies Assumption 2, then

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

Furthermore, we also have that $\max\{e_1, e_2\} \leqslant C$ by Proposition 3.1 with $\nu = 0$, and thus

$$\left\|B_{h,k} P_h - S(t_k)\right\| \leqslant C t_k^{-\rho\nu}\left(\Delta t^{\rho\nu} + h^{2\nu}\right), \quad 0 \leqslant \nu \leqslant 1/\rho,\ k \geqslant 1,\ h > 0. \tag{3.9}$$

Next, for $t \in (t_{k-1}, t_k]$, $k \geqslant 1$, we have by (3.3) and (3.9), that

$$\begin{aligned} \|E_{h,N}(t)\| &\leqslant \|B_{h,k}P_h - S(t_k)\| + \|S(t_k) - S(t)\| \\ &\leqslant C t_k^{-\rho\nu}\big(\Delta t^{\rho\nu} + h^{2\nu}\big) + C t_k^{-\rho\nu}\Delta t^{\rho\nu} \\ &\leqslant C t^{-\rho\nu}\big(\Delta t^{\rho\nu} + h^{2\nu}\big), \quad 0 \leqslant \nu \leqslant 1/\rho,\ k \geqslant 1,\ h > 0, \end{aligned}$$

which finishes the proof of (3.5).

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

$$\begin{aligned} \|A^\alpha E_{h,N}(t)\| &\leqslant \|E_{h,N}(t)\|^{1-2\rho\alpha}\|A^{1/(2\rho)}E_{h,N}(t)\|^{2\rho\alpha} \\ &\leqslant \|E_{h,N}(t)\|^{1-2\rho\alpha}\big(\|A^{1/(2\rho)}S(t)\|^{2\rho\alpha} + \|A^{1/2\rho}\tilde{B}_{h,N}(t)\|^{2\rho\alpha}\big). \end{aligned}$$

Setting $\alpha = \frac{1/\rho-\nu}{2}$, $0 \leqslant \nu \leqslant 1/\rho$, and using Proposition 3.1(i), (3.4) and (3.5) all with $\nu = 1/\rho$ the estimate in (3.6) follows. $\square$

4. Error representation

Our main assumptions concerning φ, depending on the initial data, are

$$\varphi \in C(H, \mathbb{R}), \qquad D\varphi \in C(H, H) \quad \text{and} \quad D^2\varphi \in C_b\big(H, \mathcal{B}(H)\big) \tag{4.1}$$

or

$$\varphi \in C(H, \mathbb{R}), \qquad D\varphi \in C_b(H, H) \quad \text{and} \quad D^2\varphi \in C_b\big(H, \mathcal{B}(H)\big), \tag{4.2}$$

where $C(X,Y)$ and $C_b(X,Y)$ denote the space of continuous resp. continuous and bounded functions $f : X \to Y$ and D denotes the Fréchet derivative. The next lemma and its corollary are needed in the proof of Theorem 4.3 to accommodate random initial data and test functions φ that are unbounded with possibly an unbounded first derivative. For bounded test functions φ the next result can be found, in for example, [4, Proposition 1.12].

Lemma 4.1. *Let $\varphi : H \to \mathbb{R}$ be measurable such that $|\varphi(x)| \leqslant p_N(\|x\|)$ where p_N is a real polynomial of degree N. Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G} \subset \mathcal{F}$ is a sub sigma-algebra of $\mathcal{F}$. Let $\xi_1, \xi_2 \in L^N(\Omega, H)$ be H-valued random variables such that ξ_1 is $\mathcal{G}$-measurable and ξ_2 is independent of $\mathcal{G}$. If we define $u : H \to \mathbb{R}$ by $u(x) = \mathbf{E}(\varphi(x + \xi_2))$, $x \in H$, then, almost surely, $u(\xi_1) = \mathbf{E}(\varphi(\xi_1 + \xi_2)|\mathcal{G})$.*

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

$$\big|u_n(x) - u(x)\big| \leqslant \big|\mathbf{E}\big(\varphi_n(\eta) - \varphi(\eta)\big)\big| \leqslant \big\|\varphi_n(\eta) - \varphi(\eta)\big\|_{L^1(\Omega,\mathbb{R})} \to 0$$

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

$$\begin{aligned} \big|u_n(x)\big| &\leqslant \mathbf{E}\big|\big(\varphi_n(x + \xi_2)\big)\big| \leqslant \mathbf{E}\big(p_N\big(\|x + \xi_2\|\big)\big) \\ &\leqslant C\big(p_N\big(\|x\|\big) + \mathbf{E}\big(p_N\big(\|\xi_2\|\big)\big)\big) \\ &\leqslant C\big(p_N\big(\|x\|\big) + \|\xi_2\|_{L^N(\Omega,H)}\big), \end{aligned}$$

and hence

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

Therefore,

$$u_n(\xi_1) \to u(\xi_1) \quad \text{in } L^1(\Omega,\mathbb{R}) \tag{4.3}$$

as $n \to \infty$ by dominated convergence. Since φ_n is a bounded and measurable function it follows from [4, Proposition 1.12] that $u_n(\xi_1) = \mathbf{E}(\varphi(\xi_1+\xi_2)|\mathcal{G})$. By taking $\eta = \xi_1+\xi_2$ it follows as above that $\varphi_n(\xi_1+\xi_2) \to \varphi(\xi_1+\xi_2)$ in $L^1(\Omega,\mathbb{R})$ and thus by the dominated convergence theorem for conditional expectations we conclude that

$$u_n(\xi_1) = \mathbf{E}\left(\varphi_n(\xi_1+\xi_2)|\mathcal{G}\right) \to \mathbf{E}\left(\varphi(\xi_1+\xi_2)|\mathcal{G}\right) \quad \text{in } L^1(\Omega,H)$$

as $n \to \infty$ which finishes the proof in view of (4.3). $\square$

For any $x \in H$ and $t \in [0,T]$, we define

$$Z(T,t,x) := x + \int_t^T S(T-s)\,\mathrm{d}W^Q(s).$$

The above stochastic integral makes sense since $SQ^{1/2} \in L^2((0,T),\mathrm{HS})$ (see Remark 2.1). Let φ satisfy (4.1) or (4.2), and define

$$u(x,t) := \mathbf{E}\left(\varphi\left(Z(T,t,x)\right)\right), \quad x \in H,\ t \in [0,T]. \tag{4.4}$$

Since $\mathrm{Tr}(S(T-\cdot)QS(T-\cdot)^*) \in L^1(0,T)$ (see Remark 2.1) and $D^2\varphi \in C_b(H,\mathcal{B}(H))$, it is well known that u is a solution of the following backward Kolmogorov equation

$$u_t(x,t) + \frac{1}{2}\mathrm{Tr}\left(u_{xx}(x,t)S(T-t)QS(T-t)^*\right) = 0, \quad x \in H,\ t \in [0,T), \tag{4.5}$$

with the terminal condition $u(x,T) = \varphi(x)$, $x \in H$.

Corollary 4.2. *Let ξ be $\mathcal{F}_t$-measurable where $\{\mathcal{F}_t\}_{t\geqslant 0}$ is the normal filtration generated by W. Let φ satisfy (4.1) and $\xi \in L^2(\Omega,H)$ or let φ satisfy (4.2) and $\xi \in L^1(\Omega,H)$. Let u defined by (4.4). Then*

$$u(\xi,t) = \mathbf{E}\left(\varphi\left(Z(T,t,\xi)\right)\middle|\mathcal{F}_t\right), \quad t \in [0,T].$$

Proof. The statement follows from Lemma 4.1 with $\xi_1 = \xi$ and $\xi_2 = \int_t^T S(T-s)\,\mathrm{d}W^Q(s)$ noting that $\xi_2 \in L^2(\Omega,H) \subset L^1(\Omega,H)$ as, by Itô's isometry,

$$\mathbf{E}\|\xi_2\|^2 = \int_t^T \left\|S(T-s)Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 \mathrm{d}s \leqslant \int_0^T \left\|S(t)Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 \mathrm{d}t < \infty. \qquad \square$$

We quote the following Itô's formula from [2].

Proposition 4.1 *(Itô's formula). Let* $f : [c,d) \times H \to \mathbb{R}$, $0 \leqslant c < d \leqslant \infty$, *such that* $f, \partial_t f, \partial_x f$ *and* $\partial^2_{xx} f$ *are continuous on* $[c,d) \times H$ *with values in the appropriate spaces. Let* $a \in L^1_{loc}(\Omega \times (c,d); H)$ *and* $\xi Q^{1/2} \in L^2_{loc}(\Omega \times (c,d), \mathrm{HS})$ *and*

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

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

$$f\big(t, X(t)\big) - f\big(c, X(c)\big) = \int_c^t \partial_t f\big(s, X(s)\big)\,\mathrm{d}s + \int_c^t \big(\partial_x f\big(s, X(s)\big), a(s)\big)\,\mathrm{d}s$$
$$+ \int_c^t \big(\partial_x f\big(s, X(s)\big), \xi(s)\,\mathrm{d}W^Q(s)\big) + \frac{1}{2}\int_c^t \mathrm{Tr}\big(\partial^2_{xx} f\big(s, X(s)\big)\xi(s) Q \xi^*(s)\big)\,\mathrm{d}s.$$

The proof of the main approximation result of the paper relies on the ability to compare the laws of two different Itô processes of the form

$$Y(t) := Y(0) + \int_0^t S(T-s)\,\mathrm{d}W^Q(s)$$

and

$$\tilde{Y}(t) := \tilde{Y}(0) + \int_0^t \tilde{S}(T-s)\,\mathrm{d}W^Q(s),$$

where $\{\tilde{S}(t)\}_{t>0}$ denotes another family of bounded operators on H such that $\tilde{S}Q^{1/2} \in L^2((0,T), \mathrm{HS})$. We have the following general error formula for

$$e(T) = \mathbf{E}\big(\varphi\big(\tilde{Y}(T)\big) - \varphi\big(Y(T)\big)\big). \tag{4.6}$$

Theorem 4.3. *Let* $\{S(t)\}_{t>0}$ *and* $\{\tilde{S}(t)\}_{t>0}$ *two families of bounded operators on* H *such that* $SQ^{1/2}, \tilde{S}Q^{1/2} \in L^2((0,T), \mathrm{HS})$. *If* φ *satisfies* (4.1) *and* $Y(0), \tilde{Y}(0) \in L^2(\Omega, H)$ *or* φ *satisfies* (4.2) *and* $Y(0), \tilde{Y}(0) \in L^1(\Omega, H)$, *then* Y *and* $\tilde{Y}$ *are well-defined and the weak error* $e(T)$ *in* (4.6) *has the representation*

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

where

$$\mathcal{O}(t) = \big(\tilde{S}(T-t) + S(T-t)\big)Q\big(\tilde{S}(T-t) - S(T-t)\big)^*,$$

or

$$\mathcal{O}(t) = \big(\tilde{S}(T-t) - S(T-t)\big)Q\big(\tilde{S}(T-t) + S(T-t)\big)^*.$$

Proof. The proof is analogous to the semigroup case in [8, Theorem 3.1] and [9, Theorem 3.1] using Proposition 4.1, Corollary 4.2 and the Kolmogorov's equation (4.5). □

5. The convergence result

In this section we apply Theorem 4.3 to the approximation scheme (1.6). In order to do so we set

$$Y(t) = S(T)X_0 + \int_0^t S(T-s)\,\mathrm{d}W^Q(s)$$

and

$$\tilde{Y}(t) = \tilde{B}_{h,N}(T)X_0 + \int_0^t \tilde{B}_{h,N}(T-s)\,\mathrm{d}W^Q(s).$$

Note that $Y(T) = X(T)$ and $\tilde{Y}(T) = X_h^N$. Our main result is stated below.

Theorem 5.1. *Let $T > 0$, $N \geqslant 1$ an integer and $\Delta t = T/N$. For any $h > 0$, let $\{X_h^n\}_{0\leqslant n\leqslant N}$ be defined by (1.6) and let $\{X(t)\}_{t\in[0,T]}$ be the unique weak solution (2.6) of (1.1). Let φ satisfy (4.1) and suppose that $X_0 \in L^2(\Omega, H)$ or let φ satisfy (4.2) and suppose $X_0 \in L^1(\Omega, H)$. If $\|A^{\frac{\nu-1/\rho}{2}}Q^{\frac12}\|_{\mathrm{HS}}^2 < \infty$, $0 \leqslant \nu \leqslant 1/\rho$, then there exists a constant $C = C(T,\nu,\varphi,X_0) > 0$ which does not depend on h and N such that for $h^{2/\rho} + \Delta t < T$,*

$$\big|\mathbf{E}\,\varphi\big(X_h^N\big) - \mathbf{E}\,\varphi\big(X(T)\big)\big| \leqslant C\ln\left(\frac{T}{h^{2/\rho}+\Delta t}\right)\big(\Delta t^{\rho\nu} + h^{2\nu}\big). \tag{5.1}$$

Proof. We use Theorem 4.3 with $Y(0) = S(T)X_0$, $\tilde{Y}(0) = \tilde{B}_{h,N}(T)X_0$, and $\tilde{S}(t) = \tilde{B}_{h,N}(t)$. Using (2.2) and (3.4) with $\nu = 0$, we have that

$$\big\|\tilde{B}_{h,N}(t)\big\|_{\mathrm{HS}} = \big\|P_h\tilde{B}_{h,N}(t)\big\|_{\mathrm{HS}} \leqslant C\|P_h\|_{\mathrm{HS}}, \quad t \in (0,T).$$

Therefore, as V_h is finite dimensional, it follows that $\tilde{S}Q^{1/2} \in L^2((0,T),\mathrm{HS})$. Furthermore,

$$\int_0^T \big\|S(t)Q^{\frac12}\big\|_{\mathrm{HS}}^2\,\mathrm{d}t \leqslant C\big\|A^{-\frac{1}{2\rho}}Q^{\frac12}\big\|_{\mathrm{HS}}^2 \leqslant C\big\|A^{-\frac{\nu}{2}}\big\|_{\mathcal{B}(H)}\big\|A^{\frac{\nu-1/\rho}{2}}Q^{\frac12}\big\|_{\mathrm{HS}}^2 < \infty,$$

where the first inequality is shown in the proof of [10, Theorem 3.6] and the second inequality follows from (2.2). Thus, Theorem 4.3 is applicable.

We estimate the trace term first. Using that the operators $A, \tilde{B}_{h,N}$, and S are self-adjoint, and taking inequalities (2.1) and (2.2) into account, we have

$$\begin{aligned}
&\left|\mathbf{E}\int_0^T \mathrm{Tr}\big(u_{xx}\big(\tilde{Y}(t),t\big)\big[\tilde{B}_{h,N}(T-t)+S(T-t)\big]Q\big[\tilde{B}_{h,N}(T-t)-S(T-t)\big]^*\big)\,\mathrm{d}t\right| \\
&\quad= \bigg|\mathbf{E}\int_0^T \mathrm{Tr}\big(u_{xx}\big(\tilde{Y}(t),t\big)\big[\tilde{B}_{h,N}(T-t)+S(T-t)\big]^* \\
&\qquad\times A^{\frac{1/\rho-\nu}{2}}A^{\frac{\nu-1/\rho}{2}}Q^{\frac12}Q^{\frac12}A^{\frac{\nu-1/\rho}{2}}A^{\frac{1/\rho-\nu}{2}}E_{h,N}(T-t)\big)\,\mathrm{d}t\bigg|
\end{aligned}$$

$$
\begin{aligned}
&= \left| \mathbf{E} \int_0^T \operatorname{Tr}\bigl(u_{xx}\bigl(\tilde{Y}(t), t\bigr)\bigl(A^{\frac{1/\rho-\nu}{2}}\bigl[\tilde{B}_{h,N}(T-t) + S(T-t)\bigr]\bigr)^* \right. \\
&\quad \left. \times A^{\frac{\nu-1/\rho}{2}} Q^{\frac{1}{2}} Q^{\frac{1}{2}} A^{\frac{\nu-1/\rho}{2}} A^{\frac{1/\rho-\nu}{2}} E_{h,N}(T-t)\bigr)\,\mathrm{d}t \right| \\
&\leqslant \mathbf{E} \int_0^T \bigl\| u_{xx}\bigl(\tilde{Y}(t), t\bigr)\bigl(A^{\frac{1/\rho-\nu}{2}}\bigl[\tilde{B}_{h,N}(T-t) + S(T-t)\bigr]\bigr)^* A^{\frac{\nu-1/\rho}{2}} Q^{\frac{1}{2}} \bigr\|_{\mathrm{HS}} \\
&\quad \times \bigl\| Q^{\frac{1}{2}} A^{\frac{\nu-1/\rho}{2}} A^{\frac{1/\rho-\nu}{2}} E_{h,N}(T-t)\bigr\|_{\mathrm{HS}}\,\mathrm{d}t \\
&\leqslant \sup_{(x,t)\in H\times[0,T]} \bigl\| u_{xx}(x,t)\bigr\|_{\mathcal{B}(H)} \bigl\| A^{\frac{\nu-1/\rho}{2}} Q^{\frac{1}{2}} \bigr\|_{\mathrm{HS}}^2 \\
&\quad \times \int_0^T \bigl\| A^{\frac{1/\rho-\nu}{2}}\bigl(\tilde{B}_{h,N}(t) + S(t)\bigr)\bigr\|_{\mathcal{B}(H)} \bigl\| A^{\frac{1/\rho-\nu}{2}} E_{h,N}(t)\bigr\|_{\mathcal{B}(H)}\,\mathrm{d}t. \qquad (5.2)
\end{aligned}
$$

Next we split the integral from 0 to $\Delta t + h^{2/\rho}$ and from $h^{2/\rho} + \Delta t$ to T. Then, using Proposition 3.1(i) and (3.4),

$$
\begin{aligned}
&\int_0^{\Delta t + h^{2/\rho}} \bigl\| A^{\frac{1/\rho-\nu}{2}}\bigl(\tilde{B}_{h,N}(t) + S(t)\bigr)\bigr\|_{\mathcal{B}(H)} \bigl\| A^{\frac{1/\rho-\nu}{2}} E_{h,N}(t)\bigr\|_{\mathcal{B}(H)}\,\mathrm{d}t \\
&\quad \leqslant 2 \int_0^{\Delta t + h^{2/\rho}} \bigl(\bigl\| A^{\frac{1/\rho-\nu}{2}} \tilde{B}_{h,N}(t)\bigr\|_{\mathcal{B}(H)}^2 + \bigl\| A^{\frac{1/\rho-\nu}{2}} S(t)\bigr\|_{\mathcal{B}(H)}^2\bigr)\,\mathrm{d}t \\
&\quad \leqslant C \int_0^{\Delta t + h^{2/\rho}} t^{-1+\rho\nu}\,\mathrm{d}t \leqslant C\bigl(\Delta^{\rho\nu} + h^{2\nu}\bigr).
\end{aligned}
$$

Furthermore, by 3.1(i), (3.4) and (3.6), it follows that

$$
\begin{aligned}
&\int_{\Delta t + h^{2/\rho}}^T \bigl\| A^{\frac{1/\rho-\nu}{2}}\bigl(\tilde{B}_{h,N}(t) + S(t)\bigr)\bigr\|_{\mathcal{B}(H)} \bigl\| A^{\frac{1/\rho-\nu}{2}} E_{h,N}(t)\bigr\|_{\mathcal{B}(H)}\,\mathrm{d}t \\
&\quad \leqslant C \int_{\Delta t + h^{2/\rho}}^T t^{-1/2+\rho\nu/2}\bigl(\Delta^{\rho\nu} + h^{2\nu}\bigr) t^{-1/2-\rho\nu/2}\,\mathrm{d}t = C \ln\biggl(\frac{T}{\Delta t + h^{2/\rho}}\biggr)\bigl(\Delta^{\rho\nu} + h^{2\nu}\bigr).
\end{aligned}
$$

This finishes the estimate of the trace term considering that

$$
\sup_{(x,t)\in H\times[0,T]} \bigl\| u_{xx}(x,t)\bigr\|_{\mathcal{B}(H)} \leqslant \sup_{x\in H} \bigl\| D^2\varphi(x)\bigr\|_{\mathcal{B}(H)}.
$$

To estimate the initial error term first assume φ satisfies (4.1) and $\mathbf{E}\|X_0\|^2 < \infty$. Note that under assumption (4.1) it follows from Taylor's Formula that

$$
\bigl|\varphi(x) - \varphi(y)\bigr| \leqslant \bigl\| D\varphi(y)\bigr\| \cdot \|x - y\| + C\|x - y\|^2,
$$

where $C = \sup_{x \in H} \|D^2\varphi(x)\|_{\mathcal{B}(H)}$ and that $\|D\varphi(x)\| \leqslant K(1+\|x\|)$ where $K = \max\{C, \|D\varphi(0)\|\}$. Therefore,

$$\left|\varphi(x) - \varphi(y)\right| \leqslant C\left(1 + \|y\|\right) \cdot \|x - y\| + C\|x - y\|^2.$$

Then, using the law of double expectations, and noting that

$$X(T) = S(T)X_0 + \int_0^T S(T - s)\,\mathrm{d}W(s),$$

we have, using Corollary 4.2, Proposition 2.1 with $\nu = 0$ and (3.5), that

$$\begin{aligned}
&\left|\mathbf{E}\left(u\left(\tilde{Y}(0), 0\right) - u\left(Y(0), 0\right)\right)\right| \\
&\quad = \left|\mathbf{E}\left(u\left(\tilde{B}_{h,N}(T)X_0, 0\right) - u\left(S(T)X_0, 0\right)\right)\right| \\
&\quad = \left|\mathbf{E}\left(\mathbf{E}\left(\left(\varphi\left(\tilde{B}_{h,N}(T)X_0 + \int_0^T S(T - s)\,\mathrm{d}W^Q(s)\right) - \varphi\left(S(T)X_0 + \int_0^T S(T - s)\,\mathrm{d}W^Q(s)\right)\right)\middle|\mathcal{F}_0\right)\right)\right| \\
&\quad \leqslant C\mathbf{E}\left(\left\|\tilde{B}_{h,N}(T)X_0 - S(T)X_0\right\| \cdot \left(1 + \left\|X(T)\right\|\right)\right) + C\mathbf{E}\left(\left\|\tilde{B}_{h,N}(T)X_0 - S(T)X_0\right\|^2\right) \\
&\quad \leqslant CT^{-\rho\nu}\left(\Delta t^{\rho\nu} + h^{2\nu}\right)\left(1 + \mathbf{E}\left(\left\|X_0\right\|^2 + \left\|X(T)\right\|^2\right)\right) + CT^{-2\rho\nu}\left(\Delta t^{2\rho\nu} + h^{4\nu}\right)\mathbf{E}\|X_0\|^2.
\end{aligned}$$

Finally, if φ satisfies (4.2) and $\mathbf{E}\|X_0\| < \infty$, then, again by Taylor's Formula, $|\varphi(x) - \varphi(y)| \leqslant C\|x - y\|$ with $C = \sup_{x \in H} \|D\varphi(x)\|$. Thus, similarly to the above calculation, we have that

$$\begin{aligned}
&\left|\mathbf{E}\left(u\left(\tilde{Y}(0), 0\right) - u\left(Y(0), 0\right)\right)\right| \\
&\quad = \left|\mathbf{E}\left(u\left(\tilde{B}_{h,N}(T)X_0, 0\right) - u\left(S(T)X_0, 0\right)\right)\right| \\
&\quad = \left|\mathbf{E}\left(\mathbf{E}\left(\left(\varphi\left(\tilde{B}_{h,N}(T)X_0 + \int_0^T S(T - s)\,\mathrm{d}W^Q(s)\right) - \varphi\left(S(T)X_0 + \int_0^T S(T - s)\,\mathrm{d}W^Q(s)\right)\right)\middle|\mathcal{F}_0\right)\right)\right| \\
&\quad \leqslant C\mathbf{E}\left(\left\|\tilde{B}_{h,N}(T)X_0 - S(T)X_0\right\|\right) \leqslant CT^{-\rho\nu}\left(\Delta t^{\rho\nu} + h^{2\nu}\right)\mathbf{E}\|X_0\|,
\end{aligned}$$

and the proof is complete. □

Remark 5.2. Below we give examples of the rate of convergence obtained in Theorem 5.1 in some typical cases.

(i) If $Q = I$ (white noise), then, as mentioned in the introduction, we must have $d = 1$ and the rate of weak convergence in time is $(1 - \rho/2)_-$ and in space it is $(2/\rho - 1)_-$.
(ii) If Q is of trace class, then we may take $\nu = 1/\rho$ and recover the finite dimensional order; that is, 1_- in time and 2_- in space.
(iii) Suppose that there exists some real numbers κ and $\alpha > 0$ such that $A^\kappa Q \in \mathcal{B}(H)$, $\mathrm{Tr}(A^{-\alpha}) < \infty$ and $\alpha - 1/\rho < \kappa \leqslant \alpha$. Then, since

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

we recover a space weak order of convergence $(2/\rho - 2(\alpha - \kappa))_-$ and a time weak order of $(1 - \rho(\alpha - \kappa))_-$. These are twice the strong orders (modulo the logarithmic term) respectively in space and in time found in [10].

Acknowledgments

The authors wish to thank Dr. Fredrik Lindgren for valuable discussions.

References

[1] A. Andersson, S. Larsson, Weak convergence for a spatial approximation of the nonlinear stochastic heat equation, arXiv:1212.5564.

[2] Z. Brzeźniak, Some remarks on Itô and Stratonovich integration in 2-smooth Banach spaces, in: Probabilistic Methods in Fluids, World Sci. Publ., River Edge, NJ, 2003, pp. 48–69.

[3] P. Clément, G. Da Prato, J. Prüss, White noise perturbation of the equations of linear parabolic viscoelasticity, Rend. Istit. Mat. Univ. Trieste XXIX (1997) 207–220.

[4] G. Da Prato, J. Zabczyk, Stochastic Equations in Infinite Dimensions, Cambridge University Press, Cambridge, 1992, xviii+454 pp.

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

[6] M. Geissert, M. Kovács, S. Larsson, Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise, BIT 49 (2009) 343–356.

[7] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Appl. Math., vol. 23, Springer-Verlag, Berlin, 1992, xxxvi+632 pp.

[8] M. Kovács, S. Larsson, F. Lindgren, Weak convergence of finite element approximations of stochastic evolution equations with additive noise, BIT 52 (2012) 85–108.

[9] M. Kovács, S. Larsson, F. Lindgren, Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes, BIT 53 (2013) 497–525.

[10] M. Kovács, J. Printems, Strong order of convergence of a fully discrete approximation of a linear stochastic Volterra type evolution equation, Math. Comp. (2013), in press, arXiv:1205.5601.

[11] F. Lindner, R.L. Schilling, Weak order for the discretization of the stochastic heat equation driven by impulsive noise, Potential Anal. 38 (2013) 345–379.

[12] C. Lubich, Convolution quadrature and discretized operational calculus. I, Numer. Math. 52 (1988) 129–145.

[13] C. Lubich, Convolution quadrature and discretized operational calculus. II, Numer. Math. 52 (1988) 413–425.

[14] C. Lubich, I. Sloan, V. Thomée, Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term, Math. Comp. 65 (1996) 1–17.

[15] W. McLean, V. Thomée, Numerical solution of an evolution equation with a positive-type memory term, J. Aust. Math. Soc. Ser. B 35 (1993) 23–70.

[16] G.N. Milstein, M.V. Tretyakov, Stochastic Numerics for Mathematical Physics, Sci. Comput., Springer-Verlag, Berlin, 2004, xx+594 pp.

[17] J. Prüss, Evolutionary Integral Equations and Applications, Birkhäuser Verlag, Basel, 1993, xxvi+366 pp.

[18] D. Talay, Probabilistic numerical methods for partial differential equations: elements of analysis, in: Probabilistic Models for Nonlinear Partial Differential Equations, Montecatini Terme, 1995, in: Lecture Notes in Math., vol. 1627, Springer-Verlag, Berlin, 1996, pp. 148–196.

[19] X. Wang, S. Gan, Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise, J. Math. Anal. Appl. 398 (2013) 151–169.