IMA Journal of Numerical Analysis (2022) **42**, 2118–2150
https://doi.org/10.1093/imanum/drab033
Advance Access publication on 20 May 2021

A linear implicit Euler method for the finite element discretization of a controlled stochastic heat equation

PETER BENNER
Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany

TONY STILLFJORD*
Centre for Mathematical Sciences, Lund University, PO Box 118, 221 00 Lund, Sweden
*Corresponding author: tony.stillfjord@math.lth.se

AND

CHRISTOPH TRAUTWEIN
Institute for Mathematics, Friedrich Schiller University Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany

[Received on 9 June 2020; revised on 22 February 2021]

We consider a numerical approximation of a linear quadratic control problem constrained by the stochastic heat equation with nonhomogeneous Neumann boundary conditions. This involves a combination of distributed and boundary control, as well as both distributed and boundary noise. We apply the finite element method for the spatial discretization and the linear implicit Euler method for the temporal discretization. Due to the low regularity induced by the boundary noise, convergence orders above $1/2$ in space and $1/4$ in time cannot be expected. We prove such optimal convergence orders for our full discretization when the distributed noise and the initial condition are sufficiently smooth. Under less smooth conditions the convergence order is further decreased. Our results only assume that the related (deterministic) differential Riccati equation can be approximated with a certain convergence order, which is easy to achieve in practice. We confirm these theoretical results through a numerical experiment in a two-dimensional domain.

Keywords: optimal control; stochastic partial differential equation; full discretization; convergence analysis; heat conduction.

1. Introduction

This paper is devoted to a numerical scheme for a linear quadratic control problem constrained by the stochastic heat equation with nonhomogeneous Neumann boundary conditions. We prove optimal convergence orders for a full discretization, which combines a linear implicit Euler method in time and a finite element discretization in space.

For time-dependent heat distributions considered in a bounded domain, noise terms in the sense of random heating or cooling phenomena arise due to imperfect insulation and other uncertain environmental effects. In engineering applications this might lead to undesired behavior. To keep a desired heat profile it is therefore necessary to regulate the system. This task can be formulated as a linear quadratic control problem constrained by the stochastic heat equation, where controls and additive noise terms are located inside the domain as well as on the boundary. Here, we treat the case of noise terms defined by Q-Wiener processes. In stochastic control theory it is well known that the concept of mild solutions is useful to include nonhomogeneous boundary conditions; see Debussche *et al.* (2007), Fabbri & Goldys (2009), Fabbri *et al.* (2017), Guatteri & Masiero (2011), Yu & Liu (2011). In this context we also refer to related deterministic control problems; see Benner & Mena (2018), Bensoussan *et al.* (2007) and the references therein. Typically, optimal controls as solutions of stochastic linear quadratic control problems are characterized by a feedback law; see Ahmed (1981), Benner & Trautwein (2018), Curtain & Pritchard (1978), Duncan *et al.* (2012), Hu & Tang (2018). These feedback laws often involve the solution to a suitable operator-valued differential Riccati equation. In this paper the Riccati equation is deterministic resulting from the fact that only additive noise terms are included. As a consequence the optimal heat distribution fulfills a system of a linear stochastic partial differential equations (SPDEs), referred to as the controlled stochastic heat equation, which is coupled to the operator-valued differential Riccati equation. The main obstacle is that the controlled stochastic heat equation as well as the Riccati equation cannot be solved explicitly. For that reason we analyze a numerical approximation of the system describing the optimal heat distribution.

For the spatial discretization we use the finite element method as introduced in Thomeé (2006), where only parabolic equations with homogeneous boundary conditions are considered. The case of nonhomogeneous boundary conditions is studied in, e.g., Lasiecka (1986). Here, we need a generalization of this theory since the system includes Q-Wiener processes. Numerical simulations for Q-Wiener processes with values in Hilbert spaces as well as for some specific SPDEs are demonstrated in Lord *et al.* (2014).

Temporal discretization of SPDEs has become an active research area within recent years. Equations driven by additive noise terms are considered in Wang (2017), and Kruse (2014), Lord & Tambue (2013), Tambue & Mukam (2019) also consider the case of multiplicative noise terms. These papers have in common that the linear implicit Euler method is used for the temporal discretization. This is essentially the usual implicit Euler method but with the noise terms treated explicitly, since treating them implicitly makes no sense. The stochastic part of the equation is therefore treated explicitly and the deterministic part implicitly. We follow the same approach in this paper. The error analyses are mainly based on the fact that the underlying equation involves a closed operator generating an analytic semigroup, such that fractional powers of this closed operator are well defined.

The shortcoming of the papers mentioned above is that they only consider equations with homogeneous boundary conditions. We will extend these results by including instead nonhomogeneous Neumann boundary conditions. Because this leads to a less regular solution, the convergence order is decreased. However, the theory of fractional powers to closed operators can still be applied, and we use this to prove optimal convergence orders under the assumption that the associated Riccati equation can be well approximated. We make such an assumption mainly because there is a lack of temporal error analyses applicable to the current situation, and providing such a proof is out of the scope of this paper. We refer to Lasiecka & Triggiani (2000) for related results on deterministic linear quadratic control problems and their corresponding Riccati equations.

In order to illustrate our theoretical results we implement our method in MATLAB and perform a numerical experiment that shows the expected convergence orders. We also confirm that achieving the assumed convergence orders for the approximation of the Riccati equation is straightforward in practice.

The paper is organized as follows. In Section 2 we introduce the linear quadratic control problem constrained by the stochastic heat equation. We state the optimal controls and derive the resulting system describing the optimal heat distribution. Section 3 is devoted to the numerical scheme of the controlled stochastic heat equation and the Riccati equation. We also state the main result concerning the convergence order. In order to prepare for the proof of this theorem we derive several auxiliary results on continuity, consistency and stability in Section 4. The proof of the main result then follows in Section 5. Finally, in Section 6 we discuss the implementation and illustrate the theoretical results through a numerical experiment.

2. A linear quadratic control problem constrained by the stochastic heat equation

Throughout this paper, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space endowed with a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ satisfying $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \geqslant 0$ and $\mathcal{F}_0$ contains all sets of $\mathcal{F}$ with $\mathbb{P}$-measure 0. We use $\mathbb{E}$ to denote the expectation with respect to this probability space. Moreover, we assume that $\mathcal{D} \subset \mathbb{R}^n$ for $n \geqslant 1$ is either a bounded domain with sufficiently smooth boundary $\partial\mathcal{D}$ or a bounded and convex domain.

First, we introduce some basic notation and we state properties of operators frequently used in the remaining part. For $s \geqslant 0$ let $H^s(\mathcal{D})$ denote the usual Sobolev space. We set $H = L^2(\mathcal{D})$ and let I denote the identity operator on H. We introduce the Neumann realization of the Laplace operator $A\colon D(A) \subset H \to H$ defined by

$$Ay = \Delta y$$

for every $y \in D(A)$ with

$$D(A) = \left\{ y \in H^2(\mathcal{D}) : \tfrac{\partial}{\partial\nu} y = 0 \text{ on } \partial\mathcal{D} \right\}.$$

The characterization of the domain follows from existence and uniqueness results of the corresponding elliptic problem; see Grisvard (1985). The operator A is the infinitesimal generator of an analytic semigroup $\left(e^{At}\right)_{t\geqslant 0}$ of contractions such that for $\lambda > 0$, fractional powers of $\lambda - A$ denoted by $(\lambda - A)^\alpha$ with $\alpha \in \mathbb{R}$ are well defined. For more details in a more general framework we refer to Pazy (1983), Vrabie (2003), but we have also collected the main properties that we need in Section 4.

For $\alpha \in \mathbb{R}$ the space $D((\lambda - A)^\alpha)$ equipped with the inner product

$$\langle y, z \rangle_\alpha = \left\langle (\lambda - A)^\alpha y, (\lambda - A)^\alpha z \right\rangle_H$$

becomes a Hilbert space. The corresponding norm is denoted by $\|\cdot\|_\alpha$. In general the domain of $(\lambda - A)^\alpha$ for $\alpha \in (0, 1)$ can be expressed explicitly by interpolation of the spaces H and $D(A)$; see Lions & Magenes (1972). In the case that $\mathcal{D}$ is bounded with sufficiently smooth boundary we have

$$D((\lambda - A)^\alpha) = \begin{cases} H^{2\alpha}(\mathcal{D}) & \text{for } \alpha \in (0, 3/4)\,, \\ \left\{ y \in H^{2\alpha}(\mathcal{D}) : \tfrac{\partial}{\partial\nu} y = 0 \text{ on } \partial\mathcal{D} \right\} & \text{for } \alpha \in (3/4, 1)\,, \end{cases}$$

where we refer to Fujiwara (1967). We set $H_b = L^2(\partial\mathcal{D})$ and introduce the Neumann operator $N\colon H_b \to H$ given by $g = Nh$ with

$$\begin{cases} \Delta g(x) = \lambda\, g(x) & \text{in } \mathcal{D}, \\ \dfrac{\partial}{\partial\nu} g(x) = h(x) & \text{on } \partial\mathcal{D}, \end{cases} \tag{2.1}$$

where $\lambda > 0$. If $\mathcal{D}$ is bounded with sufficiently smooth boundary the result $N \in \mathcal{L}\left(H_b; H^{3/2}(\mathcal{D})\right)$ has been proved in Lions & Magenes (1972). In this case we can therefore conclude that $N \in \mathcal{L}(H_b; D((\lambda - A)^{\alpha}))$ for $\alpha \in (0, 3/4)$, which means that the operator $(\lambda - A)^{\alpha} N$ is linear and bounded by the closed graph theorem. If $\mathcal{D}$ is instead bounded and convex then $\mathcal{D}$ has a Lipschitz boundary and satisfies the cone property; see Grisvard (1985). We therefore again obtain $N \in \mathcal{L}(H_b; D((\lambda - A)^{\alpha}))$ for $\alpha \in (0, 3/4)$; see Lasiecka (1980).

Next we introduce the controlled stochastic heat equation with nonhomogeneous Neumann boundary conditions as an evolution equation. Here, we include distributed and boundary controls as well as distributed and boundary noise. Let U contain all $\mathcal{F}_t$-adapted processes $(u(t))_{t\in[0,T]}$ with values in an arbitrary Hilbert space $\bar{U}$ satisfying $\mathbb{E}\int_0^T \|u(t)\|_{\bar{U}}^2 \,\mathrm{d}t < \infty$ and let V contain all $\mathcal{F}_t$-adapted processes $(v(t))_{t\in[0,T]}$ with values in $\bar{V} \subset H_b$ satisfying $\mathbb{E}\int_0^T \|v(t)\|_{H_b}^2 \,\mathrm{d}t < \infty$. We consider the following controlled system in H for $t \in [0, T]$ and $\lambda > 0$:

$$\begin{cases} \mathrm{d}y(t) = [Ay(t) + Bu(t) + (\lambda - A)Nv(t)]\,\mathrm{d}t + G\,\mathrm{d}W(t) + (\lambda - A)N\,\mathrm{d}W_b(t), \\ y(0) = \xi, \end{cases} \tag{2.2}$$

where $(u(t))_{t\in[0,T]}$ and $(v(t))_{t\in[0,T]}$ represent the distributed and the boundary controls. We assume that $u \in U$, $B \in \mathcal{L}(\bar{U}; H)$ and $v \in V$. The processes $(W(t))_{t\geqslant 0}$ and $(W_b(t))_{t\geqslant 0}$ are independent and $\mathcal{F}_t$-adapted Q-Wiener processes with values in H and H_b, respectively. The corresponding covariance operators are denoted by $Q \in \mathcal{L}(H)$ and $Q_b \in \mathcal{L}(H_b)$. We make the following assumptions.

ASSUMPTION 2.1 The initial value $\xi \in L^2(\Omega; D((\lambda - A)^{\beta/2}))$ with $\beta \in (0, 2)$ is $\mathcal{F}_0$-measurable.

REMARK 2.2 The results shown in this section also hold for an $\mathcal{F}_0$-measurable initial value $\xi \in L^2(\Omega; H)$. We make the additional regularity requirement due to the main result stated in the following section.

ASSUMPTION 2.3 We assume that G is a square-integrable random variable with values in the space of Hilbert–Schmidt operators mapping $Q^{1/2}(H)$ into $D((\lambda - A)^{(\beta-1)/2})$ denoted by $\mathcal{L}_{\mathrm{HS}}(Q^{1/2}(H); D((\lambda - A)^{(\beta-1)/2}))$, where $\beta \in (0, 2)$ arises from Assumption 2.1.

DEFINITION 2.4 A predictable process $(y(t))_{t\in[0,T]}$ with values in H is a mild solution of system (2.2) if

$$\sup_{t\in[0,T]} \mathbb{E}\|y(t)\|_H^2 < \infty$$

and for all $t \in [0, T]$ and $\mathbb{P}$-a.s.,

$$y(t) = e^{At}\xi + \int_0^t e^{A(t-s)}Bu(s)\,\mathrm{d}s + \int_0^t (\lambda - A)e^{A(t-s)}Nv(s)\,\mathrm{d}s + \int_0^t e^{A(t-s)}G\,\mathrm{d}W(s)$$
$$+ \int_0^t (\lambda - A)e^{A(t-s)}N\,\mathrm{d}W_b(s).$$

For an existence and uniqueness result for a mild solution to system (2.2) we refer to Benner & Trautwein (2018). Next we introduce the cost functional $J\colon U \times V \to \mathbb{R}$ defined by

$$J(u,v) = \mathbb{E}\left[\int_0^T \langle C\,y(t), C\,y(t)\rangle_Z + \langle R\,u(t), u(t)\rangle_H + \langle R_b\,v(t), v(t)\rangle_{H_b} \mathrm{d}t\right],$$

where $C \in \mathcal{L}(H;Z)$ represents an observation operator mapping H into an arbitrary Hilbert space Z. The operators $R \in \mathcal{L}(H)$ and $R_b \in \mathcal{L}(H_b)$ are given scaling factors for the costs of the controls and are assumed to be invertible. The aim is to find controls $\overline{u} \in U$ and $\overline{v} \in V$ such that

$$J(\overline{u},\overline{v}) = \inf_{u\in U, v\in V} J(u,v).$$

The controls $\overline{u} \in U$ and $\overline{v} \in V$ are called optimal controls. In Ahmed (1981), Benner & Trautwein (2018), Curtain & Pritchard (1978), Duncan *et al.* (2012), Fabbri & Goldys (2009), Hu & Tang (2018), similar control problems are considered with the result that the optimal controls satisfy a feedback law. We follow the same approach here and therefore introduce the following Riccati equation in $\mathcal{L}(H)$:

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}(t) = A\mathcal{P}(t) + \mathcal{P}(t)A - \mathcal{P}(t)BR^{-1}B^*\mathcal{P}(t) - \mathcal{H}^*(t)\mathcal{G}R_b^{-1}\mathcal{G}^*\mathcal{H}(t) + C^*C, \\ \mathcal{P}(T) = 0, \end{cases} \tag{2.3}$$

where $\mathcal{H}(t) = (\lambda - A)^{1-\alpha}\mathcal{P}(t)$, $\mathcal{G} = (\lambda - A)^{\alpha}N$ with $\alpha \in (1/2, 3/4)$. We make the following definition, where $\Sigma(H)$ denotes the space of all symmetric operators on H and $C([0,T];\Sigma(H))$ is endowed with the topology of uniform convergence.

Definition 2.5 The process $(\mathcal{P}(t))_{t\in[0,T]}$ is a mild solution of (2.3) if

- $\mathcal{P} \in C([0,T];\Sigma(H))$,
- $\mathcal{P}(t)y \in D((\lambda - A)^{1-\alpha})$ for every $y \in H$ and all $t \in [0,T)$,
- $(\lambda - A)^{1-\alpha}\mathcal{P} \in C([0,T);\mathcal{L}(H))$,
- $\lim_{t\to 0} t^{1-\alpha}(\lambda - A)^{1-\alpha}\mathcal{P}(t)y = 0$ for every $y \in H$,

and for all $t \in [0,T]$ and every $y \in H$,

$$\begin{aligned}\mathcal{P}(t)y = &-\int_t^T e^{A(s-t)}\mathcal{P}(s)BR^{-1}B^*\mathcal{P}(s)e^{A(s-t)}y\,\mathrm{d}s \\ &-\int_t^T e^{A(s-t)}[\mathcal{H}^*(s)\mathcal{G}R_b^{-1}\mathcal{G}^*\mathcal{H}(s) - C^*C]e^{A(s-t)}y\,\mathrm{d}s.\end{aligned} \tag{2.4}$$

In Bensoussan *et al.* (2007, Part IV), existence and uniqueness results for a mild solution to system (2.3) are shown for some special cases. The ideas of these proofs are easily adapted to the current situation, and therefore there exists a unique mild solution of system (2.3).

REMARK 2.6 Equation (2.4) can be written equivalently as

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathcal{P}(t)y, z\rangle_H = \langle \mathcal{P}(t)y, Az\rangle_H + \langle \mathcal{P}(t)Ay, z\rangle_H - \langle R^{-1}B^*\mathcal{P}(t)y, B^*\mathcal{P}(t)z\rangle_H$$
$$- \langle R_b^{-1}\mathcal{G}^*\mathcal{H}(t)y, \mathcal{G}^*\mathcal{H}(t)z\rangle_H + \langle Cy, Cz\rangle_Z$$

for every $y, z \in D(A)$; see Bensoussan *et al.* (2007).

The optimal controls $\overline{u} \in U$ and $\overline{v} \in V$ satisfy a.e. on $[0, T]$ and $\mathbb{P}$-a.s.,

$$\overline{u}(t) = -R^{-1}B^*\mathcal{P}(t)y(t), \qquad \overline{v}(t) = -R_b^{-1}\mathcal{G}^*\mathcal{H}(t)y(t).$$

Plugging these formulas into (2.2) results in the following controlled system in H:

$$\begin{cases} \mathrm{d}y(t) = \Big[Ay(t) - BR^{-1}B^*\mathcal{P}(t)y(t) - (\lambda - A)NR_b^{-1}\mathcal{G}^*\mathcal{H}(t)y(t)\Big]\,\mathrm{d}t \\ \qquad\qquad + G\,\mathrm{d}W(t) + (\lambda - A)N\,\mathrm{d}W_b(t), \\ y(0) = \xi. \end{cases} \tag{2.5}$$

3. A linear implicit Euler method for the finite element discretization

In this section we introduce a fully discrete scheme for system (2.5). We denote by $\mathcal{T}_h$ a triangulation of the domain $\mathcal{D}$ with mesh width $h \in (0, 1]$. Let $Y_h \subset Y = D((\lambda - A)^{1/2})$ be the set of continuous functions that are piecewise linear over $\mathcal{T}_h$. We introduce the L^2-projection $P_h \colon H \to Y_h$ defined by

$$\langle P_h y, z\rangle_H = \langle y, z\rangle_H$$

for every $y \in H$ and every $z \in Y_h$. Then we have the basic estimate

$$\|y - P_h y\|_H \leqslant Ch^\rho \|y\|_{\rho/2} \tag{3.1}$$

for a constant $C > 0$ and every $y \in D((\lambda - A)^{\rho/2})$ with $\rho \in [0, 2]$; see Thomeé (2006). Moreover, let $R_h \colon Y \to Y_h$ be the Y-projection given by

$$\langle (\lambda - A)R_h y, z\rangle_H = \langle (\lambda - A)y, z\rangle_H$$

for every $y \in Y$ and every $z \in Y_h$. We have the following relation between the L^2-projection P_h and the Y-projection R_h:

$$(\lambda - A_h)R_h y = P_h(\lambda - A)y \tag{3.2}$$

for every $y \in D(A)$; see Lord & Tambue (2013, Lemma 3.1). We consider the following semidiscrete version of system (2.5) in Y_h:

$$\begin{cases} \mathrm{d}y_h(t) = \Big[A_h y_h(t) - B_h R^{-1} B_h^* \mathcal{P}_h(t) y_h(t) - B_h^b R_b^{-1} \left(B_h^b\right)^* \mathcal{P}_h(t) y_h(t)\Big]\,\mathrm{d}t \\ \qquad\qquad + P_h G\,\mathrm{d}W(t) + B_h^b\,\mathrm{d}W_b(t), \\ y^h(0) = P_h \xi, \end{cases} \tag{3.3}$$

where the operator $A_h\colon Y_h \to Y_h$ satisfies for every $y, z \in Y_h$,

$$\langle A_h y, z\rangle_H = \langle Ay, z\rangle_H$$

and $B_h = P_h B$. As a consequence of inequality (3.1) we get

$$\|B^* y - B_h^* y\|_H \leqslant C h^\rho \|y\|_{\rho/2} \tag{3.4}$$

for a constant $C > 0$ and every $y \in D((\lambda - A)^{\rho/2})$ with $\rho \in [0, 2]$. Moreover, we have $B_h^b = (\lambda - A_h) R_h N \in \mathcal{L}(H_b; H)$ and $(\mathcal{P}_h(t))_{t\in[0,T]}$ with $\mathcal{P}_h(t) \in \mathcal{L}(Y_h)$ is the solution of the semidiscrete version of system (2.3) given by

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_h(t) = A_h \mathcal{P}_h(t) + \mathcal{P}_h(t) A_h - \mathcal{P}_h(t) B_h R^{-1} B_h^* \mathcal{P}_h(t) \\ \qquad\qquad - \mathcal{P}_h(t) B_h^b R_b^{-1} \left(B_h^b\right)^* \mathcal{P}_h(t) + C_h^* C_h, \\ \mathcal{P}_h(T) = 0, \end{cases} \tag{3.5}$$

where $C_h = C P_h$. By definition the operator A_h is again the infinitesimal generator of an analytic semigroup $(e^{A_h t})_{t\geqslant 0}$ on Y_h such that fractional powers of $\lambda - A_h$ with $\lambda > 0$ are well defined. We can therefore introduce the solutions of system (3.3) and (3.5) in a mild sense analogously to Definitions 2.4 and 2.5. We note that the mild solution of system (3.5) coincides again with the weak solution according to Remark 2.6.

Next, let $t_0, t_1, \ldots, t_M$ be a partition of the time interval $[0, T]$ such that $0 = t_0 < t_1 < \cdots < T_M = T$. We assume that $t_m - t_{m-1} = \Delta t$ for each $m = 1, \ldots, M$ with $\Delta t \in (0, 1]$. Applying a linear implicit Euler method to system (3.3) gives us the following fully discrete system in Y_h for $m = 1, \ldots, M$:

$$\begin{cases} y_h^m = S_{h,\Delta t} y_h^{m-1} - \Delta t S_{h,\Delta t} B_h R^{-1} B_h^* \mathcal{P}_h^{m-1} y_h^{m-1} - \Delta t S_{h,\Delta t} B_h^b R_b^{-1} \left(B_h^b\right)^* \mathcal{P}_h^{m-1} y_h^{m-1} \\ \qquad + S_{h,\Delta t} P_h G\,\delta W_m + S_{h,\Delta t} B_h^b\,\delta W_{b,m}, \\ y_h^0 = P_h \xi, \end{cases} \tag{3.6}$$

where $S_{h,\Delta t} = (I - \Delta t A_h)^{-1}$, $\delta W_{m-1} = W(t_m) - W(t_{m-1})$ and $\delta W_{b,m} = W_b(t_m) - W_b(t_{m-1})$. The operator $\mathcal{P}_h^m \in \mathcal{L}(Y_h)$ results from a time discretization of system (3.5). We make the following assumption.

ASSUMPTION 3.1 We require for each $m = 0, 1, \ldots, M-1$,

$$\|\mathcal{P}(t_m) - \mathcal{P}_h^m P_h\|_{\mathcal{L}(H)} \leqslant c\,(h^2 + \Delta t),$$
$$\|\mathcal{G}^*\mathcal{H}\,(t_m) - \left(B_h^b\right)^* \mathcal{P}_h^m P_h\|_{\mathcal{L}(H)} \leqslant c\,(h + \Delta t^{1/4}),$$

where $c > 0$ is a constant.

REMARK 3.2 Note that we can at least write formally $\mathcal{G}^*\mathcal{H}(t) = \left(B_h^b\right)^* P(t)$ for all $t \in [0, T)$. Hence, Assumption 3.1 especially provides the convergence rate for the operator $\left(B_h^b\right)^* P(t_m) - \left(B_h^b\right)^* \mathcal{P}_h^m P_h$ for each $m = 0, 1, \ldots, M-1$. For some convergence results we refer to Lasiecka & Triggiani (2000). Here we will verify the convergence rates by a numerical experiment in Section 6.

We are now in a position to state the main result of the paper.

THEOREM 3.3 Let $(y(t))_{t\in[0,T]}$ be the mild solution of system (2.5) and let y_h^m satisfy the fully discrete system (3.6) for $m = 0, 1, \ldots, M-1$. If Assumptions 2.1, 2.3 and 3.1 are fulfilled then there exists a constant $c > 0$ such that for sufficiently small $\varepsilon > 0$,

$$\|y(t_m) - y_h^m\|_{L^2(\Omega;H)} \leqslant c\left(h^{\min\{1/2-\varepsilon,\beta\}} + \Delta t^{\min\{1/4-\varepsilon,\beta/2\}}\right).$$

The proof of this theorem will be provided in Section 5. In order to prepare we will first collect and derive a number of lemmata in the following section.

4. Auxiliary results

We start by collecting some well-known properties of fractional powers of operators. For a proof see, e.g., Pazy (1983), Vrabie (2003).

LEMMA 4.1 Let $A\colon D(A) \subset H \to H$ be the Neumann realization of the Laplace operator. Then

(i) for $\alpha \leqslant 0$ the operator $(\lambda - A)^\alpha$ is linear and bounded, and for $\alpha > 0$ the operator $(\lambda - A)^\alpha$ is linear and closed;

(ii) $\alpha \geqslant \beta \geqslant 0$ implies $D\left((\lambda - A)^\alpha\right) \subset D((\lambda - A)^\beta)$ and for every $y \in D\left((\lambda - A)^\alpha\right)$,

$$\|(\lambda - A)^\beta y\|_H \leqslant M_0\|(\lambda - A)^\alpha y\|_H;$$

(iii) $D\left((\lambda - A)^\alpha\right)$ with $\alpha > 0$ is dense in H;

(iv) $(\lambda - A)^{\alpha+\beta} y = (\lambda - A)^\alpha (\lambda - A)^\beta y$ if $y \in D\left((\lambda - A)^\gamma\right)$, where $\gamma = \max\{\alpha, \beta, \alpha + \beta\}$;

(v) for $\alpha > 0$ and $t > 0$, we have $e^{At}\colon H \to D\left((\lambda - A)^\alpha\right)$ and $(\lambda - A)^\alpha e^{At} y = e^{At}(\lambda - A)^\alpha y$ if $y \in D\left((\lambda - A)^\alpha\right)$;

(vi) the operator $(\lambda - A)^\alpha e^{At}$ is linear and bounded for $\alpha > 0$ and $t > 0$, and moreover we have for every $y \in H$,

$$\|(\lambda - A)^\alpha e^{At} y\|_H \leqslant M_\alpha t^{-\alpha}\|y\|_H;$$

(vii) we have for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in (0,1]$ and all $t > 0$,

$$\|e^{At}y - y\|_H \leqslant c_\alpha t^\alpha \|(\lambda - A)^\alpha y\|_H.$$

4.1 *Continuity of mild solutions to the controlled system*

Next we show some useful properties of the exact mild solution y to the controlled system (2.5). In the following we use $c > 0$ as a generic constant, which may take different values at different points.

LEMMA 4.2 Let $(y(t))_{t\in[0,T]}$ be the mild solution of system (2.5). If Assumptions 2.1 and 2.3 hold then there exists a constant $c > 0$ such that for all $t \in [0,T]$,

$$\|y(t)\|_{L^2(\Omega;H)} \leqslant c\left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right).$$

Proof. Similar results are well known in the theory of SPDEs with homogeneous boundary conditions; see Lord & Tambue (2013), Tambue & Mukam (2019), Wang (2017). Indeed, the mild solution $(y(t))_{t\in[0,T]}$ of system (2.5) involves mostly linear and bounded operators. By using Lemma 4.1 we can introduce suitable powers $(\lambda - A)^\alpha$ in front of the remaining unbounded operators to make them bounded, if we pay by also multiplying the exponential terms by $(\lambda - A)^{-\alpha}$. For all $\alpha \in (1/2, 3/4)$ this leads to

$$\|y(t)\|_{L^2(\Omega;H)} \leqslant c\left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right) + c\int_0^t (t-s)^{\alpha-1}\|y(s)\|_{L^2(\Omega;H)}\,\mathrm{d}s.$$

The claim follows by applying a generalized Grönwall inequality; see Ye *et al.* (2007, Corollary 2). □

LEMMA 4.3 Let $(y(t))_{t\in[0,T]}$ be the mild solution of system (2.5). If Assumptions 2.1 and 2.3 hold then there exists a constant $c > 0$ such that for all $\tau_1, \tau_2 \in [0,T]$ with $\tau_1 < \tau_2$ and all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\|y(\tau_2) - y(\tau_1)\|_{L^2(\Omega;H)} \leqslant c\,(\tau_2 - \tau_1)^\gamma \left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right).$$

Proof. Similar results are well known in the theory of SPDEs with homogeneous boundary conditions; see Kruse (2014), Lord & Tambue (2013), Wang (2017). Here, we state the main steps of the proof. By definition we get

$$\|y(\tau_2) - y(\tau_1)\|_{L^2(\Omega;H)} \leqslant \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4 + \mathcal{I}_5, \tag{4.1}$$

where

$$\mathcal{I}_1 = \left\| \left[e^{A\tau_2} - e^{A\tau_1} \right] \xi \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_2 = \left\| \int_0^{\tau_1} \left[e^{A(\tau_2-s)} - e^{A(\tau_1-s)} \right] BR^{-1}B^*\mathcal{P}(s)y(s)\,\mathrm{d}s + \int_{\tau_1}^{\tau_2} e^{A(\tau_2-s)} BR^{-1}B^*\mathcal{P}(s)y(s)\,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_3 = \left\| \int_0^{\tau_1} (\lambda - A) \left[e^{A(\tau_2-s)} - e^{A(\tau_1-s)} \right] NR_b^{-1}\mathcal{G}^*\mathcal{H}(s)y(s)\,\mathrm{d}s \right.$$
$$\left. + \int_{\tau_1}^{\tau_2} (\lambda - A) e^{A(\tau_2-s)} NR_b^{-1}\mathcal{G}^*\mathcal{H}(s)y(s)\,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_4 = \left\| \int_0^{\tau_1} \left[e^{A(\tau_2-s)} - e^{A(\tau_1-s)} \right] G\,\mathrm{d}W(s) + \int_{\tau_1}^{\tau_2} e^{A(\tau_2-s)} G\,\mathrm{d}W(s) \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_5 = \left\| \int_0^{\tau_1} (\lambda - A) \left[e^{A(\tau_2-s)} - e^{A(\tau_1-s)} \right] N\,\mathrm{d}W_b(s) + \int_{\tau_1}^{\tau_2} (\lambda - A) e^{A(\tau_2-s)} N\,\mathrm{d}W_b(s) \right\|_{L^2(\Omega;H)}.$$

Recall that the semigroup $(e^{At})_{t\geqslant 0}$ is a contraction. By Lemma 4.1(ii), (vii) we obtain

$$\mathcal{I}_1 \leqslant c_\gamma (\tau_2 - \tau_1)^\gamma \, \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}. \tag{4.2}$$

Because the operators $B, R^{-1}, \mathcal{P}(t)$ are linear and bounded, using Lemma 4.1(i), (iv)–(vii) and Lemma 4.2 shows that

$$\mathcal{I}_2 \leqslant c\,(\tau_2 - \tau_1)^\gamma \left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right). \tag{4.3}$$

The operators $(\lambda - A)^\alpha N, R_b^{-1}, \mathcal{G}^*, \mathcal{H}(t)$ with $\alpha \in (0, 3/4)$ are also linear and bounded. Lemma 4.1(iv)–(vii) and Lemma 4.2 therefore give us for all $\alpha \in (\gamma, 3/4)$ that

$$\mathcal{I}_3 \leqslant c\,(\tau_2 - \tau_1)^\gamma \left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right). \tag{4.4}$$

The Itô isometry and Lemma 4.1(ii), (iv)–(vii) yield for $\beta < 1$,

$$\mathcal{I}_4^2 \leqslant c\,(\tau_2 - \tau_1)^{2\gamma}\, \mathbb{E}\,\|G\|^2_{\mathcal{L}_{\mathrm{HS}}(Q^{1/2}(H);D((\lambda-A)^{(\beta-1)/2}))}. \tag{4.5}$$

For $\beta \in [1, 2)$ the above inequality holds by a similar argument involving the fact that the operator $(\lambda - A)^{-(\beta-1)/2}$ is bounded. Finally, using the Itô isometry and Lemma 4.1(iv)–(vii) we get for all

$\alpha \in (1/2+\gamma, 3/4)$ that

$$\mathcal{I}_5^2 \leqslant c\,(\tau_2-\tau_1)^{2\gamma}. \tag{4.6}$$

Substituting inequalities (4.2)–(4.6) into (4.1) yields the result. □

4.2 *Continuity of mild solutions to the Riccati equation*

We also need similar continuity properties of the mild solution $\mathcal{P}$ to the Riccati equation (2.3) and the transformed version $\mathcal{H} = (\lambda - A)^{1-\alpha}\mathcal{P}$. In the following we use $c > 0$ as a generic constant that may change from time to time.

LEMMA 4.4 Let $(\mathcal{P}(t))_{t\in[0,T]}$ be the mild solution of system (2.3). Then there exists a constant $c > 0$ such that for all $\tau_1, \tau_2 \in [0, T]$ with $\tau_1 < \tau_2$ and all $\gamma \in (0, 1)$,

$$\|\mathcal{P}(\tau_2) - \mathcal{P}(\tau_1)\|_{\mathcal{L}(H)} \leqslant c\,(\tau_2-\tau_1)^{\gamma}.$$

Proof. Let $y \in H$. We set, for all $t \in [0, T]$,

$$\mathcal{J}(t) = \mathcal{P}(t)BR^{-1}B^*\mathcal{P}(t), \qquad \mathcal{K}(t) = \mathcal{H}^*(t)\mathcal{G}R_b^{-1}\mathcal{G}^*\mathcal{H}(t) - C^*C.$$

Note that the operators $\mathcal{J}(t)$ and $\mathcal{K}(t)$ are linear and bounded. By definition and Lemma 4.1(iv), (v) we get

$$\|\mathcal{P}(\tau_2)y - \mathcal{P}(\tau_1)y\|_H \leqslant \mathcal{I}_1 + \mathcal{I}_2, \tag{4.7}$$

where

$$\begin{aligned}\mathcal{I}_1 = &\int_{\tau_2}^{T} \left\| \left[I - e^{A(\tau_2-\tau_1)}\right] e^{A(s-\tau_2)}\mathcal{J}(s)e^{A(s-\tau_2)}y\right\|_H \mathrm{d}s + \int_{\tau_2}^{T} \left\| e^{A(s-\tau_1)}\mathcal{J}(s)\left[I - e^{A(\tau_2-\tau_1)}\right] e^{A(s-\tau_2)}y\right\|_H \mathrm{d}s\\ &+ \int_{\tau_1}^{\tau_2} \left\|(\lambda - A)^{1-\gamma} e^{A(s-\tau_1)}(\lambda-A)^{\gamma-1}\mathcal{J}(s)e^{A(s-\tau_1)}y\right\|_H \mathrm{d}s\end{aligned}$$

and

$$\begin{aligned}\mathcal{I}_2 = &\int_{\tau_2}^{T} \left\| \left[I - e^{A(\tau_2-\tau_1)}\right] e^{A(s-\tau_2)}\mathcal{K}(s)e^{A(s-\tau_2)}y\right\|_H \mathrm{d}s + \int_{\tau_2}^{T} \left\| e^{A(s-\tau_1)}\mathcal{K}(s)\left[I - e^{A(\tau_2-\tau_1)}\right] e^{A(s-\tau_2)}y\right\|_H \mathrm{d}s\\ &+ \int_{\tau_1}^{\tau_2} \left\|(\lambda - A)^{1-\gamma} e^{A(s-\tau_1)}(\lambda-A)^{\gamma-1}\mathcal{K}(s)e^{A(s-\tau_1)}y\right\|_H \mathrm{d}s.\end{aligned}$$

Recall that the semigroup $(e^{At})_{t\geqslant 0}$ is a contraction and that the operators $\mathcal{P}(t)$, B and R^{-1} are linear and bounded for all $t \in [0, T]$. Lemma 4.1(i), (vi), (vii) give us

$$\mathcal{I}_1 \leqslant c\left[(\tau_2-\tau_1)^{\gamma}\int_{\tau_2}^{T}(s-\tau_2)^{-\gamma}\,\mathrm{d}s+\int_{\tau_1}^{\tau_2}(s-\tau_1)^{\gamma-1}\,\mathrm{d}s\right]\|y\|_H \leqslant c\,(\tau_2-\tau_1)^{\gamma}\,\|y\|_H. \tag{4.8}$$

Recall that the operators $\mathcal{H}(t)$, $\mathcal{G}$, R_b^{-1} and C are linear and bounded for all $t \in [0, T]$. Similarly to the above we obtain

$$\mathcal{I}_2 \leqslant c\,(\tau_2-\tau_1)^{\gamma}\,\|y\|_H. \tag{4.9}$$

Substituting inequalities (4.8) and (4.9) into (4.7) yields the result. □

LEMMA 4.5 Let $(\mathcal{H}(t))_{t\in[0,T]}$ be given by

$$\mathcal{H}(t) = (\lambda - A)^{1-\alpha}\mathcal{P}(t)$$

for $\alpha \in (1/2, 3/4)$, where $(\mathcal{P}(t))_{t\in[0,T]}$ is the mild solution of system (2.3). Then there exists a constant $c > 0$ such that, for all $\tau_1, \tau_2 \in [0, T)$ with $\tau_1 < \tau_2$ and all $\gamma \in (0, \alpha)$,

$$\|\mathcal{H}(\tau_2) - \mathcal{H}(\tau_1)\|_{\mathcal{L}(H)} \leqslant c\,(\tau_2 - \tau_1)^{\gamma}.$$

Proof. Let $y \in H$. We set, for all $t \in [0, T]$,

$$\mathcal{J}(t) = \mathcal{P}(t)BR^{-1}B^*\mathcal{P}(t), \qquad \mathcal{K}(t) = \mathcal{H}^*(t)\mathcal{G}R_b^{-1}\mathcal{G}^*\mathcal{H}(t) - C^*C.$$

Note that the operators $\mathcal{J}(t)$ and $\mathcal{K}(t)$ are linear and bounded. By definition and Lemma 4.1(iv), (v) we get

$$\|\mathcal{H}(\tau_2)y - \mathcal{H}(\tau_1)y\|_H = \|(\lambda - A)^{1-\alpha}\mathcal{P}(\tau_2)y - (\lambda - A)^{1-\alpha}\mathcal{P}(\tau_1)y\|_H \leqslant \mathcal{I}_1 + \mathcal{I}_2, \tag{4.10}$$

where

$$\begin{aligned}\mathcal{I}_1 = &\int_{\tau_2}^{T}\left\|\left[I - e^{A(\tau_2-\tau_1)}\right](\lambda - A)^{1-\alpha}e^{A(s-\tau_2)}\mathcal{J}(s)e^{A(s-\tau_2)}y\right\|_H \mathrm{d}s\\ &+\int_{\tau_2}^{T}\left\|(\lambda - A)^{1-\alpha}e^{A(s-\tau_1)}\mathcal{J}(s)\left[I - e^{A(\tau_2-\tau_1)}\right]e^{A(s-\tau_2)}y\right\|_H \mathrm{d}s\\ &+\int_{\tau_1}^{\tau_2}\left\|(\lambda - A)^{1-\alpha}e^{A(s-\tau_1)}\mathcal{J}(s)(\lambda - A)^{\alpha-\gamma}e^{A(s-\tau_1)}(\lambda - A)^{\gamma-\alpha}y\right\|_H \mathrm{d}s\end{aligned}$$

and

$$\mathcal{I}_2 = \int_{\tau_2}^{T} \left\| \left[I - e^{A(\tau_2-\tau_1)} \right] (\lambda - A)^{1-\alpha} e^{A(s-\tau_2)} \mathcal{K}(s) e^{A(s-\tau_2)} y \right\|_H \mathrm{d}s$$
$$+ \int_{\tau_2}^{T} \left\| (\lambda - A)^{1-\alpha} e^{A(s-\tau_1)} \mathcal{K}(s) \left[I - e^{A(\tau_2-\tau_1)} \right] e^{A(s-\tau_2)} y \right\|_H \mathrm{d}s$$
$$+ \int_{\tau_1}^{\tau_2} \left\| (\lambda - A)^{1-\alpha} e^{A(s-\tau_1)} \mathcal{K}(s) (\lambda - A)^{\alpha-\gamma} e^{A(s-\tau_1)} (\lambda - A)^{\gamma-\alpha} y \right\|_H \mathrm{d}s.$$

Recall that the semigroup $(e^{At})_{t\geqslant 0}$ is a contraction and that the operators $\mathcal{P}(t)$, B and R^{-1} are linear and bounded for all $t \in [0, T]$. Lemma 4.1(i), (vi), (vii) show that

$$\mathcal{I}_1 \leqslant c \left[(\tau_2 - \tau_1)^{\gamma} \int_{\tau_2}^{T} (s - \tau_2)^{\alpha-1-\gamma} \, \mathrm{d}s \right.$$
$$\left. + (\tau_2 - \tau_1)^{-\alpha+1+\gamma} \int_{\tau_2}^{T} (s - \tau_1)^{\alpha-1} (s - \tau_2)^{\alpha-1-\gamma} \, \mathrm{d}s + \int_{\tau_1}^{\tau_2} (s - \tau_1)^{\gamma-1} \, \mathrm{d}s \right] \|y\|_H$$
$$\leqslant c \, (\tau_2 - \tau_1)^{\gamma} \, \|y\|_H. \tag{4.11}$$

Since the operators $\mathcal{H}(t)$, $\mathcal{G}$, R_b^{-1} and C are linear and bounded for all $t \in [0, T]$ a very similar argument leads to

$$\mathcal{I}_2 \leqslant c \, (\tau_2 - \tau_1)^{\gamma} \, \|y\|_H. \tag{4.12}$$

Substituting inequalities (4.11) and (4.12) into (4.10) yields the result. □

4.3 *Discretized solution operators*

Finally, we collect some results that compare the spatially discretized solution operator $e^{A_h t}$ to the exact solution operator e^{At} and to the fully discretized time stepping operator $S_{h,\Delta t}$.

LEMMA 4.6 There exists a constant $c > 0$ such that

(i) for every $y \in D((\lambda - A)^{\rho/2})$ with $\rho, r \in [0, 2]$ satisfying $\rho \leqslant r$ and all $t > 0$,

$$\left\| e^{At} y - e^{A_h t} P_h y \right\|_H \leqslant c \, h^r t^{-(r-\rho)/2} \|y\|_{\rho/2};$$

(ii) for every $y \in D((\lambda - A)^{-\rho/2})$ with $\rho \in [0,1]$ and all $t > 0$,

$$\left\| e^{At}y - e^{A_h t}P_h y \right\|_H \leqslant c\, h^{2-\rho} t^{-1} \|y\|_{-\rho/2};$$

(iii) for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ and all $t > 0$,

$$\left\| (\lambda - A)e^{At}y - e^{A_h t}(\lambda - A_h)R_h y \right\|_H \leqslant c\, h^{2\alpha} t^{-1} \|y\|_{\alpha}.$$

Proof. A proof of (i) can be found in Lord & Tambue (2013, Lemma 3.1) for $r \in \{1, 2\}$. For $r = 0$ the inequality is an immediate consequence of the fact that the semigroups $(e^{At})_{t\geqslant 0}$ and $(e^{A_h t})_{t\geqslant 0}$ are contractions. The result holds for all $r \in [0, 2]$ applying interpolation techniques, which is demonstrated in Thomeé (2006, Theorem 3.5). For assertion (ii) we can follow Tambue & Mukam (2019, Lemma 3.2 (iii)). It remains to show (iii). Let us first assume that $y \in D(A)$. By Equation (3.2), Lemma 4.1(iv), (v) and claim (ii) with $\rho = 2 - 2\alpha$ we obtain

$$\begin{aligned}
\left\| (\lambda - A)e^{At}y - e^{A_h t}(\lambda - A_h)R_h y \right\|_H &= \left\| e^{At}(\lambda - A)y - e^{A_h t}P_h(\lambda - A)y \right\|_H \\
&\leqslant c\, h^{2\alpha} t^{-1} \left\| (\lambda - A)^{\alpha - 1}(\lambda - A)y \right\|_H \\
&= c\, h^{2\alpha} t^{-1} \|y\|_{\alpha}.
\end{aligned}$$

The above inequality holds also for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ by standard density arguments. Indeed, for every $y \in D((\lambda - A)^{\alpha})$, there exists a sequence $(y_k)_{k\in\mathbb{N}} \subset D(A)$ such that $y_k \to y$ in $D((\lambda - A)^{\alpha})$ as $k \to \infty$ resulting from Lemma 4.1(ii), (iii). Due to Lemma 4.1(iv)–(vi) we get for each $k \in \mathbb{N}$,

$$\begin{aligned}
&\left\| (\lambda - A)e^{At}y - e^{A_h t}(\lambda - A_h)R_h y \right\|_H \\
&\quad = \left\| (\lambda - A)e^{At}(y - y_k) - e^{A_h t}(\lambda - A_h)R_h(y - y_k) + (\lambda - A)e^{At}y_k - e^{A_h t}(\lambda - A_h)R_h y_k \right\|_H \\
&\quad \leqslant \left\| (\lambda - A)^{1-\alpha}e^{At}(\lambda - A)^{\alpha}(y - y_k) \right\|_H + \left\| e^{A_h t}(\lambda - A_h)R_h(y - y_k) \right\|_H \\
&\qquad + \left\| (\lambda - A)e^{At}y_k - e^{A_h t}(\lambda - A_h)R_h y_k \right\|_H \\
&\quad \leqslant \left(M_{1-\alpha}\, t^{\alpha-1} + c\, M_0 \right) \|y - y_k\|_{\alpha} + c\, h^{2\alpha} t^{-1} \left\| y_k \right\|_{\alpha} \\
&\quad \leqslant \left(M_{1-\alpha}\, t^{\alpha-1} + c\, M_0 \right) \|y - y_k\|_{\alpha} + c\, h^{2\alpha} t^{-1} \left\| y_k - y \right\|_{\alpha} + c\, h^{2\alpha} t^{-1} \|y\|_{\alpha}.
\end{aligned}$$

Hence, the result follows as $k \to \infty$. □

LEMMA 4.7 (Fujita & Mizutani, 1976, Theorem 6.1). For each $m = 1, \ldots, M$ we have

$$\left\| S^m_{h,\Delta t} \right\|_{\mathcal{L}(H)} \leqslant 1,$$

where $S^m_{h,\Delta t}$ denotes the composition of $S_{h,\Delta t}$ with itself m times.

LEMMA 4.8 There exists a constant $c > 0$ such that

(i) for every $y \in D((\lambda - A)^{\rho/2})$ with $\rho \in [0, 2]$ and each $m = 0, 1, \dots, M$,

$$\left\| e^{A_h t_m} P_h y - S^m_{h,\Delta t} P_h y \right\|_H \leqslant c\, \Delta t^{\rho/2} \|y\|_{\rho/2};$$

(ii) for every $y \in D((\lambda - A)^{-\rho/2})$ with $\rho \in [0, 1]$ and each $m = 1, \dots, M$,

$$\left\| e^{A_h t_m} P_h y - S^m_{h,\Delta t} P_h y \right\|_H \leqslant c\, t_m^{-1}\, \Delta t^{(2-\rho)/2} \|y\|_{-\rho/2};$$

(iii) for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ and each $m = 1, \dots, M$,

$$\left\| e^{A_h t_m} (\lambda - A_h) R_h y - S^m_{h,\Delta t} (\lambda - A_h) R_h y \right\|_H \leqslant c\, t_m^{-1}\, \Delta t^{\alpha} \|y\|_{\alpha}.$$

Proof. Claims (i) and (ii) are proven in Tambue & Mukam (2019, Lemma 3.3). It remains to show (iii). Let us first assume that $y \in D(A)$. Using Equation (3.2) and (ii) with $\rho = 2 - 2\alpha$ we get

$$\begin{aligned}\left\| e^{A_h t_m} (\lambda - A_h) R_h y - S^m_{h,\Delta t} (\lambda - A_h) R_h y \right\|_H &= \left\| e^{A_h t_m} P_h (\lambda - A) y - S^m_{h,\Delta t} P_h (\lambda - A) y \right\|_H \\ &\leqslant c\, t_m^{-1}\, \Delta t^{\alpha} \left\| (\lambda - A)^{\alpha - 1} (\lambda - A) y \right\|_H \\ &= c\, t_m^{-1}\, \Delta t^{\alpha} \|y\|_{\alpha}.\end{aligned}$$

The above inequality holds also for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ by standard density arguments as demonstrated in Lemma 4.6(iii). □

LEMMA 4.9 There exists a constant $c > 0$ such that

(i) for every $y \in D((\lambda - A)^{-\rho/2})$ with $\rho \in [0, 1]$ and all $t > 0$,

$$\left\| \int_0^t e^{As} y - e^{A_h s} P_h y \, \mathrm{d}s \right\|_H \leqslant c\, h^{2-\rho} \|y\|_{-\rho/2};$$

(ii) for every $y \in D((\lambda - A)^{(\mu-1)/2})$ with $\mu \in [0, 2]$ and all $t > 0$,

$$\left(\int_0^t \left\| e^{As} y - e^{A_h s} P_h y \right\|_H^2 \mathrm{d}s \right)^{1/2} \leqslant c\, h^{\mu} \|y\|_{(\mu-1)/2};$$

(iii) for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ and all $t > 0$,

$$\left\| \int_0^t (\lambda - A)e^{As}y - e^{A_h s}(\lambda - A_h)R_h y \,\mathrm{d}s \right\|_H \leqslant c\, h^{2\alpha} \|y\|_{\alpha};$$

(iv) for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 3/2]$ and all $t > 0$,

$$\left(\int_0^t \left\| (\lambda - A)e^{As}y - e^{A_h s}(\lambda - A_h)R_h y \right\|_H^2 \mathrm{d}s \right)^{1/2} \leqslant c\, h^{2\alpha-1} \|y\|_{\alpha}.$$

Proof. Claims (i) and (ii) are shown in Tambue & Mukam (2019, Lemma 3.2). It remains to show (iii) and (iv). First, we assume that $y \in D(A)$. Using Lemma 4.1(iv), (v), Equation (3.2) and (i) with $\rho = 2 - 2\alpha$ we get

$$\begin{aligned} \left\| \int_0^t (\lambda - A)e^{As}y - e^{A_h s}(\lambda - A_h)R_h y \,\mathrm{d}s \right\|_H &= \left\| \int_0^t e^{As}(\lambda - A)y - e^{A_h s}P_h(\lambda - A)y \,\mathrm{d}s \right\|_H \\ &\leqslant c\, h^{2\alpha} \|(\lambda - A)^{\alpha-1}(\lambda - A)y\|_H \\ &= c\, h^{2\alpha} \|y\|_{\alpha}. \end{aligned} \tag{4.13}$$

Using Lemma 4.1(iv), (v), Equation (3.2) and (ii) with $\mu = 2\alpha - 1$ we have

$$\begin{aligned} \left(\int_0^t \left\| (\lambda - A)e^{As}y - e^{A_h s}(\lambda - A_h)R_h y \right\|_H^2 \mathrm{d}s \right)^{1/2} &= \left(\int_0^t \left\| e^{As}(\lambda - A)y - e^{A_h s}P_h(\lambda - A)y \right\|_H^2 \mathrm{d}s \right)^{1/2} \\ &\leqslant c\, h^{2\alpha-1} \|(\lambda - A)^{\alpha-1}(\lambda - A)y\|_H \\ &= c\, h^{2\alpha-1} \|y\|_{\alpha}. \end{aligned} \tag{4.14}$$

Inequality (4.13) holds for every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ and inequality (4.14) holds for every $y \in D((\lambda - A)^{\alpha}$ with $\alpha \in [1/2, 3/2]$ by standard density arguments as shown in Lemma 4.6(iii). □

LEMMA 4.10 There exists a constant $c > 0$ such that

(i) for arbitrary small $\varepsilon > 0$, every $y \in D((\lambda - A)^{-\rho/2})$ with $\rho \in [0, 1]$ and each $m = 1, \ldots, M$,

$$\left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} e^{A_h s}P_h y - S_{h,\Delta t}^{k+1}P_h y \,\mathrm{d}s \right\|_H \leqslant c\, \Delta t^{(2-\rho)/2-\varepsilon} \|y\|_{-\rho/2};$$

(ii) for arbitrary small $\varepsilon > 0$, every $y \in D((\lambda - A)^{(\mu-1)/2})$ with $\mu \in [0, 2]$ and each $m = 1, \ldots, M$,

$$\left(\sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left\| e^{A_h s} P_h y - S_{h,\Delta t}^{k+1} P_h y \right\|_H^2 \mathrm{d}s \right)^{1/2} \leqslant c\, \Delta t^{\mu/2-\varepsilon} \|y\|_{(\mu-1)/2};$$

(iii) for arbitrary small $\varepsilon > 0$, every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 1]$ and each $m = 1, \ldots, M$,

$$\left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} e^{A_h s}(\lambda - A_h) R_h y - S_{h,\Delta t}^{k+1}(\lambda - A_h) R_h y \,\mathrm{d}s \right\|_H \leqslant c\, \Delta t^{\alpha-\varepsilon} \|y\|_{\alpha};$$

(iv) for arbitrary small $\varepsilon > 0$, every $y \in D((\lambda - A)^{\alpha})$ with $\alpha \in [1/2, 3/2]$ and each $m = 1, \ldots, M$,

$$\left(\sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left\| e^{A_h s}(\lambda - A_h) R_h y - S_{h,\Delta t}^{k+1}(\lambda - A_h) R_h y \right\|_H^2 \mathrm{d}s \right)^{1/2} \leqslant c\, \Delta t^{(2\alpha-1)/2-\varepsilon} \|y\|_{\alpha}.$$

Proof. Assertions (i) and (ii) are proven in Tambue & Mukam (2019, Lemma 3.5). Claims (iii) and (iv) can be obtained similarly to Lemma 4.9(iii), (iv). □

5. Proof of Theorem 3.3

After all the preparation in the previous section we can now prove the main result.

Proof of Theorem 3.3. The mild solution of system (2.5) can be rewritten $\mathbb{P}$-a.s.,

$$y(t_m) = e^{At_m}\xi - \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} e^{A(t_m-s)} BR^{-1}B^* \mathcal{P}(s) y(s)\,\mathrm{d}s - \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (\lambda - A) e^{A(t_m-s)} NR_b^{-1} \mathcal{G}^* \mathcal{H}(s) y(s)\,\mathrm{d}s$$
$$+ \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} e^{A(t_m-s)} G\,\mathrm{d}W(s) + \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (\lambda - A) e^{A(t_m-s)} N\,\mathrm{d}W_b(s).$$

Similarly, the fully discrete scheme (3.6) can be rewritten $\mathbb{P}$-a.s.,

$$y_h^m = S_{h,\Delta t}^m P_h \xi - \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h R^{-1} B_h^* \mathcal{P}_h^k y_h^k\,\mathrm{d}s - \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h^b R_b^{-1} \left(B_h^b\right)^* \mathcal{P}_h^k y_h^k\,\mathrm{d}s$$
$$+ \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} P_h G(t_{m-1})\,\mathrm{d}W(s) + \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h^b\,\mathrm{d}W_b(s).$$

Therefore, we obtain

$$\|y(t_m) - y_h^m\|_{L^2(\Omega;H)} \leqslant \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4 + \mathcal{I}_5, \tag{5.1}$$

where

$$\mathcal{I}_1 = \left\| e^{At_m}\xi - S_{h,\Delta t}^m P_h \xi \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_2 = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} e^{A(t_m-s)} BR^{-1}B^*\mathcal{P}(s)y(s)\,\mathrm{d}s - \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h R^{-1} B_h^* \mathcal{P}_h^k y_h^k \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_3 = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (\lambda - A) e^{A(t_m-s)} NR_b^{-1}\mathcal{G}^*\mathcal{H}(s)y(s)\,\mathrm{d}s - \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h^b R_b^{-1} \left(B_h^b\right)^* \mathcal{P}_h^k y_h^k \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_4 = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} e^{A(t_m-s)} G\,\mathrm{d}W(s) - \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} P_h G\,\mathrm{d}W(s) \right\|_{L^2(\Omega;H)} \quad \text{and}$$

$$\mathcal{I}_5 = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (\lambda - A) e^{A(t_m-s)} N\,\mathrm{d}W_b(s) - \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h^b\,\mathrm{d}W_b(s) \right\|_{L^2(\Omega;H)}.$$

Lemma 4.6(i) with $r = \rho = \beta$ and Lemma 4.8(i) with $\rho = \beta$ give us

$$\begin{aligned}\mathcal{I}_1 &\leqslant \left\| e^{At_m}\xi - e^{A_h t_m} P_h \xi \right\|_{L^2(\Omega;H)} + \left\| e^{A_h t_m} P_h \xi - S_{h,\Delta t}^m P_h \xi \right\|_{L^2(\Omega;H)} \\ &\leqslant c\,(h^\beta + \Delta t^{\beta/2}) \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}.\end{aligned} \tag{5.2}$$

Recall that $B_h = P_h B$. We have

$$\mathcal{I}_2 \leqslant \mathcal{I}_{2,1} + \mathcal{I}_{2,2} + \mathcal{I}_{2,3} + \mathcal{I}_{2,4} + \mathcal{I}_{2,5} + \mathcal{I}_{2,6}, \tag{5.3}$$

where

$$\mathcal{I}_{2,1} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A(t_m-s)} - e^{A_h(t_m-s)} P_h \right] BR^{-1}B^* \mathcal{P}(s) y(s) \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_{2,2} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h(t_m-s)} P_h - S_{h,\Delta t}^{m-k} P_h \right] BR^{-1}B^* \mathcal{P}(s) y(s) \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_{2,3} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h R^{-1} [B^* - B_h^*] \mathcal{P}(s) y(s) \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_{2,4} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h R^{-1} B_h^* [\mathcal{P}(s) y(s) - \mathcal{P}(t_k) y(t_k)] \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_{2,5} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h R^{-1} B_h^* \left[\mathcal{P}(t_k) - \mathcal{P}_h^k P_h \right] y(t_k) \,\mathrm{d}s \right\|_{L^2(\Omega;H)},$$

$$\mathcal{I}_{2,6} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k} B_h R^{-1} B_h^* \mathcal{P}_h^k P_h \left[y(t_k) - y_h^k \right] \,\mathrm{d}s \right\|_{L^2(\Omega;H)}.$$

Recall that the operator $\mathcal{P}(t)$ is linear and bounded for all $t \in [0, T]$. Using Lemmas 4.2–4.4 there exists a constant $c > 0$ such that for all $\tau_1, \tau_2 \in [0, T]$ with $\tau_1 < \tau_2$ and all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\begin{aligned} \left\| \mathcal{P}(\tau_2) y(\tau_2) - \mathcal{P}(\tau_1) y(\tau_1) \right\|_{L^2(\Omega;H)} &\leqslant \left\| \mathcal{P}(\tau_2) \left[y(\tau_2) - y(\tau_1) \right] \right\|_{L^2(\Omega;H)} \\ &\quad + \left\| \left[\mathcal{P}(\tau_2) - \mathcal{P}(\tau_1) \right] y(\tau_1) \right\|_{L^2(\Omega;H)} \\ &\leqslant c\,(\tau_2 - \tau_1)^{\gamma} \left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))} \right). \end{aligned} \tag{5.4}$$

We set for all $t \in [0, T]$ and $\mathbb{P}$-a.s.,

$$\tilde{y}(t) = BR^{-1}B^* \mathcal{P}(t) y(t).$$

By a change of variables we get

$$\mathcal{I}_{2,1} \leqslant \left\| \int_0^{t_m} \left[e^{As} - e^{A_h s} P_h \right] \left(\tilde{y}(t_m - s) - \tilde{y}(t_m) \right) \mathrm{d}s \right\|_{L^2(\Omega;H)} + \left\| \int_0^{t_m} \left[e^{As} - e^{A_h s} P_h \right] \tilde{y}(t_m) \,\mathrm{d}s \right\|_{L^2(\Omega;H)}.$$

Recall that the operators B and R^{-1} are linear and bounded. Due to Lemmas 4.2, 4.3, 4.6(ii) with $\rho = 0$, 4.9(i) with $\rho = 0$ and inequality (5.4) we obtain, for all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\begin{aligned}
\mathcal{I}_{2,1} &\leqslant c\, h^2 \int_0^{t_m} s^{-1} \left\|\tilde{y}(t_m - s) - \tilde{y}(t_m)\right\|_{L^2(\Omega;H)} \mathrm{d}s + c\, h^2 \left\|\tilde{y}(t_m)\right\|_{L^2(\Omega;H)} \\
&\leqslant c\, h^2 \left[\int_0^{t_m} s^{\gamma-1} \mathrm{d}s + 1\right] \left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right) \\
&\leqslant c\, h^2 \left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right). \qquad (5.5)
\end{aligned}$$

We have

$$\mathcal{I}_{2,2} \leqslant \mathcal{I}_{2,2}^{(1)} + \mathcal{I}_{2,2}^{(2)}, \qquad (5.6)$$

where

$$\begin{aligned}
\mathcal{I}_{2,2}^{(1)} &= \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h(t_m-s)} P_h - S_{h,\Delta t}^{m-k} P_h\right] \left(\tilde{y}(s) - \tilde{y}(t_m)\right) \mathrm{d}s \right\|_{L^2(\Omega;H)} \quad \text{and} \\
\mathcal{I}_{2,2}^{(2)} &= \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h(t_m-s)} P_h - S_{h,\Delta t}^{m-k} P_h\right] \tilde{y}(t_m) \,\mathrm{d}s \right\|_{L^2(\Omega;H)}.
\end{aligned}$$

By a change of variables we obtain

$$\begin{aligned}
\mathcal{I}_{2,2}^{(1)} \leqslant{}& \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[I - e^{A_h(t_{k+1}-s)}\right] e^{A_h s} P_h \left(\tilde{y}(t_m - s) - \tilde{y}(t_m)\right) \mathrm{d}s \right\|_{L^2(\Omega;H)} \\
&+ \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h t_{k+1}} P_h - S_{h,\Delta t}^{k+1} P_h\right] \left(\tilde{y}(t_m - s) - \tilde{y}(t_m)\right) \mathrm{d}s \right\|_{L^2(\Omega;H)} \quad \text{and} \\
\mathcal{I}_{2,2}^{(2)} ={}& \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h s} P_h - S_{h,\Delta t}^{k+1} P_h\right] \tilde{y}(t_m) \,\mathrm{d}s \right\|_{L^2(\Omega;H)}.
\end{aligned}$$

Note that the properties from Lemma 4.1 hold also for the operator A_h and for the corresponding semigroup $(e^{A_h t})_{t \geqslant 0}$. Moreover, the operator B_h is linear and bounded. Using Lemmas 4.3, 4.8(ii) with

$\rho = 0$ and inequality (5.4) we get, for all $\gamma \in (0, 1/4)$ with $\gamma \leqslant \beta/2$,

$$\begin{aligned}\mathcal{I}_{2,2}^{(1)} &\leqslant c \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_{k+1} - s) s^{-1} \left\| P_h \left(\tilde{y}(t_m - s) - \tilde{y}(t_m)\right) \right\|_{L^2(\Omega; H)} \mathrm{d}s \\ &\quad + c\,\Delta t \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} t_{k+1}^{-1} \left\| \tilde{y}(t_m - s) - \tilde{y}(t_m) \right\|_{L^2(\Omega; H)} \mathrm{d}s \\ &\leqslant c\,\Delta t \int_0^{t_m} s^{\gamma - 1} \mathrm{d}s \\ &\leqslant c\,\Delta t. \end{aligned} \tag{5.7}$$

Due to Lemmas 4.2 and 4.10(i) with $\rho = 0$ we have

$$\mathcal{I}_{2,2}^{(2)} \leqslant c\,\Delta t^{1-\varepsilon} \left\| \tilde{y}(t_m) \right\|_{L^2(\Omega; H)} \leqslant c\,\Delta t^{1-\varepsilon} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))}\right). \tag{5.8}$$

Substituting inequalities (5.7) and (5.8) into (5.6) yields

$$\mathcal{I}_{2,2} \leqslant c\,\Delta t^{1-\varepsilon} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))}\right). \tag{5.9}$$

Using Lemmas 4.2, 4.7 and inequality (3.4) with $\rho < 1$ we obtain

$$\mathcal{I}_{2,3} \leqslant c h^{\rho} \int_0^{t_m} \left\| (\lambda - A)^{\rho/2} \mathcal{P}(s) y(s) \right\|_{L^2(\Omega; H)} \mathrm{d}s \leqslant c h^{\rho} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))}\right). \tag{5.10}$$

Lemma 4.7 and inequality (5.4) give us, for all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\mathcal{I}_{2,4} \leqslant c \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (s - t_k)^{\gamma} \mathrm{d}s \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))}\right) \leqslant c\,\Delta t^{\gamma} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))}\right). \tag{5.11}$$

Due to Assumption 3.1 and Lemma 4.2 we get

$$\mathcal{I}_{2,5} \leqslant c\,(h^2 + \Delta t) \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left\| y(t_k) \right\|_{L^2(\Omega; H)} \mathrm{d}s \leqslant c\,(h^2 + \Delta t) \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))}\right). \tag{5.12}$$

As a consequence of Lemma 4.7 we have

$$\mathcal{I}_{2,6} \leqslant c\,\Delta t \sum_{k=0}^{m-1} \left\| y(t_k) - y_h^k \right\|_{L^2(\Omega; H)}. \tag{5.13}$$

Substituting inequalities (5.5) and (5.9)–(5.13) into (5.3) yields for sufficiently small $\varepsilon > 0$ that

$$\mathcal{I}_2 \leqslant c\,(h^{\rho} + \Delta t^{\min\{1/4-\varepsilon,\beta/2\}})\left(1 + \|\xi\|_{L^2(\Omega;\,D((\lambda-A)^{\beta/2}))}\right) + c\,\Delta t \sum_{k=0}^{m-1} \left\| y(t_k) - y_h^k \right\|_{L^2(\Omega;\,H)}. \tag{5.14}$$

Recall that $B_h^b = (\lambda - A_h)R_h N$. Similarly to above we get

$$\mathcal{I}_3 \leqslant \mathcal{I}_{3,1} + \mathcal{I}_{3,2} + \mathcal{I}_{3,3} + \mathcal{I}_{3,4} + \mathcal{I}_{3,5}, \tag{5.15}$$

where

$$\mathcal{I}_{3,1} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[(\lambda - A)e^{A(t_m-s)} - e^{A_h(t_m-s)}(\lambda - A_h)R_h \right] N R_b^{-1} \mathcal{G}^* \mathcal{H}(s) y(s)\,\mathrm{d}s \right\|_{L^2(\Omega;\,H)},$$

$$\mathcal{I}_{3,2} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h(t_m-s)}(\lambda - A_h)R_h - S_{h,\Delta t}^{m-k}(\lambda - A_h)R_h \right] N R_b^{-1} \mathcal{G}^* \mathcal{H}(s) y(s)\,\mathrm{d}s \right\|_{L^2(\Omega;\,H)},$$

$$\mathcal{I}_{3,3} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k}(\lambda - A_h)R_h N R_b^{-1} \mathcal{G}^* \left[\mathcal{H}(s) y(s) - \mathcal{H}(t_k) y(t_k) \right] \mathrm{d}s \right\|_{L^2(\Omega;\,H)},$$

$$\mathcal{I}_{3,4} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k}(\lambda - A_h)R_h N R_b^{-1} \left[\mathcal{G}^* \mathcal{H}(t_k) - \left(B_h^b\right)^* \mathcal{P}_h^k P_h \right] y(t_k)\,\mathrm{d}s \right\|_{L^2(\Omega;\,H)} \quad \text{and}$$

$$\mathcal{I}_{3,5} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} S_{h,\Delta t}^{m-k}(\lambda - A_h)R_h N R_b^{-1} \left(B_h^b\right)^* \mathcal{P}_h^k P_h \left[y(t_k) - y_h^k \right] \mathrm{d}s \right\|_{L^2(\Omega;\,H)}.$$

Recall that the operator $\mathcal{H}(t)$ is linear and bounded for all $t \in [0,T)$. Using Lemmas 4.2, 4.3 and 4.5 there exists a constant $c > 0$ such that for all $\tau_1, \tau_2 \in [0,T)$ with $\tau_1 < \tau_2$ and all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\begin{aligned}
\left\| \mathcal{H}(\tau_2) y(\tau_2) - \mathcal{H}(\tau_1) y(\tau_1) \right\|_{L^2(\Omega;\,H)} &\leqslant \left\| \mathcal{H}(\tau_2) \left[y(\tau_2) - y(\tau_1) \right] \right\|_{L^2(\Omega;\,H)} \\
&\quad + \left\| \left[\mathcal{H}(\tau_2) - \mathcal{H}(\tau_1) \right] y(\tau_1) \right\|_{L^2(\Omega;\,H)} \\
&\leqslant c\,(\tau_2 - \tau_1)^{\gamma} \left(1 + \|\xi\|_{L^2(\Omega;\,D((\lambda-A)^{\beta/2}))}\right).
\end{aligned} \tag{5.16}$$

We set, for all $t \in [0,T)$ and $\mathbb{P}$-a.s.,

$$\overline{y}(t) = NR_b^{-1}\mathcal{G}^*\mathcal{H}(t)y(t).$$

By a change of variables we obtain

$$\begin{aligned}\mathcal{I}_{3,1} \leqslant & \left\| \int_0^{t_m} \left[(\lambda - A)e^{As} - e^{A_h s}(\lambda - A_h)R_h \right] \left(\overline{y}(t_m - s) - \overline{y}(t_m) \right) \mathrm{d}s \right\|_{L^2(\Omega;H)} \\ & + \left\| \int_0^{t_m} \left[(\lambda - A)e^{As} - e^{A_h s}(\lambda - A_h)R_h \right] \overline{y}(t_m) \, \mathrm{d}s \right\|_{L^2(\Omega;H)}.\end{aligned}$$

Recall that the operators $(\lambda - A)^\alpha N, R_b^{-1}, \mathcal{G}^*$ are linear and bounded for all $\alpha \in (0, 3/4)$. Lemmas 4.2, 4.6(iii) with $\alpha \in [1/2, 3/4)$, inequality (5.16) and Lemma 4.9(iii) with $\alpha \in [1/2, 3/4)$ give us, for all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\begin{aligned}\mathcal{I}_{3,1} &\leqslant c\, h^{2\alpha} \int_0^{t_m} s^{-1} \left\| (\lambda - A)^\alpha \left(\overline{y}(t_m - s) - \overline{y}(t_m) \right) \right\|_{L^2(\Omega;H)} \mathrm{d}s + c\, h^{2\alpha} \left\| (\lambda - A)^\alpha \overline{y}(t_m) \right\|_{L^2(\Omega;H)} \\ &\leqslant c\, h^{2\alpha} \left[\int_0^{t_m} s^{\gamma - 1} \mathrm{d}s + 1 \right] \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right) \\ &\leqslant c\, h^{2\alpha} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right). \end{aligned} \tag{5.17}$$

We have

$$\mathcal{I}_{3,2} \leqslant \mathcal{I}_{3,2}^{(1)} + \mathcal{I}_{3,2}^{(2)}, \tag{5.18}$$

where

$$\mathcal{I}_{3,2}^{(1)} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h(t_m - s)}(\lambda - A_h)R_h - S_{h,\Delta t}^{m-k}(\lambda - A_h)R_h \right] \left(\overline{y}(s) - \overline{y}(t_m) \right) \mathrm{d}s \right\|_{L^2(\Omega;H)} \quad \text{and}$$

$$\mathcal{I}_{3,2}^{(2)} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h(t_m - s)}(\lambda - A_h)R_h - S_{h,\Delta t}^{m-k}(\lambda - A_h)R_h \right] \overline{y}(t_m) \, \mathrm{d}s \right\|_{L^2(\Omega;H)}.$$

By a change of variables we get

$$\mathcal{I}_{3,2}^{(1)} \leqslant \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[I - e^{A_h(t_{k+1}-s)} \right] e^{A_h s} (\lambda - A_h) R_h \left(\bar{y}(t_m - s) - \bar{y}(t_m) \right) \mathrm{d}s \right\|_{L^2(\Omega;H)}$$
$$+ \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h t_{k+1}} (\lambda - A_h) R_h - S_{h,\Delta t}^{k+1} (\lambda - A_h) R_h \right] \left(\bar{y}(t_m - s) - \bar{y}(t_m) \right) \mathrm{d}s \right\|_{L^2(\Omega;H)} \quad \text{and}$$
$$\mathcal{I}_{3,2}^{(2)} = \left\| \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left[e^{A_h s} (\lambda - A_h) R_h - S_{h,\Delta t}^{k+1} (\lambda - A_h) R_h \right] \bar{y}(t_m) \, \mathrm{d}s \right\|_{L^2(\Omega;H)}.$$

Recall that the operators $(\lambda - A_h), R_h$ are linear and bounded. Lemma 4.8(iii) with $\alpha \in [1/2, 3/4)$ and inequality (5.16) yield, for all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\begin{aligned}
\mathcal{I}_{3,2}^{(1)} &\leqslant c \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (t_{k+1} - s) s^{-1} \left\| (\lambda - A_h) R_h \left(\bar{y}(t_m - s) - \bar{y}(t_m) \right) \right\|_{L^2(\Omega;H)} \mathrm{d}s \\
&\quad + c \, \Delta t^{\alpha} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} t_{k+1}^{-1} \left\| (\lambda - A)^{\alpha} N R_b^{-1} \mathcal{G}^* \left(\bar{y}(t_m - s) - \bar{y}(t_m) \right) \right\|_{L^2(\Omega;H)} \mathrm{d}s \\
&\leqslant c \left[\Delta t \int_0^{t_m} s^{\gamma-1} \mathrm{d}s + \Delta t^{\alpha} \int_0^{t_m} s^{\gamma-1} \mathrm{d}s \right] \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right) \\
&\leqslant c \, \Delta t^{\alpha} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right).
\end{aligned} \tag{5.19}$$

Due to Lemmas 4.2 and 4.10(iii) with $\alpha \in [1/2, 3, 4)$ we have

$$\mathcal{I}_{3,2}^{(2)} \leqslant c \, \Delta t^{\alpha - \varepsilon} \left\| (\lambda - A)^{\alpha} N R_b^{-1} \mathcal{G}^* \mathcal{H}(t_m) y(t_m) \right\|_{L^2(\Omega;H)} \leqslant c \, \Delta t^{\alpha - \varepsilon} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right). \tag{5.20}$$

Substituting inequalities (5.19) and (5.20) into (5.18) yields

$$\mathcal{I}_{3,2} \leqslant c \, \Delta t^{\mu} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right) \tag{5.21}$$

with $\mu \in (0, 3/4)$. By Lemma 4.7 and inequality (5.16) we get, for all $\gamma \in (0, 1/4)$ with $\gamma < \beta/2$,

$$\mathcal{I}_{3,3} \leqslant c \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} (s - t_k)^{\gamma} \mathrm{d}s \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right) \leqslant c \, \Delta t^{\gamma} \left(1 + \|\xi\|_{L^2(\Omega; D((\lambda - A)^{\beta/2}))} \right). \tag{5.22}$$

Using Lemmas 4.2, 4.7 and Assumption 3.1 we have

$$\mathcal{I}_{3,4} \leqslant c\left(h + \Delta t^{1/4}\right)\left(1 + \|\xi\|_{L^2(\Omega; D((\lambda-A)^{\beta/2}))}\right). \tag{5.23}$$

Lemma 4.7 gives us

$$\mathcal{I}_{3,5} \leqslant c\, \Delta t \sum_{k=0}^{m-1} \left\| y(t_k) - y_h^k \right\|_{L^2(\Omega; H)}. \tag{5.24}$$

Substituting inequalities (5.17) and (5.21)–(5.24) into (5.15) yields for sufficiently small $\varepsilon > 0$ that

$$\mathcal{I}_3 \leqslant c\left(h + \Delta t^{\min\{1/4-\varepsilon, \beta/2\}}\right)\left(1 + \|\xi\|_{L^2(\Omega; D((\lambda-A)^{\beta/2}))}\right) + c\, \Delta t \sum_{k=0}^{m-1} \left\| y(t_k) - y_h^k \right\|_{L^2(\Omega; H)}. \tag{5.25}$$

We set $S(t) = S_{h,\Delta t}^k$ if $t \in [t_{k-1}, t_k)$ for each $k = 1, \ldots, M$. The Itô isometry and a change of variables give us

$$\begin{aligned}\mathcal{I}_4 \leqslant{}& \left(\mathbb{E} \int_0^{t_m} \left\| \left[e^{As} - e^{A_h s} P_h\right] G \right\|^2_{\mathcal{L}_{\mathrm{HS}}(Q^{1/2}(H); H)} \mathrm{d}s\right)^{1/2} \\ &+ \left(\mathbb{E} \sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left\| \left[e^{A_h s} P_h - S_{h,\Delta t}^{k+1} P_h\right] G \right\|^2_{\mathcal{L}_{\mathrm{HS}}(Q^{1/2}(H); H)} \mathrm{d}s\right)^{1/2}.\end{aligned}$$

By Lemma 4.9(ii) with $\mu = \beta$ and Lemma 4.10(ii) with $\mu = \beta - 2\varepsilon$ we obtain

$$\begin{aligned}\mathcal{I}_4 &\leqslant c\left(h^\beta + \Delta t^{\beta/2}\right)\left(\mathbb{E}\left\|(\lambda - A)^{-\varepsilon}(\lambda - A)^{(\beta-1)/2} G\right\|^2_{\mathcal{L}_{\mathrm{HS}}(Q^{1/2}(H); H)}\right)^{1/2} \\ &\leqslant c\left(h^\beta + \Delta t^{\beta/2}\right)\left(\mathbb{E}\,\|G\|^2_{\mathcal{L}_{\mathrm{HS}}(Q^{1/2}(H); D((\lambda-A)^{(\beta-1)/2}))}\right)^{1/2}.\end{aligned} \tag{5.26}$$

Similarly, we have

$$\begin{aligned}\mathcal{I}_5 \leqslant{}& \left(\int_0^{t_m} \left\| \left[(\lambda - A)e^{As} - e^{A_h s}(\lambda - A_h) R_h\right] N \right\|^2_{\mathcal{L}_{\mathrm{HS}}(Q_b^{1/2}(H_b); H)} \mathrm{d}s\right)^{1/2} \\ &+ \left(\sum_{k=0}^{m-1} \int_{t_k}^{t_{k+1}} \left\| \left[e^{A_h s}(\lambda - A_h) R_h - S_{h,\Delta t}^k (\lambda - A_h) R_h\right] N \right\|^2_{\mathcal{L}_{\mathrm{HS}}(Q_b^{1/2}(H_b); H)} \mathrm{d}s\right)^{1/2},\end{aligned}$$

resulting from the Itô isometry and a change of variables. Lemmas 4.9(iv) and 4.10(iv), both with $\alpha \in [1/2, 3, 4)$, give us for sufficiently small $\varepsilon > 0$ that

$$\mathcal{I}_5 \leqslant c\left(h^{1/2-\varepsilon} + \Delta t^{1/4-\varepsilon}\right) \left\|(\lambda - A)^{\alpha} N\right\|_{\mathcal{L}_{\mathrm{HS}}(Q_b^{1/2}(H_b);H)} \leqslant c\left(h^{1/2-\varepsilon} + \Delta t^{1/4-\varepsilon}\right). \tag{5.27}$$

Substituting inequalities (5.2), (5.14), (5.25), (5.26) and (5.27) into (5.1) yields, for sufficiently small $\varepsilon > 0$,

$$\begin{aligned}\|y(t_m) - y_h^m\|_{L^2(\Omega;H)} \leqslant{}& c\left(h^{\min\{1/2-\varepsilon,\beta\}} + \Delta t^{\min\{1/4-\varepsilon,\beta/2\}}\right)\left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right)\\ &+ c\,\Delta t \sum_{k=0}^{m-1} \left\|y(t_k) - y_h^k\right\|_{L^2(\Omega;H)}.\end{aligned}$$

By applying a discrete version of the Grönwall inequality (see Clark, 1987) we therefore get

$$\|y(t_m) - y_h^m\|_{L^2(\Omega;H)} \leqslant c\left(h^{\min\{1/2-\varepsilon,\beta\}} + \Delta t^{\min\{1/4-\varepsilon,\beta/2\}}\right)\left(1 + \|\xi\|_{L^2(\Omega;D((\lambda-A)^{\beta/2}))}\right)$$

for sufficiently small $\varepsilon > 0$. □

6. Numerical experiments

In order to illustrate the proposed method and the bounds given in Theorem 3.3 we have implemented the algorithm in MATLAB[1] and performed a number of numerical experiments on a two-dimensional linear quadratic control problem with noise. We ran all the experiments on one node of the Mechthild computing cluster at the Max Planck Institute Magdeburg. Such a node consists of two Intel Xeon Skylake Silver 4110 processors with 8 cores/CPU, a clockrate of 2.1 GHz and 384 GB RAM.

6.1 *Implementation*

Let $\{\phi_k^h\}_{k=1}^{N_h}$ be the standard finite element basis of Y_h, consisting of the piecewise linear so-called hat functions. These take the value 1 at the kth node of $\mathcal{T}_h$ and 0 at all other nodes. Then for $y_h \in Y_h$ we have $y_h = \sum_{k=1}^{N_h} \boldsymbol{y}_k \phi_k^h$ for some coefficients $\{\boldsymbol{y}_k\}_{k=1}^{N_h}$. Similarly, let the distributed noise $P_h G\,\delta W^m$ with $G = I$ and the boundary noise $B_h^b\,\delta W_b^m$ be represented by the coefficient vectors $\delta \boldsymbol{W}^m$ and $\delta \boldsymbol{W}_b^m$, respectively. Using these representations in (3.6) and testing with ϕ_j^h shows that (3.6) is equivalent to

$$\begin{aligned}\sum_{k=1}^{N_h} \boldsymbol{y}_k^m \langle (I - \Delta t A_h)\phi_k^h, \phi_j^h\rangle = {}& \sum_{k=1}^{N_h} \boldsymbol{y}_k^{m-1}\Big(\langle \phi_k^h, \phi_j^h\rangle - \Delta t \langle B_h R^{-1} B_h^* \mathcal{P}_h^{m-1} \phi_k^h, \phi_j^h\rangle\\ &- \Delta t \langle B_h^b R_b^{-1} (B_h^b)^* \mathcal{P}_h^{m-1} \phi_k^h, \phi_j^h\rangle\Big)\\ &+ \sum_{k=1}^{N_h} \left((\delta \boldsymbol{W}^m)_k + (\delta \boldsymbol{W}_b^m)_k\right)\langle \phi_k^h, \phi_j^h\rangle,\end{aligned} \tag{6.1}$$

[1] Full code available at www.tonystillfjord.net.

for $j,k = 1,\dots,N_h$. To simplify this we introduce the mass matrix $\boldsymbol{M}$, the stiffness matrix $\boldsymbol{A}$, the distributed and boundary input matrices $\boldsymbol{B}$ and $\boldsymbol{B}^b$, the output matrix $\boldsymbol{C}$ and the weighting matrices $\boldsymbol{R}$, $\boldsymbol{R}^b$ and $\boldsymbol{Q}$, satisfying

$$\begin{aligned}
&\boldsymbol{M}_{i,j} = \langle \phi_j^h, \phi_i^h \rangle, && \boldsymbol{A}_{i,j} = \langle A_h \phi_j^h, \phi_i^h \rangle, && \\
&\boldsymbol{B}_{i,j} = \langle B_h \phi_j^h, \phi_i^U \rangle, && \boldsymbol{B}^b_{i,j} = \langle (\lambda - A_h) R_h N \phi_j^h, \phi_i^V \rangle, && \boldsymbol{C}_{i,j} = \langle C_h \phi_j^h, \phi_i^Z \rangle, \\
&\boldsymbol{R}_{i,j} = \langle \phi_j^U, \phi_i^U \rangle, && \boldsymbol{R}^b_{i,j} = \langle \phi_j^V, \phi_i^V \rangle, && \boldsymbol{Q}_{i,j} = \langle \phi_j^Z, \phi_i^Z \rangle.
\end{aligned}$$

Here, $\{\phi_i^U\}$, $\{\phi_i^V\}$ and $\{\phi_i^Z\}$ denote orthonormal bases for the input and output spaces $\bar{U}$, $\bar{V}$ and Z, respectively. We omit the dependency on h to reduce notational clutter.

The matrices given above were all generated using the FreeFEM++ library[2] (see Hecht, 2012) and then imported into MATLAB. With these at hand we can first rewrite (3.5) as the matrix-valued equation

$$\begin{aligned}
\boldsymbol{M}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{P}(t)\boldsymbol{M} &= -\boldsymbol{A}\boldsymbol{P}(t)\boldsymbol{M} - \boldsymbol{M}\boldsymbol{P}(t)\boldsymbol{A} + \boldsymbol{M}\boldsymbol{P}(t)\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}(t)\boldsymbol{M} \\
&\quad + \boldsymbol{M}\boldsymbol{P}(t)\boldsymbol{B}^b(\boldsymbol{R}^b)^{-1}(\boldsymbol{B}^b)^{\mathrm{T}}\boldsymbol{P}(t)\boldsymbol{M} - \boldsymbol{C}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{C}, \\
\boldsymbol{P}(T) &= 0,
\end{aligned} \tag{6.2}$$

where $\boldsymbol{P}$ denotes the matrix representation of $\mathcal{P}_h$ satisfying

$$\mathcal{P}_h z = \sum_{i,j=1}^{N_h} \boldsymbol{P}_{i,j} \langle z, \phi_j^h \rangle \phi_i^h;$$

see, e.g., Målqvist *et al.* (2018). Further, denote the coefficients at time t_m by $\boldsymbol{P}^m$. Then we can rewrite (6.1) as an equation for the coefficients $\boldsymbol{y}_k^m$ as

$$\begin{aligned}
\boldsymbol{y}^{m+1} = \boldsymbol{S}_{h,\Delta t}\boldsymbol{M}\boldsymbol{y}^m &- \Delta t \boldsymbol{S}_{h,\Delta t}\boldsymbol{B}\boldsymbol{R}^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{P}^m\boldsymbol{M}\boldsymbol{y}^m \\
&- \Delta t \boldsymbol{S}_{h,\Delta t}\boldsymbol{B}^b(\boldsymbol{R}^b)^{-1}(\boldsymbol{B}^b)^{\mathrm{T}}\boldsymbol{P}^m\boldsymbol{M}\boldsymbol{y}^m \\
&+ \boldsymbol{S}_{h,\Delta t}\boldsymbol{M}\big(\delta\boldsymbol{W}^m + \delta\boldsymbol{W}_b^m\big),
\end{aligned}$$

where $\boldsymbol{S}_{h,\Delta t} = (\boldsymbol{M} - \Delta t\boldsymbol{A})^{-1}$. Note the similarity to (3.6), with $\boldsymbol{M}$ taking on the role of the identity operator.

Since (6.2) is matrix-valued, numerically approximating its solution for reasonably fine spatial discretizations is unfeasible unless we can utilize features such as low-rank structure. For this reason we assume that the input operators are of the form $\mathbb{R}^{n_u} \ni u \mapsto u_1\Psi_1 + \cdots + u_{n_u}\Psi_{n_u}$ with $\Psi_j \in H$ and $n_u \ll N_h$. Similarly, we assume that the output operator $C\colon H \mapsto \mathbb{R}^{n_z}$ with $n_z \ll N_h$. This means

[2] Available at https://freefem.org/.

that $\boldsymbol{B}$, $\boldsymbol{B}^b$ and $\boldsymbol{C}$ are rectangular matrices, which typically leads to a solution $\boldsymbol{P}$ of low numerical rank. See, e.g., Stillfjord (2018b) for supporting theory in the operator-valued setting. In order to approximate the solution we apply the Strang splitting scheme (Stillfjord, 2018a) available in the MATLAB package DREsplit.[3] We note that Strang splitting is a second-order method, which means that we get a more accurate approximation in time than what we need according to Assumption 3.1. It is, however, essentially as cheap to apply as the corresponding first-order scheme, which is why we use it.

Generating the noise can be done in many ways. Since we only consider rectangular domains in our experiments we compute samples of the distributed noise using FFT techniques as outlined in Lord *et al.* (2014, Chapter 10). In particular we assume that the eigenvalues $\lambda_{j,k}$ and corresponding eigenvectors $\varphi_{j,k}$ of the covariance operator Q are given by

$$\lambda_{j,k} = (j^2 + k^2)^{-\beta-\epsilon} \quad \text{and} \quad \varphi_{j,k}(x_1, x_2) = \cos(j\pi x_1)\cos(k\pi x_2)$$

with $\beta = 1$ and $\epsilon = 10^{-4}$. Then the increments $\delta W^m = W(t_m) - W(t_{m-1})$ are given by

$$\delta W^m \approx \sqrt{\Delta t} \sum_{j,k=0}^{N} \sqrt{\lambda_{j,k}} \varphi_{j,k} \xi_j^m,$$

where ξ_j^n are the i.i.d. increments of an N(0,1) Gaussian distribution (Lord *et al.*, 2014). This leads to noise satisfying Assumption 2.3. We note that the sum should actually go to infinity, and the truncation to $(N+1)^2$ terms represents a discretization. We use $N = N_h$ in our experiments, which means that the truncation does not affect the convergence order (Yan, 2005).

A similar procedure could conceivably be followed for the boundary noise. However, we found it simpler to express the one-dimensional noise $\delta W_{b,m}$ on each of the edges as $\sqrt{\Delta t}\sum_{k=0}^{N} \lambda_k \cos(k\pi x)$ with $x \in [0,1]$ and $\lambda_k = k^{-\beta-\epsilon}$. Then the map N can be explicitly constructed by using the observation that the function

$$\rho(x_1, x_2) = -\frac{\cos(k\pi x_1)\cosh(c(1-x_2))}{c\sinh(c)} \quad \text{with} \quad c = \sqrt{\lambda + k^2\pi^2}$$

satisfies $\frac{\mathrm{d}}{\mathrm{d}x_1}\rho = 0$ at $x_1 = 0$ and $x_1 = 1$, $\frac{\mathrm{d}}{\mathrm{d}x_2}\rho = 0$ at $x_2 = 1$ and $\frac{\mathrm{d}}{\mathrm{d}x_2}\rho = \cos(k\pi x_1)$ at $x_2 = 0$. Further, it satisfies $\lambda\rho = \Delta\rho$ in the interior of the domain. The constructions for the other parts of the boundary are similar. Summing the four parts then gives the solution of (2.1). We then computed the Ritz projections of these functions in FreeFem++ by solving $\langle R_h\rho, \phi\rangle_Y = \langle \rho, \phi\rangle_Y$ for $\phi \in Y_h$. Finally, the resulting coefficient vectors were multiplied by $\lambda\boldsymbol{M} - \boldsymbol{A}$.

The latter construction was also used for the boundary input operator, by computing N applied to the constant function 1 on the boundary. This requires no further calculations, since it corresponds to the first eigenvector.

[3] Available at www.tonystillfjord.net.

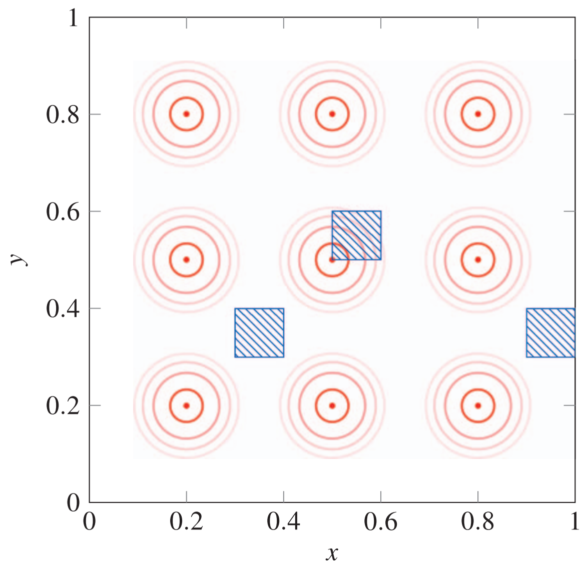

FIG. 1. Locations of distributed inputs (red) and outputs (blue, shaded) in the test problem. The red lines indicate intensities of 0.8, 0.4, 0.2 and 0.1, respectively.

6.2 *Test problem*

For simplicity we consider the problem on the unit square $\mathcal{D} = [0,1]^2$. We let the distributed control operator $B : \mathbb{R}^{n_u} \mapsto L^2(\mathcal{D})$ be defined by

$$Bu = u_1 \Psi_{p^1} + \cdots + u_{n_u} \Psi_{p^{n_u}},$$

where $p^j = (p_1^j, p_2^j)$ are points in the plane and $\Psi_{p^j}(x_1, x_2) = e^{-200(x_1 - p_1^j)^2 - 200(x_2 - p_2^j)^2}$. The interpretation of this is that we have heat sources with high intensity at p^j that tapers off exponentially as we move away radially from p^j. The locations of these points are illustrated in Fig. 1. We note that $B \in \mathcal{L}(\mathbb{R}^{n_u}, L^2(\mathcal{D}))$. For this example we picked $n_u = 9$. For the boundary control we consider a single boundary condition $\frac{\partial}{\partial \nu} y(t,x) = v$ with $v \in \mathbb{R}$.

As output we take the operator

$$Cy = 10^2 \int_{\mathcal{D}} y(x)\chi_{T_1}(x) + \cdots + y(x)\chi_{T_{n_z}}(x)\,\mathrm{d}x,$$

where χ_S denotes the characteristic function of the set S and the T_j denote different areas, illustrated in Fig. 1. Thus we attempt to control the mean value of the solution in these areas. We note that $C \in \mathcal{L}(H, \mathbb{R}^{n_z})$. Here $n_z = 3$.

Finally, we use a diffusion coefficient of 10^{-2}, $\lambda = 1$, and the scaling factors $R = 10^{-2}$ and $R_b = 25$. The latter was chosen such that the distributed and boundary controls influence the solution to a similar extent.

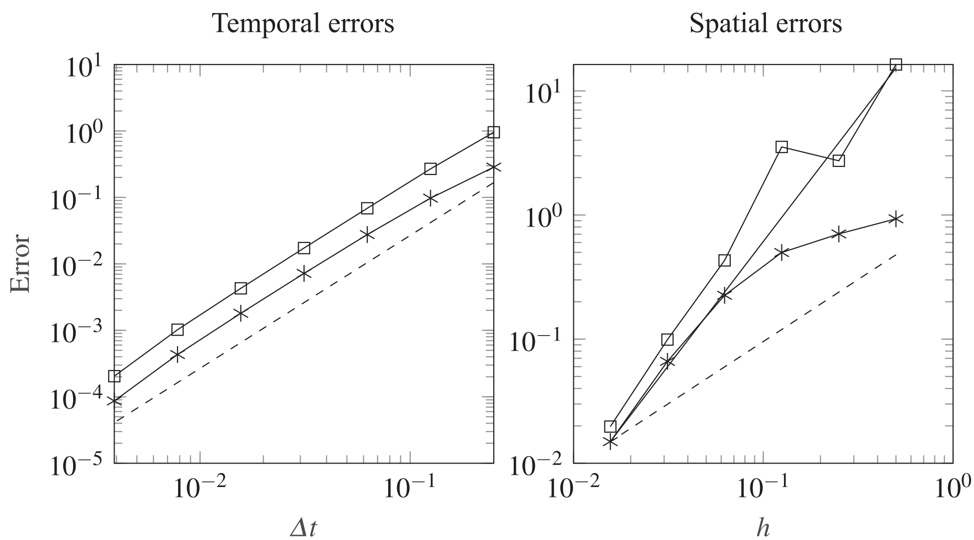

FIG. 2. The errors $\|P_h^{\mathrm{ref}}(0) - P_h^0\|_{L(H)}$ (—✳—) and $\|(B_h^b)^* P_h^{\mathrm{ref}}(0) - (B_h^b)^* P_h^0\|_{L(H)}$ (—⊟—) for the various discretizations outlined in Section 6.3. Reference lines: $\mathcal{O}(\Delta t^2)$ (- - -), $\mathcal{O}(h)$ (- - -), $\mathcal{O}(h^2)$ (——).

6.3 *Results*

We first verify Assumption 3.1 by computing the errors

$$\|P_h^{\mathrm{ref}}(0) - P_h^0\|_{L(H)} \quad \text{and} \quad \|(B_h^b)^* P_h^{\mathrm{ref}}(0) - (B_h^b)^* P_h^0\|_{L(H)}$$

for different choices of h and Δt. We first choose $h = 2^{-6}$ and $\Delta t = 2^{j+2}$, $j = 1, \ldots, 7$, with the reference solution P_h^{ref} having the same h and $\Delta t = 2^{10}$. The result is shown in Fig. 2 (left) and shows clear second-order temporal convergence, as expected. We then choose $\Delta t = 2^9$ and take $h = 2^j$, $j = 1, \ldots, 6$, with the reference solution P_h^{ref} having the same Δt and $h = 2^7$. The result is shown in Fig. 2 (right) and also demonstrates second-order spatial convergence except for the first few coarse discretizations.

Next we check Theorem 3.3. By choosing $h = \Delta t^2$ the expected error is $\mathcal{O}\big(\Delta t^{1/4}\big)$ and there is only one parameter to adjust. We therefore choose $h = 2^{-2j}$ and $\Delta t = 2^{-j}$ for $j = 1, \ldots, 6$ and compute a reference solution with $j = 7$. We start by computing the noise for the finest discretization first. Then for each coarser discretization we add up the temporal increments and compute the L^2-projection onto the coarser space. In this way we use the same noise for all the discretizations of each of the 100 sample paths. The resulting errors measured at $t = T$ are shown in Fig. 3, both for the controlled system and for the corresponding uncontrolled system where $b = v = 0$. We can observe that they decrease with a rate that is decidedly less than 1/2 and close to 1/4. Since our theoretical bound is for the worst-case situation this is fully in line with Theorem 3.3.

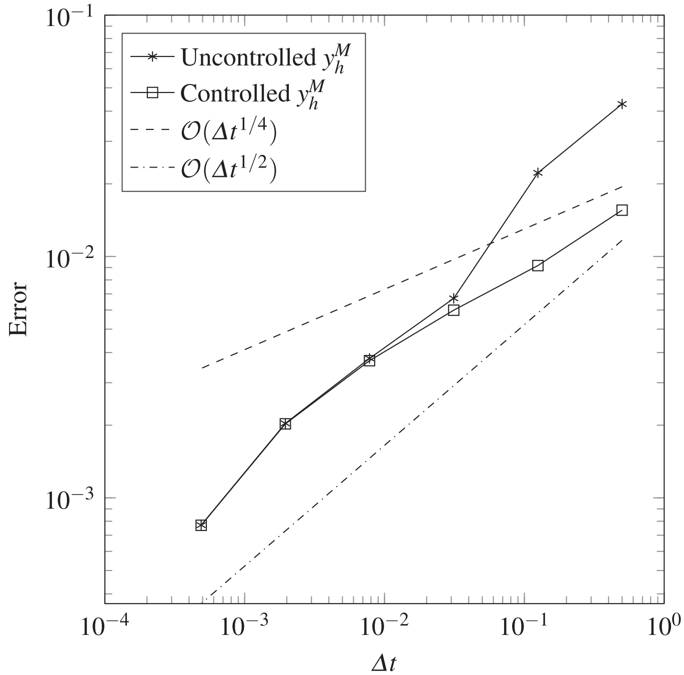

FIG. 3. The computed errors for the experiment outlined in Section 6.3. They are in line with the $\mathcal{O}(\Delta t^{1/4})$-prediction of Theorem 3.3.

7. Conclusions

We have proved convergence with optimal orders of a numerical scheme for an optimal control problem with both distributed and boundary control, as well as distributed and boundary Q-Wiener noise. Due to the irregularity of the noise we can expect at most order $1/4$ in time and order $1/2$ in space. A numerical experiment confirms that this bound is optimal.

Acknowledgements

Parts of this work were completed while the second and third authors were with the Max Planck Institute Magdeburg.

Funding

The work of T.S. was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

REFERENCES

AHMED, N. U. (1981) Stochastic control on Hilbert space for linear evolution equations with random operator-valued coefficients. *SIAM J. Control Optim.*, **19**, 401–430.

BENNER, P. & MENA, H. (2018) Numerical solution of the infinite-dimensional LQR problem and the associated Riccati differential equations. *J. Numer. Math.*, **1**, 1–20.

Benner, P. & Trautwein, C. (2018) Optimal control of a stochastic heat equation driven by Q-Wiener processes. *J. Math. Anal. Appl.*, **457**, 776–802.

Bensoussan, A., Da Prato, G., Delfour, M. C. & Mitter, S. K. (2007). *Representation and Control of Infinite Dimensional Systems*. Boston: Birkhäuser.

Clark, D. S. (1987) Short proof of a discrete Gronwall inequality. *Discrete Appl. Math.*, **16**, 279–281.

Curtain, R. F. & Pritchard, A. J. (1978) *Infinite Dimensional Linear Systems Theory*. Berlin, Heidelberg: Springer-Verlag.

Debussche, A., Fuhrman, M. & Tessitore, G. (2007) Optimal control of a stochastic heat equation with boundary-noise and boundary-control. *ESAIM Control Optim. Calc. Var.*, **13**, 178–205.

Duncan, T. E., Maslowski, B. & Pasik-Duncan, B. (2012) Linear-quadratic control for stochastic equations in a Hilbert space with fractional Brownian motions. *SIAM J. Control Optim.*, **50**, 507–531.

Fabbri, G. & Goldys, B. (2009) An LQ problem for the heat equation on the halfline with Dirichlet boundary control and noise. *SIAM J. Control Optim.*, **48**, 1473–1488.

Fabbri, G., Gozzi, F. & Swiech, A. (2017) *Stochastic Optimal Control in Infinite Dimension: Dynamic Programming and HJB Equations*. Cham: Springer International Publishing.

Fujita, H. & Mizutani, A. (1976) On the finite element method for parabolic equations, I; approximation of holomorphic semigroups. *J. Math. Soc. Japan*, **28**, 749–771.

Fujiwara, D. (1967) Concrete characterization of the domains of fractional powers of some elliptic differential operators of the second order. *Proc. Japan Acad.*, **43**, 82–86.

Grisvard, P. (1985) *Elliptic Problems in Nonsmooth Domains*. Boston: Pitman Advanced Pub. Program.

Guatteri, G. & Masiero, F. (2011) On the existence of optimal controls for SPDEs with boundary-noise and boundary-control. *SIAM J. Control Optim.*, **51**, 1909–1939.

Hecht, F. (2012) New development in freefem++. *J. Numer. Math.*, **20**, 251–265.

Hu, Y. and Tang, S. (2018). Stochastic LQ and associated Riccati equation of PDEs driven by state-and control-dependent white noise. arXiv:1809.05308.

Kruse, R. (2014) Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise. *IMA J. Numer. Anal.*, **34**, 217–251.

Lasiecka, I. (1980) Unified theory for abstract parabolic boundary problems: a semigroup approach. *Appl. Math. Optim.*, **6**, 287–333.

Lasiecka, I. (1986) Galerkin approximations of abstract parabolic boundary value problems with rough boundary data–L_p theory. *Math. Comp.*, **47**, 55–75.

Lasiecka, I. and Triggiani, R. (2000). *Control Theory for Partial Differential Equations: Volume 1, Abstract Parabolic Systems: Continuous and Approximation Theories*. Cambridge: Cambridge University Press.

Lions, J. L. and Magenes, E. (1972). *Non-Homogeneous Boundary Value Problems and Applications*, vol. 1. Berlin: Springer.

Lord, G. J., Powell, C. E. & Shardlow, T. (2014) *An Introduction to Computational Stochastic PDEs*. Cambridge Texts in Applied Mathematics. New York: Cambridge University Press.

Lord, G. J. & Tambue, A. (2013) Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise. *IMA J. Numer. Anal.*, **33**, 515–543.

Målqvist, A., Persson, A. & Stillfjord, T. (2018) Multiscale differential Riccati equations for linear quadratic regulator problems. *SIAM J. Sci. Comput.*, **56**, 3598–3618.

Pazy, A. (1983) *Semigroups of Linear Operators and Applications to Partial Differential Equations*. New York: Springer.

Stillfjord, T. (2018a) Adaptive high-order splitting schemes for large-scale differential Riccati equations. *Numer. Algorithms*, **78**, 1129–1151.

Stillfjord, T. (2018b) Singular value decay of operator-valued differential Lyapunov and Riccati equations. *SIAM J. Control Optim.*, **56**, 3598–3618.

Tambue, A. & Mukam, J. D. (2019) Strong convergence of the linear implicit Euler method for the finite element discretization of semilinear SPDEs driven by multiplicative or additive noise. *Appl. Math. Comput.*, **346**, 23–40.

Thomeé, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*. Berlin: Springer.

Vrabie, I. (2003) *C_0-Semigroups and Applications*. Mathematics Studies no. 191. Amsterdam: Elsevier.

Wang, X. (2017) Strong convergence rates of the linear implicit Euler method for the finite element discretization of SPDEs with additive noise. *IMA J. Numer. Anal.*, **37**, 965–984.

Yan, Y. (2005) Galerkin finite element methods for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal.*, **43**, 1363–1384.

Ye, H., Gao, J. & Ding, Y. (2007) A generalized Gronwall inequality and its application to a fractional differential equation. *J. Math. Anal. Appl.*, **328**, 1075–1081.

Yu, H. & Liu, H. (2011) Properties of value function and existence of viscosity solution of HJB equation for stochastic boundary control problems. *J. Franklin Inst. B*, **348**, 2108–2127.