Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 130 (2020) 4968–5005
stochastic processes and their applications
www.elsevier.com/locate/spa

Strong convergence of a stochastic Rosenbrock-type scheme for the finite element discretization of semilinear SPDEs driven by multiplicative and additive noise

Jean Daniel Mukam[a], Antoine Tambue[b,c,d,*]

[a] Fakultät für Mathematik, Technische Universität Chemnitz, 09126 Chemnitz, Germany
[b] Department of Computer science, Electrical engineering and Mathematical sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway
[c] Center for Research in Computational and Applied Mechanics (CERECAM), and Department of Mathematics and Applied Mathematics, University of Cape Town, 7701 Rondebosch, South Africa
[d] The African Institute for Mathematical Sciences(AIMS) of South Africa, 6-8 Melrose Road, Muizenberg 7945, South Africa

Received 23 May 2018; received in revised form 26 January 2020; accepted 17 February 2020
Available online 29 February 2020

Abstract

This paper aims to investigate the numerical approximation of a general second order parabolic stochastic partial differential equation(SPDE) driven by multiplicative and additive noise. Our main interest is on such SPDEs where the nonlinear part is stronger than the linear part, usually called stochastic dominated transport equations. Most standard numerical schemes lose their good stability properties on such equations, including the current linear implicit Euler method. We discretize the SPDE in space by the finite element method and propose a novel scheme called stochastic Rosenbrock-type scheme for temporal discretization. Our scheme is based on the local linearization of the semi-discrete problem obtained after space discretization and is more appropriate for such equations. We provide a strong convergence of the new fully discrete scheme toward the exact solution for multiplicative and additive noise and obtain optimal rates of convergence. Numerical experiments to sustain our theoretical results are provided.

Keywords: Rosenbrock-type scheme; Stochastic partial differential equations; Multiplicative & additive noise; Strong convergence; Finite element method

* Corresponding author at: Department of Computer science, Electrical engineering and Mathematical sciences, Western Norway University of Applied Sciences, Inndalsveien 28, 5063 Bergen, Norway.
E-mail addresses: jean.d.mukam@aims-senegal.org (J.D. Mukam), antonio@aims.ac.za (A. Tambue).

https://doi.org/10.1016/j.spa.2020.02.008
0304-4149/

1. Introduction

We consider the numerical approximation of the following SPDE

$$dX(t) + [AX(t) + F(X(t))]dt = B(X(t))dW(t), \quad X(0) = X_0, \quad t \in (0, T], \tag{1}$$

in the Hilbert space $L^2(\Lambda)$, where $\Lambda \subset \mathbb{R}^d$, $d = 1, 2, 3$ is bounded with smooth boundary, $T > 0$ is the final time, F and B are nonlinear functions, X_0 is the initial data which is random and A is a linear operator, unbounded, not necessary self-adjoint. Precise assumptions on F, B, X_0 and A will be given in the next section. Equations of type (1) are used to model many real world phenomena in different fields such as biology, chemistry, physics [3,28,29,32]. In many cases explicit solutions of SPDEs are unknown, therefore numerical approximations are powerful tools to provide realistic approximations. Numerical approximation of SPDE (1) is therefore an active research area and has attracted a lot of attentions since two decades (see e.g. [8,9,11,12,16,27,28,34–37]). Due to the time step restriction of the explicit Euler method, linear implicit Euler method is used in many situations. Linear implicit Euler method has been largely investigated in the literature (see e.g. [12,17,33,34] and the references therein). The resolvent operator $(\mathbf{I} + \Delta t A_h)^{-1}$ plays a key role to stabilize the linear implicit Euler method, where A_h is the discrete version of A, obtained after the space discretization. Such approach is justified when the linear operator A is strong. Indeed, when A is stronger than F, the linear operator A drives the SPDE (1) and the good stability properties of the linear implicit Euler method and exponential integrators are guaranteed. In more concrete applications, the nonlinear function F can be stronger. Typical examples are stochastic reaction equations with stiff reaction term. For such equations, both linear implicit Euler method [12,17,33,34] and exponential integrators [8,16,35] behave like the standard explicit Euler method (see Section 2.3) and therefore lose their good stabilities properties. For such problems in the deterministic context, exponential Rosenbrock-type methods [7,31] and Rosenbrock-type methods [21,22,31] were proved to be efficient. Recently, the exponential Rosenbrock method was extended to the case of stochastic partial differential equations [20] and was proved to be very stable for stochastic reactive dominated transport equations. However the computation of the stochastic exponential matrix functions involved was far to be efficient. Since solving linear systems are more straightforward than computing the exponential of a matrix, it is important to develop alternative methods based on the resolution of linear systems, which may be more efficient if the appropriate preconditioners are used. In this paper, we propose a novel scheme based on the combination of the Rosenbrock-type method and the linear implicit Euler method. The resulting numerical scheme that we call stochastic Rosenbrock-type scheme(SROS) is stable and efficient in contrast to the exponential scheme in [20], which is only stable. The space discretization is performed using the finite element method and our novel scheme is based on the local linearization of the nonlinear drift part of the semi-discrete problem obtained after spatial discretization. The local linearization therefore weakens the nonlinear part of the drift function so that the linearized semi-discrete problem is driven by its new linear part, which changes at each time step. The standard linear implicit Euler method [12,34] is then applied to the linearized semi-discrete problem. This combination yields our novel SROS scheme. We analyze the strong convergence of the novel fully discrete scheme toward the exact solution in the root-mean-square L^2-norm. The main challenge here comes from the fact that the resolvent operator $S^m_{h,\Delta t}(\omega)$ appearing in the numerical scheme (33) is not constant as it changes at each time step. Furthermore the operator $S^m_{h,\Delta t}(\omega)$ is a random operator. To address those challenges, we provide in Section 3.1 novel stability estimates to handle the composition of the perturbed random resolvent operators, useful in our convergence analysis. The results indicate how the convergence orders depend on the regularity of the initial data and the noise. More precisely, we achieve the optimal convergence orders $\mathcal{O}\left(h^{\beta}+\Delta t^{\frac{\min(\beta,1)}{2}}\right)$ for multiplicative noise and the optimal convergence orders $\mathcal{O}\left(h^{\beta}+\Delta t^{\frac{\beta}{2}-\epsilon}\right)$ for additive noise, where β is the regularity's parameter of the noise (see Assumption 2.2) and $\epsilon>0$ is an arbitrary number small enough.

The rest of this paper is organized as follows. Section 2 deals with the well posedness problem, the numerical scheme and the main results. In Section 3, we provide some error estimates for the deterministic homogeneous problem as preparatory results along with proof of the main results. Section 4 provides some numerical experiments to sustain the theoretical findings. Those numerical experiments show the efficiency of the novel scheme comparing to the exponential scheme developed in [20].

2. Mathematical setting and main results

2.1. Main assumptions and well posedness problem

Let us define functional spaces, norms and notations that will be used in the rest of the paper. Let $(H,\langle.,.\rangle_H,\|.\|)$ be a separable Hilbert space. For all $p\geq 2$ and for a Hilbert space U, we denote by $L^p(\Omega,U)$ the Banach space of all equivalence classes of p integrable U-valued random variables. We denote by $L(U,H)$ the space of bounded linear mappings from U to H endowed with the usual operator norm $\|.\|_{L(U,H)}$. By $\mathcal{L}_2(U,H):=HS(U,H)$, we denote the space of Hilbert–Schmidt operators from U to H. We equip $\mathcal{L}_2(U,H)$ with the norm

$$\|l\|^2_{\mathcal{L}_2(U,H)}:=\sum_{i=1}^{\infty}\|l\psi_i\|^2,\quad l\in\mathcal{L}_2(U,H),\tag{2}$$

where $(\psi_i)_{i=1}^{\infty}$ is an orthonormal basis of U. Note that (2) is independent of the orthonormal basis of U. For simplicity, we use the notations $L(U,U)=:L(U)$ and $\mathcal{L}_2(U,U)=:\mathcal{L}_2(U)$. It is well known that for all $l\in L(U,H)$ and $l_1\in\mathcal{L}_2(U)$, $ll_1\in\mathcal{L}_2(U,H)$ and

$$\|ll_1\|_{\mathcal{L}_2(U,H)}\leq\|l\|_{L(U,H)}\|l_1\|_{\mathcal{L}_2(U)}.\tag{3}$$

In the rest of this paper, we take $H=L^2(\Lambda)$. In order to ensure the existence and the uniqueness of the solution of (1), and for the purpose of the convergence analysis, we make the following assumptions.

Assumption 2.1 (*Linear Operator* A)**.** $-A:\mathcal{D}(A)\subset H\longrightarrow H$ is the generator of an analytic semigroup $S(t)=:e^{-At}$ on $L^2(\Lambda)$, i.e. $S(t)$ is given as follows [1,2,6,24]

$$S(t)=\frac{1}{2\pi i}\int_{\mathcal{C}}e^{-t\lambda}(\lambda\mathbf{I}-A)^{-1}d\lambda,\quad t>0,$$

where $\mathcal{C}$ denotes a path that surrounds the spectrum of $-A$.

Assumption 2.2 (*Initial Value* X_0)**.** The initial value X_0 belongs to $L^p\left(\Omega,\mathcal{D}\left((A)^{\frac{\beta}{2}}\right)\right)$, for some $\beta\in(0,2]$ and some $p\in[2,\infty)$.

As in the current literature for deterministic Rosenbrock-type methods [21,22], deterministic exponential Rosenbrock-type method [7,19] and stochastic exponential Rosenbrock-type methods [20], we make the following assumption on the nonlinear drift term.

Assumption 2.3 (*Nonlinear Term F*)**.** The nonlinear map $F : H \longrightarrow H$ is Fréchet differentiable with bounded derivative, i.e. there exists a constant $b > 0$ such that

$$\|F'(u)\|_{L(H)} \leq b, \quad u \in H. \tag{4}$$

Moreover, as in [14, Page 6] for deterministic Rosenbrock-type method, we assume that the resolvent set of $-A - F'(u)$ contains $(0, \infty)$ for all $u \in H$.

Remark 2.1. Inequality (4) together with the mean value theorem shows that there exists a constant $C_F = C_F(b) \geq 0$ such that

$$\|F(u) - F(v)\| \leq C_F \|u - v\|, \quad u, v \in H. \tag{5}$$

In addition, if $\|F(0)\| < \infty$, then from (5) there exists a constant $C = (C_F, \|F(0)\|) \geq 0$ such that

$$\|F(u)\| \leq \|F(0)\| + \|F(u) - F(0)\| \leq \|F(0)\| + C_F\|u\| \leq C(1 + \|u\|), \quad u \in H.$$

Remark 2.2. An illustrative example for which the resolvent set of $-A-F'(u)$ contains $(0, \infty)$ is obtained when A generates a contraction semigroup and the derivative of the nonlinear drift term F satisfies the following coercivity condition

$$\left\langle F'(u)v, v\right\rangle_H \geq 0, \quad u, v \in H. \tag{6}$$

In fact, it follows from (6) that $-F'(u)$ is an relatively A-bounded and dissipative operator with A-bound $a_0 = 0$ (see e.g. [1, Chapter III, Definition 2.1]). Therefore, from [1, Chapter III, Theorem 2.7], it follows that $-A - F'(u)$ is a generator of a contraction semigroup. Hence, for all $u \in H$ $(0, \infty) \subset \rho\left(-A - F'(u)\right)$.

Remark 2.3. The condition $(0, \infty) \subset \rho\left(-A - F'(u)\right)$ on Assumption 2.3 can be relaxed, but the drawback is that the resolvent set of the perturbed semigroup is smaller than that of the initial semigroup.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{\mathcal{F}_t\}_{t\in[0,T]}$ a normal filtration on $(\Omega, \mathcal{F}, \mathbb{P})$, that is $\{\mathcal{F}_t\}_{t\in[0,T]}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the following (see e.g. [26, Definition 2.1.11]):

- $\mathcal{F}_0$ contains all elements $O \in \mathcal{F}$ with $\mathbb{P}(O) = 0$,
- $\mathcal{F}_t = \mathcal{F}_{t^+} := \bigcap_{s>t} \mathcal{F}_s$ for all $t \in [0, T]$.

Let $Q : H \longrightarrow H$ be a linear selfadjoint and positive operator. In this work, the noise $W(t) = W(x, t)$ is assumed to be an H-valued Q-Wiener process defined in the filtered probability space $\left(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\geq 0}\right)$. Let us recall below the definition of a Q-Wiener process.

Definition 2.1 (*Q-Wiener Process [26, Definition 2.1.12]*)**.** An H-valued stochastic process $\{W(t) : t \geq 0\}$ is called Q-Wiener process if

(i) $W(0) = 0$ almost surely (a.s).
(ii) The application $t \longmapsto W(t, \omega)$ is continuous from $\mathbb{R}^+$ to H for every $\omega \in \Omega$.
(iii) $W(t)$ is $\mathcal{F}_t$-adapted and $W(t) - W(s)$ is independent of $\mathcal{F}_s$ for $s < t$.
(iv) For all $0 \leq s \leq t$, the random variable $W(t) - W(s)$ follows a normal distribution with mean 0 and covariance operator $(t - s)Q$. We write $W(t) - W(s) \sim \mathcal{N}(0, (t - s)Q)$.

It is well known that if Q has finite trace,[1] then the Q-Wiener process $W(t)$ can be represented as follows [26, Proposition 2.1.10]

$$W(x,t)=\sum_{i\in\mathbb{N}}\sqrt{q_i}e_i(x)\beta_i(t),\quad t\in[0,T],\quad x\in\Lambda,\tag{7}$$

where q_i, e_i, $i\in\mathbb{N}$ are respectively the eigenvalues and the eigenfunctions of the covariance operator Q, and β_i are independent and identically distributed standard Brownian motions.

The space of Hilbert–Schmidt operators from $Q^{\frac{1}{2}}(H)$ to H is denoted by $L_2^0=\mathcal{L}_2(Q^{\frac{1}{2}}(H),H)=:HS(Q^{\frac{1}{2}}(H),H)$ with the corresponding norm $\|.\|_{L_2^0}$ defined by

$$\|l\|_{L_2^0}:=\|lQ^{\frac{1}{2}}\|_{HS}=\left(\sum_{i=1}^{\infty}\|lQ^{\frac{1}{2}}e_i\|^2\right)^{\frac{1}{2}},\quad l\in L_2^0,\tag{8}$$

where $(e_i)_{i=1}^{\infty}$ is an orthonormal basis of H. Note that (8) is also independent of the orthonormal basis of H. Following [25, Chapter 7] or [10,12,16,37], we make the following assumption on the diffusion term.

Assumption 2.4 (*Diffusion Term*)**.** The operator $B:H\longrightarrow L_2^0$ satisfies the global Lipschitz condition, i.e. there exists a positive constant C_B such that

$$\|B(0)\|_{L_2^0}\leq C_B,\quad \|B(u)-B(v)\|_{L_2^0}\leq C_B\|u-v\|,\quad u,v\in H.$$

As a consequence of Assumption 2.4, it holds that

$$\|B(u)\|_{L_2^0}\leq\|B(0)\|_{L_2^0}+\|B(u)-B(0)\|_{L_2^0}\leq\|B(0)\|_{L_2^0}+C_B\|u\|\leq C_B(1+\|u\|),\quad u\in H.$$

We equip $V_\alpha:=\mathcal{D}(A^{\frac{\alpha}{2}})$, $\alpha\in\mathbb{R}$ with the norm $\|v\|_\alpha:=\|A^{\frac{\alpha}{2}}v\|$, for all $v\in V_\alpha$. It is well known that $(V_\alpha,\|.\|_\alpha)$ is a Banach space [6].

To establish our root-mean-square L^2 strong convergence result when dealing with multiplicative noise, we will also need the following further assumption on the diffusion term when $\beta\in[1,2)$, which was also used in [10,13] to achieve optimal regularity rates in space and time, and in [12,16,20] to achieve optimal strong convergence rates.

Assumption 2.5. There exists a positive constant $c\geq 0$ such that

$$B\left(\mathcal{D}\left(A^{\frac{(\beta-1)}{2}}\right)\right)\subset HS\left(Q^{\frac{1}{2}}(H),\mathcal{D}\left(A^{\frac{(\beta-1)}{2}}\right)\right)\text{ and }\left\|A^{\frac{(\beta-1)}{2}}B(v)\right\|_{L_2^0}\leq c\left(1+\|v\|_{\beta-1}\right)$$

for all $v\in\mathcal{D}\left(A^{\frac{(\beta-1)}{2}}\right)$, where β comes from Assumption 2.2.

Typical examples fulfilling Assumption 2.5 are stochastic reaction diffusion equations (see e.g. [10, Section 4]).

When dealing with additive noise (i.e. when $B=\mathbf{I}$), the strong convergence proof will make use of the following assumption, also used in [20,34,35].

Assumption 2.6. The covariance operator Q satisfies the following estimate

$$\left\|A^{\frac{\beta-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathcal{L}_2(H)}\leq C_Q,\tag{9}$$

where β comes from Assumption 2.2 and C_Q is a positive constant.

[1] In this case $W(t)$ is called trace class noise.

When dealing with additive noise, to achieve convergence order greater than $\frac{1}{2}$ in time, we use the following further assumption on the nonlinear function, also used in [20,34,35].

Assumption 2.7. The deterministic mapping $F : H \longrightarrow H$ is twice differentiable and there exist two constants $L \geq 0$ and $\eta \in (0, 2)$ such that

$$\|F'(u)v\| \leq L\|v\|, \quad \|F''(u)(v_1, v_2)\|_{-\eta} \leq L\|v_1\|.\|v_2\|, \quad u, v, v_1, v_2 \in H.$$

The following proposition will be useful in the rest of the paper.

Proposition 2.1 (*Smoothing Properties of the Semigroup [6]*). *Let $\alpha > 0$, $\delta \geq 0$ and $0 \leq \gamma \leq 1$, then there exists a constant $C > 0$ such that*

$$\|A^{\delta}S(t)\|_{L(H)} \leq Ct^{-\delta}, \quad t > 0; \quad \|A^{-\gamma}(\mathbf{I} - S(t))\|_{L(H)} \leq Ct^{\gamma}, \quad t \geq 0;$$
$$A^{\delta}S(t) = S(t)A^{\delta} \quad on \quad \mathcal{D}(A^{\delta}) \quad and \quad \|D_t^l S(t)v\|_{\delta} \leq Ct^{-l-\frac{(\delta-\alpha)}{2}}\|v\|_{\alpha}, \quad t > 0,$$
$$v \in \mathcal{D}(A^{\frac{\alpha}{2}});$$

where $l = 0, 1$, and $D_t^l = \dfrac{d^l}{dt^l}$. Moreover, if $\delta \geq \gamma$ then $\mathcal{D}(A^{\delta}) \subset \mathcal{D}(A^{\gamma})$.

The well posedness result is given in the following theorem.

Theorem 2.1 (*[25, Theorem 7.2]*). *Let Assumptions 2.1, 2.3 and 2.4 be satisfied. If X_0 is a $\mathcal{F}_0$-measurable H-valued random variable, then there exists a unique mild solution X of (1), which has the following representation*

$$X(t) = S(t)X_0 - \int_0^t S(t-s)F(X(s))ds + \int_0^t S(t-s)B(X(s))dW(s), \quad t \in (0, T] \quad (10)$$

and satisfies

$$\mathbb{P}\left[\int_0^T \|X(s)\|^2 ds < \infty\right] = 1.$$

Moreover, for any $p \geq 2$, there exists a constant $C = C(p, T) > 0$ such that

$$\sup_{t\in[0,T]} \mathbb{E}\|X(t)\|^p \leq C(1 + \mathbb{E}\|X_0\|^p).$$

2.2. Finite element discretization

In the rest of the paper, to simplify the presentation, we consider the linear operator A to be of second-order. More precisely, we consider the SPDE (1) to be a second-order semilinear parabolic SPDE of the following form

$$dX(t,x)+[-\nabla\cdot(\mathbf{D}\nabla X(t,x))+\mathbf{q}\cdot\nabla X(t,x)]dt+f(x,X(t,x))dt = b(x,X(t,x))dW(t,x), \quad (11)$$

for $x \in \Lambda$ and $t \in (0, T]$, where the functions $f : \Lambda \times \mathbb{R} \longrightarrow \mathbb{R}$ and $b : \Lambda \times \mathbb{R} \longrightarrow \mathbb{R}$ are continuously differentiable with globally bounded derivatives. In the abstract framework (1), the linear operator A is the $L^2(\Lambda)$ realization [2, p. 812] of the following differential operator

$$\mathcal{A}u = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(D_{ij}(x)\frac{\partial u}{\partial x_j}\right) + \sum_{i=1}^{d} q_i(x)\frac{\partial u}{\partial x_i}, \quad \mathbf{D} := \left(D_{i,j}\right)_{1\leq i,j\leq d}, \quad \mathbf{q} := (q_i)_{1\leq i\leq d}. \quad (12)$$

where $D_{ij} \in L^{\infty}(\Lambda)$, $q_i \in L^{\infty}(\Lambda)$ and there exists a constant $c_1 > 0$ such that

$$\sum_{i,j=1}^{d} D_{ij}(x)\xi_i\xi_j \geq c_1|\xi|^2, \quad \xi \in \mathbb{R}^d, \quad x \in \overline{\Lambda}.$$

The functions $F : H \longrightarrow H$ and $B : H \longrightarrow HS\left(Q^{\frac{1}{2}}(H), H\right)$ are defined respectively by

$$(F(v))(x) = f(x, v(x)), \quad (B(v)u)(x) = b(x, v(x)).u(x), \quad x \in \Lambda, \ v \in H,$$
$$u \in Q^{\frac{1}{2}}(H). \tag{13}$$

For an appropriate family of eigenfunctions $(e_i)_{i\in\mathbb{N}}$ such that $\sup_{i\in\mathbb{N}^d}\left[\sup_{x\in\Lambda}\|e_i(x)\|\right] < \infty$, it is well known that the Nemytskii operator F related to f and the multiplication operator B associated to the function b defined in (13) satisfy Assumptions 2.3–2.5, see e.g. [10, Section 4]. As in [2,16] we introduce two spaces $\mathbb{H}$ and V, such that $\mathbb{H} \subset V$; the two spaces depend on the boundary conditions and the domain of the operator A. For Dirichlet (or first-type) boundary conditions we take

$$V = \mathbb{H} = H_0^1(\Lambda) = \{v \in H^1(\Lambda) : v = 0 \quad \text{on} \quad \partial\Lambda\}.$$

For Robin (third-type) boundary condition and Neumann (second-type) boundary condition, which is a special case of Robin boundary condition, we take $V = H^1(\Lambda)$

$$\mathbb{H} = \{v \in H^2(\Lambda) : \partial v/\partial \mathrm{v}_{\mathcal{A}} + \alpha_0 v = 0, \quad \text{on} \quad \partial\Lambda\}, \quad \alpha_0 \in \mathbb{R},$$

where $\partial v/\partial \mathrm{v}_{\mathcal{A}}$ is the normal derivative of v and v_A is the exterior pointing normal at $n = (n_i)$ to the boundary of $\mathcal{A}$, given by

$$\partial v/\partial \mathrm{v}_{\mathcal{A}} = \sum_{i,j=1}^{d} n_i(x)D_{ij}(x)\frac{\partial v}{\partial x_j}, \qquad x \in \partial\Lambda.$$

Using Green's formula and the boundary conditions, the corresponding bilinear form associated to $\mathcal{A}$ and A is given by

$$a(u, v) = \int_{\Lambda}\left(\sum_{i,j=1}^{d} D_{ij}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} + \sum_{i=1}^{d} q_i\frac{\partial u}{\partial x_i}v\right)dx, \quad u, v \in V,$$

for Dirichlet and Neumann boundary conditions, and

$$a(u, v) = \int_{\Lambda}\left(\sum_{i,j=1}^{d} D_{ij}\frac{\partial u}{\partial x_i}\frac{\partial v}{\partial x_j} + \sum_{i=1}^{d} q_i\frac{\partial u}{\partial x_i}v\right)dx + \int_{\partial\Lambda}\alpha_0 uv dx, \quad u, v \in V,$$

for Robin boundary conditions. Using Gårding's inequality (see e.g. [29]), it holds that there exist two constants c_0 and $\lambda_0 > 0$ such that

$$a(v, v) \geq \lambda_0\|v\|^2_{H^1(\Lambda)} - c_0\|v\|^2, \quad v \in V. \tag{14}$$

By adding and subtracting c_0Xdt in both sides of (1), we have a new linear operator still denoted by A, and the corresponding bilinear form is also still denoted by a. Therefore, the following coercivity property holds

$$a(v, v) \geq \lambda_0\|v\|_1^2, \quad v \in V. \tag{15}$$

Note that the expression of the nonlinear term F has changed as we included the term $c_0 X$ in the new nonlinear term that we still denote by F. The coercivity property (15) implies that $-A$ is sectorial on $L^2(\Lambda)$, i.e. there exist C_1, $\theta \in (\frac{1}{2}\pi, \pi)$ such that

$$\|(\lambda I + A)^{-1}\|_{L(L^2(\Lambda))} \leq \frac{C_1}{|\lambda|}, \qquad \lambda \in S_\theta,$$

where $S_\theta := \left\{\lambda \in \mathbb{C} : \lambda = \rho e^{i\phi},\ \rho > 0,\ 0 \leq |\phi| \leq \theta\right\}$ (see e.g. [6]). The coercivity property (15) implies that $-A$ is the infinitesimal generator of a contraction semigroup $S(t) = e^{-tA}$ on $L^2(\Lambda)$. The coercivity property (15) also implies that A is positive and its fractional powers are well defined for any $\alpha > 0$, by

$$\begin{cases} A^{-\alpha} &= \frac{1}{\Gamma(\alpha)} \displaystyle\int_0^\infty t^{\alpha-1} \mathrm{e}^{-tA} dt, \\ A^{\alpha} &= (A^{-\alpha})^{-1}, \end{cases} \tag{16}$$

where $\Gamma(\alpha)$ is the Gamma function (see [6]). Let us now turn our attention to the space discretization of our problem (1). We start by splitting the domain Λ in finite triangles. Let $\mathcal{T}_h$ be the triangulation with maximal length h satisfying the usual regularity assumptions, and $V_h \subset V$ be the space of continuous functions that are piecewise linear over the triangulation $\mathcal{T}_h$. We consider the projection P_h from $H = L^2(\Lambda)$ to V_h defined for every $u \in H$ by

$$\langle P_h u, \chi \rangle_H = \langle u, \chi \rangle_H, \quad \chi \in V_h. \tag{17}$$

The discrete operator $A_h : V_h \longrightarrow V_h$ is defined by

$$\langle A_h \phi, \chi \rangle_H = \langle A^{1/2} \phi, A^{*1/2} \chi \rangle_H = a(\phi, \chi), \quad \phi, \chi \in V_h, \tag{18}$$

Like $-A$, $-A_h$ is also a generator of a bounded analytic semigroup $S_h(t)$ on V_h, given by (see e.g. [2, Chapter II, (7.14)] or [15])

$$S_h(t) = e^{-tA_h} = \frac{1}{2\pi i} \int_{\mathcal{C}} e^{-t\lambda} (\lambda \mathbf{I} - A_h)^{-1} d\lambda, \quad t > 0,$$

where $\mathcal{C}$ is a path that surrounds the spectrum of $-A_h$. Let K be a constant satisfying

$$\|S_h(t)\|_{L(H)} \leq K, \quad t \geq 0. \tag{19}$$

As any semigroup and its generator, $-A_h$ and $S_h(t)$ satisfy the smoothing properties of Proposition 2.1 with a uniform constant C (i.e. independent of h). Following [2,15,16], we characterize the domain of the operator $A^{\frac{\gamma}{2}}$, $1 \leq \gamma \leq 2$ as follows:

$$\mathcal{D}(A^{\frac{\gamma}{2}}) = \mathbb{H} \cap H^\gamma(\Lambda) \text{ for Dirichlet boundary conditions,}$$
$$\mathcal{D}(A) = \mathbb{H}, \quad \mathcal{D}(A^{\frac{1}{2}}) = H^1(\Lambda) \text{ for Robin boundary conditions.}$$

The semi-discrete version of (1) consists to find $X^h(t) \in V_h$, $t \in (0, T]$ such that

$$dX^h(t) + [A_h X^h(t) + P_h F(X^h(t))]dt = P_h B(X^h(t)) dW(t), \quad X^h(0) = P_h X_0, \\ t \in (0, T]. \tag{20}$$

The proof of the following lemma can be found in [20, Lemma 4 & Lemma 5]. Its provides the space and time regularities of the mild solution $X^h(t)$ of (20).

Lemma 2.1.

(i) Let Assumptions 2.1 (with $\beta \in [0,1)$), 2.2–2.4 be fulfilled. Then the mild solution $X^h(t)$ of (20) satisfies the following regularity estimates

$$\left\| A_h^{\frac{\beta}{2}} X^h(t) \right\|_{L^p(\Omega,H)} \leq C\left(1 + \left\| A^{\frac{\beta}{2}} X_0 \right\|_{L^p(\Omega,H)}\right), \quad t \in [0,T], \tag{21}$$

$$\begin{aligned} &\left\| X^h(t_2) - X^h(t_1) \right\|_{L^p(\Omega,H)} \\ &\quad \leq C|t_2 - t_1|^{\frac{\beta}{2}} \left(1 + \left\| A^{\frac{\beta}{2}} X_0 \right\|_{L^p(\Omega,H)}\right), \; t_1, t_2 \in [0,T]. \end{aligned} \tag{22}$$

Moreover, if $\beta \in [1,2)$ and if Assumption 2.5 is fulfilled, then

$$\left\| X^h(t_2) - X^h(t_1) \right\|_{L^p(\Omega,H)} \leq C|t_2 - t_1|^{\frac{1}{2}} \left(1 + \left\| A^{\frac{\beta}{2}} X_0 \right\|_{L^p(\Omega,H)}\right), \; t_1, t_2 \in [0,T].$$

(ii) Let Assumptions 2.1–2.3 and 2.7 be fulfilled with $\beta \in [0,2)$. Then in the case of additive noise, the mild solution $X^h(t)$ of (20) satisfies the following regularity estimates

$$\left\| A_h^{\frac{\beta}{2}} X^h(t) \right\|_{L^p(\Omega,H)} \leq C\left(1 + \left\| A^{\frac{\beta}{2}} X_0 \right\|_{L^p(\Omega,H)}\right), \quad t \in [0,T], \tag{23}$$

$$\begin{aligned} &\left\| X^h(t_2) - X^h(t_1) \right\|_{L^p(\Omega,H)} \leq C|t_2 - t_1|^{\frac{\min(\beta,1)}{2}} \left(1 + \left\| A^{\frac{\beta}{2}} X_0 \right\|_{L^p(\Omega,H)}\right), \\ &t_1, t_2 \in [0,T]. \end{aligned} \tag{24}$$

Here $C = C(C_F, C_B, C_Q, \|F(0)\|, T, \beta)$ is a positive constant, independent of h, t, t_1 and t_2.

Corollary 2.1. *As a consequence of Lemma 2.1, it holds that*

$$\|X^h(t)\|_{L^p(\Omega,H)} \leq C, \quad \left\| F\left(X^h(t)\right) \right\|_{L^p(\Omega,H)} \leq C, \quad \left\| B\left(X^h(t)\right) \right\|_{L^p(\Omega,H)} \leq C, \quad t \in [0,T].$$

2.3. Standard linear implicit Euler method and stability properties

Let us recall that the linear implicit Euler scheme applied to the semi-discrete problem (30) is given by

$$Z_{m+1}^h = S_{h,\Delta t} Z_m^h + \Delta t S_{h,\Delta t} P_h F(Z_m^h) + S_{h,\Delta t} P_h B(Z_m^h), \tag{25}$$

$$S_{h,\Delta t} := (\mathbf{I} + \Delta t A_h)^{-1}, \quad Z_0^h = P_h X_0. \tag{26}$$

If the linear operator A tends to the null[2] operator, its corresponding discrete version A_h tends to the null operator and $S_{h,\Delta t}$ tends to the identity operator $\mathbf{I}$. In this case, the numerical scheme (25) and the standard exponential integrator [16] behave like the unstable Euler–Maruyama scheme. See also [20, Section 2.3] for more details. For a simple illustration of the stability properties of such problems, let us consider the following deterministic linear differential equation

$$y' = ay + cy, \quad a > 0, c < 0, \quad \text{such that} \quad c < -a. \tag{27}$$

[2] Think for instance of the Laplace operator $A = \alpha\Delta$, with $\alpha \longrightarrow 0$. Here the null operator is understood in the sense of $Au = 0$ for all $u \in \mathcal{D}(A)$.

The linear implicit Euler method applied to (27) by considering $F(y) = cy$ as the nonlinear part is given by

$$y_{n+1} = \frac{1 + c\Delta t}{1 - a\Delta t} y_n, \quad n \geq 0. \tag{28}$$

The numerical scheme (28) is stable [23,31] if and only if $\Delta t < \frac{2}{a-c}$. Note that when a is small enough and $|c|$ large enough, the numerical scheme (28) behaves like the explicit Euler method and the stability region becomes very small. Rosenbrock-type methods were proved to be efficient in such situations and were studied in [4,5,23] for ordinary differential equations. Applying the Rosenbrock–Euler method to the linear problem (27) yields

$$y_{n+1} = \frac{1}{1 - (a + c)\Delta t} y_n, \quad n \geq 0. \tag{29}$$

Note that (29) coincides with the full implicit method with $F(y) = cy$. Rosenbrock–Euler method (29) is unconditionally stable (A-stable). This demonstrates the strong stability property of Rosenbrock-type methods for stiff problems. Authors of [21,22] extended Rosenbrock-type methods to parabolic partial differential equations and the methods were proved to be efficient for solving transport equations in porous media [31]. To the best of our knowledge, the case of stiff stochastic partial differential equations is not yet studied in the scientific literature and will be the aim of this paper.

2.4. Novel fully discrete scheme and main results

Let us build a more stable scheme, robust when the linear operator A tends to null operator. For the time discretization, we consider the one-step method which provides the numerical approximated solution X_m^h of $X^h(t_m)$ at discrete time $t_m = m\Delta t$, $m = 0, \ldots, M$. The method is based on the continuous linearization of (20). More precisely, we linearize (20) at each time step as follows

$$dX^h(t) + [A_h X^h(t) + J_m^h X^h(t)]dt = G_m^h\left(X^h(t)\right)dt + P_h B\left(X^h(t)\right)dW(t), \; t_m \leq t \leq t_{m+1}, \tag{30}$$

where J_m^h is the Fréchet derivative of $P_h F$ at X_m^h and G_m^h is the remainder at X_m^h. Both J_m^h and G_m^h are random variables and are defined for all $\omega \in \Omega$ by

$$J_m^h(\omega) := (P_h F)'(X_m^h(\omega)) = P_h F'(X_m^h(\omega)), \tag{31}$$

$$G_m^h(\omega)(X^h(t)) := -P_h F(X^h(t)) + J_m^h(\omega) X^h(t). \tag{32}$$

Applying the linear implicit Euler method to (30) yields the following fully discrete scheme, called stochastic Rosenbrock-type scheme (SROS)

$$\begin{cases} X_0^h = P_h X_0, \\ X_{m+1}^h = S_{h,\Delta t}^m X_m^h + \Delta t S_{h,\Delta t}^m G_m^h(X_m^h) + S_{h,\Delta t}^m P_h B(X_m^h)\Delta W_m, \end{cases} \tag{33}$$

where ΔW_m and $S_{h,\Delta t}$ are defined respectively by

$$\Delta W_m := W_{t_{m+1}} - W_{t_m}, \qquad S_{h,\Delta t}^m(\omega) := \left(\mathbf{I} + \Delta t A_{h,m}(\omega)\right)^{-1}, \tag{34}$$

and the linear operator $A_{h,m}$ is given by

$$A_{h,m}(\omega) := A_h + J_m^h(\omega), \quad \omega \in \Omega. \tag{35}$$

In the numerical scheme (33), the resolvent operator (defined in (34)) is random and changes at each time step. Having the numerical method (33) in hand, our goal is to analyze its strong convergence toward the exact solution in the root-mean-square L^2 norm for multiplicative and additive noise.

Throughout this paper we take $t_m = m\Delta t \in [0, T]$, where $\Delta t = \frac{T}{M}$ for $m, M \in \mathbb{N}$, $m \leq M$, C is a generic constant that may change from one place to another but is independent of both Δt and h. The main results of this paper are formulated in the following theorems.

Theorem 2.2 (*Multiplicative Noise*). *Let $X(t_m)$ and X^h_m be respectively the mild solution given by (10) and the numerical approximation given by (33) at $t_m = m\Delta t$. Let Assumptions 2.1 and 2.2 (with $p = 2$), 2.3 and 2.4 be fulfilled.*

(i) If $0 < \beta < 1$, then the following error estimate holds

$$\|X(t_m) - X^h_m\|_{L^2(\Omega,H)} \leq C\left(h^\beta + \Delta t^{\frac{\beta}{2}}\right).$$

(ii) If $1 \leq \beta \leq 2$ and if Assumption 2.5 is fulfilled, then the following error estimate holds

$$\|X(t_m) - X^h_m\|_{L^2(\Omega,H)} \leq C\left(h^\beta + \Delta t^{\frac{1}{2}}\right),$$

where $C = C(C_F, C_B, T, \|F(0)\|, c, X_0)$ is a positive constant independent of h, M and Δt.

Theorem 2.3 (*Additive Noise*). *When dealing with additive noise (i.e. when $B = \mathbf{I}$), let Assumptions 2.1, 2.2 with $p = 4$, 2.3, 2.6 and 2.7 be fulfilled. Then the following error estimate holds for the mild solution $X(t)$ of (1) and the numerical approximation (33)*

$$\|X(t_m) - X^h_m\|_{L^2(\Omega,H)} \leq C\left(h^\beta + \Delta t^{\frac{\beta}{2}-\epsilon}\right), \tag{36}$$

where $C = C(C_F, C_Q, T, \|F(0)\|, X_0)$ is a positive constant independent of h, M and Δt.

3. Proof of the main results

The proofs of the main results require some preparatory results.

3.1. Preparatory results

For non commutative operators H_j in a Banach space, we introduce the following notation, which will be used in the rest of the paper.

$$\prod_{j=l}^{k} H_j := \begin{cases} H_k H_{k-1} \cdots H_l, & \text{if} \quad k \geq l, \\ \mathbf{I}, & \text{if} \quad k < l. \end{cases}$$

Lemma 3.1. *[20, Lemma 10] Let Assumption 2.2 be fulfilled. Then for all $\omega \in \Omega$ the following estimate holds*

$$\left\| \left(\prod_{j=l}^{m} e^{\Delta t A_{h,j}(\omega)} \right) A_h^\gamma \right\|_{L(H)} \leq C t_{m+1-l}^{-\gamma}, \quad 0 \leq l \leq m, \quad 0 \leq \gamma < 1. \tag{37}$$

Lemma 3.2 (*[20, Lemma 5]*). *For all $m \in \mathbb{N}$ and all $\omega \in \Omega$, the random linear operator $A_h + J^h_m(\omega)$ is the generator of an analytic semigroup $S^h_m(\omega)(t) =: e^{\left(A_h + J^h_m(\omega)\right)t}$, called random* *(or stochastic) perturbed semigroup, which is uniformly bounded on* $[0, T]$*, i.e. there exists a positive constant* $C_1 = C_1(b, T)$ *independent of* h*,* m*,* Δt *and the sample* ω *such that*

$$\left\| e^{\left(A_h + J^h_m(\omega)\right)t} \right\|_{L(H)} \begin{aligned} &\leq K e^{Kbt}, \quad t \geq 0 \\ &\leq C_1, \qquad 0 \leq t \leq T. \end{aligned}$$

The following lemma is an analogous of [18, (3.31)], but here our semigroup is not constant. In fact, it is random and further its changes at each time step.

Lemma 3.3. *Let Assumptions 2.1 and 2.3 be fulfilled.*

(i) For all $\alpha \in [0, 1]$*,* $n > 1$*,* $j \geq 0$ *and all* $\omega \in \Omega$*, it holds that*

$$\left\| A^\alpha_h \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-n} \right\|_{L(H)} \leq C((n-1)t)^{-\alpha} \leq C(nt)^{-\alpha}, \quad t > 0. \tag{38}$$

(ii) For all $\alpha \in [0, 1)$*,* $j \geq 0$ *and* $\omega \in \Omega$*, it holds that*

$$\left\| A^\alpha_h \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-1} \right\|_{L(H)} \leq Ct^{-\alpha}, \quad t > 0. \tag{39}$$

(iii) For all $n, j \in \mathbb{N}$*, it holds that*

$$\left\| \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-n} \right\|_{L(H)} \leq C, \quad t > 0, \tag{40}$$

where $C = C(b, T, \alpha)$ *is a positive constant independent of* h*,* j *and* Δt*.*

Proof. Note that for all $n \geq 2$, $\frac{1}{2}nt \leq (n-1)t$. Therefore $((n-1)t)^{-\alpha} \leq C(nt)^{-\alpha}$. It remains to prove the first inequality of (38). Using the interpolation theory, we only need to prove (38) for $\alpha = 0$ and $\alpha = 1$. Since $\frac{1}{t} > 0$ and the resolvent set of $-A_{h,j}$ contains $(0, \infty)$,[3] it follows from [24, (5.23)] that

$$\begin{aligned} \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-n} v &= t^{-n} \left(\frac{1}{t}\mathbf{I} + A_{h,j}(\omega)\right)^{-n} v \\ &= \frac{t^{-n}}{(n-1)!} \int_0^\infty s^{n-1} e^{-\frac{1}{t}s} S^h_j(\omega)(s) v ds, \quad v \in H. \end{aligned} \tag{41}$$

Taking the norm in both sides of (41) and using the uniformly boundedness of $S^h_j(\omega)$ (see Lemma 3.2) yields

$$\left\| \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-n} v \right\| \leq \frac{Ct^{-n}}{(n-1)!} \int_0^\infty s^{n-1} e^{-\frac{1}{t}s} \|v\| ds. \tag{42}$$

Using the change of variable $r = \frac{s}{t}$ yields

$$\left\| \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-n} v \right\| \leq \frac{C}{(n-1)!} \int_0^\infty r^{n-1} e^{-r} \|v\| dr \leq C\|v\|. \tag{43}$$

This shows that (38) holds for $\alpha = 0$. Pre-multiplying both sides of (41) by A_h yields

$$A_h \left(\mathbf{I} + tA_{h,j}(\omega)\right)^{-n} v = \frac{t^{-n}}{(n-1)!} \int_0^\infty s^{n-1} e^{-\frac{1}{t}s} A_h S^h_j(\omega)(s) v ds. \tag{44}$$

[3] Since Assumption 2.3 is fulfilled.

Taking the norm in both sides of (44) and using [20, Lemma 9 (iii)] yields

$$\left\|\left(\mathbf{I}+tA_{h,j}(\omega)\right)^{-n}v\right\| \leq \frac{Ct^{-n}}{(n-1)!}\int_0^{\infty} s^{n-2}e^{-\frac{1}{t}s}\|v\|ds. \tag{45}$$

Using the change of variable $r=\frac{s}{t}$ yields

$$\begin{aligned}\left\|\left(\mathbf{I}+tA_{h,j}(\omega)\right)^{-n}v\right\| &\leq \frac{Ct^{-1}}{(n-1)!}\int_0^{\infty} u^{n-2}e^{-r}\|v\|dr\\ &\leq \frac{Ct^{-1}(n-2)!}{(n-1)!}\|v\| = C\left((n-1)t\right)^{-1}\|v\|.\end{aligned} \tag{46}$$

This proves that (38) holds for $\alpha=1$, and the proof of (38) is completed by interpolation theory. The proofs of (39) and (40) follow from the integral equation (41). ■

The following lemma will be useful in our convergence analysis.

Lemma 3.4. *Let Assumptions 2.1 and 2.3 be fulfilled.*

(i) For all $\alpha\in(0,1]$ it holds that

$$\left\|A_h^{\alpha}\left(\prod_{j=i}^{m}S_{h,\Delta t}^{j}(\omega)\right)\right\|_{L(H)} \leq Ct_{m-i+1}^{-\alpha},\quad 0\leq i\leq m\leq M,\quad 0\leq k\leq M.$$

(ii) For all $\alpha_1,\alpha_2\in[0,1)$ it holds that

$$\left\|A_h^{\alpha_1}\left(\prod_{j=i}^{m}S_{h,\Delta t}^{j}(\omega)\right)A_h^{-\alpha_2}\right\|_{L(H)} \leq Ct_{m-i+1}^{-\alpha_1+\alpha_2},\quad 0\leq i\leq m\leq M,\quad 0\leq k\leq M,$$

where $C=C(b,T,\alpha,\alpha_1,\alpha_2)$ is a positive constant independent of h, i, m, M and Δt.

Proof. Note that the proof of the lemma in the case $i=m$ is straightforward from Lemma 3.3. We only concentrate on the case $i<m$.

(i) Using Lemma 3.3 it holds that

$$\left\|A_h^{\alpha}(\mathbf{I}+\Delta tA_{h,i}(\omega))^{-(m-i+1)}\right\|_{L(H)} \leq Ct_{m-i+1}^{-\alpha}. \tag{47}$$

It remains to estimate $A_h^{\alpha}\Delta_{m,i}^{h}(\omega)$, where

$$\Delta_{m,i}^{h}(\omega) := \prod_{j=i}^{m}S_{h,\Delta t}^{j}(\omega)-\left(S_{h,\Delta t}^{i}(\omega)\right)^{m-i+1}. \tag{48}$$

One can easily check that the following identity holds

$$\begin{aligned}&\left(\mathbf{I}+\Delta tA_{h,j+1}(\omega)\right)^{-1}-(\mathbf{I}+\Delta tA_{h,i}(\omega))^{-1}\\ &= \Delta t(\mathbf{I}+\Delta tA_{h,j+1}(\omega))^{-1}\left(A_{h,i}(\omega)-A_{h,j+1}(\omega)\right)(\mathbf{I}+\Delta tA_{h,i}(\omega))^{-1}\\ &= \Delta t(\mathbf{I}+\Delta tA_{h,j+1}(\omega))^{-1}\left(J_i^h(\omega)-J_{j+1}^h(\omega)\right)(\mathbf{I}+\Delta tA_{h,i}(\omega))^{-1}.\end{aligned} \tag{49}$$

Using the telescopic sum, it holds that

$$\Delta^h_{m,i}(\omega) = \sum_{j=0}^{m-i-1}\left(\prod_{k=j+i+1}^{m} S^k_{h,\Delta t}(\omega)\right)\left(\mathbf{I}+\Delta t A_{h,j+i+1}(\omega)\right)$$
$$\left[\left(\mathbf{I}+\Delta t A_{h,j+i+1}(\omega)\right)^{-1}-(\mathbf{I}+\Delta t A_{h,i}(\omega))^{-1}\right]\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-1}. \quad (50)$$

Substituting the identity (49) in (50) yields

$$\begin{aligned}
&\Delta^h_{m,i}(\omega)\\
&= \Delta t\sum_{j=0}^{m-i-1}\left(\prod_{k=j+i+1}^{m} S^k_{h,\Delta t}(\omega)\right)\left(J^h_i(\omega)-J^h_{j+i+1}(\omega)\right)\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-2}\\
&= \Delta t\sum_{j=0}^{m-i-1}\left(\mathbf{I}+\Delta t A_{h,j+i+1}(\omega)\right)^{-(m-j-i)}\left(J^h_i(\omega)-J^h_{j+i+1}(\omega)\right)\\
&\quad\times\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-2}\\
&\quad+\Delta t\sum_{j=0}^{m-i-1}\Delta^h_{m,j+i+1}(\omega)\left(J^h_i(\omega)-J^h_{j+i+1}(\omega)\right)\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-2}. \qquad (51)
\end{aligned}$$

Therefore we have

$$\begin{aligned}
&A^\alpha_h\Delta^h_{m,i}(\omega)\\
&= \Delta t\sum_{j=0}^{m-i-1}A^\alpha_h\left(\mathbf{I}+\Delta t A_{h,j+i+1}(\omega)\right)^{-(m-j-i)}\left(J^h_i(\omega)-J^h_{j+i+1}(\omega)\right)\\
&\quad\times\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-1}\\
&\quad+\Delta t\sum_{j=0}^{m-i-1}A^\alpha_h\Delta^h_{m,j+i+1}(\omega)\left(J^h_i(\omega)-J^h_{j+i+1}(\omega)\right)\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-1}. \qquad (52)
\end{aligned}$$

Taking the norm in both sides of (52), using triangle inequality and Lemma 3.3 yields

$$\begin{aligned}
\|A^\alpha_h\Delta^h_{m,i}(\omega)\|_{L(H)} &\le C\Delta t\sum_{j=0}^{m-i-1}t^{-\alpha}_{m-j-i}+C\Delta t\sum_{j=0}^{m-i-1}\|A^\alpha_h\Delta^h_{m,j+i+1}(\omega)\|_{L(H)}\\
&\le C+C\Delta t\sum_{j=i+1}^{m}\|A^\alpha_h\Delta^h_{m,j}(\omega)\|_{L(H)}. \qquad (53)
\end{aligned}$$

Applying the discrete Gronwall‘s lemma to (53) yields

$$\|A^\alpha_h\Delta^h_{m,i}(\omega)\|_{L(H)} \le C.$$

This completes the proof of (i).

(ii) Following the same lines as in Lemma 3.3, we can show that

$$\left\|A^{\alpha_1}_h\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-(m-i+1)}A^{-\alpha_2}_{h,i}\right\|_{L(H)} \le Ct^{-\alpha_1+\alpha_2}_{m-i+1}. \qquad (54)$$

It remains to bound $A_h^{\alpha_1}\Delta^h_{m,i}(\omega)A_h^{-\alpha_2}$, where $\Delta^h_{m,i}(\omega)$ is defined by (48). From (51), it holds that

$$
\begin{aligned}
A_h^{\alpha_1}\Delta^h_{m,i}(\omega)A_h^{-\alpha_2} = {}& \Delta t\sum_{j=0}^{m-i-1} A_h^{\alpha_1}\left(\mathbf{I}+\Delta t A_{h,j+i+1}(\omega)\right)^{-(m-j-i)}\\
&\times\left(J_i^h(\omega)-J^h_{j+i+1}(\omega)\right)\\
&\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-1}A_h^{-\alpha_2}\\
&+\Delta t\sum_{j=0}^{m-i-1} A_h^{\alpha_1}\Delta^h_{m,j+i+1}(\omega)\left(J_i^h(\omega)-J^h_{j+i+1}(\omega)\right)\\
&\left(\mathbf{I}+\Delta t A_{h,i}(\omega)\right)^{-j-1}A_h^{-\alpha_2}.
\end{aligned}
\tag{55}
$$

Taking the norm in both sides of (55), using triangle inequality, Lemmas 3.3 and 3.4(i) yields

$$
\begin{aligned}
\|A_h^{\alpha_1}\Delta^h_{m,i}(\omega)A_h^{-\alpha_2}\|_{L(H)} &\leq C\Delta t\sum_{j=0}^{m-i-1}\left\|A_h^{\alpha_1}\Delta^h_{m,j+i+1}(\omega)\right\|_{L(H)}\\
&\quad + C\Delta t\sum_{j=0}^{m-i-1}\left\|A_h^{\alpha_1}\left(\mathbf{I}+\Delta t A_{h,j+i+1}(\omega)\right)^{-(m-j-i)}\right\|_{L(H)}\\
&\leq C\Delta t\sum_{j=0}^{m-i-1} + C\Delta t\sum_{j=0}^{m-i-1} t_{m-j-i}^{-\alpha_1}\\
&\leq C.
\end{aligned}
$$

This proves (ii) and the proof of the lemma is completed. ∎

The following lemma will be useful in our convergence analysis.

Lemma 3.5. *Let Assumptions 2.1 and 2.3 be fulfilled.*

(i) For all $\alpha_1,\alpha_2\in(0,1]$, $0\leq j\leq M$ and $\omega\in\Omega$ the following estimate holds

$$
\left\|A_h^{-\alpha_1}\left(e^{A_{h,j}(\omega)\Delta t}-S^j_{h,\Delta t}(\omega)\right)A_h^{-\alpha_2}\right\|_{L(H)}\leq C\Delta t^{\alpha_1+\alpha_2}.
\tag{56}
$$

(ii) For all $\alpha_1\in[0,1]$, $\alpha_2\in(0,1)$, $0\leq j\leq M$ and $\omega\in\Omega$ the following estimate holds

$$
\left\|A_h^{\alpha_1}\left(e^{A_{h,j}(\omega)\Delta t}-S^j_{h,\Delta t}(\omega)\right)A_h^{-\alpha_2}\right\|_{L(H)}\leq C\Delta t^{-\alpha_1+\alpha_2},
\tag{57}
$$

where $C=C(b,T,\alpha_1,\alpha_2)$ is a positive constant independent of h, j, M and Δt.

Proof. We only prove (56) since the proof of (57) is similar. Let us set

$$
K^j_{h,\Delta t}(\omega):=e^{A_{h,j}(\omega)\Delta t}-S^j_{h,\Delta t}(\omega).
$$

One can easily check that

$$
\begin{aligned}
-K^j_{h,\Delta t}(\omega) &= \int_0^{\Delta t} \frac{d}{ds}\left(\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-1} e^{-(\Delta t-s)A_{h,j}(\omega)}\right) ds \\
&= \int_0^{\Delta t} sA^2_{h,j}(\omega)\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} e^{-(\Delta t-s)A_{h,j}(\omega)} ds \\
&= \int_0^{\Delta t} sA_{h,j}(\omega)\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} A_{h,j}(\omega) e^{-(\Delta t-s)A_{h,j}(\omega)} ds. \qquad (58)
\end{aligned}
$$

From (58) it holds that

$$
\begin{aligned}
-A_h^{-\alpha_1} K^j_{h,\Delta t}(\omega) A_h^{-\alpha_2} &= \int_0^{\Delta t} sA_h^{-\alpha_1} A_{h,j}(\omega)\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} \\
&\quad \times e^{-(\Delta t-s)A_{h,j}} A_{h,j}(\omega) A_h^{-\alpha_2} ds. \qquad (59)
\end{aligned}
$$

Taking the norm in both sides of (59) yields

$$
\begin{aligned}
&\left\| -A_h^{-\alpha_1} K^j_{h,\Delta t}(\omega) A_h^{-\alpha_2} \right\|_{L(H)} \\
&\leq \int_0^{\Delta t} s \left\| A_h^{-\alpha_1} A_{h,j}(\omega)\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} \right\|_{L(H)} \left\| e^{-(\Delta t-s)A_{h,j}(\omega)} A_{h,j}(\omega) A_h^{-\alpha_2} \right\|_{L(H)} ds. \qquad (60)
\end{aligned}
$$

Using triangle inequality and Lemma 3.3, it holds that

$$
\begin{aligned}
&\left\| A_h^{-\alpha_1} A_{h,j}(\omega)\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} \right\|_{L(H)} \\
&\leq \left\| A_h^{-\alpha_1+1}\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} \right\|_{L(H)} + \left\| A_h^{-\alpha_1} J^h_j(\omega)\left(\mathbf{I}+sA_{h,j}(\omega)\right)^{-2} \right\|_{L(H)} \\
&\leq Cs^{-1+\alpha_1} + C \\
&\leq Cs^{-1+\alpha_1}. \qquad (61)
\end{aligned}
$$

Using triangle inequality and [20, Lemma 9 (ii)], it holds that

$$
\begin{aligned}
&\left\| e^{-(\Delta t-s)A_{h,j}(\omega)} A_{h,j}(\omega) A_h^{-\alpha_2} \right\|_{L(H)} \\
&\leq \left\| e^{-(\Delta t-s)A_{h,j}(\omega)} A_h^{1-\alpha_2} \right\|_{L(H)} + \left\| e^{-(\Delta t-s)A_{h,j}(\omega)} J^h_j A_h^{-\alpha_2} \right\|_{L(H)} \\
&\leq C(\Delta t-s)^{-1+\alpha_2} + C \\
&\leq C(\Delta t-s)^{-1+\alpha_2}. \qquad (62)
\end{aligned}
$$

Substituting (62) and (61) in (60) yields

$$
\left\| -A_h^{-\alpha_1} K^j_{h,\Delta t}(\omega) A_h^{-\alpha_2} \right\|_{L(H)} \leq C \int_0^{\Delta t} ss^{-1+\alpha_1}(\Delta t-s)^{-1+\alpha_2} ds \leq C\Delta t^{\alpha_1+\alpha_2}.
$$

This completes the proof of (56). The proof of (57) is similar. ■

The following lemma can be found in [15].

Lemma 3.6. *For all $\alpha_1, \alpha_2 > 0$ and $\alpha \in [0, 1]$, there exist two positive constants $C_{\alpha_1\alpha_2}$ and C_{α,α_2} such that*

$$
\Delta t \sum_{j=1}^{m} t_{m-j+1}^{-1+\alpha_1} t_j^{-1+\alpha_2} \leq C_{\alpha_1\alpha_2} t_m^{-1+\alpha_1+\alpha_2}, \qquad \Delta t \sum_{j=1}^{m} t_{m-j+1}^{-\alpha} t_j^{-1+\alpha_2} \leq C_{\alpha\alpha_2} t_m^{-\alpha+\alpha_2}. \qquad (63)
$$

Proof. The proof of the first estimate of (63) comes from the comparison with the integral

$$\int_0^t (t-s)^{-1+\alpha_1} s^{-1+\alpha_2} ds.$$

The proof of the second estimate of (63) is a consequence of the first one. ■

Lemma 3.7. *Let $0 \leq \alpha < 2$ and let Assumption 2.1 be fulfilled.*

(i) If $v \in \mathcal{D}\left((A^{\frac{\alpha}{2}})\right)$, $\omega \in \Omega$, $0 \leq i \leq M$, then the following estimate holds

$$\left\| \left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S^j_{h,\Delta t}(\omega)\right) P_h v \right\| \leq C \Delta t^{\frac{\alpha}{2}} \|v\|_\alpha.$$

(ii) For non-smooth data, i.e. for $v \in H$ and for all $\omega \in \Omega$, $0 \leq i < m \leq M$, it holds that

$$\left\| \left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S^j_{h,\Delta t}(\omega)\right) P_h v \right\| \leq C \Delta t^{\frac{\alpha}{2}} t_{m-i}^{-\frac{\alpha}{2}} \|v\|.$$

(iii) For all $\alpha_1, \alpha_2 \in [0, 1)$ such that $\alpha_1 \leq \alpha_2$, $\omega \in \Omega$ and $0 \leq i < m \leq M$, it holds that

$$\left\| \left[\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) - \left(\prod_{j=i}^{m} S^j_{h,\Delta t}(\omega)\right) \right] A_h^{\alpha_1-\alpha_2} \right\|_{L(H)} \leq C \Delta t^{\alpha_2} t_{m-i}^{-\alpha_1},$$

where $C = C(b, T, \alpha, \alpha_1, \alpha_2)$ is a positive constant independent of h, i, m, M and Δt.

Proof.

(i) Using the telescopic identity, we have

$$\begin{aligned}
&\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S^j_{h,\Delta t}(\omega)\right) P_h v \\
&= \sum_{k=1}^{m-i+1} \left(\prod_{j=i+k}^{m} e^{A_{h,j}(\omega)\Delta t}\right) \left(e^{A_{h,i+k-1}(\omega)\Delta t} - S^{i+k-1}_{h,\Delta t}(\omega)\right) \left(\prod_{j=i}^{i+k-2} S^j_{h,\Delta t}(\omega)\right) P_h v. \quad (64)
\end{aligned}$$

Writing down the first and the last terms of (64) explicitly, we obtain

$$\begin{aligned}
&\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S^j_{h,\Delta t}(\omega)\right) P_h v \\
&= \left(e^{A_{h,m}(\omega)\Delta t} - S^m_{h,\Delta t}(\omega)\right) \left(\prod_{j=i}^{m-1} S^j_{h,\Delta t}(\omega)\right) P_h v \\
&+ \left(\prod_{j=i+1}^{m} e^{A_{h,j}(\omega)\Delta t}\right) \left(e^{A_{h,i}(\omega)\Delta t} - S^i_{h,\Delta t}(\omega)\right) P_h v \\
&+ \sum_{k=2}^{m-i} \left(\prod_{j=i+k}^{m} e^{A_{h,j}(\omega)\Delta t}\right) \left(e^{A_{h,i+k-1}(\omega)\Delta t} - S^{i+k-1}_{h,\Delta t}(\omega)\right) \left(\prod_{j=i}^{i+k-2} S^j_{h,\Delta t}(\omega)\right) P_h v. \quad (65)
\end{aligned}$$

Taking the norm in both sides of (65), inserting an appropriate power of A_h and using triangle inequality yields

$$
\begin{aligned}
&\left\| \left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t} \right) P_h v - \left(\prod_{j=i}^{m} S^j_{h,\Delta t}(\omega) \right) P_h v \right\| \\
&\leq \left\| \left(e^{A_{h,m}(\omega)\Delta t} - S^m_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} A_h^{\frac{\alpha}{2}} \left(\prod_{j=i}^{m-1} S^j_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} A_h^{\frac{\alpha}{2}} P_h v \right\| \\
&+ \left\| \left(\prod_{j=i+1}^{m} e^{A_{h,j}(\omega)\Delta t} \right) \left(e^{A_{h,i}(\omega)\Delta t} - S^i_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} A_h^{\frac{\alpha}{2}} P_h v \right\| \\
&+ \sum_{k=2}^{m-i} \left\| \left(\prod_{j=i+k}^{m} e^{A_{h,j}(\omega)\Delta t} \right) A_h^{1-\epsilon} A_h^{-1+\epsilon} \left(e^{A_{h,i+k-1}(\omega)\Delta t} - S^{i+k-1}_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}-\epsilon} \right. \\
&\quad \left. .A_h^{\frac{\alpha}{2}+\epsilon} \left(\prod_{j=i}^{i+k-2} S^j_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} A_h^{\frac{\alpha}{2}} P_h v \right\| \\
&=: I_1 + I_2 + I_3.
\end{aligned} \tag{66}
$$

Using Lemmas 3.5, 3.4(ii) and [20, Lemma 1] yields

$$
\begin{aligned}
&I_1 \\
&\leq \left\| \left(e^{A_{h,m}(\omega)\Delta t} - S^m_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} \right\|_{L(H)} \left\| A_h^{\frac{\alpha}{2}} \left(\prod_{j=i}^{m-1} S^j_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} \right\|_{L(H)} \| A_h^{\frac{\alpha}{2}} P_h v \| \\
&\leq C \Delta t^{\frac{\alpha}{2}} \| v \|_{\alpha}.
\end{aligned} \tag{67}
$$

Using Lemmas 3.1, 3.5 and [20, Lemma 1] yields

$$
\begin{aligned}
I_2 &\leq \left\| \left(\prod_{j=i+1}^{m} e^{A_{h,j}(\omega)\Delta t} \right) \right\|_{L(H)} \left\| \left(e^{A_{h,i}(\omega)\Delta t} - S^i_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} \right\|_{L(H)} \| A_h^{\frac{\alpha}{2}} P_h v \| \\
&\leq C \Delta t^{\frac{\alpha}{2}} \| v \|_{\alpha}.
\end{aligned} \tag{68}
$$

Using Lemmas 3.1, 3.5, 3.4(ii), 3.6 and [20, Lemma 1] yields

$$
\begin{aligned}
I_3 &\leq \sum_{k=2}^{m-i} \left\| \left(\prod_{j=i+k}^{m} e^{A_{h,j}(\omega)\Delta t} \right) A_h^{1-\epsilon} \right\|_{L(H)} \\
&\times \left\| A_h^{-1+\epsilon} \left(e^{A_{h,i+k-1}(\omega)\Delta t} - S^{i+k-1}_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}-\epsilon} \right\|_{L(H)} \\
&\times \left\| A_h^{\frac{\alpha}{2}+\epsilon} \left(\prod_{j=i}^{i+k-2} S^j_{h,\Delta t}(\omega) \right) A_h^{-\frac{\alpha}{2}} \right\|_{L(H)} \| A_h^{\frac{\alpha}{2}} P_h v \|
\end{aligned}
$$

$$\leq C\sum_{k=2}^{m-i} t_{m+1-i-k}^{-1+\epsilon}\Delta t^{1+\frac{\alpha}{2}} t_{k-1}^{-\epsilon} = C\Delta t^{\frac{\alpha}{2}}\sum_{k=2}^{m-i} t_{m-i-k+1}^{-1+\epsilon} t_{k-1}^{-\epsilon}\Delta t$$
$$\leq C\Delta t^{\frac{\alpha}{2}}. \tag{69}$$

Substituting (69), (68) and (67) in (66) yields

$$\left\|\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S_{h,\Delta t}^{j}(\omega)\right) P_h v\right\| \leq C\Delta t^{\frac{\alpha}{2}}\|v\|_{\alpha}.$$

This completes the proof of (i).

(ii) For non-smooth initial data, taking the norm in both sides of (65) and inserting an appropriate power of A_h yields

$$\begin{aligned}
&\left\|\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S_{h,\Delta t}^{j}(\omega)\right) P_h v\right\| \\
&\leq \left\|\left(e^{A_{h,m}(\omega)\Delta t} - S_{h,\Delta t}^{m}(\omega)\right) A_h^{-\frac{\alpha}{2}}\right\|_{L(H)} \left\|A_h^{\frac{\alpha}{2}}\left(\prod_{j=i}^{m-1} S_{h,\Delta t}^{j}(\omega)\right) P_h v\right\| \\
&+ \left\|\left(\prod_{j=i+1}^{m} e^{A_{h,j}(\omega)\Delta t}\right) A_h^{\frac{\alpha}{2}}\right\|_{L(H)} \left\|A_h^{-\frac{\alpha}{2}}\left(e^{A_{h,i}(\omega)\Delta t} - S_{h,\Delta t}^{i}(\omega)\right) P_h v\right\| \\
&+ \sum_{k=2}^{m-i}\left\|\left(\prod_{j=i+k}^{m} e^{A_{h,j}(\omega)\Delta t}\right) A_h^{1-\epsilon}\right\|_{L(H)} \\
&\times \left\|A_h^{-1+\epsilon}\left(e^{A_{h,i+k-1}(\omega)\Delta t} - S_{h,\Delta t}^{i+k-1}(\omega)\right) A_h^{-1+\epsilon}\right\|_{L(H)} \\
&\times \left\|A_h^{1-\epsilon}\left(\prod_{j=i}^{i+k-2} S_{h,\Delta t}^{j}(\omega)\right) P_h v\right\|.
\end{aligned} \tag{70}$$

Using Lemmas 3.5, 3.4(i), 3.6 and 3.1, it follows that

$$\begin{aligned}
&\left\|\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right) P_h v - \left(\prod_{j=i}^{m} S_{h,\Delta t}^{j}(\omega)\right) P_h v\right\| \\
&\leq C\Delta t^{\frac{\alpha}{2}} t_{m-i}^{-\frac{\alpha}{2}}\|v\| + C\Delta t^{\frac{\alpha}{2}} t_{m-i}^{-\frac{\alpha}{2}}\|v\| + C\Delta t^{1-\epsilon}\sum_{k=2}^{m-i}\Delta t t_{m-i-k+1}^{-1+\epsilon} t_{k-1}^{-1+\epsilon}\|v\| \\
&\leq C\Delta t^{\frac{\alpha}{2}} t_{m-i-k}^{-\frac{\alpha}{2}}\|v\| + C\Delta t^{\alpha/2} t_{m-i}^{-\frac{\alpha}{2}}\|v\| + C\Delta t^{1-\epsilon} t_{m-i}^{-1+2\epsilon}\|v\| \\
&\leq C\Delta t^{\frac{\alpha}{2}} t_{m-i}^{-\frac{\alpha}{2}}\|v\|.
\end{aligned} \tag{71}$$

(iii) Taking the norm in both sides of (65) and inserting an appropriate power of A_h yields

$$
\begin{aligned}
&\left\|\left[\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right)-\left(\prod_{j=i}^{m} S_{h,\Delta t}^{j}(\omega)\right)\right] A_h^{\alpha_1-\alpha_2}\right\|_{L(H)}\\
&\leq \left\|\left(e^{A_{h,m}(\omega)\Delta t}-S_{h,\Delta t}^{m}(\omega)\right) A_h^{-\alpha_2}\right\|_{L(H)} \left\|A_h^{\alpha_2}\left(\prod_{j=i}^{m-1} S_{h,\Delta t}^{j}(\omega)\right) A_h^{\alpha_1-\alpha_2}\right\|_{L(H)}\\
&+\left\|\left(\prod_{j=i+1}^{m} e^{A_{h,j}(\omega)\Delta t}\right) A_h^{\alpha_1}\right\|_{L(H)} \left\|A_h^{-\alpha_1}\left(e^{A_{h,i}(\omega)\Delta t}-S_{h,\Delta t}^{i}(\omega)\right) A_h^{-(\alpha_2-\alpha_1)}\right\|_{L(H)}\\
&+\sum_{k=2}^{m-i}\left\|\left(\prod_{j=i+k}^{m} e^{A_{h,j}(\omega)\Delta t}\right) A_h^{\alpha_2+\epsilon}\right\|_{L(H)}\\
&\times\left\|A_h^{-\alpha_2-\epsilon}\left(e^{A_{h,i+k-1}(\omega)\Delta t}-S_{h,\Delta t}^{i+k-1}(\omega)\right) A_h^{-1+\epsilon}\right\|_{L(H)}\\
&\times\left\|A_h^{1-\epsilon}\left(\prod_{j=i}^{i+k-2} S_{h,\Delta t}^{j}(\omega)\right) A_h^{-(\alpha_2-\alpha_1)}\right\|_{L(H)}.
\end{aligned}
\tag{72}
$$

Using Lemmas 3.5, 3.4(ii), 3.6 and 3.1, it follows from (72) that

$$
\begin{aligned}
&\left\|\left[\left(\prod_{j=i}^{m} e^{A_{h,j}(\omega)\Delta t}\right)-\left(\prod_{j=i}^{m} S_{h,\Delta t}^{j}(\omega)\right)\right] A_h^{\alpha_1-\alpha_2}\right\|_{L(H)}\\
&\leq C\Delta t^{\alpha_2} t_{m-i}^{-\alpha_1}+C\Delta t^{\alpha_2} t_{m-i}^{-\alpha_1}+C\Delta t^{\alpha_2}\sum_{k=2}^{m-i}\Delta t\, t_{m-i-k+1}^{-\alpha_2-\epsilon} t_{k-1}^{-1+\epsilon+\alpha_2-\alpha_1}\\
&\leq C\Delta t^{\alpha_2} t_{m-i-k}^{-\alpha_1}+C\Delta t^{\alpha_2} t_{m-i}^{-\alpha_1}+C\Delta t^{\alpha_2} t_{m-i}^{-\alpha_1}\\
&\leq C\Delta t^{\alpha_2} t_{m-i}^{-\alpha_1}.
\end{aligned}
$$

This completes the proof of (iii). ■

Lemma 3.8.

(i) Let Assumption 2.6 be fulfilled. Then the following estimate holds

$$
\left\|(A_h)^{\frac{\beta-1}{2}} P_h Q^{\frac{1}{2}}\right\|_{\mathcal{L}_2(H)} \leq C_Q,
$$

where β comes from Assumption 2.1.

(ii) Under Assumption 2.7, for all $\omega\in\Omega$ and $m\in\mathbb{N}$, the following estimates hold

$$
\begin{aligned}
\left\|\left(G_m^h(\omega)\right)'(u)v\right\| &\leq C\|v\|, \quad u, v\in H,\\
\left\|(A_h)^{\frac{-\eta}{2}}\left(G_m^h(\omega)\right)''(u)(v_1, v_2)\right\| &\leq C\|v_1\|\|v_2\|, \quad u, v_1, v_2\in H,
\end{aligned}
$$

where η comes from Assumption 2.7 and $C = C(C_F, T, L, \eta)$ is a positive constant independent of h, ω, m, M and Δt..

Proof. The proof of (i) can be found in [20, Lemma 11] and the proof of (ii) can be found in [20, Lemma 12]. ■

With the above preparation, we are now in position to prove our main results.

3.2. *Proof of Theorem 2.2*

Iterating the numerical solution (33) at t_m by replacing X_i^h, $i = m-1, \ldots, 2, 1$ by its expression only on the first term yields

$$X_m^h = \left(\prod_{k=0}^{m-1} S_{h,\Delta t}^k\right) P_h X_0 + \Delta t S_{h,\Delta t}^{m-1} G_{m-1}^h(X_{m-1}^h) + S_{h,\Delta t}^{m-1} P_h B(X_{m-1}^h)\Delta W_{m-1}$$
$$+ \Delta t \sum_{i=2}^{m-1}\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^j\right) G_{m-i}^h(X_{m-i}^h) + \sum_{i=2}^{m-1}\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^j\right) P_h B(X_{m-i}^h)\Delta W_{m-i}. \quad (73)$$

Iterating the mild solution (30) at time t_m yields

$$X^h(t_m) = \left(\prod_{k=0}^{m-1} e^{A_{h,k}\Delta t}\right) P_h X_0 + \int_{t_{m-1}}^{t_m} e^{(t_m-s)A_{h,m-1}} G_{m-1}^h\left(X^h(s)\right) ds$$
$$+ \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) e^{(t_{m-i}-s)A_{h,m-i-1}} G_{m-i}^h\left(X^h(s)\right) ds$$
$$+ \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) e^{(t_{m-i}-s)A_{h,m-i-1}} P_h B\left(X^h(s)\right) dW(s)$$
$$+ \int_{t_{m-1}}^{t_m} e^{(t_m-s)A_{h,m-1}} P_h B\left(X^h(s)\right) dW(s). \quad (74)$$

Subtracting (74) from (73), taking the L^2 norm and using triangle inequality yields

$$\left\|X^h(t_m) - X_m^h\right\|_{L^2(\Omega,H)}^2 \leq 25\sum_{i=0}^{4}\|II_i\|_{L^2(\Omega,H)}^2, \quad (75)$$

where

$$II_0 = \left(\prod_{j=0}^{m-1} e^{A_{h,j}\Delta t}\right) P_h X_0 - \left(\prod_{j=0}^{m-1} S_{h,\Delta t}^j\right) P_h X_0,$$
$$II_1 = \int_{t_{m-1}}^{t_m}\left(e^{(t_m-s)A_{h,m-1}} G_{m-1}^h\left(X^h(s)\right) - S_{h,\Delta t}^{m-1} G_{m-1}^h\left(X_{m-1}^h\right)\right) ds,$$
$$II_2 = \int_{t_{m-1}}^{t_m}\left(e^{(t_m-s)A_{h,m-1}} P_h B\left(X^h(s)\right) - S_{h,\Delta t}^{m-1} P_h B\left(X_{m-1}^h\right)\right) dW(s),$$
$$II_3 = \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) e^{(t_{m-i}-s)A_{h,m-i-1}} G_{m-i}^h\left(X^h(s)\right) ds$$
$$- \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^j\right) G_{m-i}^h\left(X_{m-i}^h\right) ds,$$

$$
\begin{aligned}
II_4 = & \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) e^{(t_{m-i}-s)A_{h,m-i-1}} P_h B\left(X^h(s)\right) dW(s) \\
& - \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} S^j_{h,\Delta t}\right) P_h B\left(X^h_{m-i}\right) dW(s).
\end{aligned}
$$

In the following sections we estimate II_i, $i = 0, \ldots, 4$ separately.

3.2.1. Estimation of II_0, II_1 and II_2

Using Lemma 3.7(i) with $\alpha = \beta$, it holds that

$$
\begin{aligned}
\|II_0\|_{L^2(\Omega,H)} &\leq \left(\mathbb{E}\left[\left\|\left(\prod_{j=0}^{m-1} e^{A_{h,j}\Delta t}\right) P_h X_0 - \left(\prod_{j=0}^{m-1} S^j_{h,\Delta t}\right) P_h X_0\right\|^2\right]\right)^{\frac{1}{2}} \\
&\leq C\Delta t^{\frac{\beta}{2}}\left(\left[\mathbb{E}\|X_0\|^2_\beta\right]\right)^{\frac{1}{2}} \leq C\Delta t^{\frac{\beta}{2}}.
\end{aligned} \tag{76}
$$

The term II_1 can be recast in three terms as follows:

$$
\begin{aligned}
II_1 = & \int_{t_{m-1}}^{t_m} e^{(t_m-s)A_{h,m-1}}\left(G^h_{m-1}\left(X^h(s)\right) - G^h_{m-1}\left(X^h(t_{m-1})\right)\right) ds \\
& + \int_{t_{m-1}}^{t_m}\left(e^{(t_m-s)A_{h,m-1}} - S^{m-1}_{h,\Delta t}\right) G^h_{m-1}\left(X^h(t_{m-1})\right) ds \\
& + \int_{t_{m-1}}^{t_m} S^{m-1}_{h,\Delta t}\left(G^h_{m-1}\left(X^h(t_{m-1})\right) - G^h_{m-1}\left(X^h_{m-1}\right)\right) ds \\
:= & \; II_{11} + II_{12} + II_{13}.
\end{aligned} \tag{77}
$$

Therefore using triangle inequality we obtain

$$
\|II_1\|_{L^2(\Omega,H)} \leq \|II_{11}\|_{L^2(\Omega,H)} + \|II_{12}\|_{L^2(\Omega,H)} + \|II_{13}\|_{L^2(\Omega,H)}. \tag{78}
$$

Using Corollary 2.1 yields

$$
\begin{aligned}
\|II_{11}\|_{L^2(\Omega,H)} &\leq C\int_{t_{m-1}}^{t_m}\left\|G^h_{m-1}\left(X^h(s)\right)\right\|_{L^2(\Omega,H)} ds \\
&\quad + \int_{t_{m-1}}^{t_m}\left\|G^h_{m-1}\left(X^h(t_{m-1})\right)\right\|_{L^2(\Omega,H)} ds \\
&\leq C\int_{t_{m-1}}^{t_m}\left(1 + \|X_0\|_{L^2(\Omega,H)}\right) ds \leq C\Delta t.
\end{aligned} \tag{79}
$$

Using Lemma 3.7(i) with $\alpha = 0$ and Corollary 2.1, it holds that

$$
\begin{aligned}
\|II_{12}\|_{L^2(\Omega,H)} &\leq C\int_{t_{m-1}}^{t_m}\left\|G^h_{m-1}\left(X^h(t_{m-1})\right)\right\|_{L^2(\Omega,H)} ds \leq C\int_{t_{m-1}}^{t_m}\left(1 + \|X_0\|_{L^2(\Omega,H)}\right) ds \\
&\leq C\Delta t.
\end{aligned} \tag{80}
$$

Using Lemma 3.4(i) with $\alpha = 0$ and Assumption 2.3, it holds that

$$
\|II_{13}\|_{L^2(\Omega,H)} \leq C\Delta t\left\|X^h(t_{m-1}) - X^h_{m-1}\right\|_{L^2(\Omega,H)}. \tag{81}
$$

Substituting (81), (80) and (79) in (78) yields

$$\|II_1\|_{L^2(\Omega,H)} \leq C\Delta t + C\Delta t \left\|X^h(t_{m-1}) - X^h_{m-1}\right\|_{L^2(\Omega,H)}. \tag{82}$$

We can recast II_2 as follows:

$$\begin{aligned} II_2 &= \int_{t_{m-1}}^{t_m} e^{(t_m-s)A_{h,m-1}} \left(P_hB\left(X^h(s)\right) - P_hB\left(X^h(t_{m-1})\right)\right) dW(s) \\ &+ \int_{t_{m-1}}^{t_m} \left(e^{(t_m-s)A_{h,m-1}} - S^{m-1}_{h,\Delta t}\right) P_hB\left(X^h(t_{m-1})\right) dW(s) \\ &+ \int_{t_{m-1}}^{t_m} S^{m-1}_{h,\Delta t} \left(P_hB\left(X^h(t_{m-1})\right) - P_hB\left(X^h_{m-1}\right)\right) dW(s) \\ &:= II_{21} + II_{22} + II_{23}. \end{aligned} \tag{83}$$

Therefore using triangle inequality we obtain

$$\|II_2\|^2_{L^2(\Omega,H)} \leq 9\|II_{21}\|^2_{L^2(\Omega,H)} + 9\|II_{22}\|^2_{L^2(\Omega,H)} + 9\|II_{23}\|^2_{L^2(\Omega,H)}. \tag{84}$$

Using Itô-isometry, [20, Lemma 5], Assumption 2.4 and Lemma 3.2, it holds that

$$\begin{aligned} \|II_{21}\|^2_{L^2(\Omega,H)} &= \int_{t_{m-1}}^{t_m} \left\|e^{(t_m-s)A_{h,m-1}} \left(P_hB\left(X^h(s)\right) - P_hB\left(X^h(t_{m-1})\right)\right)\right\|^2_{L^2(\Omega,H)} ds \\ &\leq C \int_{t_{m-1}}^{t_m} (s-t_{m-1})^{\min(\beta,1)} ds \leq C\Delta t^{\min(\beta+1,2)}. \end{aligned} \tag{85}$$

Using again Itô-isometry, Lemma 3.7(i) with $\alpha = 0$ and Corollary 2.1 yields

$$\begin{aligned} \|II_{22}\|^2_{L^2(\Omega,H)} &= \int_{t_{m-1}}^{t_m} \left\|\left(e^{(t_m-s)A_{h,m-1}} - S^{m-1}_{h,\Delta t}\right) P_hB\left(X^h(t_{m-1})\right)\right\|^2_{L^2(\Omega,H)} ds \\ &\leq C \int_{t_{m-1}}^{t_m} \left(1 + \|X_0\|^2_{L^2(\Omega,H)}\right) ds \leq C\Delta t. \end{aligned} \tag{86}$$

The Itô-isometry together with Lemma 3.4(i) (with $\alpha = 0$) and Assumption 2.4 yields

$$\begin{aligned} \|II_{23}\|^2_{L^2(\Omega,H)} &= \int_{t_{m-1}}^{t_m} \left\|S^{m-1}_{h,\Delta t} \left(P_hB\left(X^h(t_{m-1})\right) - P_hB\left(X^h_{m-1}\right)\right)\right\|^2_{L^2(\Omega,H)} ds \\ &\leq C\Delta t \left\|X^h(t_{m-1}) - X^h_{m-1}\right\|^2_{L^2(\Omega,H)}. \end{aligned} \tag{87}$$

Substituting (87), (86) and (85) in (84) yields

$$\|II_2\|^2_{L^2(\Omega,H)} \leq C\Delta t + C\Delta t \left\|X^h(t_{m-1}) - X^h_{m-1}\right\|^2_{L^2(\Omega,H)}. \tag{88}$$

3.2.2. Estimation of II_3

We can recast II_3 in four terms as follows:

$$\begin{aligned} II_3 &= \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) \left(e^{(t_{m-i}-s)A_{h,m-i-1}} - \mathbf{I}\right) G^h_{m-i}\left(X^h(s)\right) ds \\ &+ \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) \left(G^h_{m-i}\left(X^h(s)\right) - G^h_{m-i}\left(X^h(t_{m-i})\right)\right) ds \end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left[\left(\prod_{j=m-i}^{m-1}e^{\Delta t A_{h,j}}\right)-\left(\prod_{j=m-i}^{m-1}S^j_{h,\Delta t}\right)\right]G^h_{m-i}\left(X^h(t_{m-i})\right)ds\\
&+\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1}S^j_{h,\Delta t}\right)\left(G^h_{m-i}\left(X^h(t_{m-i})\right)-G^h_{m-i}\left(X^h_{m-i}\right)\right)ds\\
&:= II_{31}+II_{32}+II_{33}+II_{34}.
\end{aligned}
$$

Therefore, using triangle inequality we obtain

$$
\|II_3\|_{L^2(\Omega,H)}\leq\|II_{31}\|_{L^2(\Omega,H)}+\|II_{32}\|_{L^2(\Omega,H)}+\|II_{33}\|_{L^2(\Omega,H)}+\|II_{34}\|_{L^2(\Omega,H)}. \tag{89}
$$

Inserting an appropriate power of A_h, using Lemma 3.1 and Corollary 2.1 yields

$$
\begin{aligned}
\|II_{31}\|_{L^2(\Omega,H)}&\leq\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1}e^{\Delta t A_{h,j}}\right)A_h^{1-\epsilon}\right\|^2_{L(H)}\right.\\
&\quad\left.\times\left\|A_h^{-1+\epsilon}\left(e^{(t_{m-i}-s)A_{h,m-i-1}}-\mathbf{I}\right)\right\|^2_{L(H)}\left\|G^h_{m-i}\left(X^h(s)\right)\right\|^2ds\right]^{\frac{1}{2}}\\
&\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}t_i^{-1+\epsilon}(t_{m-i}-s)^{1-\epsilon}ds\\
&\leq C\Delta t^{1-\epsilon}\sum_{i=2}^{m-1}\Delta t\ t_i^{-1+\epsilon}\leq C\Delta t^{1-\epsilon}.
\end{aligned} \tag{90}
$$

Using triangle inequality, Lemma 3.1, Assumption 2.3 and Lemma 2.1 yields

$$
\begin{aligned}
&\|II_{32}\|_{L^2(\Omega,H)}\\
&\leq\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1}e^{\Delta t A_{h,j}}\right)\right\|^2_{L(H)}\|G^h_{m-i}\left(X^h(s)\right)-G^h_{m-i}\left(X^h(t_{m-i})\right)\|^2\right]^{\frac{1}{2}}ds\\
&\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left\|X^h(s)-X^h(t_{m-i})\right\|_{L^2(\Omega,H)}ds\\
&\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}(t_{m-i}-s)^{\frac{\min(\beta,1)}{2}}ds\leq C\Delta t^{\frac{\min(\beta,1)}{2}}.
\end{aligned} \tag{91}
$$

Using triangle inequality, Lemma 3.7(ii) and Corollary 2.1, it holds that

$$
\begin{aligned}
\|II_{33}\|_{L^2(\Omega,H)}&\leq\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1}e^{\Delta t A_{h,j}}\right)-\left(\prod_{j=m-i}^{m-1}S^j_{h,\Delta t}\right)\right\|^2_{L(H)}\right.\\
&\quad\left.\times\left\|G^h_{m-i}\left(X^h(t_{m-i})\right)\right\|^2\right]^{\frac{1}{2}}ds\\
&\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}t_{i-1}^{-\frac{\beta}{2}}\Delta t^{\frac{\beta}{2}}ds\leq C\Delta t^{\frac{\beta}{2}}.
\end{aligned} \tag{92}
$$

Using Lemma 3.4(i) with $\alpha = 0$ and Assumption 2.3 yields

$$\begin{aligned}
&\|II_{34}\|_{L^2(\Omega,H)}\\
&\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^{j}\right)\right\|^2_{L(H)}\right.\\
&\quad \left.\times \left\|\left(G^h_{m-i}\left(X^h(t_{m-i})\right) - G^h_{m-i}\left(X^h_{m-i}\right)\right)\right\|^2_{L^2(\Omega,H)}\right]^{\frac{1}{2}} ds\\
&\leq C\Delta t \sum_{i=2}^{m-1} \left\|X^h(t_{m-i}) - X^h_{m-i}\right\|_{L^2(\Omega,H)}. \qquad (93)
\end{aligned}$$

Substituting (93), (92), (91) and (90) in (89) yields

$$\|II_3\|_{L^2(\Omega,H)} \leq C\Delta t^{\frac{\min(\beta,1)}{2}} + C\Delta t \sum_{i=2}^{m-1} \left\|X^h(t_{m-i}) - X^h_{m-i}\right\|_{L^2(\Omega,H)}. \qquad (94)$$

3.2.3. Estimation of II_4

We can recast II_4 in four terms as follows.

$$\begin{aligned}
II_4 &= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(e^{(t_{m-i}-s)A_{h,m-i-1}} - \mathbf{I}\right) P_h B\left(X^h(s)\right) dW(s)\\
&\quad + \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(P_h B\left(X^h(s)\right) - P_h B\left(X^h(t_{m-i})\right)\right) dW(s)\\
&\quad + \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) - \left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^{j}\right)\right] P_h B\left(X^h(t_{m-i})\right) dW(s)\\
&\quad + \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^{j}\right)\left[P_h B\left(X^h(t_{m-i})\right) - P_h B\left(X^h_{m-i}\right)\right] dW(s)\\
&:= II_{41} + II_{42} + II_{43} + II_{44}. \qquad (95)
\end{aligned}$$

Therefore using triangle inequality we have

$$\|II_4\|^2_{L^2(\Omega,H)} \leq 16\|II_{41}\|^2_{L^2(\Omega,H)} + 16\|II_{42}\|^2_{L^2(\Omega,H)} + 16\|II_{43}\|^2_{L^2(\Omega,H)} + 16\|II_{44}\|^2_{L^2(\Omega,H)}. \qquad (96)$$

Using Itô isometry, inserting an appropriate power of A_h, using Lemma 3.1 and Corollary 2.1 yields

$$\begin{aligned}
&\|II_{41}\|^2_{L^2(\Omega,H)}\\
&= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(e^{(t_{m-i}-s)A_{h,m-i-1}} - \mathbf{I}\right) P_h B\left(X^h(s)\right)\right\|^2_{L^0_2}\right] ds
\end{aligned}$$

$$\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) A_h^{\frac{1-\epsilon}{2}}\right\|_{L(H)}^2 \right.$$
$$\left. \times \left\|A_h^{\frac{-1+\epsilon}{2}}\left(e^{(t_{m-i}-s)A_{h,m-i-1}}-\mathbf{I}\right)\right\|_{L(H)}^2 \left\|P_h B\left(X^h(s)\right)\right\|_{L_2^0}^2\right] ds$$
$$\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} t_i^{-1+\epsilon}(t_{m-i}-s)^{1-\epsilon} ds \leq C\Delta t^{1-\epsilon}\sum_{i=2}^{m-1}\Delta t\ t_i^{-1+\epsilon} \leq C\Delta t^{1-\epsilon}. \quad (97)$$

Using Itô isometry, Lemma 3.1, Assumption 2.4 and Lemma 2.1 yields

$$\|II_{42}\|^2_{L^2(\Omega,H)}$$
$$= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(P_h B\left(X^h(s)\right) - P_h B\left(X^h(t_{m-i})\right)\right)\right\|_{L_2^0}^2\right] ds$$
$$\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\right\|_{L(H)}^2 \left\|P_h B\left(X^h(s)\right) - P_h B\left(X^h(t_{m-i})\right)\right\|_{L_2^0}^2\right] ds$$
$$\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \left\|X^h(s) - X^h(t_{m-i})\right\|^2_{L^2(\Omega,H)} ds$$
$$\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} (t_{m-i}-s)^{\min(\beta,1)} ds \leq C\Delta t^{\min(\beta,1)}. \quad (98)$$

Using Itô isometry, Lemma 3.7(ii) with $\alpha = \frac{1-\epsilon}{2}$ and Corollary 2.1, it holds that

$$\|II_{43}\|^2_{L^2(\Omega,H)}$$
$$= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) - \left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^j\right)\right] P_h B\left(X^h(t_{m-i})\right)\right\|_{L_2^0}^2\right] ds$$
$$\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) - \left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^j\right)\right\|_{L(H)}^2 \left\|P_h B\left(X^h(t_{m-i})\right)\right\|_{L_2^0}^2\right] ds$$
$$\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} t_{i-1}^{1-\epsilon}\Delta t^{1-\epsilon} ds \leq C\Delta t^{1-\epsilon}. \quad (99)$$

Using Itô isometry, Lemma 3.4(i) with $\alpha = 0$ and Assumption 2.4 yields

$$\|II_{44}\|^2_{L^2(\Omega,H)}$$
$$= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^j\right)\left(P_h B\left(X^h(t_{m-i})\right) - P_h B\left(X^h_{m-i}\right)\right)\right\|_{L_2^0}^2\right] ds$$

$$\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} S^j_{h,\Delta t}\right)\right\|^2_{L(H)} \left\|P_hB\left(X^h(t_{m-i})\right)-P_hB\left(X^h_{m-i}\right)\right\|^2_{L^0_2}\right]ds$$
$$\leq C\Delta t\sum_{i=2}^{m-1}\left\|X^h(t_{m-i})-X^h_{m-i}\right\|^2_{L^2(\Omega,H)}. \tag{100}$$

Substituting (100), (99), (98) and (97) in (96) yields

$$\|II_4\|^2_{L^2(\Omega,H)} \leq C\Delta t^{\min(\beta,1-\epsilon)} + C\Delta t\sum_{i=2}^{m-1}\|X^h(t_{m-i})-X^h_{m-i}\|^2_{L^2(\Omega,H)}. \tag{101}$$

Substituting (101), (94), (88), (82) and (76) in (75) yields

$$\left\|X^h(t_m)-X^h_m\right\|^2_{L^2(\Omega,H)} \leq C\Delta t^{\min(\beta,1-\epsilon)} + C\Delta t\sum_{i=1}^{m-1}\left\|X^h(t_i)-X^h_i\right\|^2_{L^2(\Omega,H)}. \tag{102}$$

Applying the discrete Gronwall's lemma to (102) yields

$$\left\|X^h(t_m)-X^h_m\right\|_{L^2(\Omega,H)} \leq C\Delta t^{\frac{\min(\beta,1-\epsilon)}{2}}.$$

This completes the proof of Theorem 2.2.

3.3. *Proof of Theorem 2.3*

Let us recall that

$$\|X^h(t_m)-X^h_m\|^2_{L^2(\Omega,H)} \leq 25\sum_{i=0}^{4}\|III_i\|^2_{L^2(\Omega,H)}, \tag{103}$$

where III_0 and III_1 are exactly the same as II_0 and II_1 respectively. Therefore from (76) and (82) we have

$$\|III_0\|_{L^2(\Omega,H)} + \|III_1\|_{L^2(\Omega,H)} \leq C\Delta t^{\frac{\beta}{2}} + C\Delta t\|X^h(t_{m-1})-X^h_{m-1}\|_{L^2(\Omega,H)}. \tag{104}$$

It remains to re-estimate III_3 in order to achieve higher order convergence rate. We also need to re-estimate the terms involving the noise III_2 and III_4, which are given below

$$III_2 = \int_{t_{m-1}}^{t_m}\left(e^{(t_m-s)A_{h,m-1}} - S^{m-1}_{h,\Delta t}\right)P_h dW(s),$$
$$III_3 = \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) e^{(t_{m-i}-s)A_{h,m-i-1}} G^h_{m-i}\left(X^h(s)\right)ds$$
$$-\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} S^j_{h,\Delta t}\right) G^h_{m-i}\left(X^h_{m-i}\right)ds$$
$$III_4 = \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right) e^{(t_{m-i}-s)A_{h,m-i-1}} P_h dW(s)$$
$$-\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} S^j_{h,\Delta t}\right) P_h dW(s).$$

3.3.1. Estimation of III_2

We can split III_2 in two terms as follows:

$$III_2 = \int_{t_{m-1}}^{t_m} \left[e^{(t_m-s)A_{h,m-1}} - e^{\Delta t A_{h,m-1}}\right] P_h dW(s) + \int_{t_{m-1}}^{t_m} \left[e^{\Delta t A_{h,m-1}} - S_{h,\Delta t}^{m-1}\right] P_h dW(s)$$
$$:= III_{21} + III_{22}. \tag{105}$$

Using itô isometry, Lemma 3.2, [20, Lemma 9 (i) & (ii)] and Lemma 3.8(i), it holds that

$$\begin{aligned}
&\|III_{21}\|^2_{L^2(\Omega,H)}\\
&= \int_{t_{m-1}}^{t_m} \mathbb{E}\left[\left\|\left(e^{(t_m-s)A_{h,m-1}} - e^{(t_m-t_{m-1})A_{h,m-1}}\right) P_h Q^{1/2}\right\|^2_{\mathcal{L}_2(H)}\right] ds\\
&\leq \int_{t_{m-1}}^{t_m} \mathbb{E}\left[\left\|e^{(t_m-s)A_{h,m-1}}\left(\mathbf{I} - e^{(s-t_{m-1})A_{h,m-1}}\right) A_h^{\frac{1-\beta}{2}}\right\|^2_{L(H)} \left\|A_h^{\frac{\beta-1}{2}} P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\right] ds\\
&\leq \int_{t_{m-1}}^{t_m} \mathbb{E}\left[\left\|e^{(t_m-s)A_{h,m-1}} A_h^{\frac{1-\epsilon}{2}}\right\|^2_{L(H)} \left\|A_h^{\frac{-1+\epsilon}{2}}\left(\mathbf{I} - e^{(s-t_{m-1})A_{h,m-1}}\right) A_h^{\frac{1-\beta}{2}}\right\|^2_{L(H)}\right.\\
&\quad \left.\times \left\|A_h^{\frac{\beta-1}{2}} P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\right] ds\\
&\leq C\int_{t_{m-1}}^{t_m} (t_m-s)^{-1+\epsilon}(s-t_{m-1})^{\beta-\epsilon} ds \leq C\Delta t^{\beta-\epsilon}\int_{t_{m-1}}^{t_m} (t_m-s)^{-1+\epsilon} ds \leq C\Delta t^{\beta}.
\end{aligned} \tag{106}$$

Applying Itô isometry, using Lemma 3.7(i) and Lemma 3.8(i) yields

$$\begin{aligned}
\|III_{22}\|^2_{L^2(\Omega,H)} &= \int_{t_{m-1}}^{t_m} \mathbb{E}\left[\left\|\left(e^{\Delta t A_{h,m-1}} - S_{h,\Delta t}^{m-1}\right) P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\right] ds\\
&\leq C\int_{t_{m-1}}^{t_m} \Delta t^{\beta-1}\left\|A_h^{\frac{\beta-1}{2}} P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)} ds \leq C\Delta t^{\beta}.
\end{aligned} \tag{107}$$

Substituting (107), (106) in (105) yields

$$\|III_2\|^2_{L^2(\Omega,H)} \leq 2\|III_{21}\|^2_{L^2(\Omega,H)} + 2\|III_{22}\|^2_{L^2(\Omega,H)} \leq C\Delta t^{\beta}. \tag{108}$$

3.3.2. Estimation of III_3

Since III_3 is the same as II_3, it follows from (89) that

$$III_3 = III_{31} + III_{32} + III_{33} + III_{34}, \tag{109}$$

where III_{31} III_{32}, III_{33} and III_{34} are respectively the same as II_{31} II_{32}, II_{33} and II_{34}. Therefore from (90), (92) and (93) we have

$$\begin{aligned}
&\|III_{31}\|_{L^2(\Omega,H)} + \|III_{33}\|_{L^2(\Omega,H)} + \|III_{34}\|_{L^2(\Omega,H)}\\
&\leq C\Delta t^{\beta} + C\Delta t\sum_{i=2}^{m-1}\|X^h(t_{m-i}) - X^h_{m-i}\|_{L^2(\Omega,H)}.
\end{aligned} \tag{110}$$

To achieve convergence order greater than $\frac{1}{2}$ we need to re-estimate III_{32} by using Assumption 2.7. Recall that III_{32} is given by

$$III_{32} = \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) \left(G^h_{m-i}\left(X^h(s)\right) - G^h_{m-i}\left(X^h(t_{m-i})\right)\right) ds. \tag{111}$$

Using Taylor's formula in Banach space yields

$$\begin{aligned}
&III_{32}\\
&= \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) \left(e^{(s-t_{m-i-1})A_{h,m-i}} - \mathbf{I}\right) X^h(t_{m-i}) ds\\
&+ \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) \left(G^h_{m-i}\right)' \left(X^h(t_{m-i})\right)\\
&\quad \times \int_{t_{m-i-1}}^{s} e^{(s-\sigma)A_{h,m-i-1}} \left(G^h_{m-i}(X^h(\sigma))\right) d\sigma ds\\
&+ \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) \left(G^h_{m-i}\right)' \left(X^h(t_{m-i})\right) \int_{t_{m-i-1}}^{s} e^{(s-\sigma)A_{h,m-i-1}} P_h dW(\sigma) ds\\
&+ \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) R^h_{m-i} ds\\
&=: III_{32}^{(1)} + III_{32}^{(2)} + III_{32}^{(3)} + III_{32}^{(4)},
\end{aligned} \tag{112}$$

where the remainder R^h_{m-i} is given by

$$\begin{aligned}
R^h_{m-i} := \int_0^1 &\left(G^h_{m-i}\right)'' \left(X^h(t_{m-i}) + \lambda \left(X^h(s) - X^h(t_{m-i})\right)\right)\\
&\left(X^h(s) - X^h(t_{m-i}), X^h(s) - X^h(t_{m-i})\right) (1-\lambda) d\lambda.
\end{aligned}$$

Inserting an appropriate power of A_h, using Lemma 3.1 and Corollary 2.1, it holds that

$$\begin{aligned}
\| III_{32}^{(1)} \|_{L^2(\Omega,H)} &\leq \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \left[\mathbb{E} \left\| \left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) A_h^{1-\epsilon} \right\|^2_{L(H)} \right.\\
&\qquad \left. \times \left\| A_h^{-1+\epsilon} \left(e^{(s-t_{m-i-1})A_{h,m-i}} - \mathbf{I}\right) \right\|_{L(H)} \| X^h(t_{m-i}) \| \right]^{\frac{1}{2}} ds\\
&\leq C \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} t_i^{-1+\epsilon} (s - t_{m-i-1})^{1-\epsilon} ds\\
&\leq C \Delta t^{1-\epsilon} \sum_{i=2}^{m-1} t_i^{-1+\epsilon} \Delta t \leq C \Delta t^{1-\epsilon}.
\end{aligned} \tag{113}$$

Using Lemmas 3.1, 3.8(ii) and Corollary 2.1 yields

$$\begin{aligned}\|III_{32}^{(2)}\|^2_{L^2(\Omega,H)} &\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left\|\int_{t_{m-i-1}}^{s} e^{(s-\sigma)A_{h,m-i-1}}\left(G^h_{m-i}\right)(X^h(\sigma))d\sigma\right\|_{L^2(\Omega,H)} ds\\ &\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}(s-t_{m-i-1})ds \leq C\Delta t. \qquad (114)\end{aligned}$$

Since the expectation of the cross-product vanishes, using Itô isometry, triangle inequality, Hölder inequality and Lemma 3.1 yields

$$\begin{aligned}&\|III_{32}^{(3)}\|^2_{L^2(\Omega,H)}\\ &= \mathbb{E}\left[\left\|\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)\right.\right.\\ &\quad\left.\left.\times\int_{t_{m-i-1}}^{s} e^{(s-\sigma)A_{h,m-i-1}}P_h dW(\sigma)ds\right\|^2\right]\\ &= \sum_{i=2}^{m-1}\mathbb{E}\left[\left\|\int_{t_{m-i-1}}^{t_{m-i}}\int_{t_{m-i-1}}^{s}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)\right.\right.\\ &\quad\left.\left.\times\ e^{(s-\sigma)A_{h,m-i-1}}P_h dW(\sigma)ds\right\|^2\right]\\ &\leq \Delta t\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\int_{t_{m-i-1}}^{s}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)\right.\right.\\ &\quad\left.\left.\times\ e^{(s-\sigma)A_{h,m-i-1}}P_h dW(\sigma)\right\|^2\right]ds\\ &\leq \Delta t\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\int_{t_{m-i-1}}^{s}\mathbb{E}\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)\right.\\ &\quad\left.\times\ e^{(s-\sigma)A_{h,m-i-1}}P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)} d\sigma ds\\ &\leq \Delta t\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\int_{t_{m-i-1}}^{s}\mathbb{E}\left\|\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)e^{(s-\sigma)A_{h,m-i-1}}P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)} d\sigma ds. \qquad (115)\end{aligned}$$

Using Lemma 3.8 yields

$$\begin{aligned}
&\mathbb{E}\left\|\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)e^{(s-\sigma)A_{h,m-i-1}}P_hQ^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\\
&=\mathbb{E}\left\|\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)e^{(s-\sigma)A_{h,m-i-1}}A_h^{\frac{1-\beta}{2}}A_h^{\frac{\beta-1}{2}}P_hQ^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\\
&\leq\mathbb{E}\left\|\left(G^h_{m-i}\right)'\left(X^h(t_{m-i})\right)e^{(s-\sigma)A_{h,m-i-1}}A_h^{\frac{1-\beta}{2}}\right\|^2_{L(H)}\left\|A_h^{\frac{\beta-1}{2}}P_hQ^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\\
&\leq\mathbb{E}\left\|e^{(s-\sigma)A_{h,m-i-1}}A_h^{\frac{1-\beta}{2}}\right\|^2_{L(H)}\left\|A_h^{\frac{\beta-1}{2}}P_hQ^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\\
&\leq C(s-\sigma)^{\min(-1+\beta,0)}.
\end{aligned}\tag{116}$$

Substituting (116) in (115) yields

$$\|III^{(3)}_{32}\|^2_{L^2(\Omega,H)}\leq C\Delta t\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\int_{t_{m-i-1}}^{s}(s-\sigma)^{\min(-1+\beta,0)}d\sigma ds\leq C\Delta t^{\min(1+\beta,2)}.\tag{117}$$

Using Lemma 3.8(ii) and 2.1 yields

$$\begin{aligned}
\left\|A_h^{-\frac{\eta}{2}}R^h_{h,m-i}\right\|_{L^2(\Omega,H)}&\leq C\left\|\left\|X^h(s)-X^h(t_{m-i})\right\|^2\right\|_{L^2(\Omega,H)}\\
&\leq C\left\|X^h(s)-X^h(t_{m-i})\right\|^2_{L^4(\Omega,H)}\leq C\Delta t^{\min(\beta,1)}.
\end{aligned}\tag{118}$$

Therefore we obtain the following estimate for $III^{(4)}_{32}$

$$\begin{aligned}
\|III^{(4)}_{32}\|_{L^2(\Omega,H)}&\leq C\Delta t^{\min(\beta,1)}\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1}e^{\Delta tA_{h,j}}\right)A_h^{\frac{\eta}{2}}\right\|^2_{L(H)}\right]ds\\
&\leq C\Delta t^{\min(\beta,1)}\sum_{i=2}^{m-1}t_i^{-\frac{\eta}{2}}\Delta t\leq C\Delta t^{\min(\beta,1)}.
\end{aligned}\tag{119}$$

Substituting (119), (117), (114) and (113) in (112) yields

$$\begin{aligned}
\|III_{32}\|_{L^2(\Omega,H)}&\leq\|III^{(1)}_{32}\|_{L^2(\Omega,H)}+\|III^{(2)}_{32}\|_{L^2(\Omega,H)}\\
&\quad+\|III^{(3)}_{32}\|_{L^2(\Omega,H)}+\|III^{(4)}_{32}\|_{L^2(\Omega,H)}\\
&\leq C\Delta t^{\frac{\beta}{2}-\epsilon}.
\end{aligned}\tag{120}$$

Substituting (120) and (95) in (111) yields

$$\|III_{32}\|_{L^2(\Omega,H)}\leq\|III^{(1)}_{32}\|_{L^2(\Omega,H)}+\|III^{(2)}_{32}\|_{L^2(\Omega,H)}\leq C\Delta t^{\frac{\beta}{2}-\epsilon}.\tag{121}$$

Substituting (121) and (110) in (109) yields

$$\|III_3\|_{L^2(\Omega,H)}\leq C\Delta t^{\frac{\beta}{2}-\epsilon}.\tag{122}$$

3.3.3. Estimation of III_4

We can recast III_4 in two terms as follows

$$
\begin{aligned}
III_4 &= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(e^{(t_{m-i}-s)A_{h,m-i}}-\mathbf{I}\right)P_h dW(s)\\
&\quad+\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)-\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^{j}\right)\right]P_h dW(s)\\
&=: III_{41}+III_{42}. \qquad (123)
\end{aligned}
$$

Using Itô isometry, Lemma 3.8(i), [20, Lemma 9 (i) & (iv)], Lemma 3.1 and [20, Lemma 10] yields

$$
\begin{aligned}
&\|III_{41}\|^2_{L^2(\Omega,H)}\\
&= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(e^{(t_{m-i}-s)A_{h,m-1}}-\mathbf{I}\right)P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\right]ds\\
&\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)\left(e^{(t_{m-i}-s)A_{h,m-1}}-\mathbf{I}\right)A_h^{\frac{1-\beta}{2}}\right\|^2_{L(H)}\right.\\
&\quad\times\left.\left\|A_h^{\frac{\beta-1}{2}}P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\right]ds\\
&\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)A_h^{\frac{1-\epsilon}{2}}\right\|^2_{L(H)}\left\|\left(e^{(t_{m-i}-s)A_{h,m-1}}-\mathbf{I}\right)A_h^{\frac{-\beta+\epsilon}{2}}\right\|^2_{L(H)}\right]ds\\
&\leq C\sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}} t_i^{-1+\epsilon}(t_{m-i}-s)^{\beta-\epsilon}ds \leq C\Delta t^{\beta-\epsilon}\sum_{i=2}^{m-1} t_i^{-1+\epsilon}\Delta t \leq C\Delta t^{\beta-\epsilon}. \qquad (124)
\end{aligned}
$$

Using Itô isometry and Lemma 3.8(i) yields

$$
\begin{aligned}
&\|III_{42}\|^2_{L^2(\Omega,H)}\\
&= \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)-\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^{j}\right)\right]P_h Q^{\frac{1}{2}}\right\|^2_{\mathcal{L}_2(H)}\right]ds\\
&\leq \sum_{i=2}^{m-1}\int_{t_{m-i-1}}^{t_{m-i}}\mathbb{E}\left[\left\|\left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}}\right)-\left(\prod_{j=m-i}^{m-1} S_{h,\Delta t}^{j}\right)\right]A_h^{\frac{1-\beta}{2}}\right\|^2_{L(H)}\right.
\end{aligned}
$$

$$\times \left\| A_h^{\frac{\beta-1}{2}} P_h Q^{\frac{1}{2}} \right\|^2_{\mathcal{L}_2(H)} \Bigg] ds$$

$$\leq C \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\| \left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) - \left(\prod_{j=m-i}^{m-1} S^j_{h,\Delta t} \right) \right] A_h^{\frac{1-\beta}{2}} \right\|^2_{L(H)} \right] ds. \tag{125}$$

If $0 < \beta < 1$ then applying Lemma 3.7(ii) yields

$$\begin{aligned} \|III_{42}\|^2_{L^2(\Omega,H)} &\leq C \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \mathbb{E}\left[\left\| \left[\left(\prod_{j=m-i}^{m-1} e^{\Delta t A_{h,j}} \right) \right.\right.\right. \\ &\qquad \left.\left.\left. - \left(\prod_{j=m-i}^{m-1} S^j_{h,\Delta t} \right) \right] A_h^{\frac{1-\epsilon}{2}} \right\|^2_{L(H)} \right] ds \\ &\leq C \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \Delta t^{1-\epsilon} t_{i-1}^{-1+\epsilon} ds \leq C \Delta t^{1-\epsilon} \sum_{i=2}^{m-1} t_{i-1}^{-1+\epsilon} \Delta t \leq C \Delta t^{1-\epsilon}. \end{aligned} \tag{126}$$

If $\beta \in [1, 2]$ then applying Lemma 3.7(iii) yields

$$\|III_{42}\|^2_{L^2(\Omega,H)} \leq C \sum_{i=2}^{m-1} \int_{t_{m-i-1}}^{t_{m-i}} \Delta t^{\beta-\epsilon}\, t_{i-1}^{-1+\epsilon} ds \leq C \Delta t^{\beta-\epsilon}. \tag{127}$$

Therefore for all $\beta \in (0, 2]$ it holds that

$$\|III_{42}\|^2_{L^2(\Omega,H)} \leq C \Delta t^{\beta-\epsilon}. \tag{128}$$

Substituting (128) and (124) in (123) yields

$$\|III_4\|^2_{L^2(\Omega,H)} \leq 2\|III_{41}\|^2_{L^2(\Omega,H)} + 2\|III_{42}\|^2_{L^2(\Omega,H)} \leq C \Delta t^{\beta-\epsilon}. \tag{129}$$

Substituting (129), (122), (108) and (104) in (103) yields

$$\left\| X^h(t_m) - X^h_m \right\|^2_{L^2(\Omega,H)} \leq C \Delta t^{\beta-2\epsilon} + C \Delta t \sum_{i=2}^{m-1} \left\| X^h(t_{m-i}) - X^h_{m-i} \right\|^2_{L^2(\Omega,H)}.$$

Applying the discrete Gronwall's lemma yields

$$\left\| X^h(t_m) - X^h_m \right\|_{L^2(\Omega,H)} \leq C \Delta t^{\frac{\beta}{2}-\epsilon}.$$

This completes the proof of Theorem 2.3.

4. Numerical simulations

We consider the following stochastic reactive dominated advection diffusion equation with constant diagonal diffusion tensor

$$dX = \left[\nabla \cdot (\mathbf{D} \nabla X) - \nabla \cdot (\mathbf{q} X) - \frac{10X}{X+1} \right] dt + b(X) dW, \quad \mathbf{D} = \begin{pmatrix} 10^{-1} & 0 \\ 0 & 10^{-2} \end{pmatrix}. \tag{130}$$

with mixed Neumann–Dirichlet boundary conditions on $\Lambda = [0, L_1] \times [0, L_2]$. The Dirichlet boundary condition is $X = 1$ at $\Gamma = \{(x, y) : x = 0\}$ and we use the homogeneous Neumann boundary conditions elsewhere. The eigenfunctions $\{e_{i,j}\} = \{e_i^{(1)} \otimes e_j^{(2)}\}_{i,j\geq 0}$ of the covariance operator Q are the same as for Laplace operator $-\Delta$ with homogeneous boundary condition and are given by

$$e_0^{(l)}(x) = \sqrt{\frac{1}{L_l}}, \qquad e_i^{(l)}(x) = \sqrt{\frac{2}{L_l}} \cos\left(\frac{i\pi}{L_l}x\right), \quad l \in \{1, 2\}, \; x \in \Lambda, \quad i \in \mathbb{N}.$$

We assume that the noise can be represented as

$$W(x,t) = \sum_{i\in\mathbb{N}^2} \sqrt{\lambda_{i,j}} e_{i,j}(x)\beta_{i,j}(t), \tag{131}$$

where $\beta_{i,j}(t)$ are independent and identically distributed standard Brownian motions, $\lambda_{i,j}$, $(i,j) \in \mathbb{N}^2$ are the eigenvalues of Q, with

$$\lambda_{i,j} = \left(i^2 + j^2\right)^{-(\beta+\epsilon)}, \; \beta > 0, \tag{132}$$

in the representation (131) for some small $\epsilon > 0$. When dealing with additive noise, we take $b(u) = 1$, so Assumption 2.6 is obviously satisfied for any $\beta \in (0, 2]$. When dealing with multiplicative noise, we take $b(u) = u$ in (13), Therefore, from [10, Section 4] it follows that the operators B defined by (13) fulfills obviously Assumptions 2.4 and 2.5. For both additive and multiplicative noise, the function $F(X) = -\dfrac{10X}{1+X}$ obviously satisfies the global Lipschitz condition in Assumptions 2.3 and 2.7. We obtain the Darcy velocity field $\mathbf{q} = (q_i)$ by solving the following system

$$\nabla \cdot \mathbf{q} = 0, \qquad \mathbf{q} = -\frac{\mathbf{k}}{\mu}\nabla p, \tag{133}$$

with Dirichlet boundary conditions on $\Gamma_D^1 = \{0, L_1\} \times [0, L_2]$ and Neumann boundary conditions on $\Gamma_N^1 = (0, L_1) \times \{0, L_2\}$ such that

$$p = \begin{cases} 1 & \text{in} \quad \{0\} \times [0, L_2] \\ 0 & \text{in} \quad \{L_1\} \times [0, L_2] \end{cases}$$

and $-\mathbf{k}\nabla p(\mathbf{x}, t) \cdot \mathbf{n} = 0$ in Γ_N^1. Note that $\mathbf{k}$ is the permeability tensor. We use a random permeability field as in [32] and take $\mu = 10$. The finite volume method viewed as a finite element method (see [30]) is used for the advection and the finite element method is used for the remainder. In the legends of our graphs, we use the following notations:

1. 'Rosenbrock-A-noise' is used for graphs from our Rosenbrock scheme with additive noise.
2. 'Rosenbrock-M-noise' is used for graphs from our Rosenbrock scheme with multiplicative noise.
3. 'Expo-Rosenbrock-A-noise' is used for graphs of stochastic exponential Rosenbrock scheme presented in [20] with additive noise.
4. 'Expo-Rosenbrock-M-noise' is used for graphs of stochastic exponential Rosenbrock scheme presented in [20] with multiplicative noise.

We take $L_1 = 2$ and $L_2 = 2$ and our reference solutions samples are numerical solutions with time step $\Delta t = 1/2048$. The errors are computed at the final time $T = 1$. The initial solution is $X_0 = 0$, so we can therefore expect high orders convergence, which depend only on the noise term. For both additive and multiplicative noise, we use $\beta = 2$ and

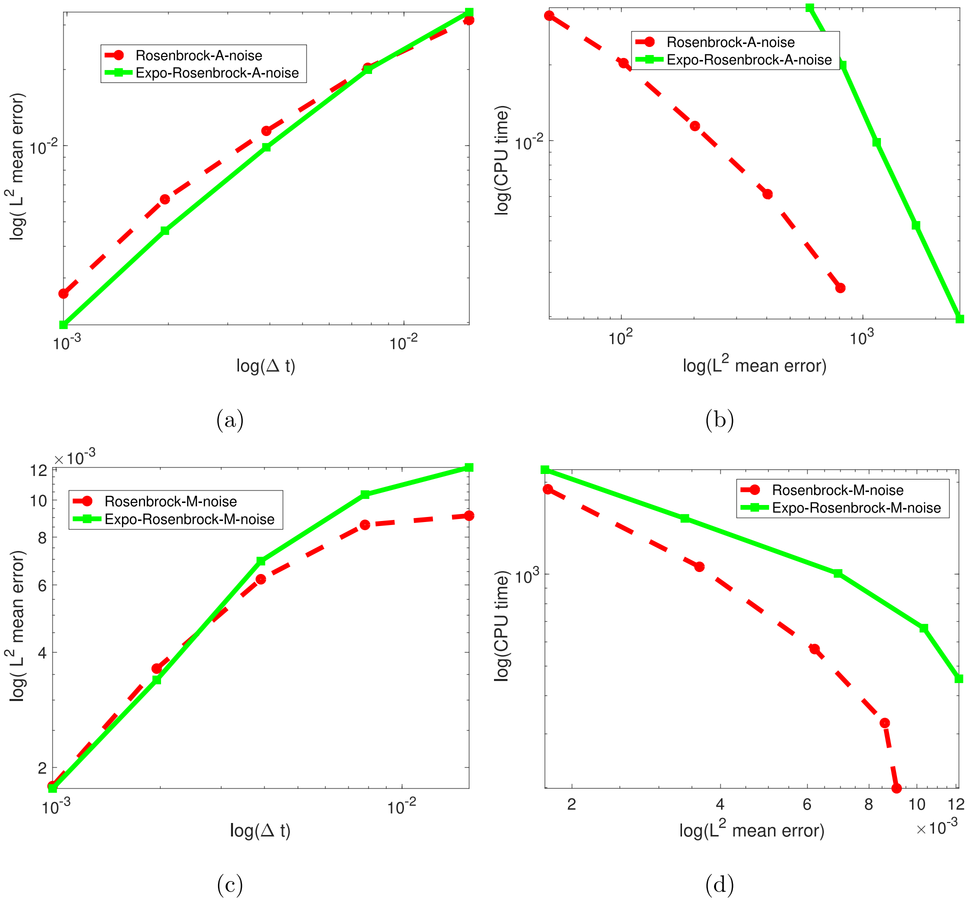
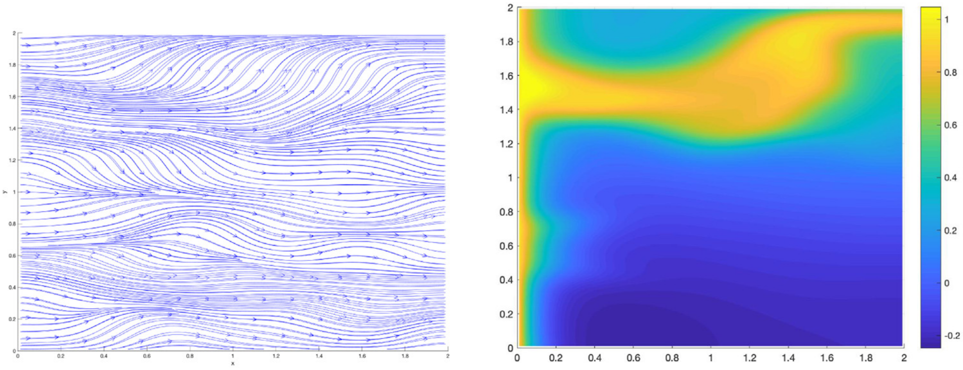


Fig. 1. Convergence in the root mean square L^2 norm at $T = 1$ as a function of Δt for additive noise (a) and multiplicative noise (c). We take $\beta = 2$, and $\epsilon = 10^{-1}$ in relation (132) and use 80 realizations. Graph (b) and graph (d) show the CPU time per sample versus the root mean square L^2 errors for additive noise and multiplicative noise respectively.

$\epsilon = 10^{-1}$. The streamline of velocity is given at Fig. 2(a) while a sample of the numerical solution with the stochastic Rosenbrock scheme for additive noise is given at Fig. 2(b). In Fig. 1(a) and (c), the graphs of strong errors versus the time steps are plotted for stochastic Rosenbrock scheme and exponential Rosenbrock for additive noise and multiplicative noise respectively. The orders of convergence are 0.59 (exponential Rosenbrock scheme) and 0.55 (Rosenbrock scheme) for multiplicative noise, 1.03 (exponential Rosenbrock scheme) and 0.92 (Rosenbrock scheme) for additive noise, which are close to 0.5 and 1 in our theoretical results in Theorems 2.2 and 2.3 respectively. The implementation of the stochastic Rosenbrock-type scheme is straightforward and only needs the resolution of a linear system of equations at each time step. For efficiency, all linear systems are solved using the Matlab function bicgstab coupled with ILU(0) preconditioners with no fill-in. The ILU(0) are done on the deterministic part of the matrix A_h, that is $(I + \Delta t A_h)$, at each time step. Fig. 1(b) and (d) show the mean of



(a) (b)

Fig. 2. (a) The streamline of velocity (a) and a sample of the numerical solution with the stochastic Rosenbrock scheme for additive noise.

CPU time per sample versus the root mean square L^2 errors corresponding for Fig. 1(a)(additive noise) and Fig. 1(c)(multiplicative noise) respectively. As we can observe, the novel stochastic Rosenbrock scheme is more efficient than the stochastic exponential Rosenbrock scheme, thanks to the preconditioners.

Declaration of competing interest

None.

Acknowledgments

J. D. Mukam acknowledges the support of the TU Chemnitz and thanks Prof. Dr. Peter Stollmann for his constant encouragement. A. Tambue was supported by the "Robert Bosch Stiftung", Germany through the "AIMS ARETE Chair programme" (Grant No 11.5.8040.0033.0). We would like to thank Prof. Dr. Thomas Kalmes for very useful discussions. We would also like to thank the reviewers for their careful readings which helped to improve this paper.

References

[1] K.J. Engel, R. Nagel, One-Parameter Semigroup for Linear Evolution Equations, Springer-Verlag, New York, Inc, 2000.

[2] F. Fujita, T. Suzuki, Evolution problems (Part 1), in: P.G. Ciarlet, J.L. Lions (Eds.), HandBook of Numerical Analysis, Vol. 2, North-Holland, Amsterdam, The Netherlands, 1991, pp. 789–928.

[3] S. Geiger, G.J. Lord, A. Tambue, Exponential time integrators for stochastic partial differential equations in 3D reservoir simulation, Comput. Geosci. 16 (2) (2012) 323–334, http://dx.doi.org/10.1007/s10596-011-9273-z.

[4] E. Hairer, Ch. Lubich, M. Roche, Error of Rosenbrock methods for stiff problems studied via differential algebraic equations, BIT 29 (1989) 77–90, http://dx.doi.org/10.1007/BF01932707.

[5] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II. Siff and Differential–Algebraic Problems, Springer-Verlag, Berlin, Heidelberg, 1991.

[6] D. Henry, Geometric Theory of Semilinear Parabolic Equations, in: Lecture Notes in Mathematics, vol. 840, Springer, Berlin, 1981.

[7] M. Hochbruck, A. Ostermann, J. Schweitzer, Exponential Rosenbrock-type methods, SIAM J. Numer. Anal. 47 (1) (2009) 786–803, http://dx.doi.org/10.1137/080717717.

[8] A. Jentzen, P.E. Kloeden, Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space–time noise, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 465 (2102) (2009) 649–667, http://dx.doi.org/10.1098/rspa.2008.0325.
[9] A. Jentzen, P.E. Kloeden, G. Winkel, Efficient simulation of nonlinear parabolic SPDEs with additive noise, Ann. Appl. Probab. 21 (3) (2011) 908–950, https://www.jstor.org/stable/23033359.
[10] A. Jentzen, M. Röckner, Regularity analysis for stochastic partial differential equations with nonlinear multiplicative trace class noise, J. Differential Equations 252 (1) (2012) 114–136, http://dx.doi.org/10.1016/j.jde.2011.08.050.
[11] M. Kovács, S. Larsson, F. Lindgren, Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise, Numer. Algorithms 53 (2010) 309–320, http://dx.doi.org/10.1007/s11075-009-9281-4.
[12] R. Kruse, Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise, IMA J. Numer. Anal. 34 (1) (2014) 217–251, http://dx.doi.org/10.1093/imanum/drs055.
[13] R. Kruse, S. Larsson, Optimal regularity for semilinear stochastic partial differential equations with multiplicative noise, Electron. J. Probab. 17 (65) (2012) 1–19, http://dx.doi.org/10.1214/EJP.v17-2240.
[14] J. Lang, Adaptive multilevel solution of nonlinear parabolic PDE systems, in: Theory, Algorithm, and Applications, in: Lecture Notes in Computational Sciences and Engineering, vol. 16, Springer Verlag, 2000.
[15] S. Larsson, Nonsmooth Data Error Estimates with Applications to the Study of the Long-Time Behavior of Finite Element Solutions of Semilinear Parabolic Problems, Preprint 1992-36, Department of Mathematics, Chalmers University of Technology, 1992, Available at: http://citeseerx.ist.psu.edu/viewdoc/summary/doi=10.1.1.28.1250.
[16] G.J. Lord, A. Tambue, Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise, IMA J. Numer. Anal. 2 (2013) 515–543, http://dx.doi.org/10.1093/imanum/drr059.
[17] G.J. Lord, A. Tambue, A modified semi-implicit Euler–Maruyama scheme for finite element discretization of SPDEs with additive noise, Appl. Math. Comput. 332 (2018) 105–122, http://dx.doi.org/10.1016/j.amc.2018.03.014.
[18] Ch. Lubich, O. Nevanlina, On the resolvent conditions and stability estimates, BIT 31 (1991) 293–313, http://dx.doi.org/10.1007/BF01931289.
[19] J.D. Mukam, A. Tambue, A note on exponential Rosenbrock-Euler method for the finite element discretization of a semilinear parabolic partial differential equation, Comput. Math. Appl. 76 (7) (2018) 1719–1738, http://dx.doi.org/10.1016/j.camwa.2018.07.025.
[20] J.D. Mukam, A. Tambue, Strong convergence analysis of the stochastic exponential rosenbrock scheme for the finite element discretization of semilinear SPDEs driven by multiplicative and additive noise, J. Sci. Comput. 74 (2018) 937–978, http://dx.doi.org/10.1007/s10915-017-0475-y.
[21] A. Ostermann, M. Roche, Rosenbrock methods for partial differential equations and fractional orders of convergence, SIAM J. Numer. 30 (4) (1993) 1084–1098, http://dx.doi.org/10.1137/0730056.
[22] A. Ostermann, M. Thalhammer, Non-smooth data error estimates for linearly implicit Runge–Kutta methods, IMA J. Numer. Anal. 20 (2000) 167–184, http://dx.doi.org/10.1093/imanum/20.2.167.
[23] K. Paps, G. Wanner, A study of Rosenbrock-type methods of higher order, Numer. Math. 38 (1981) 279–298, http://dx.doi.org/10.1007/BF01397096.
[24] A. Pazy, Semigroups of Linear Operators and Applications To Partial Differential Equations, in: Applied Mathematical Sciences, vol. 44, Springer-Verlag, New York, 1983.
[25] D. Prato, G.J. Zabczyk, Stochastic Equations in Infinite Dimensions, Vol. 152, Cambrige University Press, Cambridge, United Kingdom, 2014.
[26] C. Prévôt, M. Röckner, A Concise Course on Stochastic Partial Differential Equations, in: Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
[27] J. Printems, On the discretization in time of parabolic stochastic partial differential equations, Math. Modelling Numer. Anal. 35 (6) (2001) 1055–1078, http://dx.doi.org/10.1051/m2an:2001148.
[28] T. Shardlow, Numerical simulation of stochastic PDEs for excitable media, J. Comput. Appl. Math. 175 (2) (2005) 429–446, http://dx.doi.org/10.1016/j.cam.2004.06.020.
[29] A. Tambue, Efficient Numerical Schemes for Porous Media Flow (Ph.D. thesis), in: Department of Mathematics, Heriot–Watt University, 2010.
[30] A. Tambue, An exponential integrator for finite volume discretization of a reaction-advection-diffusion equation, Comput. Math. Appl. 71 (9) (2016) 1875–1897, http://dx.doi.org/10.1016/j.camwa.2016.03.001.

[31] A. Tambue, I. Berre, J.M. Nordbotten, Efficient simulation of geothermal processes in heterogeneous porous media based on the exponential Rosenbrock–Euler and Rosenbrock-type methods, Adv. Water Resour. 53 (2013) 250–262, http://dx.doi.org/10.1016/j.advwatres.2012.12.004.

[32] A. Tambue, G.J. Lord, S. Geiger, An exponential integrator for advection-dominated reactive transport in heterogeneous porous media, J. Comput. Phys. 229 (10) (2010) 3957–3969, http://dx.doi.org/10.1016/j.jcp.2010.01.037.

[33] A. Tambue, J.D. Mukam, Strong convergence of the linear implicit Euler method for the finite element discretization of semilinear SPDEs driven by multiplicative and additive noise, Appl. Math. Comput. 346 (2019) 23–40, http://dx.doi.org/10.1016/j.amc.2018.09.073.

[34] X. Wang, Strong convergence rates of the linear implicit Euler method for the finite element discretization of SPDEs with additive noise, IMA J. Numer. Anal. 37 (2) (2017) 965–984, http://dx.doi.org/10.1093/imanum/drw016.

[35] X. Wang, Q. Ruisheng, A note on an accelerated exponential Euler method for parabolic SPDEs with additive noise, Appl. Math. Lett. 46 (2015) 31–37, http://dx.doi.org/10.1016/j.aml.2015.02.001.

[36] Y. Yan, Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise, BIT 44 (4) (2004) 829–847, http://dx.doi.org/10.1007/s10543-004-3755-5.

[37] Y. Yan, Galerkin finite element methods for stochastic parabolic partial differential equations, SIAM J. Num. Anal. 43 (4) (2005) 1363–1384, http://dx.doi.org/10.1137/040605278.