Approximation of SPDE covariance operators by finite elements: a semigroup approach

(article starts on next page)

Approximation of SPDE covariance operators by finite elements: a semigroup approach

MIHÁLY KOVÁCS

Department of Differential Equations, Faculty of Natural Sciences, Budapest University of Technology and Economics, Muőegyetem rkp. 3., H-1111 Budapest, Hungary; Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, P.O. Box 278, H-1444 Budapest, Hungary; and Department of Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, S–412 96 Göteborg, Sweden
kovacs.mihaly@itk.ppke.hu

ANNIKA LANG

Department of Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, S–412 96 Göteborg, Sweden
annika.lang@chalmers.se

AND

ANDREAS PETERSSON*

Department of Mathematics, The Faculty of Mathematics and Natural Sciences, University of Oslo, Postboks 1053, Blindern, 0316 Oslo, Norway
*Corresponding author: andreep@math.uio.no

The problem of approximating the covariance operator of the mild solution to a linear stochastic partial differential equation is considered. An integral equation involving the semigroup of the mild solution is derived and a general error decomposition is proven. This formula is applied to approximations of the covariance operator of a stochastic advection-diffusion equation and a stochastic wave equation, both on bounded domains. The approximations are based on finite element discretizations in space and rational approximations of the exponential function in time. Convergence rates are derived in the trace class and Hilbert–Schmidt norms with numerical simulations illustrating the results.

Keywords: stochastic partial differential equations; integral equations; covariance operators; finite element method; stochastic advection-diffusion equations; stochastic wave equations.

1. Introduction

This paper considers stochastic partial differential equations (SPDEs) formulated as linear stochastic evolution equations on a Hilbert space H. That is to say, equations of the form

$$
\begin{aligned}
\mathrm{d}X(t) + AX(t)\,\mathrm{d}t &= FX(t)\,\mathrm{d}t + B\,\mathrm{d}W(s) \text{ for } t \in (0,T], T < \infty, \\
X(0) &= \xi.
\end{aligned}
\tag{1.1}
$$

Here, X is an H-valued stochastic process, F and B are linear operators, W is a Wiener process in H with covariance operator Q and $-A$ is the generator of a C_0-semigroup $S = (S(t))_{t\in[0,T]}$ of linear operators on H. Since (1.1) is linear and the noise term is additive and Gaussian, the solution $X(t)$ to (1.1) at time $t \in (0,T]$ is an H-valued Gaussian random variable when the initial value ξ is Gaussian. The distribution of $X(t)$ is therefore completely determined by its mean value $\mathbb{E}[X(t)]$ and covariance (operator) $\mathrm{Cov}(X(t)) = \mathbb{E}[(X(t) - \mathbb{E}[X(t)]) \otimes (X(t) - \mathbb{E}[X(t)])]$. Computing these quantities is therefore vital for understanding of $X(t)$. In general, there are no analytic solutions, so numerical approximations are needed. In this paper, we focus on the approximation of $\mathrm{Cov}(X(t))$.

The literature on the numerical analysis of approximations of SPDE covariance operators is sparse. We are only aware of Lang *et al.* (2013), Kirchner *et al.* (2017) and Kirchner (2020). Therein, the authors consider SPDEs of parabolic type and solve a tensorized equation related to the concept of a weak solution to (1.1). They assume the operator A to be self-adjoint. In this paper, we take a different approach. We work with the mild solution of an SPDE to derive an operator-valued integral equation for the covariance, expressed in terms of the semigroup S.

This approach allows us to treat parabolic SPDEs where A is not self-adjoint. One example of such an equation comes from the modeling of sea surface temperature dynamics; see Herterich & Hasselmann (1987). The equation, posed in the space $H = L^2(\mathscr{D})$ of square integrable functions on some domain $\mathscr{D} \subset \mathbb{R}^2$, is given by

$$\mathrm{d}X(t,x) + \mathscr{A}X(t,x)\,\mathrm{d}t = \mathrm{d}W(t,x) \text{ for } t \in (0,T], x \in \mathscr{D}. \tag{1.2}$$

Here, $AX(t)$ corresponds to $\mathscr{A}X(t,\cdot) = \delta X(t,\cdot) - D\Delta X(t,\cdot) - \mathbf{a} \cdot \nabla X(t,\cdot)$, where $X(t,x)$ is the sea surface temperature at time t and point $x \in \mathscr{D}$, $\mathbf{a} \in \mathbb{R}^2$ is the velocity vector field of the upper ocean layer, $D > 0$ is a diffusion coefficient and $\delta \in \mathbb{R}$ is a feedback parameter. When $\mathbf{a} \neq 0$, the operator A ceases to be self-adjoint. The Wiener process W models small time scale fluctuations in the heat flux across the ocean-atmosphere interface. Its covariance operator Q is an integral operator with a kernel q having small correlation length. Similar models have recently been considered for reconstructing the evolution of cloud systems from discrete measurements; see Mider *et al.* (2017).

Moreover, our approach extends the parabolic setting to any SPDE that has a mild solution in terms of a semigroup. This includes hyperbolic SPDEs, such as a stochastic equation for the vertical displacement $U(t,x)$ of a strand of DNA suspended in a liquid at time t and space $x \in \mathscr{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$ from Dalang (2009). It is given by

$$\mathrm{d}\dot{U}(t,x) - \Delta U(t,x)\,\mathrm{d}t = -(Q U)(t,x)\,\mathrm{d}t + \mathrm{d}W(t,x) \text{ for } t \in (0,T], x \in \mathscr{D}. \tag{1.3}$$

The first term on the right-hand side model's friction due to viscosity of the fluid, while the Wiener process term W corresponds to random bombardment of the DNA strand by the fluid's molecules. Writing $X = [U, \dot{U}]^\top$, the equation can be put in the form of (1.1) by considering it on a product space; see Section 3.2. Fig. 1 shows realizations of the solutions to (1.2) and (1.3) for the domain $\mathscr{D} = (0,1)$; see Examples 3.2 and 3.5.

Let us now outline the content of the paper. In Section 2, we formulate an operator-valued integral equation for the covariance of the mild solution X to (1.1) in an abstract Hilbert space framework and give assumptions that ensure that it has a unique solution. We confirm that X is Gaussian and that the process $[0,T] \ni t \mapsto \mathrm{Cov}(X(t))$ is a solution to the integral equation. The mild Itô formula in Da Prato *et al.* (2019) is key for this. We finish the section by giving an abstract error decomposition formula for approximations of this process. The error is, for $t \in [0,T]$, analyzed with respect to the norms $\|\cdot\|_{\mathscr{L}_1(H)}$
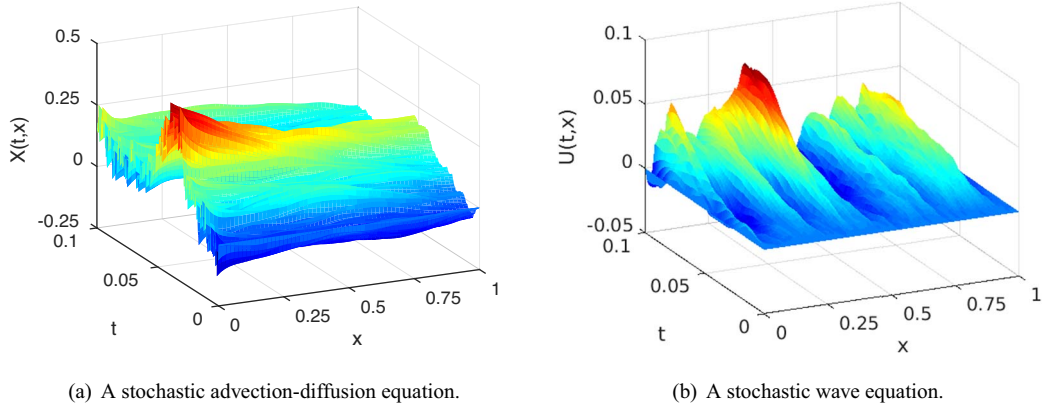
(a) A stochastic advection-diffusion equation.

(b) A stochastic wave equation.

Fig. 1. Realizations of the solutions X and U to (1.2) and (1.3) when $\mathscr{D} = (0, 1)$.

and $\|\cdot\|_{\mathscr{L}_2(H)}$. Here, $\mathscr{L}_1(H)$ and $\mathscr{L}_2(H)$ denote the spaces of trace class and Hilbert–Schmidt operators, respectively. The first norm is a natural choice since, if $(Q_j)_{j=1}^{\infty}$ is a sequence of covariances of some Gaussian H-valued random variables $(X_j)_{j=1}^{\infty}$ with zero mean, then $Q_j \to \text{Cov}(X(t))$ in $\mathscr{L}_1(H)$ if and only if $X_j \to X(t)$ weakly, i.e., $\mathbb{E}[f(X_j)] \to \mathbb{E}[f(X(t))]$ for all continuous and bounded functionals f on H; see Chevet (1983). The norm of $\mathscr{L}_2(H)$ is weaker. It has a natural meaning when $H = L^2(\mathscr{D})$: if $X(t) = X(t, \cdot)$ is $\mathscr{F} \times \mathscr{B}(\mathscr{D})$-measurable,
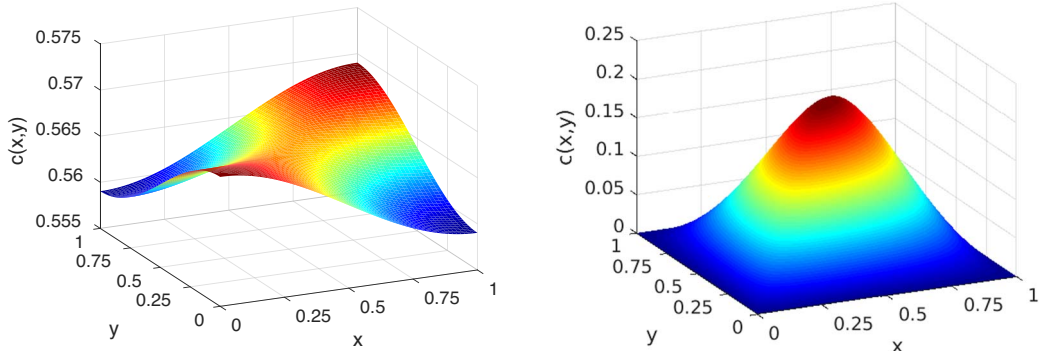
$$\langle \text{Cov}(X(t))u, v \rangle_H = \int_{\mathscr{D} \times \mathscr{D}} \text{Cov}(X(t,x), X(t,y))u(x)v(y)\,\mathrm{d}x\,\mathrm{d}y$$

and $\|\text{Cov}(X(t))\|_{\mathscr{L}_2(H)}^2 = \|\text{Cov}(X(t,\cdot), X(t,\cdot))\|_{L^2(\mathscr{D} \times \mathscr{D})}^2$. Therefore, we may, formally at least, view convergence in $\mathscr{L}_2(H)$ as convergence in $L^2(\mathscr{D} \times \mathscr{D})$ of underlying covariance functions on $\mathscr{D}$. Fig. 2 shows the covariance functions for the solutions to (1.2) and (1.3) at $T = 0.1$.

In Section 3, we apply our abstract framework to the two concrete equations (1.2) and (1.3). In both cases, the covariance integral equations are discretized by finite elements in space and rational approximations of the driving semigroup in time. The resulting approximations are expressed as integral equations based on a fully discrete approximation $\tilde{S}$ of S.

Consider a discrete approximation $\tilde{X}(t)$ of $X(t)$, $t \in [0, T]$. Suppose, without loss of generality, that $\mathbb{E}[X(t)] = \mathbb{E}[\tilde{X}(t)] = 0$. Then, by properties of the norms $\|\cdot\|_{\mathscr{L}_i(H)}$, $i \in \{1, 2\}$, and the Hölder inequality,

$$\|\text{Cov}(X(t)) - \text{Cov}(\tilde{X}(t))\|_{\mathscr{L}_i(H)}$$

$$= \|\mathbb{E}[X(t) \otimes X(t)] - \mathbb{E}[\tilde{X}(t) \otimes \tilde{X}(t)]\|_{\mathscr{L}_i(H)}$$

$$= \frac{1}{2} \left\| \mathbb{E}\left[(X(t) + \tilde{X}(t)) \otimes (X(t) - \tilde{X}(t)) \right] + \mathbb{E}\left[(X(t) - \tilde{X}(t)) \otimes (X(t) + \tilde{X}(t)) \right] \right\|_{\mathscr{L}_i(H)}$$

$$\leqslant \max\left(\mathbb{E}[\|X(t)\|_H^2]^{\frac{1}{2}}, \mathbb{E}[\|\tilde{X}(t)\|_H^2]^{\frac{1}{2}} \right) \mathbb{E}\left[\|X(t) - \tilde{X}(t)\|_H^2 \right]^{\frac{1}{2}}. \tag{1.4}$$

(a) Covariance function for a stochastic advection-diffusion equation.

(b) Covariance function for a stochastic wave equation.

Fig. 2. Plot of the covariance functions $c(x, y) = \text{Cov}(X(0.1, x), X(0.1, y))$ and $c(x, y) = \text{Cov}(U(0.1, x), U(0.1, y))$, $x, y \in \mathscr{D} = (0, 1)$, for the solutions X and U to (1.2) and (1.3).

Therefore, if $\text{Cov}(\tilde{X}(t))$ can be calculated, we obtain an approximation scheme for $\text{Cov}(X(t))$ for which the error can be bounded by the strong error $\mathbb{E}\left[\|X(t) - \tilde{X}(t)\|_H^2\right]^{1/2}$. In Section 3.1, where a stochastic advection-diffusion equation is considered, we demonstrate that this bound is suboptimal. More precisely, we compare the convergence rate for a covariance approximation based on replacing S by a fully discrete approximation $\tilde{S}$ in the operator-valued integral equation with the rate for the strong error with respect to an approximation $\tilde{X}$ of X based on the same discretization $\tilde{S}$. It turns out that the convergence rate for our approximation is typically higher than the strong error; see Remark 3.1. We also demonstrate, in theory and by a numerical simulation, that there are cases when the $\mathscr{L}_2$ error decays faster than the stronger $\mathscr{L}_1$ error.

In Section 3.2, we work with the stochastic wave equation and provide, to the best of our knowledge, the first results on convergence rates for the approximation of covariance operators of hyperbolic SPDEs. We first consider a temporally semidiscrete approximation, which is then used for analyzing a fully discrete approximation. Numerical simulations finish the section.

Throughout the paper, we adopt the notion of generic constants, which may vary from occurrence to occurrence and are independent of any parameter of interest, such as spatial or temporal step sizes. By $a \lesssim b$, we denote the existence of a generic constant such that $a \leqslant Cb$.

2. Covariance operators of stochastic evolution equations

In this section, we prove existence and uniqueness of the solution to an operator-valued integral equation. We then show that the covariance of the mild solution of an SPDE is a solution of this equation. First, however, we introduce our setting and reiterate some facts from operator theory and probability theory in Hilbert spaces.

2.1 *Operator theory*

Let $(H, \langle \cdot, \cdot \rangle_H)$ and $(U, \langle \cdot, \cdot \rangle_U)$ be real and separable Hilbert spaces. We denote by $\mathscr{L}(H, U)$ the space of bounded linear operators from H to U equipped with the usual operator norm, and by $\mathscr{L}_1(H, U)$ and

$\mathcal{L}_2(H, U)$ the spaces of trace class and Hilbert–Schmidt operators, respectively. We use the shorthand notations $\mathcal{L}(H)$, $\mathcal{L}_1(H)$ and $\mathcal{L}_2(H)$ when $U = H$. Additionally, we denote by $\Sigma(H) \subset \mathcal{L}(H)$ the set of symmetric bounded operators on H. An operator $\Gamma \in \mathcal{L}_1(H, U)$ if and only if there are two orthonormal sequences $(e_j)_{j=1}^{\infty} \subset H$, $(f_j)_{j=1}^{\infty} \subset U$ and a sequence $(\mu_j)_{j=1}^{\infty} \in \ell^1$ such that

$$\Gamma x = \sum_{j=1}^{\infty} \mu_j \langle x, e_j \rangle f_j \text{ for } x \in H. \tag{2.1}$$

The space $\mathcal{L}_1(H, U)$ is a separable Banach space with norm

$$\|\Gamma\|_{\mathcal{L}_1(H,U)} = \inf_{\substack{(a_j) \subset H \\ (b_j) \subset U}} \left\{ \sum_{j=1}^{\infty} \|a_j\|_H \|b_j\|_U : \Gamma = \sum_{j=1}^{\infty} \langle \cdot, a_j \rangle_H b_j \right\}, \tag{2.2}$$

where the infimum is taken over all sequences $(a_j)_{j=1}^{\infty} \subset H, (b_j)_{j=1}^{\infty} \subset U$, see Trèves (1967, Sections 47 and 48). Moreover, $\mathcal{L}_2(H, U)$ is a separable Hilbert space with an inner product, for an arbitrary orthonormal basis $(e_j)_{j=1}^{\infty}$ of H, given by

$$\langle \Gamma_1 \Gamma_2 \rangle_{\mathcal{L}_2(H,U)} = \sum_{j=1}^{\infty} \langle \Gamma_1 e_j, \Gamma_2 e_j \rangle \text{ for } \Gamma_1, \Gamma_2 \in \mathcal{L}_2(H, U).$$

We have $\mathcal{L}_1(H, U) \subset \mathcal{L}_2(H, U)$ and $\Gamma \in \mathcal{L}_i(H, U)$ if and only if $\Gamma^* \in \mathcal{L}_i(U, H)$ with

$$\|\Gamma\|_{\mathcal{L}_i(H,U)} = \|\Gamma^*\|_{\mathcal{L}_i(U,H)}, \ i \in \{1, 2\}. \tag{2.3}$$

We identify the spaces $\mathcal{L}_2(H, U)$ and $U \otimes H$, the Hilbert tensor product, with equivalent norms. The tensor $u \otimes v$ is regarded as an element of $\mathcal{L}(H, U)$ by the relation $(u \otimes v)w = \langle v, w \rangle_H u$ for $v, w \in H$ and $u \in U$. It can be seen that $u \otimes v \in \mathcal{L}_1(H, U)$ with $\|u \otimes v\|_{\mathcal{L}_1(H,U)} = \|u \otimes v\|_{\mathcal{L}_2(H,U)} = \|u\|_U \|v\|_H$. Moreover,

$$\langle \Gamma, u \otimes v \rangle_{\mathcal{L}_2(H,U)} = \langle \Gamma v, u \rangle_U \text{ for } \Gamma \in \mathcal{L}_2(H, U). \tag{2.4}$$

If V and G are two other real and separable Hilbert spaces then

$$\Gamma_1 u \otimes \Gamma_2 v = \Gamma_1 (u \otimes v) \Gamma_2^* \tag{2.5}$$

for $u \in U, v \in H, \Gamma_1 \in \mathcal{L}(U, V)$ and $\Gamma_2 \in \mathcal{L}(H, G)$, with Γ_2^* denoting the adjoint of Γ_2. The spaces $\mathcal{L}_i(H, U)$, $i \in \{1, 2\}$, are operator ideals: if $\Gamma_1 \in \mathcal{L}(G, V)$, $\Gamma_2 \in \mathcal{L}_i(U, G)$ and $\Gamma_3 \in \mathcal{L}(H, U)$ then $\Gamma_3 \Gamma_2 \Gamma_1 \in \mathcal{L}_i(H, V)$ with

$$\|\Gamma_1 \Gamma_2 \Gamma_3\|_{\mathcal{L}_i(H,V)} \leqslant \|\Gamma_1\|_{\mathcal{L}(G,V)} \|\Gamma_2\|_{\mathcal{L}_i(U,G)} \|\Gamma_3\|_{\mathcal{L}(H,U)}. \tag{2.6}$$

If $\Gamma_1 \in \mathscr{L}_2(U, V)$ and $\Gamma_2 \in \mathscr{L}_2(H, U)$ then $\Gamma_1 \Gamma_2 \in \mathscr{L}_1(H, V)$ and

$$\|\Gamma_1 \Gamma_2\|_{\mathscr{L}_1(H,V)} \leqslant \|\Gamma_1\|_{\mathscr{L}_2(U,V)} \|\Gamma_2\|_{\mathscr{L}_2(H,U)}. \tag{2.7}$$

The trace of $\Gamma \in \mathscr{L}_1(H)$ is, for an arbitrary orthonormal basis $(e_j)_{j=1}^{\infty}$ of H, defined by

$$\mathrm{Tr}(\Gamma) = \sum_{j=1}^{\infty} \langle \Gamma e_j, e_j \rangle_H.$$

If $\Gamma \in \Sigma^+(H) \subset \Sigma(H)$, the space of all positive semidefinite operators then $\mathrm{Tr}(\Gamma) = \|\Gamma\|_{\mathscr{L}_1(H)}$.

2.2 *Probability theory in Hilbert spaces*

Below, we work on the bounded interval $[0, T]$, $T < \infty$. Let $(\Omega, \mathscr{A}, (\mathscr{F}_t)_{t \in [0,T]}, P)$ be a complete filtered probability space satisfying the usual conditions, which is to say that $\mathscr{F}_0$ contains all P-null sets and $\mathscr{F}_t = \cap_{s>t} \mathscr{F}_s$ for all $t \in [0, T]$. By $L^p(\Omega, H)$, $p \in [1, \infty)$, we denote the space of all H-valued random variables Y with norm $\|Y\|_{L^p(\Omega,H)} = (\mathbb{E}[\|Y\|_H^p])^{1/p}$. For $Y, Y' \in L^2(\Omega, H)$, the cross-covariance (operator) of Y, Y' is defined by $\mathrm{Cov}(Y, Y') = \mathbb{E}[(Y - \mathbb{E}[Y]) \otimes (Y' - \mathbb{E}[Y'])] \in \mathscr{L}_1(H)$ and the covariance of Y by $\mathrm{Cov}(Y) = \mathrm{Cov}(Y, Y) \in \mathscr{L}_1(H) \cap \Sigma^+(H)$. Note that $\mathrm{Cov}(Y, Y')$ is uniquely determined by $\langle \mathrm{Cov}(Y, Y')u, v \rangle_H = \mathrm{Cov}(\langle Y, v \rangle_H, \langle Y', u \rangle_H)$.

An H-valued random variable Y is said to be Gaussian if $\langle Y, v \rangle_H$ is a Gaussian real-valued random variable for all $v \in H$. Then $Y \in L^p(\Omega, H)$ for all $p \geq 1$. A pair Y, Y' of H-valued random variables is said to be jointly Gaussian if $Y \oplus Y'$ is an $H \oplus H$-valued Gaussian random variable. Then Y and Y' are independent if and only if $\mathrm{Cov}(Y, Y') = 0$; cf. Mandelbaum (1984).

Let W be a generalized Wiener process in U (see Da Prato & Zabczyk, 2014, Chapter 4) with covariance $Q \in \Sigma^+(U)$, not necessarily of trace class. The Hilbert space $(Q^{1/2}(U), \langle Q^{-1/2} \cdot, Q^{-1/2} \cdot \rangle_U)$ is denoted by U_0, where $Q^{1/2}$ is the unique positive semidefinite square root of Q and $Q^{-1/2}$ its pseudoinverse.

2.3 *Covariance integral equations and mild solutions to SPDEs*

Let $\hat{H}$ be another Hilbert space such that $H \hookrightarrow \hat{H}$ continuously and densely. The main topic of study in this paper are integral equations of the form

$$K(t) = S(t)Q_\xi S(t)^* + \int_0^t S(t-s)FK(s)S(t-s)^* + S(t-s)K(s)(S(t-s)F)^* \, \mathrm{d}s$$

$$+ \int_0^t S(t-s)B(S(t-s)B)^* \, \mathrm{d}s, \tag{2.8}$$

taking values in the space $\mathscr{L}_1(H)$, with the adjoint being taken with respect to the inner product of H. Here, $S = (S(t))_{t \in [0,T]}$ is a family of $\mathscr{L}(\hat{H}, H)$-valued operators, not yet assumed to be a semigroup, while $F \in \mathscr{L}(H, \hat{H})$, $B \in \mathscr{L}_2(U_0, \hat{H})$ and $Q_\xi \in \mathscr{L}_1(H) \cap \Sigma^+(H)$. We assume that for any $v \in H$, the mapping $t \mapsto S(t)v \in H$ is continuous on $[0, T]$ and the mappings $t \mapsto S(t)Fv \in H$ and $t \mapsto S(t)B \in \mathscr{L}_2(U_0, H)$ are continuous on $(0, T]$. Below, we make an assumption on the boundedness of these mappings, which is used to deduce existence and uniqueness of solutions to (2.8) and the

stochastic evolution equation

$$X(t) = S(t)X(0) + \int_0^t S(t-s)FX(s)\,\mathrm{d}s + \int_0^t S(t-s)B\,\mathrm{d}W(s), \text{ for } t \in (0, T]. \tag{2.9}$$

When S is a semigroup, this is the mild solution of (1.1). Here, $X(0) = \xi$ is a Gaussian (possibly deterministic) $\mathscr{F}_0$-measurable H-valued random variable. The stochastic integral is of the Itô kind (Da Prato & Zabczyk, 2014, Chapter 4).

ASSUMPTION 2.1 There is a constant $C < \infty$ and functions $a \in L^1([0, T], \mathbb{R}), b \in L^2([0, T], \mathbb{R})$ such that $\|S(t)\|_{\mathscr{L}(H)} \leqslant C$ for all $t \in [0, T]$ and $\|S(t)F\|_{\mathscr{L}(H)} \leqslant a(t), \|S(t)B\|_{\mathscr{L}_2(U_0, H)} \leqslant b(t)$ for all $t \in (0, T]$.

We look for a solution K to (2.8) in the space $\mathscr{C}([0, T], \mathscr{L}_1(H))$ of continuous mappings with values in $\mathscr{L}_1(H)$. This is a Banach space with norm $\|f\|_{\infty, \mathscr{L}_1(H)} = \sup_{t \in [0,T]} \|f(t)\|_{\mathscr{L}_1(H)}$.

PROPOSITION 2.1 Under Assumption 2.1, there is a unique solution $K \in \mathscr{C}([0, T], \mathscr{L}_1(H))$ to (2.8) such that $K(t) \in \Sigma(H)$ for all $t \in [0, T]$.

Proof. First, we note that for $v \in H$ fixed, we may write

$$S(t)Q_\xi S(t)^* v - S(s)Q_\xi S(s)^* v = \sum_{j=1}^\infty (\langle v, (S(t) - S(s))e_j \rangle_H) S(t)\mu_j e_j$$
$$+ \sum_{j=1}^\infty \langle v, S(s)e_j \rangle_H (S(s) - S(t))\mu_j e_j, \tag{2.10}$$

where $(e_j)_{j=1}^\infty$ is an orthonormal eigenbasis of Q_ξ with corresponding eigenvalues $(\mu_j)_{j=1}^\infty$. Since $(\mu_j)_{j=1}^\infty \in \ell^1$, both sums are well-defined operators in $\mathscr{L}(H)$ applied to v. By (2.2), we get

$$\|S(t)Q_\xi S(t)^* - S(s)Q_\xi S(s)^*\|_{\mathscr{L}_1(H)} \leqslant \sum_{j=1}^\infty \|(S(t) - S(s))e_j\|_H \|\mu_j S(t)e_j\|_H$$
$$+ \sum_{j=1}^\infty \|(S(t) - S(s))\mu_j e_j\|_H \|S(s)e_j\|_H.$$

Hence, $S(\cdot)Q_\xi S(\cdot)^* \in \mathscr{C}([0, T], \mathscr{L}_1(H))$ as a consequence of Assumption 2.1 and the dominated convergence theorem. By (2.7), the mapping $[0, t) \ni s \mapsto S(t-s)B(S(t-s)B)^*$ takes values in the separable Banach space $\mathscr{L}_1(H)$ and it is continuous so that the Bochner integral of it is well defined. Similarly, for $\mathscr{K} \in \mathscr{C}([0, T], \mathscr{L}_1(H))$, the mapping $[0, t) \ni s \mapsto S(t-s)F\mathscr{K}(s)S(t-s)^* + S(t-s)\mathscr{K}(s)(S(t-s)F)^*$ takes values in $\mathscr{L}_1(H)$ and it can be seen to be continuous by applying (2.1) along

with a calculation similar to (2.10). The mapping

$$\mathcal{K} \mapsto S(\cdot)Q_\xi S(\cdot)^* + \int_0^\cdot S(\cdot - s)F\mathcal{K}(s)S(\cdot - s)^* + S(\cdot - s)\mathcal{K}(s)(S(\cdot - s)F)^* \, ds$$

$$+ \int_0^\cdot S(\cdot - s)B(S(\cdot - s)B)^* \, ds \qquad (2.11)$$

from $\mathscr{C}([0,T], \mathscr{L}_1(H))$ into itself is therefore well defined. By Assumption 2.1,

$$\Big\| \int_0^t S(t-s)F\mathcal{K}_1(s)S(t-s)^* + S(t-s)\mathcal{K}_1(s)(S(t-s)F)^* \, ds$$

$$- \int_0^t S(t-s)F\mathcal{K}_2(s)S(t-s)^* + S(t-s)\mathcal{K}_2(s)(S(t-s)F)^* \, ds \Big\|_{\mathscr{L}_1(H)}$$

$$\lesssim e^{\sigma t} \int_0^t a(t-s) \, e^{-\sigma(t-s)} \left(e^{-\sigma s} \|\mathcal{K}_1(s) - \mathcal{K}_2(s)\|_{\mathscr{L}_1(H)} \right) \, ds$$

for arbitrary $\sigma \in \mathbb{R}$ and $\mathcal{K}_1, \mathcal{K}_2 \in \mathscr{C}([0,T], \mathscr{L}_1(H))$. Therefore, since $\lim_{\sigma \to \infty} \int_0^T a(s) \, e^{-\sigma s} \, ds = 0$ by the dominated convergence theorem, the mapping (2.11) is a contraction with respect to the norm defined by $\sup_{t \in [0,T]} e^{-\sigma t} \|\mathcal{K}(t)\|_{\mathscr{L}_1(H)}$ for sufficiently large $\sigma \geq 0$. This norm is equivalent to $\|\cdot\|_{\infty, \mathscr{L}_1(H)}$, so the Banach fixed point theorem yields existence and uniqueness of K as the limit of the sequence $(K_n)_{n=0}^\infty \subset \mathscr{C}([0,T], \mathscr{L}_1(H))$. Here, $K_0 = 0$ and K_n, $n \geq 1$, is given by

$$K_n(t) = S(t)Q_\xi S(t)^* + \int_0^t S(t-s)FK_{n-1}(s)S(t-s)^* + S(t-s)K_{n-1}(s)(S(t-s)F)^* \, ds$$

$$+ \int_0^t S(t-s)B(S(t-s)B)^* \, ds, \, t \in [0,T].$$

Clearly, $K_0(t) \in \Sigma(H)$ for all $t \in [0,T]$. By induction, $K_n(t) \in \Sigma(H)$ for all $n \in \mathbb{N}$, $t \in [0,T]$. Since convergence of $(K_n)_{n=0}^\infty$ in $\mathscr{C}([0,T], \mathscr{L}_1(H))$ yields convergence of $(K_n(t))_{n=0}^\infty$ in $\mathscr{L}(H)$, we have $\langle K(t)u, v \rangle_H = \lim_n \langle K_n(t)u, v \rangle_H$ for all $t \in [0,T]$ and $u, v \in H$. Therefore, $K(t) \in \Sigma(H)$ for all $t \in [0,T]$. $\qquad \square$

REMARK 2.1   Since the mapping $t \mapsto S(t)Fv \in H$ is allowed to be discontinuous at 0, the advection term in (1.2) could be treated as a linear perturbation F, at least in the case of Dirichlet boundary conditions, see Kruse (2014, Example 2.22). In this case, S is the semigroup generated by the elliptic operator of (1.2). For notational convenience and to easily treat more general boundary conditions, we instead choose to, in Section 3.1, treat the advection term in (1.2) as part of an elliptic operator.

The next proposition confirms that the solution X to the stochastic evolution equation (2.9) is Gaussian at all times $t \in [0,T]$. As a consequence, $K(t)$ therefore determines the distribution of $X(t)$ when S is a semigroup, since by Theorem 2.3 below, $K(t) = \text{Cov}(X(t))$ for $t \in [0,T]$.

PROPOSITION 2.2   Under Assumption 2.1, there is a unique solution $X \in \mathscr{C}([0,T], L^2(\Omega, H))$ to (2.9) and $X(t)$ is Gaussian for all $t \in [0,T]$.

Proof. In the case that S is a semigroup, ξ is deterministic and $F = 0$, the result is well known; see, e.g., Da Prato & Zabczyk (2014, Theorem 5.2). We only sketch the proof in our general case. Existence and uniqueness of a solution to (2.9) follow from a Banach fixed point theorem as in Proposition 2.1, using the Itô isometry for the stochastic integral (Kruse, 2014, Theorem 2.25). In particular, we have existence and uniqueness of the process

$$X_1 = S(\cdot)\xi + \int_0^{\cdot} S(\cdot - s)FX_1(s)\,\mathrm{d}s$$

in $\mathscr{C}([0,T], L^2(\Omega, H))$ and of the process

$$X_2 = \int_0^{\cdot} S(\cdot - s)FX_2(s)\,\mathrm{d}s + \int_0^{\cdot} S(\cdot - s)B\,\mathrm{d}W(s)$$

in $\mathscr{C}([0,T], L^2(\Omega, H))$ as limits of iterative sequences $(X_1^n)_{n=0}^{\infty}$ and $(X_2^n)_{n=0}^{\infty}$ as in the proof of Proposition 2.1, with $X_1^0 = X_2^0 = 0$. Since $X_1^n(t)$ is obtained from a linear and bounded transformation of ξ, it is Gaussian for each $t \in [0,T]$. For X_2^n, one can use an inductive argument along with the stochastic Fubini theorem (Kruse, 2014, Theorem 4.18) to see that there is a function $\psi_n \colon [0,T] \times [0,T] \to \mathscr{L}_2(U_0, H)$, continuous in each argument, with $\sup_{t \in [0,T]} \|\psi_n(t, \cdot)\|_{L^2([0,t], \mathscr{L}_2(U_0, H))} < \infty$ such that $X_2^n(t) = \int_0^t \psi_n(t, s)\,\mathrm{d}W(s)$ for all $t \in [0,T]$. Therefore, $X_2^n(t)$ is also Gaussian for each $t \in [0,T]$. Since $\lim_n X_1^n(t) \oplus X_2^n(t) = X_1(t) \oplus X_2(t)$ in $L^2(\Omega, H \oplus H)$, (X_1, X_2) is a jointly Gaussian pair, from which the result follows. $\qquad\square$

We now prove our main result, connecting the equations (2.8) and (2.9). For this, we need to assume that S is a C_0-semigroup. This is required by the main tool of our proof, the mild Itô formula (Da Prato *et al.*, 2019, Theorem 1). In our setting, this formula gives that for a twice continuously Fréchet differentiable functional φ and $t \in [0,T]$,

$$\varphi(X(t)) = \varphi(S(t)\xi) + \int_0^t \langle \varphi'(S(t-s)X(s)), S(t-s)FX(s)\rangle_H \,\mathrm{d}s$$

$$+ \int_0^t \langle \varphi'(S(t-s)X(s)), S(t-s)B\cdot\rangle_H \,\mathrm{d}W(s)$$

$$+ \frac{1}{2}\sum_{j=1}^{\infty} \int_0^t \langle \varphi''(S(t-s)X(s))S(t-s)Bf_j, S(t-s)Bf_j\rangle_H.$$

Here, the Fréchet derivatives φ' and φ'' take, by the Riesz representation theorem, values in H and $\Sigma(H)$, respectively, and $(f_j)_{j=1}^{\infty}$ is an orthonormal basis of U_0. The formula is proven by applying the standard Itô formula to (2.9) with t fixed in the integrand, and then using the semigroup property of S to relate the process with t fixed to the original process X.

THEOREM 2.3 Let $\mathrm{Cov}(\xi) = Q_\xi$, and let S be a C_0-semigroup satisfying Assumption 2.1. Then, the process $K = (\mathrm{Cov}(X(t)))_{t\in[0,T]}$, where X is given by (2.9), is the solution of (2.8).

Proof. We first suppose that $X(0) = \xi = 0$ so that $\mathrm{Cov}(X(t)) = \mathbb{E}[X(t) \otimes X(t)]$. By an argument analogous to (1.4), the continuity of X implies that $\mathrm{Cov}(X(\cdot)) \in \mathscr{C}([0,T], \mathscr{L}_1(H))$. Below, we will

make several interchanges of integration, summation and expectation. These are allowed by Fubini's theorem, using Assumption 2.1 and the fact that $\sup_{t\in[0,T]}\|X(t)\|_{L^2(\Omega,H)} < \infty$. Let $(e_j)_{j=1}^\infty$ and $(f_j)_{j=1}^\infty$ be orthonormal bases of H and U_0, respectively. By Assumption 2.1, the mild Itô formula is applicable with the functional φ given by $\varphi(x) = \langle x, e_i\rangle_H \langle x, e_j\rangle_H$ for $x \in H$, $i,j \in \mathbb{N}$. We have $\varphi'(x) = \langle x, e_i\rangle_H e_j + \langle x, e_j\rangle_H e_i$ and $\varphi''(x) = e_i \otimes e_j + e_j \otimes e_i$. Using also the zero expectation property of the Itô integral and (2.4), we find that

$$\langle \mathrm{Cov}(X(t))e_i \otimes e_j\rangle_{\mathscr{L}_2} = \mathbb{E}[\langle X(t), e_i\rangle_H \langle X(t), e_j\rangle_H]$$

$$= \int_0^t \mathbb{E}[\langle S(t-s)FX(s), e_i\rangle_H \langle S(t-s)X(s), e_j\rangle_H]\,\mathrm{d}s$$

$$+ \int_0^t \mathbb{E}[\langle S(t-s)X(s), e_i\rangle_H \langle S(t-s)FX(s), e_j\rangle_H]\,\mathrm{d}s$$

$$+ \sum_{n=1}^\infty \int_0^t \langle S(t-s)Bf_n, e_i\rangle_H \langle S(t-s)Bf_n, e_j\rangle_H\,\mathrm{d}s,$$

for $t \in [0,T]$, $i,j \in \mathbb{N}$. By (2.4), (2.5) and the definition of $\langle \cdot,\cdot\rangle_{U_0}$, this is equal to

$$\int_0^t \langle \mathbb{E}[S(t-s)FX(s) \otimes S(t-s)X(s)], e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s$$

$$+ \int_0^t \langle \mathbb{E}[S(t-s)X(s) \otimes S(t-s)FX(s)], e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s$$

$$+ \int_0^t \langle (S(t-s)B)^*e_i, (S(t-s)B)^*e_j\rangle_{U_0}\,\mathrm{d}s$$

$$= \int_0^t \langle S(t-s)F\mathrm{Cov}(X(s))S(t-s)^*, e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s$$

$$+ \int_0^t \langle S(t-s)\mathrm{Cov}(X(s))(S(t-s)F)^*, e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s$$

$$+ \int_0^t \langle S(t-s)B(S(t-s)B)^*, e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s.$$

By applying $\langle \cdot, e_i \otimes e_j\rangle_H$ to (2.8) and using (2.4), we obtain similarly that

$$\langle K(t), e_i \otimes e_j\rangle_{\mathscr{L}_2(H)} = \int_0^t \langle S(t-s)FK(s)S(t-s)^*, e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s$$

$$+ \int_0^t \langle S(t-s)K(s)(S(t-s)F)^*, e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s$$

$$+ \int_0^t \langle S(t-s)B(S(t-s)B)^*, e_i \otimes e_j\rangle_{\mathscr{L}_2(H)}\,\mathrm{d}s.$$

Combining this with the previous result yields

$$\langle K(t) - \mathrm{Cov}(X(t)), e_i \otimes e_j \rangle_{\mathscr{L}_2(H)} = \int_0^t \langle S(t-s)F\left(K(s) - \mathrm{Cov}(X(s))\right)S(t-s)^*, e_i \otimes e_j \rangle_{\mathscr{L}_2(H)} \, \mathrm{d}s$$

$$+ \int_0^t \langle S(t-s)\left(K(s) - \mathrm{Cov}(X(s))\right)(S(t-s)F)^*, e_i \otimes e_j \rangle_{\mathscr{L}_2(H)} \, \mathrm{d}s.$$

This shows that

$$K(t) - \mathrm{Cov}(X(t)) = \sum_{i,j=1}^{\infty} \langle K(t) - \mathrm{Cov}(X(t)), e_i \otimes e_j \rangle_{\mathscr{L}_2(H)} e_i \otimes e_j$$

$$= \int_0^t S(t-s)F\left(K(s) - \mathrm{Cov}(X(s))\right)S(t-s)^*$$

$$+ S(t-s)\left(K(s) - \mathrm{Cov}(X(s))\right)(S(t-s)F)^* \, \mathrm{d}s$$

so that $K(t) = \mathrm{Cov}(X(t))$ for all $t \in [0, T]$ by uniqueness of K in (2.8).

For the general case $\xi \neq 0$, we write $X = X_1 + X_2$ as in the proof of Proposition 2.2. From the fact that $X_1(t) \oplus X_2(t) = \lim_n X_1^n(t) \oplus X_2^n(t)$ in $L^2(\Omega, H \oplus H)$, we find that $\mathrm{Cov}(\langle X_1(t), u \rangle_H, \langle X_2(t), v \rangle_H) = \lim_n \mathrm{Cov}(\langle X_1^n(t), u \rangle_H, \langle X_2^n(t), v \rangle_H) = 0$. This implies that $\mathrm{Cov}(X_1(t), X_2(t)) = \mathrm{Cov}(X_2(t), X_1(t)) = 0$ for arbitrary $t \in [0, T]$. Since $X_2(0) = 0$,

$$\mathrm{Cov}(X_2(t)) = \int_0^t S(t-s)F\mathrm{Cov}(X_2(s))S(t-s)^* + S(t-s)\mathrm{Cov}(X_2(s))(S(t-s)F)^* \, \mathrm{d}s$$

$$+ \int_0^t S(t-s)B(S(t-s)B)^* \, \mathrm{d}s$$

as a consequence of what we have already shown. A similar argument using the mild Itô formula yields

$$\mathrm{Cov}(X_1(t)) = S(t)Q_\xi S(t)^*$$

$$+ \int_0^t S(t-s)F\mathrm{Cov}(X_1(s))S(t-s)^* + S(t-s)\mathrm{Cov}(X_1(s))(S(t-s)F)^* \, \mathrm{d}s$$

for all $t \in [0, T]$. The proof is completed by noting that

$$\mathrm{Cov}(X_1(t) + X_2(t)) = \mathrm{Cov}(X_1(t)) + \mathrm{Cov}(X_1(t), X_2(t)) + \mathrm{Cov}(X_2(t), X_1(t)) + \mathrm{Cov}(X_2(t)).$$

$\square$

REMARK 2.2 As a consequence of the theorem, $K(t) \in \Sigma^+(H)$ for all $t \in [0, T]$.

We finish this section with a general error decomposition formula with respect to the integral equation (2.8) and an approximation of the semigroup S. For this, we consider a family $\tilde{S} = (\tilde{S}(t))_{t \in [0,T]}$ of operators in $\mathscr{L}(\hat{H}, H)$ such that the mappings $t \mapsto \tilde{S}(t)v \in H$, $t \mapsto \tilde{S}(t)Fv \in H$ and $t \mapsto \tilde{S}(t)B \in \mathscr{L}_2(U_0, H)$ are continuous almost everywhere on $[0, T]$ for all $v \in H$. If $\tilde{S}$ also satisfies Assumption 2.1

and we consider a function $\hat{K} : [0, T] \to \mathscr{L}_1(H) \cap \Sigma(H))$ that is continuous almost everywhere on $[0, T]$ then $\tilde{K}(t)$, given by

$$\tilde{K}(t) = \tilde{S}(t)Q_\xi \tilde{S}(t)^* + \int_0^t \tilde{S}(t-s)F\hat{K}(s)\tilde{S}(t-s)^* + \tilde{S}(t-s)\hat{K}(s)(\tilde{S}(t-s)F)^* \, \mathrm{d}s$$

$$+ \int_0^t \tilde{S}(t-s)B(\tilde{S}(t-s)B)^* \, \mathrm{d}s, \tag{2.12}$$

is well defined since the integrands are measurable mappings with values in the separable Banach space $\mathscr{L}_1(H)$. In the error decomposition formula, we consider $\tilde{K}(t)$, defined by (2.12), as an approximation of $K(t)$, $t \in [0, T]$.

PROPOSITION 2.4   Let, for $t \in [0, T]$, $K(t)$ be given by (2.8) and $\tilde{K}(t)$ by (2.12). Then, with $\mathscr{O}^+(s) = S(s) + \tilde{S}(s)$ and $\mathscr{O}^-(s) = S(s) - \tilde{S}(s)$ for $s \in [0, t]$,

$$\|K(t) - \tilde{K}(t)\|_{\mathscr{L}_i(H)} \leqslant \|\mathscr{O}^-(t)Q_\xi \mathscr{O}^+(t)\|_{\mathscr{L}_i(H)}$$

$$+ 2\int_0^t \|S(t-s)F(K(s) - \hat{K}(s))S(t-s)^*\|_{\mathscr{L}_i(H)} \, \mathrm{d}s$$

$$+ \int_0^t \|\mathscr{O}^-(t-s)F\hat{K}(s)(\mathscr{O}^+(t-s))^*\|_{\mathscr{L}_i(H)} \, \mathrm{d}s$$

$$+ \int_0^t \|\mathscr{O}^+(t-s)F\hat{K}(s)(\mathscr{O}^-(t-s))^*\|_{\mathscr{L}_i(H)} \, \mathrm{d}s$$

$$+ \int_0^t \|\mathscr{O}^-(t-s)B(\mathscr{O}^+(t-s)B)^*\|_{\mathscr{L}_i(H)} \, \mathrm{d}s \text{ for } i \in \{1, 2\}.$$

Proof.   The proposition is a straightforward consequence of the triangle inequality, (2.3), the fact that $K(t), \tilde{K}(t) \in \Sigma(H)$ for $t \in [0, T]$ and the identity

$$\Gamma_1 \tilde{\Gamma} \Gamma_1^* - \Gamma_2 \tilde{\Gamma} \Gamma_2^* = \frac{1}{2} \left((\Gamma_1 + \Gamma_2)\tilde{\Gamma}(\Gamma_1 - \Gamma_2)^* + (\Gamma_1 - \Gamma_2)\tilde{\Gamma}(\Gamma_1 + \Gamma_2)^* \right) \tag{2.13}$$

for $\Gamma_1, \Gamma_2, \tilde{\Gamma} \in \mathscr{L}(H)$. $\square$

3. Applications

We apply the theory of the previous section to two concrete stochastic equations, a stochastic advection-diffusion equation and the stochastic wave equation. Fully discrete approximation schemes are analyzed and numerical simulations are provided for illustration.

3.1 *A stochastic advection–diffusion equation*

Let $\mathscr{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$ be a bounded domain. The SPDE we consider in this section is formally given by

$$dX(t,x) + \mathscr{A}X(t,x)\,dt = dW(t,x) \text{ for } t \in (0,T], x \in \mathscr{D},$$
$$X(0,x) = \xi(x), \qquad \text{for } x \in \mathscr{D},$$

(3.1)

for a random initial condition $X(0) = \xi$ and an operator

$$\mathscr{A} = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} a_{i,j} \frac{\partial}{\partial x_j} + \sum_{j=1}^{d} a_j \frac{\partial}{\partial x_j} + a_0.$$

We consider either Dirichlet, Neumann or Robin boundary conditions and we let W be a generalized Wiener process in $H = L^2(\mathscr{D})$ with covariance operator $Q \in \Sigma^+(H)$. We assume that ξ is an H-valued $\mathscr{F}_0$-measurable Gaussian random variable with covariance Q_ξ. The coefficients $a_{i,j}, a_i$, $i,j = 1, \ldots, d$, and a_0 are functions on $\bar{\mathscr{D}}$ fulfilling $a_{i,j} = a_{j,i}$. We assume that for there is some $\lambda_0 > 0$ such that $\sum_{i,j=1}^{d} a_{i,j}(x) y_i y_j \geq \lambda_0 |y|^2$ for all $y \in \mathbb{R}^d$ and $x \in \bar{\mathscr{D}}$, so that $\mathscr{A}$ is elliptic.

To put (3.1) into our framework, we follow Fujita & Suzuki (1991) and introduce the spaces V and $\mathbb{H}$ as subspaces of the Sobolev spaces $H^1 = H^1(\mathscr{D})$ and $H^2 = H^2(\mathscr{D})$, respectively. In the Dirichlet case, we set $V = \mathbb{H} = H_0^1 = \{u \in H^1 : u = 0 \text{ on } \partial\mathscr{D}\}$. In the Robin case (Neumann boundary conditions being a special case thereof), we set $V = H^1(\mathscr{D})$ and $\mathbb{H} = \{u \in H^2 : \partial u/\partial v_\Lambda + \sigma u = 0 \text{ on } \partial\mathscr{D}\}$, where $\sigma : \partial\mathscr{D} \to \mathbb{R}$ is a sufficiently smooth function and

$$\frac{\partial u}{\partial v_\Lambda} = \sum_{i,j=1}^{d} n_i a_{i,j} \frac{\partial u}{\partial x_j},$$

with $\mathbf{n} = (n_1, \ldots, n_d)$ being the outward unit normal to $\partial\mathscr{D}$. Here and below, the boundary conditions should be understood in terms of trace operators; see Grisvard (1985, Sections 1.5 and 1.6). We define a bilinear form $\lambda : V \times V \to \mathbb{R}$ associated with $\mathscr{A}$ by

$$\lambda(u,v) = \int_{\mathscr{D}} \sum_{i,j=1}^{d} a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_{j=1}^{d} a_j \frac{\partial u}{\partial x_j} v + a_0 u v \, dx + \int_{\partial\mathscr{D}} \sigma u v \, dx,$$

where the last term is dropped in the Dirichlet case. If the coefficients are bounded then $|\lambda(u,v)| \leq \|u\|_V \|v\|_V$ so that we may associate an operator $\Lambda : V \to V^*$ to λ by $\lambda(u,v) = {}_{V^*}\langle \Lambda u, v \rangle_V$. By Riesz's representation theorem, we obtain a Gelfand triple $V \subset H \subset V^*$. We restrict Λ to $\text{dom}(\Lambda) = \{u \in V : \Lambda u \in H\}$ without changing notation. When $\mathscr{D}$ has Lipschitz boundary, one may show (using the trace inequality Grisvard, 1985, Theorem 1.5.1.3 if necessary) that there are two constants $\tilde{\lambda}_0 > 0$, $c_0 \geq 0$ such that $\lambda(u,u) \geq \tilde{\lambda}_0 \|u\|_{H^1}^2 - c_0 \|u\|_H^2$ for $u \in V$. We add the term $c_0 X(t,x)$ to both sides of (3.1). Then the associated bilinear form $a(\cdot,\cdot) = \lambda(\cdot,\cdot) + c_0 \langle\cdot,\cdot\rangle_H$ is coercive. In the case that $\sigma \geq 0$, the constant

c_0 may, for example, be chosen as any number fulfilling

$$c_0 > \frac{\sup_{x\in\mathscr{D}} \sum_{j=1}^{d} |a_j(x)|}{4\lambda_0 \epsilon} - \inf_{x\in\mathscr{D}} a_0(x), \tag{3.2}$$

for an arbitrary $\epsilon \in (0,1)$. In the Dirichlet case, we may pick $c_0 = 0$ if $a_j = 0$ for all $j = 1,\ldots,d$ and $a_0(x) \geq 0$. With this, the SPDE (3.1) is put into the form of (1.1) by letting $A = \Lambda + c_0 I$ and $F = c_0 I$. The adjoint operator A^* is defined by associating it with the bilinear form a^*, given by $a^*(u,v) = a(v,u)$ for $u, v \in V$. For smooth coefficients, we use Green's formula to see that A^* can be identified with the formal adjoint of $\mathscr{A}$ perturbed by c_0 (cf. Yagi, 2010, Section 2.1.3). It is given by

$$\mathscr{A}^* = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} a_{i,j} \frac{\partial}{\partial x_j} - \sum_{j=1}^{d} a_j \frac{\partial}{\partial x_j} - \sum_{j=1}^{d} \frac{\partial a_j}{\partial x_j} + a_0 + c_0.$$

In the Robin case, the boundary conditions of $\mathscr{A}^*$ change to the ones of the space

$$\mathbb{H}^* = \left\{ u \in H^2 : \frac{\partial u}{\partial \nu_\Lambda} + \left(\sum_{j=1}^{d} a_j n_j + \sigma \right) u = 0 \text{ on } \partial\mathscr{D} \right\},$$

while in the Dirichlet case, $\mathbb{H}^* = \mathbb{H}$. We also need the symmetrized operator A_0 associated with the bilinear form $a_0 = (a + a^*)/2$. Like A^*, A_0 is identified with a differential operator

$$\mathscr{A}_0 = -\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i} a_{i,j} \frac{\partial}{\partial x_j} - \frac{1}{2} \sum_{j=1}^{d} \frac{\partial a_j}{\partial x_j} + a_0 + c_0.$$

We set

$$\mathbb{H}_0 = \left\{ u \in H^2 : \frac{\partial u}{\partial \nu_\Lambda} + \left(\frac{1}{2} \sum_{j=1}^{d} a_j n_j + \sigma \right) u = 0 \text{ on } \partial\mathscr{D} \right\},$$

in the Robin case and $\mathbb{H}_0 = \mathbb{H}$ in the Dirichlet case. With these notions in place, we introduce an assumption of elliptic regularity.

ASSUMPTION 3.1   The coefficients $a_{i,j}, a_j, a_0, \sigma$, $i,j = 1,\ldots,d$, are sufficiently smooth and $\mathscr{D}$ is sufficiently regular, with $\partial\mathscr{D}$ at least Lipschitz, to guarantee that $\mathrm{dom}(A) = H^2 \cap \mathbb{H}$, $\mathrm{dom}(A^*) = H^2 \cap \mathbb{H}^*$ and $\mathrm{dom}(A_0) = H^2 \cap \mathbb{H}_0$. The equalities hold with equivalence of $\|\cdot\|_{H^2}$ and the graph norms $\|A\cdot\|_H$, $\|A^*\cdot\|_H$ and $\|A_0\cdot\|_H$, respectively.

We refer to Grisvard (1985) for details on when this assumption holds. It is satisfied when $\mathscr{D}$ is convex and $a_{i,j}, a_j, a_0, \sigma$, $i,j = 1,\ldots,d$ are infinitely differentiable; see Fujita & Suzuki (1991).

Since a is coercive, A is a sectorial operator. Negative fractional powers of A are therefore well defined as elements of $\mathscr{L}(H)$ given by

$$A^{-\frac{s}{2}} = \frac{1}{2\pi i} \int_\gamma \lambda^{-\frac{s}{2}} (\lambda - A)^{-1} \, d\lambda,$$

where γ is a counterclockwise oriented contour surrounding the spectrum of A. Positive fractional powers are densely defined closed operators on H defined by $A^{s/2} = (A^{-s/2})^{-1}$ (Yagi, 2010, Section 2.1.7). We note that $(A^*)^{s/2} = (A^{s/2})^*$ for all $s \in \mathbb{R}$. Moreover, by Kato (1961, Theorem 3.1) and Lions (1962, Théorème 6.1) (applicable since $\mathscr{D}$ has Lipschitz boundary), we have $\mathrm{dom}(A^{s/2}) = \mathrm{dom}((A^*)^{s/2}) = \mathrm{dom}(A_0^{s/2})$ for all $s \in [0,1]$ with norm equivalence. For $s = 1$, these spaces can also be identified with V.

It is convenient to express our regularity assumptions on Q not in terms of fractional powers of A, which is usually the case when A is self-adjoint, but of A_0. As A_0 is positive definite with a compact inverse (a consequence of Yagi, 2010, Theorem 1.38 since $\mathrm{dom}(A_0) \subset H^2$), its fractional powers can be characterized in a simple way by the spectral theorem; cf. Kruse (2014, Appendix B.2). For $s \geq 0$, we write $\dot{H}^s$ for the Hilbert space $\mathrm{dom}(A_0^{s/2})$. Moreover, Hilbert spaces $\dot{H}^{-s}$ are well defined as completions of sequences in H with respect to $\|\cdot\|_{\dot{H}^{-s}} = \|A_0^{-s/2}\cdot\|_H$. In this way, we obtain a set $(\dot{H}^s)_{s\in\mathbb{R}}$ of Hilbert spaces, with $\dot{H}^\alpha \hookrightarrow \dot{H}^s$ for $\alpha > s$, continuously and densely. Lemma 2.1 in Bolin *et al.* (2020) allows us to, for all $\alpha, s \in \mathbb{R}$, extend $A_0^{s/2}$ to an operator in $\mathscr{L}(\dot{H}^\alpha, \dot{H}^{\alpha-s})$, and we do so without changing notation. Note that for $v \in H$ and $s \in [0,1]$,

$$\|A_0^{-\frac{s}{2}} v\|_H = \sup_{\substack{w\in H \\ \|w\|_H=1}} \left| \langle A_0^{-\frac{s}{2}} v, w \rangle_H \right| = \sup_{\substack{w\in H \\ \|w\|_H=1}} \left| \langle A^{-\frac{s}{2}} v, (A^{\frac{s}{2}})^* A_0^{-\frac{s}{2}} w \rangle_H \right|$$

$$\leqslant \|A^{-\frac{s}{2}} v\|_H \sup_{\substack{w\in H \\ \|w\|_H=1}} \|(A^{\frac{s}{2}})^* A_0^{-\frac{s}{2}} w\|_H \lesssim \|A^{-\frac{s}{2}} v\|_H,$$

by the equivalence $\mathrm{dom}(A_0^{s/2}) = \mathrm{dom}((A^*)^{s/2})$. Similarly,

$$\|A^{-\frac{s}{2}} v\|_H \lesssim \|A_0^{-\frac{s}{2}} v\|_H. \tag{3.3}$$

This is true also for A^*. Therefore, $A^{s/2}$ and $(A^*)^{s/2}$ can be considered as operators in $\mathscr{L}(\dot{H}^s, H)$ for $s \in [-1,1]$.

The operator $-A$ is the infinitesimal generator of a uniformly bounded analytic semigroup S on H (see Fujita & Suzuki, 1991), which yields a mild solution (2.9) to the SPDE (3.1). The semigroup maps into $\mathrm{dom}(A^s)$ for all $s \geq 0$, with S and A^s commutative on $\mathrm{dom}(A^s)$. The stability estimate

$$\|A^s S(t)\|_{\mathscr{L}(H)} \lesssim t^{-s} \tag{3.4}$$

is satisfied for $t > 0$. Moreover, as seen in Yagi (2010, Section 2.7.7),

$$\|A^{-s}(S(t) - I)\|_{\mathscr{L}(H)} \lesssim t^s, s \in [0,1], t \geq 0. \tag{3.5}$$

We now make the following assumption on Q.

ASSUMPTION 3.2   There is a constant $r \in (0, 1]$ such that

$$\|A_0^{\frac{r-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)} = \|Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H, \dot{H}^{r-1})} < \infty.$$

With this assumption in place, we have for $t > 0$ that

$$\|S(t)\|_{\mathscr{L}_2(U_0, H)} = \|A^{\frac{1-r}{2}} S(t) A^{\frac{r-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)} \leqslant \|A^{\frac{1-r}{2}} S(t)\|_{\mathscr{L}(H)} \|A^{\frac{r-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)} \lesssim t^{\frac{r-1}{2}}.$$

The last inequality is a consequence of (3.3) and (3.4). To have Assumption 2.1 fulfilled, we set $B = I$ and $F = c_0 I$. Here, the inclusion B is regarded as an operator in $U_0 = Q^{1/2}(H)$ to $\hat{H} = \dot{H}^{r-1}$, F as an operator from H to $\dot{H}^{r-1}$ and $S(t)$ as an operator from $\dot{H}^{r-1}$ to H. Continuity of $(0, T] \ni t \mapsto S(t)B \in \mathscr{L}_2(U_0, H)$ is a consequence of the fact $\|S(t)\|_{\mathscr{L}_2(U_0, H)} < \infty$ for all $t \in (0, T]$ along with continuity of the mapping $[0, T] \ni t \mapsto S(t)v$ for all $v \in U_0 = Q^{1/2}(H) \subset H$. With this Assumption 2.1 is fulfilled, so that by Proposition 2.2 and the fact that S is a semigroup, a predictable mild solution to (1.1) exists. Proposition 2.1 and Theorem 2.3 yield a unique solution K to (2.8) such that $K(t) = \mathrm{Cov}(X(t))$, $t \in [0, T]$.

We now move on to approximation of (2.8). For this, we consider the same approximation of S as in Fujita & Suzuki (1991) and assume from here on that $\mathscr{D}$ is a convex polygon. For the spatial discretization, we let $(V_h)_{h \in (0,1]} \subset H^1$ be a standard family of finite element spaces consisting of piecewise linear polynomials with respect to a regular family of triangulations of $\mathscr{D}$ with maximal mesh size h, vanishing on $\partial \mathscr{D}$ in the Dirichlet case. We assume the mesh to be quasi-uniform. The spaces are equipped with $\langle \cdot \cdot \rangle_{V-h} = \langle \cdot, \cdot \rangle_H$. On this space, let $A_h \colon V_h \to V_h$ be given by $\langle A_h v_h, u_h \rangle_H = a(v_h, u_h)$ for all $v_h, u_h \in V_h$. Since A_h is sectorial, fractional powers of it are defined in the same way as for A. We write $A_{0,h}$ for the operator defined in the same way using the bilinear form a_0. Note that A_h^* coincides with the operator defined using a^*. By $P_h \colon \dot{H}^{-1} \to V_h$ we denote the generalized orthogonal projector defined by $\langle P_h x, y_h \rangle_H = \langle A_0^{-1/2} x, A_0^{1/2} y_h \rangle_H$ for $x \in \dot{H}^{-1}, y_h \in V_h$. Since a_0 is a symmetric form, we have $\|A_{0,h}^{1/2} P_h v\|_H = \|A_0^{1/2} P_h v\|_H \lesssim \|P_h v\|_V$. The mesh of V_h is assumed to be quasi-uniform, so $\sup_{h \in (0,1]} \|P_h\|_{\mathscr{L}(V)} < \infty$ (Fujita & Suzuki, 1991, Proposition 3.2), (Kruse, 2014, Example 3.6). The interpolation arguments of Andersson & Larsson (2016) then imply that for $s \in [-1, 1]$, there is a constant $C < \infty$ such that for all $h \in (0, 1]$,

$$\|A_{0,h}^{\frac{s}{2}} P_h A_0^{-\frac{s}{2}}\|_{\mathscr{L}(H)} \leqslant C. \tag{3.6}$$

Moreover, by Kato (1961, Theorem 3.1) (see also the proof of Fujita & Suzuki, 1991, Theorem 5.3), for all $s \in [0, 1)$, there is a constant $C < \infty$ such that for all $h \in (0, 1]$, $\|A_h^{\frac{s}{2}} A_{0,h}^{-\frac{s}{2}} P_h\|_{\mathscr{L}(H)} \leqslant C$ and

$$\|A_h^{-\frac{s}{2}} A_{0,h}^{\frac{s}{2}} P_h\|_{\mathscr{L}(H)} \leqslant C. \tag{3.7}$$

We use the backward Euler method for the temporal discretization of S. For a time step $\Delta t \in (0, 1]$, let $(t_j)_{j \in \mathbb{N}_0} \subset \mathbb{R}$ be given by $t_j = \Delta t j$ and $N_{\Delta t} + 1 = \inf\{j \in \mathbb{N} : t_j \notin [0, T]\}$. We write $S_{h, \Delta t} = (I + \Delta t A_h)^{-1}$. The discrete family $(S_{h, \Delta t}^j)_{j \in \{0, \dots, N_{\Delta t}\}}$ of powers of $S_{h, \Delta t}$ acts as a fully discrete approximation

of S. For brevity, we write $S_{h,\Delta t}^j$ for $S_{h,\Delta t}^j P_h$. For all $s \in [0, 1)$, there is a constant $C < \infty$ such that for all $h, \Delta t \in (0, 1]$ and $j = 1, \ldots, N_{\Delta t}$, $\|A_h^s S_{h,\Delta t}^j\|_{\mathscr{L}(H)} \leqslant C t_j^{-s}$, see Fujita & Suzuki (1991, (8.7)). We define an interpolation $\tilde{S}_{h,\Delta t} \colon [0, T] \to \mathscr{L}(\mathscr{H})$ of $(S_{h,\Delta t}^j)_{j \in \{0, \ldots, N_{\Delta t}\}}$ by

$$\tilde{S}_{h,\Delta t}(t) = \chi_{\{0\}}(t) P_h + \sum_{j=1}^{N_{\Delta t}} \chi_{(t_{j-1}, t_j]}(t) S_{h,\Delta t}^j, t \in [0, T], \tag{3.8}$$

where χ denotes the indicator function. We immediately obtain that for all $s \in [0, 1)$, there is a constant $C < \infty$ such that for all $h, \Delta t \in (0, 1]$ and $t \in (0, T]$

$$\|A_h^s \tilde{S}_{h,\Delta t}(t)\|_{\mathscr{L}(H)} \leqslant C t^{-s}. \tag{3.9}$$

The next lemma gives an error bound for the semigroup approximation. In the Dirichlet case, a proof is given in Andersson *et al.* (2016, Lemma 5.1). In our general case, it is a consequence of the split

$$\|(S(t) - \tilde{S}_{h,\Delta t}(t))v\|_H \leqslant \|(S(t) - S(t_j))v\|_H + \|(S(t_j) - S_h(t_j)P_h)v\|_H + \|(S_h(t_j)P_h - \tilde{S}_{h,\Delta t}(t_j))v\|_H$$

for $t \in (t_{j-1}, t_j], j = 1, \ldots, N_{\Delta t}$, and $v \in H$. Here, S_h is the semigroup on V_h generated by A_h (Fujita & Suzuki, 1991, Section 7). The first term can be bounded by (3.4), (3.5) and (3.3), and the third by using an integral representation of the error as in the proof of Fujita & Suzuki (1991, Theorem 8.2). For the second term, one employs a similar interpolation argument as in Andersson *et al.* (2016, Lemma 5.1), making use of Fujita & Suzuki (1991, Theorem 7.1) and Kruse (2014, Lemma 3.8(iii)). The latter result is, again, proven for Dirichlet boundary conditions, but the arguments are the same in our general case; cf. Fujita & Suzuki (1991, Remark 7.2).

LEMMA 3.1   For all $\theta \in [0, 2]$ and $s \in [0, 2 - \theta] \cap [0, 1)$, there is a constant $C < \infty$ such that for all $h, \Delta t \in (0, 1]$ and $t > 0$

$$\|S(t) - \tilde{S}_{h,\Delta t}(t)\|_{\mathscr{L}(\dot{H}^{-s}, H)} = \|(S(t) - \tilde{S}_{h,\Delta t}(t))A_0^{\frac{s}{2}}\|_{\mathscr{L}(H)} \leqslant C(h^\theta + \Delta t^{\frac{\theta}{2}})t^{-\frac{\theta+s}{2}}.$$

We now define an approximation $(\tilde{K}_{h,\Delta t}(t_j))_{j=0}^{N_{\Delta t}}$ of (2.8) by $\tilde{K}_{h,\Delta t}(0) = P_h Q_\xi P_h$ and, for $j \geq 1$,

$$(I + \Delta t A_h)\tilde{K}_{h,\Delta t}(t_j)(I + \Delta t A_h)^* = \tilde{K}_{h,\Delta t}(t_{j-1}) + \Delta t F \tilde{K}_{h,\Delta t}(t_{j-1}) + \Delta t \tilde{K}_{h,\Delta t}(t_{j-1})F^* + \Delta t P_h B(P_h B)^*.$$

With our choice of F, this can equivalently be written as

$$(I + \Delta t A_h)\tilde{K}_{h,\Delta t}(t_j)(I + \Delta t A_h)^* = (1 + 2c_0 \Delta t)\tilde{K}_{h,\Delta t}(t_{j-1}) + \Delta t P_h Q P_h. \tag{3.10}$$

To implement this scheme, we have to know the value of c_0 explicitly. Several choices are possible; see (3.2) for an example. A closed form of $\tilde{K}_{h,\Delta t}(t_j), j = 0, \ldots, N_{\Delta t}$, is given by

$$\tilde{K}_{h,\Delta t}(t_j) = \tilde{S}_{h,\Delta t}(t_j)Q_\xi\tilde{S}_{h,\Delta t}(t_j)^* + \int_0^{t_j}\tilde{S}_{h,\Delta t}(t_j - s)F\tilde{K}_{h,\Delta t}(\lfloor s\rfloor_{\Delta t})\tilde{S}_{h,\Delta t}^*(t_j - s)\,\mathrm{d}s$$

$$+ \int_0^{t_j}\tilde{S}_{h,\Delta t}(t_j - s)\tilde{K}_{h,\Delta t}(\lfloor s\rfloor_{\Delta t})(\tilde{S}_{h,\Delta t}(t_j - s)F)^*\,\mathrm{d}s$$

$$+ \int_0^{t_j}\tilde{S}_{h,\Delta t}(t_j - s)B(\tilde{S}_{h,\Delta t}(t_j - s)B)^*\,\mathrm{d}s. \tag{3.11}$$

Here, $\lfloor \cdot \rfloor_{\Delta t} = \lfloor \cdot / \Delta t \rfloor \Delta t$, with $\lfloor \cdot \rfloor$ denoting the floor function. The $\mathscr{L}_1(H)$-norm of the last term can be bounded by (2.7), Assumption 3.2 and (3.9) via

$$\left\|\int_0^{t_j}\tilde{S}_{h,\Delta t}(t_j - s)B(\tilde{S}_{h,\Delta t}(t_j - s)B)^*\,\mathrm{d}s\right\|_{\mathscr{L}_1(H)}$$

$$\leqslant \int_0^{t_j}\|\tilde{S}_{h,\Delta t}(t_j - s)B\|_{\mathscr{L}_2(U_0,H)}^2\,\mathrm{d}s$$

$$= \int_0^{t_j}\|\tilde{S}_{h,\Delta t}(t_j - s)A_h^{\frac{1-r}{2}}A_h^{\frac{r-1}{2}}A_{0,h}^{\frac{1-r}{2}}A_{0,h}^{\frac{r-1}{2}}P_hA_0^{\frac{1-r}{2}}A_0^{\frac{r-1}{2}}Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}^2\,\mathrm{d}s$$

$$\lesssim \|A_h^{\frac{r-1}{2}}A_{0,h}^{\frac{1-r}{2}}\|_{\mathscr{L}(H)}^2\|A_{0,h}^{\frac{r-1}{2}}P_hA_0^{\frac{1-r}{2}}\|_{\mathscr{L}(H)}^2\|A_0^{\frac{r-1}{2}}Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}^2\int_0^{t_j}(t - s)^{r-1}\,\mathrm{d}s \lesssim t^r \lesssim T^r. \tag{3.12}$$

Here, we also made use of (3.6) and (3.7). Using this result along with (3.9) applied to the other terms of (3.11), we find that

$$\|\tilde{K}_{h,\Delta t}(t_j)\|_{\mathscr{L}_1(H)} \lesssim 1 + \sum_{k=0}^{j-1}\|\tilde{K}_{h,\Delta t}(t_k)\|_{\mathscr{L}_1(H)}.$$

Along with the fact that $Q_\xi \in \mathscr{L}_1(H)$, the discrete Gronwall lemma (see Grigorieff, 1975, 2.2 (9)) implies that

$$\sup_{h,\Delta t\in(0,1]}\sup_{j=0,\ldots,N_{\Delta t}}\|\tilde{K}_{h,\Delta t}(t_j)\|_{\mathscr{L}_1(H)} < \infty.$$

With this estimate in place, we move on to the main result of this section.

THEOREM 3.1 Let Assumption 3.2 be satisfied. For all $\theta < 2r$, there is a constant $C < \infty$ such that for all $h, \Delta t \in (0, 1]$

$$\sup_{j=0,\ldots,N_{\Delta t}}\|K(t_j) - \tilde{K}_{h,\Delta t}(t_j)\|_{\mathscr{L}_1(H)} \leqslant C(h^\theta + \Delta t^{\frac{\theta}{2}}).$$

Moreover, for all $\theta < 1 + r$, there is a constant $C < \infty$ such that for all $h, \Delta t \in (0, 1]$

$$\sup_{j=0,\ldots,N_{\Delta t}} \|K(t_j) - \tilde{K}_{h,\Delta t}(t_j)\|_{\mathscr{L}_2(H)} \leqslant C(h^\theta + \Delta t^{\frac{\theta}{2}}).$$

Proof. The result is a consequence of Proposition 2.4 applied to $\|K(t_j) - \tilde{K}_{h,\Delta t}(t_j)\|_{\mathscr{L}_i(H)}$. We start with $i = 1$. For the last term of the error decomposition in this proposition, the properties (2.6), (2.3), (2.7) and (3.9) along with Assumption 3.2, Lemma 3.1 and an argument similar to that of (3.12) imply that

$$\int_0^{t_j} \|\mathscr{O}^-(t_j - s)B(\mathscr{O}^+(t_j - s)B)^*\|_{\mathscr{L}_1(H)} \, ds$$

$$\leqslant \int_0^{t_j} \|\mathscr{O}^-(t_j - s)B\|_{\mathscr{L}_2(U_0,H)} \|\mathscr{O}^+(t_j - s)B\|_{\mathscr{L}_2(U_0,H)} \, ds$$

$$\lesssim \int_0^{t_j} \|(S(t_j - s) - \tilde{S}_{h,\Delta t}(t_j - s))A_0^{\frac{1-r}{2}}\|_{\mathscr{L}(H)} \|A_0^{\frac{r-1}{2}}Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)}^2$$

$$\times \left(\|\tilde{S}_{h,\Delta t}(t_j - s)A_h^{\frac{1-r}{2}}\|_{\mathscr{L}(H)} + \|S(t_j - s)A^{\frac{1-r}{2}}\|_{\mathscr{L}(H)} \right) \, ds$$

$$\lesssim (h^\theta + \Delta t^{\frac{\theta}{2}}) \int_0^{t_j} (t_j - s)^{r-1-\frac{\theta}{2}} \, ds \lesssim T^{r-\frac{\theta}{2}}(h^\theta + \Delta t^{\frac{\theta}{2}}). \tag{3.13}$$

Similarly, the first, third and fourth terms of the split in Proposition 2.4 can be bounded by a constant times $h^\theta + \Delta t^{\theta/2}$, noting that $F = c_0 I$. By (3.4), the second term can be bounded by

$$\int_0^{t_j} \|S(t_j - s)F(K(\lfloor s \rfloor_{\Delta t}) - \tilde{K}_{h,\Delta t}(\lfloor s \rfloor_{\Delta t}))S(t_j - s)^*\|_{\mathscr{L}_1(H)} \, ds \lesssim \Delta t \sum_{k=0}^{j-1} \|(K(t_k) - \tilde{K}_{h,\Delta t}(t_k))\|_{\mathscr{L}_1(H)}.$$

The proof for $i = 1$ is now completed by an application of the discrete Gronwall inequality.

In the case that $i = 2$, the only major difference is the calculation in (3.13). For this, let us note that since $\langle v, \mathscr{O}^+(t_j - s)Bu \rangle_H = \langle Q\mathscr{O}^+(t_j - s)^* v, u \rangle_{U_0}$ for $v \in H, u \in U_0$, we have $\mathscr{O}^-(t_j - s)B(\mathscr{O}^+(t_j - s)B)^* = \mathscr{O}^-(t_j - s)Q\mathscr{O}^+(t_j - s)^*$. The adjoint on the right-hand side is taken with respect to $\mathscr{L}(H)$. Using this observation, (3.9), Lemma 3.1 and the fact that $Q^{1/2} \in \mathscr{L}(H)$ yields

$$\int_0^{t_j} \|\mathscr{O}^-(t_j - s)B(\mathscr{O}^+(t_j - s)B)^*\|_{\mathscr{L}_2(H)} \, ds$$

$$\leqslant \int_0^{t_j} \|\mathscr{O}^-(t_j - s)Q^{\frac{1}{2}}\|_{\mathscr{L}(H)} \|\mathscr{O}^+(t_j - s)Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)} \, ds$$

$$\lesssim \int_0^{t_j} \|(S(t_j - s) - \tilde{S}_{h,\Delta t}(t_j - s))\|_{\mathscr{L}(H)} \|Q^{\frac{1}{2}}\|_{\mathscr{L}(H)}$$

$$\times \left(\|\tilde{S}_{h,\Delta t}(t_j - s)A_h^{\frac{1-r}{2}}\|_{\mathscr{L}(H)} + \|S(t_j - s)A^{\frac{1-r}{2}}\mathscr{L}(H)) \right) \|A_0^{\frac{r-1}{2}} Q^{\frac{1}{2}}\|_{\mathscr{L}_2(H)} \, ds$$

$$\lesssim (h^\theta + \Delta t^{\frac{\theta}{2}}) \int_0^{t_j} (t_j - s)^{\frac{r-1-\theta}{2}} \, ds \lesssim T^{\frac{r+1-\theta}{2}} (h^\theta + \Delta t^{\frac{\theta}{2}}).$$

$$\square$$

EXAMPLE 3.2   We conclude this section by demonstrating the results of Theorem 3.1 in the case that $\mathscr{D} = (0,1)$ and $T = 1$. We choose deterministic initial conditions so that $Q_\xi = 0$. Homogeneous Neumann boundary conditions are considered and we set $a_{1,1}(x) = 4$, $a_1(x) = \sin(2\pi x)$ and $a_0(x) = 0$ for all $x \in \mathscr{D}$. With this, we may take $c_0 = 1/8$. From (3.10), we obtain a matrix recursion

$$(\mathbf{M}_h + \Delta t\mathbf{A}_h)\mathbf{K}_{j,h,\Delta t}(\mathbf{M}_h + \Delta t\mathbf{A}_h)^* = (1 + 2c_0\Delta t)\mathbf{M}_h\mathbf{K}_{j-1,h,\Delta t}\mathbf{M}_h + \Delta t\mathbf{Q}_h.$$

Here, $\mathbf{K}_{j,h,\Delta t}$ is the matrix of coefficients $(k_{j,m,n})_{m,n=1}^{N_h}$ in the expansion

$$\tilde{K}_{h,\Delta t}(t_j) = \sum_{m,n=1}^{N_h} k_{j,m,n} \phi_m^h \otimes \phi_n^h,$$

where $(\phi_m^h)_{m=1}^{N_h}$ is the usual 'hat function' basis of V_h with $N_h = \dim(V_h)$. Furthermore, $(\mathbf{M}_h)_{i,j} = \langle \phi_i^h, \phi_j^h \rangle_H$, $(\mathbf{A}_h)_{i,j} = a(\phi_i^h, \phi_j^h)$ and $(\mathbf{Q}_h)_{i,j} = \langle Q\phi_i^h, \phi_j^h \rangle_H$ for $i,j = 1, \ldots, N_h$. We solve this system of matrix equations for two choices of Q and decreasing values of $h, \Delta t$.
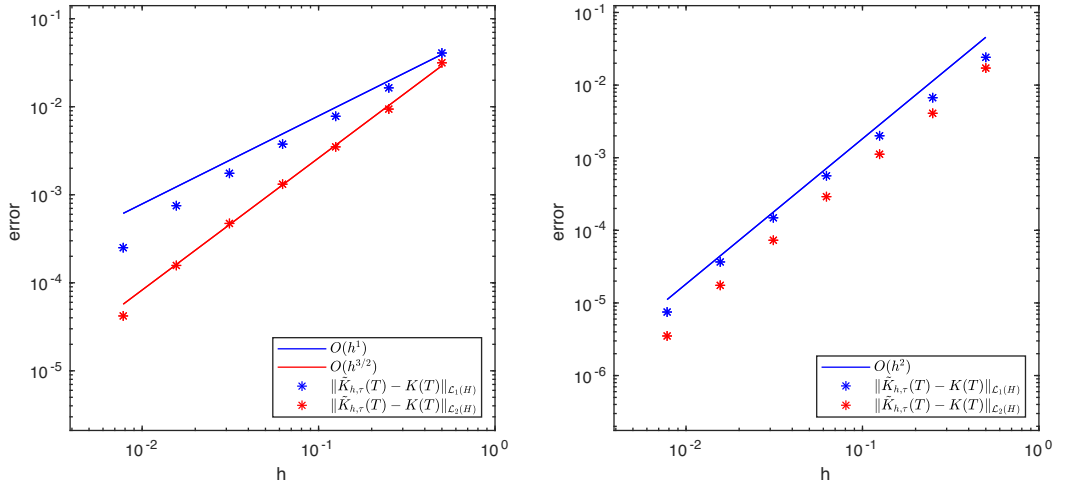
First, we consider the white noise case, i.e., $Q = I$. By our choice of b, the operator A_0 retains the Neumann boundary conditions of A. Lemma 2.3 in Kovács *et al.* (2021) then implies that Assumption 3.2 holds for all $r < 1/2$. By Theorem 3.1, we therefore expect to see a convergence rate essentially of order 1 and 3/2, respectively, if we plot the errors $\|K(T) - \tilde{K}_{h,\Delta t}(T)\|_{\mathscr{L}_i(H)}$, $i \in \{1, 2\}$, for $h = \sqrt{\Delta t} = 2^{-1}, \ldots, 2^{-7}$. This agrees with the results of Fig. 3(a).

In place of $K(T)$ we used $\tilde{K}_{h',\Delta t'}(T)$ at $h' = \sqrt{\Delta t'} = 2^{-8}$. The errors were computed by

$$\|\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)\|_{\mathscr{L}_1(H)} = \mathrm{Tr}\left(\left| \mathbf{N}_{h,h'}^{\frac{1}{2}} \begin{bmatrix} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{bmatrix} \mathbf{N}_{h,h'}^{\frac{1}{2}} \right| \right), \tag{3.14}$$

where the absolute value of an operator Γ is defined by $|\Gamma| = (\Gamma^*\Gamma)^{1/2}$, and

$$\|\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)\|_{\mathscr{L}_2(H)}^2 = \mathrm{Tr}\left((\mathbf{K}_{N_{\Delta t},h,\Delta t}\mathbf{M}_h)^2 \right) - 2\mathrm{Tr}\left(\mathbf{K}_{N_{\Delta t},h,\Delta t}\mathbf{M}_{h,h'}\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\mathbf{M}_{h',h} \right)$$

$$+ \mathrm{Tr}\left(\left(\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\mathbf{M}_{h'} \right)^2 \right). \tag{3.15}$$

(a) Errors with white noise.

(b) Errors with exponential kernel noise.

FIG. 3. Approximate errors $\|K(T) - \tilde{K}_{h,\Delta t}(T)\|_{\mathscr{L}_i(H)}$, $i \in \{1,2\}$, for the equations of Example 3.2.

Here, $(\mathbf{M}_{h,h'})_{i,j} = \langle \phi_i^h, \phi_j^{h'} \rangle_H$ for $i = 1, \ldots, N_h$, $j = 1, \ldots, N_{h'}$ and

$$\mathbf{N}_{h,h'} = \begin{bmatrix} \mathbf{M}_h & \mathbf{M}_{h,h'} \\ \mathbf{M}_{h',h} & \mathbf{M}_{h'} \end{bmatrix}.$$

The formula (3.15) is a straightforward consequence of the definition of the Hilbert–Schmidt norm while (3.14) requires a more detailed argument. We postpone this to an appendix.

Now, we consider the case that Q is an integral operator defined by

$$\langle Qu, v \rangle_H = \int_{\mathscr{D} \times \mathscr{D}} q(x,y) u(x) v(y) \, \mathrm{d}x \, \mathrm{d}y, \ u, v \in H.$$

We set $q(x,y) = \exp(-2|x-y|)$ for $x, y \in \mathscr{D}$. Fig. 1(a) shows an approximate realization of the solution to (3.1) for this q. Fig. 2(a) shows its covariance function, corresponding to $\tilde{K}_{h,\Delta t}(t)$, at $t = 0.1$. Since $\mathrm{Tr}(Q) < \infty$, Assumption 3.2 holds for $r \leqslant 1$, so we expect a rate of order 2 for both norms, in the same setup as before. Again, this agrees with Fig. 3(b).

REMARK 3.1 If one would instead directly compute an approximation $\tilde{X}_{h,\Delta t}$ based on the same semigroup approximation $\tilde{S}_{h,\Delta t}$ and compute its covariance by the method of Petersson (2020), we would directly get a bound on the covariance error $\|K(T) - \mathrm{Cov}(\tilde{X}_{h,\Delta t}(T))\|_{\mathscr{L}_i(H)}$ by the strong error $\|X(t) - \tilde{X}_{h,\Delta t}(T)\|_{L^2(\Omega,H)}$; see (1.4). Note that when $c_0 = 0$, $\mathrm{Cov}(\tilde{X}_{h,\Delta t}(T)) = \tilde{K}_{h,\Delta t}(T)$. In the first case above, when $Q = I$, this means that the error would be bounded by $h^r + k^{r/2}$ for $r < 1/2$, which is a lower rate than the estimates obtained above for both $i = 1$ and $i = 2$. See Yan (2005) for the calculations that yield this rate. Let us also note that our covariance error rate in the case that $i = 1$

coincides with that of the weak error $\left|\mathbb{E}\left[\phi\big(X(T)\big) - \phi\big(\tilde{X}_{h,\Delta t}(T)\big)\right]\right|$, where $\phi : H \to \mathbb{R}$ is a smooth functional; see Kovács *et al.* (2013). For the case $i = 2$, the covariance error rate exceeds the weak error rate.

3.2   *The stochastic wave equation*

Next, we apply our results to the stochastic wave equation perturbed by a linear inhomogeneity G, formally given by

$$\mathrm{d}\dot{U}(t,x) - \Delta U(t,x)\,\mathrm{d}t = GU(t,x)\,\mathrm{d}t + \mathrm{d}W(t,x), t \in [0,T], x \in \mathscr{D}. \qquad (3.16)$$

Here, $\dot{U}$ denotes the first time derivative of the unknown $H = L^2(\mathscr{D})$-valued stochastic process U with initial conditions $\dot{U}(0) = \psi$ and $U(0) = \chi$. With $G = -Q$, we obtain the DNA model (1.3) of the introduction. We again let W be a generalized Wiener process in H with covariance operator $Q \in \Sigma^+(H)$. We introduce precise assumptions on G, Q, u_0 and v_0 below, and write $\Lambda = -\Delta$ for the negative Laplacian with zero Dirichlet boundary conditions on H. The spaces $\dot{H}^\alpha$, along with fractional powers $\Lambda^{\alpha/2}$, can be defined as for the operator A_0 in Section 3.1.

In order to treat (3.16) in a semigroup framework, we define for $\alpha \in \mathbb{R}$ the Hilbert space $\mathscr{H}^\alpha = \dot{H}^\alpha \oplus \dot{H}^{\alpha-1}$ with inner product $\langle v, w \rangle_{\mathscr{H}^\alpha} = \langle v_1, w_1 \rangle_{\dot{H}^\alpha} + \langle v_2, w_2 \rangle_{\dot{H}^{\alpha-1}}$ for $v = [v_1, v_2]^\top$, $w = [w_1, w_2]^\top \in \mathscr{H}^\alpha$. Writing $\mathscr{H} = \mathscr{H}^0$, let $A : \mathrm{dom}(A) = \mathscr{H}^1 \to \mathscr{H}$, $B : \dot{H}^{-1} \to \mathscr{H}$ and $\Theta^{\alpha/2} : \mathscr{H}^\alpha \to \mathscr{H}, \alpha \in \mathbb{R}$, be given by

$$A = \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ I \end{bmatrix} \text{ and } \Theta^{\frac{\alpha}{2}} = \begin{bmatrix} \Lambda^{\frac{\alpha}{2}} & 0 \\ 0 & \Lambda^{\frac{\alpha}{2}} \end{bmatrix}.$$

The third operator is used to relate the norms of $\mathscr{H}^\alpha$ and $\mathscr{H}$ via $\|\cdot\|_{\mathscr{H}^\alpha} = \|\Theta^{\alpha/2}\cdot\|_{\mathscr{H}}$. Note that, since $\Lambda^{\alpha/2}$ extends to an operator in $\mathscr{L}(\dot{H}^s, \dot{H}^{s-\alpha})$ for all $\alpha, s \in \mathbb{R}$, so $\Theta^{\alpha/2}$ can be extended to an operator in $\mathscr{L}(\mathscr{H}^s, \mathscr{H}^{s-\alpha})$. We do so without changing notation. We write P_1 for the projection given by $P_1 v = v_1$ for $v = [v_1, v_2]^\top \in \mathscr{H}$. Note that $\Theta^{\alpha/2}B = B\Lambda^{\alpha/2}$. From this, we obtain the identities $\|\Theta^{\frac{\alpha}{2}}Bv\|_{\mathscr{H}} = \|\Lambda^{\frac{\alpha-1}{2}}v\|_H = \|v\|_{\dot{H}^{\alpha-1}}, v \in \dot{H}^{\alpha-1}$ and

$$\|B\|_{\mathscr{L}(\dot{H}^{-1}, \mathscr{H})} = \|B\Lambda^{\frac{1}{2}}\|_{\mathscr{L}(H, \mathscr{H})} = \|\Theta^{\frac{1}{2}}B\|_{\mathscr{L}(H, \mathscr{H})} = 1. \qquad (3.17)$$

The operator $-A$ is the generator of a C_0-semigroup (actually a group, see Lindgren, 2012) on $\mathscr{H}$ given by

$$S(t) = \begin{bmatrix} \cos(t\Lambda^{\frac{1}{2}}) & \Lambda^{-\frac{1}{2}}\sin(t\Lambda^{\frac{1}{2}}) \\ -\Lambda^{\frac{1}{2}}\sin(t\Lambda^{\frac{1}{2}}) & \cos(t\Lambda^{\frac{1}{2}}) \end{bmatrix} \text{ for } t \in \mathbb{R}.$$

It satisfies $\Theta^{\frac{\alpha}{2}}S(\cdot) = S(\cdot)\Theta^{\frac{\alpha}{2}}$ and $\Lambda^{\frac{\alpha}{2}}P_1 S(\cdot)B = P_1 S(\cdot)B\Lambda^{\frac{\alpha}{2}}$ for all $\alpha \in \mathbb{R}$. Moreover,

$$\|S(t)\|_{\mathscr{L}(\mathscr{H})} \leqslant 1 \text{ for } t \in \mathbb{R}. \qquad (3.18)$$

If we set $X(t) = [X_1(t), X_2(t)]^\top = [U(t, \cdot), \dot{U}(t, \cdot)]^\top$, (3.16) can be put in an abstract Itô form

$$\mathrm{d}X(t) + AX(t)\,\mathrm{d}t = FX(t)\,\mathrm{d}t + B\,\mathrm{d}W(t), t \in (0, T],$$

with $F = BGP_1$. We make the following assumptions on $G\colon H \to \dot{H}^{-1}$, $Q \in \Sigma^+(H)$ and the initial condition $X(0) = \xi = [\chi, \psi]^\top$.

ASSUMPTION 3.3 There is a constant $C < \infty$ such that

1. $\|G\|_{\mathscr{L}(H, \dot{H}^{-1})} = \|\Lambda^{-1/2}G\|_{\mathscr{L}(H)} \leqslant C$ and

2. $\|Q^{1/2}\|_{\mathscr{L}_2(H, \dot{H}^{-1})} = \|\Lambda^{-1/2}Q^{1/2}\|_{\mathscr{L}_2(H)} \leqslant C$.

    Moreover, ξ is an $\mathscr{H}$-valued $\mathscr{F}_0$-measurable Gaussian random variable.

    With this assumption in place, we can put the equation in the framework of Section 2. Combining (3.17) with (3.18) and Assumption 3.3, we find that $\|S(t)F\|_{\mathscr{L}(\mathscr{H})}$ and $\|S(t)B\|_{\mathscr{L}_2(U_0, \mathscr{H})}$ are bounded on $[0, T]$. By the strong continuity of S, Assumption 2.1 is fulfilled. Therefore, we can apply Propositions 2.1 and 2.2 along with Theorem 2.3 to obtain a unique solution K to (2.8) such that $K(t) = \mathrm{Cov}(X(t))$, $t \in [0, T]$.

    We now move on to approximations of K. We first consider temporally semidiscrete approximations based on a rational approximation of S. This is a stepping stone toward analyzing a fully discrete approximation. For a time step $\Delta t \in (0, 1]$, we again let $(t_j)_{j \in \mathbb{N}_0} \subset \mathbb{R}$ be given by $t_j = j\Delta t$ and $N_{\Delta t} + 1 = \inf\{j \in \mathbb{N} : t_j \notin [0, T]\}$. Let $R\colon \mathbb{C} \to \mathbb{C}$ be a rational function such that $|R(iy)| \leqslant 1$ for all $y \in \mathbb{R}$ and, for some approximation order $\rho \in \mathbb{N}$ and $C, b > 0$, $|R(iy) - e^{-iy}| \leqslant C|y|^{\rho+1}$ for all $y \in \mathbb{R}$ with $|y| \leqslant b$, where $i = \sqrt{-1}$. We write $S_{\Delta t}^n = R(\Delta t A)^n$ for the rational approximation of the operator $S(t_n)$. An interpolation $\tilde{S}_{\Delta t}\colon [0, T] \to \mathscr{L}(\mathscr{H})$ of the approximation is defined as in (3.8), and we say that $\tilde{S}_{\Delta t}$ approximates the semigroup S with order $\rho \in \mathbb{N}$. This framework covers many well-known temporal discretizations; cf. Baker & Bramble (1979, Section 4). Examples include the backward Euler scheme of the previous section (with order $\rho = 1$) as well as the Crank–Nicolson approximation (with order $\rho = 2$) given by $R(\Delta t A) = (1 - \Delta t A/2)(1 + \Delta t A/2)^{-1}$. The stability result

$$\|\tilde{S}_{\Delta t}(t)\|_{\mathscr{L}(\mathscr{H})} \leqslant 1, \tag{3.19}$$

for all $t \in [0, T]$, holds uniformly in $\Delta t \in (0, 1]$ (Kovács *et al.*, 2013, Section 4.2). Moreover, $\Theta^{\alpha/2}$ commutes with $\tilde{S}_{\Delta t}$ for all $\alpha \in \mathbb{R}$ and we have the following error estimate.

LEMMA 3.2 (Kovács *et al.*, 2013, Lemma 4.4). Assume that $\tilde{S}_{\Delta t}$ approximates S with order $\rho \in \mathbb{N}$. Then, for each $\alpha \geq 0$, there is a constant $C < \infty$ such that for all $\Delta t \in (0, 1]$,

$$\sup_{t \in [0, T]} \|(\tilde{S}_{\Delta t}(t) - S(t))\Theta^{-\frac{\alpha}{2}}\|_{\mathscr{L}(\mathscr{H})} \leqslant C\Delta t^{\min\left(\alpha\frac{\rho}{\rho+1}, 1\right)}.$$

We define a semidiscrete approximation $\tilde{K}_{\Delta t}$ of K by $\tilde{K}_{\Delta t}(0) = K(0) = Q_\xi$ and, for $j \geq 1$,

$$\begin{aligned}
\tilde{K}_{\Delta t}(t_j) &= \tilde{S}_{\Delta t}\tilde{K}_{\Delta t}(t_{j-1})\tilde{S}_{\Delta t}^* \\
&\quad + \Delta t \tilde{S}_{\Delta t}F\tilde{K}_{\Delta t}(t_{j-1})\tilde{S}_{\Delta t}^* + \Delta t \tilde{S}_{\Delta t}\tilde{K}_{\Delta t}(t_{j-1})(\tilde{S}_{\Delta t}F)^* + \Delta t \tilde{S}_{\Delta t}B(\tilde{S}_{\Delta t}B)^*.
\end{aligned} \tag{3.20}$$

This definition is then extended to arbitrary $t \in [0, T]$ by

$$
\begin{aligned}
\tilde{K}_{\Delta t}(t) = {}& \tilde{S}_{\Delta t}(t) Q_\xi \tilde{S}_{\Delta t}^*(t) \\
& + \int_0^t \tilde{S}_{\Delta t}(t-s) F \tilde{K}_{\Delta t}(\lfloor s \rfloor_{\Delta t}) \tilde{S}_{\Delta t}^*(t-s) + \tilde{S}_{\Delta t}(t-s) \tilde{K}_{\Delta t}(\lfloor s \rfloor_{\Delta t}) (\tilde{S}_{\Delta t}(t-s) F)^* \, \mathrm{d}s \\
& + \int_0^t \tilde{S}_{\Delta t}(t-s) B (\tilde{S}_{\Delta t}(t-s) B)^* \, \mathrm{d}s.
\end{aligned}
\tag{3.21}
$$

An additional assumption provides this approximation with some spatial regularity.

ASSUMPTION 3.4 There are constants $C < \infty$ and $r \geq 0$ such that, for some $i \in \{1, 2\}$,

1. $\|G\|_{\mathscr{L}(\dot{H}^r, \dot{H}^{r-1})} = \|\Lambda^{(r-1)/2} G \Lambda^{-r/2}\|_{\mathscr{L}(H)} \leqslant C$,

2. $\|Q\|_{\mathscr{L}_i(\dot{H}^1, \dot{H}^{r-1})} = \|\Lambda^{(r-1)/2} Q \Lambda^{-1/2}\|_{\mathscr{L}_i(H)} \leqslant C$ and

3. $\|Q_\xi\|_{\mathscr{L}_i(\mathscr{H}, \mathscr{H}^r)} = \|\Theta^{r/2} Q_\xi\|_{\mathscr{L}_i(\mathscr{H}, \mathscr{H})} \leqslant C$.

Before deriving the spatial regularity result in Lemma 3.3 below, we make some comments on these requirements. If Assumptions 3.4.2–3 are satisfied for $i = 1$ with a particular parameter $r > 0$, then they are also satisfied for $i = 2$ with the same parameter r. Moreover, for $i = 1$ and $r > 0$, Assumption 3.4.2 is strictly stronger than Assumption 3.3.2 (Kovács *et al.*, 2013, Lemma 4.1). Assumption 3.4.3 is fulfilled if the random variable ξ takes values in $\mathscr{H}^r$, in particular if χ and ψ with $\xi = [\chi, \psi]^\top$ are jointly Gaussian (taking values in $\dot{H}^r$ and $\dot{H}^{r-1}$, respectively) or if they are deterministic. To see this, note first that if ξ is a Gaussian $\mathscr{H}^r$-valued random variable, then it is also Gaussian in $\mathscr{H}$. Write $Q_{\xi, r}$ for the covariance of ξ in $\mathscr{H}^r$. This operator is related to the covariance of ξ in $\mathscr{H}$ by $Q_\xi = Q_{\xi, r} \Theta^{-r}$. By (2.6),

$$
\begin{aligned}
\|Q_\xi\|_{\mathscr{L}_1(\mathscr{H}^{-r}, \mathscr{H}^r)} = \|Q_{\xi, r} \Theta^{-r}\|_{\mathscr{L}_1(\mathscr{H}^{-r}, \mathscr{H}^r)} &\leqslant \|Q_{\xi, r}\|_{\mathscr{L}_1(\mathscr{H}^r, \mathscr{H}^r)} \|\Theta^{-r}\|_{\mathscr{L}(\mathscr{H}^{-r}, \mathscr{H}^r)} \\
&= \|Q_{\xi, r}\|_{\mathscr{L}_1(\mathscr{H}^r, \mathscr{H}^r)} < \infty.
\end{aligned}
$$

LEMMA 3.3 Let Assumptions 3.3 and 3.4 be satisfied for some $r \geq 0$ and $i \in \{1, 2\}$, and let $\tilde{S}_{\Delta t}$ approximate S with order $\rho \in \mathbb{N}$. Then, there is a constant $C < \infty$ such that for all $\Delta t \in (0, 1]$

$$
\sup_{t \in [0, T]} \|\tilde{K}_{\Delta t}(t)\|_{\mathscr{L}_i(\mathscr{H}, \mathscr{H}^r)} = \sup_{t \in [0, T]} \|\Theta^{\frac{r}{2}} \tilde{K}_{\Delta t}(t)\|_{\mathscr{L}_i(\mathscr{H})} \leqslant C.
$$

Proof. By applying the triangle inequality to (3.21), we obtain, for $n = 0, \ldots, N_{\Delta t}$,

$$
\|\Theta^{\frac{r}{2}} \tilde{K}_{\Delta t}(t_n)\|_{\mathscr{L}_i(\mathscr{H})} \leqslant \mathrm{I} + \Delta t \sum_{j=0}^{n-1} \left(\mathrm{II}_j + \mathrm{III}_j + \mathrm{IV}_j \right),
$$

where we have written $\mathrm{I} = \|\Theta^{r/2} \tilde{S}_{\Delta t}^n Q_\xi (\tilde{S}_{\Delta t}^n)^*\|_{\mathscr{L}_i(\mathscr{H})}$, $\mathrm{II}_j = \|\Theta^{r/2} \tilde{S}_{\Delta t}^{n-j} F K(t_j) (\tilde{S}_{\Delta t}^{n-j})^*\|_{\mathscr{L}_i(\mathscr{H})}$, $\mathrm{III}_j = \|\Theta^{r/2} \tilde{S}_{\Delta t}^{n-j} \tilde{K}_{\Delta t}(t_j) (\tilde{S}_{\Delta t}^{n-j} F)^*\|_{\mathscr{L}_i(\mathscr{H})}$ and $\mathrm{IV}_j = \|\Theta^{r/2} \tilde{S}_{\Delta t}^{n-j} B (\tilde{S}_{\Delta t}^{n-j} B)^*\|_{\mathscr{L}_i(\mathscr{H})}$. For the first term, the commutativity of $\Theta^{r/2}$ and $\tilde{S}_{\Delta t}^n$ along with (3.19) and Assumption 3.4.3 yields the existence of a

constant C, that does not depend on Δt or n, such that $\mathrm{I} \leqslant \|\tilde{S}_{\Delta t}^n\|_{\mathscr{L}(\mathscr{H})}^2 \|\Theta^{\frac{r}{2}}Q_\xi\|_{\mathscr{L}_i(\mathscr{H})} \leqslant C$. Similarly, using also (3.17), the fact that $\Theta^{r/2}B = B\Lambda^{r/2}$ and Assumption 3.4.1, we see that

$$
\begin{aligned}
\mathrm{II}_j &= \|\tilde{S}_{\Delta t}^{n-j}\Theta^{\frac{1}{2}}B\Lambda^{\frac{r-1}{2}}GP_1 K(t_j)(\tilde{S}_{\Delta t}^{n-j})^*\|_{\mathscr{L}_i(\mathscr{H})} \\
&\leqslant \|\tilde{S}_{\Delta t}^{n-j}\|_{\mathscr{L}(\mathscr{H})}^2 \|\Theta^{\frac{1}{2}}B\|_{\mathscr{L}(H,\mathscr{H})} \|\Lambda^{\frac{r-1}{2}}GP_1\tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(\mathscr{H},H)} \\
&\lesssim \|\Lambda^{\frac{r-1}{2}}G\Lambda^{-\frac{r}{2}}\|_{\mathscr{L}(H)} \|\Lambda^{\frac{r}{2}}P_1\tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(\mathscr{H},H)} \lesssim \|\Theta^{\frac{r}{2}}\tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(\mathscr{H},H)}.
\end{aligned}
$$

The term III_j is treated in the same way. Finally, we note that

$$
\mathrm{IV}_j = \|\Theta^{\frac{r}{2}}\tilde{S}_{\Delta t}^{n-j}B(\tilde{S}_{\Delta t}^{n-j}B)^*\|_{\mathscr{L}_i(\mathscr{H})} = \|\Theta^{\frac{r}{2}}\tilde{S}_{\Delta t}^{n-j}BQB^*(\tilde{S}_{\Delta t}^{n-j})^*\|_{\mathscr{L}_i(\mathscr{H})},
$$

where the adjoint of the expression involving B is taken with respect to $\mathscr{L}(U_0, \mathscr{H})$ on the left-hand side of the last equality and with respect to $\mathscr{L}(H, \mathscr{H})$ on the right-hand side. Using (2.6), (3.17), (3.19) and Assumption 3.4.2 yields the existence of a constant $C < \infty$, independent of Δt, n and j, such that

$$
\mathrm{IV}_j \lesssim \|\tilde{S}_{\Delta t}^{n-j}\|_{\mathscr{L}(\mathscr{H})}^2 \|B\Lambda^{\frac{1}{2}}\|_{\mathscr{L}(H,\mathscr{H})}^2 \|\Lambda^{\frac{r-1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathscr{L}_i(H)} \leqslant C.
$$

By an appeal to the discrete Gronwall lemma, we find that there is a constant $C < \infty$, independent of Δt, such that $\|\Theta^{\frac{r}{2}}\tilde{K}_{\Delta t}(t_n)\|_{\mathscr{L}_i(\mathscr{H})} \leqslant C$ for $n = 0, \ldots, N_{\Delta t}$. The supremum bound over $[0, T]$ follows by using this result in the representation (3.21). $\qquad\square$

Proposition 2.4 and Lemmas 3.2 and 3.3 yield a semidiscrete error bound.

THEOREM 3.3 Let Assumptions 3.3 and 3.4 be satisfied for some $r \geq 0$ and $i \in \{1, 2\}$. Let $\tilde{S}_{\Delta t}$ approximate S with order $\rho \in \mathbb{N}$. Then, there is a constant $C < \infty$ such that for all $\Delta t \in (0, 1]$

$$
\sup_{t \in [0,T]} \|K(t) - \tilde{K}_{\Delta t}(t)\|_{\mathscr{L}_i(\mathscr{H})} \leqslant C\Delta t^{\min\left(r\frac{\rho}{\rho+1}, 1\right)}.
$$

Proof. We write $\mathrm{I}, \mathrm{II}, \ldots, \mathrm{V}$ for the five terms appearing in the error decomposition of $\|K(t_n) - \tilde{K}_{\Delta t}(t_n)\|_{\mathscr{L}_i(\mathscr{H})}$ in Proposition 2.4, with $n = 0, \ldots, N_{\Delta t}$. From Lemma 3.2, (3.19), (3.18) and the commutativity of $S(t_n)$, $\tilde{S}_{\Delta t}(t_n)$ and $\Theta^{r/2}$, we find that I is bounded by

$$
\|(\tilde{S}_{\Delta t}(t_n) - S(t_n))\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} \|\Theta^{\frac{r}{2}}Q_\xi\|_{\mathscr{L}_i(\mathscr{H})} \left(\|\tilde{S}_{\Delta t}(t_n)\|_{\mathscr{L}(\mathscr{H})} + \|S(t_n)\|_{\mathscr{L}(\mathscr{H})}\right) \lesssim \Delta t^{\min\left(r\frac{\rho}{\rho+1}, 1\right)}.
$$

For the next three terms, we also use Assumption 3.4.1 (with $r = 0$) along with (3.17) and (3.18) to see that

$$
\mathrm{II} \lesssim \|B\Lambda^{\frac{1}{2}}\|_{\mathscr{L}(H,\mathscr{H})} \|\Lambda^{-\frac{1}{2}}G\|_{\mathscr{L}(H)} \Delta t \sum_{j=0}^{n-1} \|K(t_j) - \tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(\mathscr{H})}
$$

and that the terms III and IV may be bounded by a constant multiplied by

$$\sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|(\tilde{S}_{\Delta t}(t_n - s) - S(t_n - s))\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} \|\Theta^{\frac{r}{2}} \tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(\mathscr{H})} \, ds \lesssim \Delta t^{\min\left(r\frac{\rho}{\rho+1}, 1\right)},$$

where Lemmas 3.2 and 3.3 were used. The last term is treated like IV$_j$ in Lemma 3.3, so

$$\text{V} \leqslant \int_0^{t_n} \|(\tilde{S}_{\Delta t}(t_n - s) - S(t_n - s))\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} \|B\Lambda^{\frac{1}{2}}\|_{\mathscr{L}(H,\mathscr{H})}^2 \|\Lambda^{\frac{r-1}{2}} Q\Lambda^{-\frac{1}{2}}\|_{\mathscr{L}_i(H)} \, ds \lesssim \Delta t^{\min\left(r\frac{\rho}{\rho+1}, 1\right)}.$$

The proof is completed by first using the discrete Gronwall lemma and then extending the resulting bound to $[0, T]$, as in the proof of Lemma 3.3. □

Next, this result is used in a convergence analysis of a fully discrete approximation to (2.8). Let $(V_h^\kappa)_{h \in (0,1]} \subset \dot{H}^1$, $\kappa \in \{2, 3\}$ be a standard family of finite element function spaces consisting of continuous piecewise polynomials of degree $\kappa - 1$, with respect to a regular family of triangulations of $\mathscr{D}$ with maximal mesh size h, that are zero on the boundary of $\mathscr{D}$. They are equipped with the inner product $\langle \cdot \cdot \rangle_{V_h^\kappa} = \langle \cdot, \cdot \rangle_H$. On this space, let a discrete counterpart $\Lambda_h \colon V_h^\kappa \to V_h^\kappa$ to Λ be defined by $\langle \Lambda_h v_h, u_h \rangle_H = \langle \Lambda^{\frac{1}{2}} v_h, \Lambda^{\frac{1}{2}} u_h \rangle_H = \langle v_h, u_h \rangle_{\dot{H}^1}$ for all $v_h, u_h \in V_h^\kappa$. By $P_h \colon \dot{H}^{-1} \to V_h^\kappa$, we denote the generalized orthogonal projector. We define $\mathscr{V}_h^\kappa = V_h^\kappa \oplus V_h^\kappa$, equipped with the same inner product as $\mathscr{H}$. With some abuse of notation, by the expression $P_h v$, $v = [v_1, v_2]^\top \in \mathscr{H}$, we denote the element $[P_h v_1, P_h v_2]^\top \in \mathscr{V}_h^\kappa$. Let

$$A_h = \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix}$$

be a discrete counterpart to A on $\mathscr{V}_h^\kappa$, and set $S_{h,\Delta t}^k = R(\Delta t A_h)^k P_h$ for $k = 0, \ldots, n$ with a step function extension $\tilde{S}_{h,\Delta t} \colon [0, T] \to \mathscr{L}(\mathscr{H})$ defined in the same way as for $\tilde{S}_{\Delta t}$. The stability result $\|\tilde{S}_{h,\Delta t}(t)\|_{\mathscr{L}(\mathscr{H})} \leqslant 1$, $t \in [0, T]$, holds uniformly in $h, \Delta t \in (0, 1]$; see Kovács *et al.* (2013).

One could consider defining a fully discrete version of K in (2.8) by the semigroup approximation $\tilde{S}_{h,\Delta t}$ directly, obtaining from (2.5) an approximation of $\text{Cov}(U)$ via $\text{Cov}(U(t)) = P_1 \text{Cov}(X(t)) P_1^*$. The problem is that we only have access to an error bound in the first component of $\tilde{S}_{h,\Delta t}$ (see the next lemma). This means that we cannot use a Gronwall argument as in Theorem 3.3. Instead, we employ the approximation $\tilde{S}_{h,\Delta t}$ in an indirect way.

LEMMA 3.4 (Kovács *et al.*, 2013, Corollary 4.2). Let $\tilde{S}_{h,\Delta t}$ approximate S with order $\rho \in \mathbb{N}$ in time and $\kappa \in \{2, 3\}$ in space. Then, for each $\alpha \geq 0$, there is a constant $C < \infty$ such that, for all $h, \Delta t \in (0, 1]$,

$$\sup_{t \in [0,T]} \|P_1(\tilde{S}_{h,\Delta t}(t) - S(t))\Theta^{-\frac{\alpha}{2}}\|_{\mathscr{L}(\mathscr{H}, H)} \leq C \left(\Delta t^{\min\left(\alpha\frac{\rho}{\rho+1}, 1\right)} + h^{\min\left(\alpha\frac{\kappa}{\kappa+1}, \kappa\right)} \right).$$

We now define a perturbed temporally semidiscrete semigroup approximation $\hat{S}_{\Delta t} \colon [0, T] \to \mathscr{L}(\mathscr{H})$ by a step function extension (as in (3.8)) of $(R(\Delta t A)(I + \Delta t F))^n$, $n = 0, \ldots, N_{\Delta t}$. Similarly, we define a perturbed fully discrete semigroup approximation $\hat{S}_{h,\Delta t} \colon [0, T] \to \mathscr{L}(\mathscr{H})$ using

$(R(\Delta t A_h)P_h(I + \Delta t F))^n$. Note that, with $t_j \in [0, T]$, $u \in \mathscr{H}$, $\tilde{S}_{\Delta t}(t_j)u$ and $\tilde{S}_{h,\Delta t}(t_j)u$ are nothing but semidiscrete, and fully discrete approximations of $X(t_j)$ with initial value $\xi = u$ and noise covariance $Q = 0$. Using this, the following lemma is proven by a standard Gronwall argument as in, e.g., Kovács *et al.* (2020, Theorem 3.5), making use of Lemmas 3.2 and 3.4.

LEMMA 3.5 Let Assumption 3.4 be satisfied. Let $\hat{S}_{\Delta t} \colon [0, T] \to \mathscr{L}(\mathscr{H})$ and $\hat{S}_{h,\Delta t} \colon [0, T] \to \mathscr{L}(\mathscr{H})$ be step-function extensions of $(R(\Delta t A)(I + \Delta t F))^n$ and $(R(\Delta t A_h)P_h(I + \Delta t F))^n$ with approximation orders $\rho \in \mathbb{N}$ in time and $\kappa \in \{2, 3\}$ in space. For each $\alpha \in [0, r]$, there is a constant $C < \infty$ such that, for all $h, \Delta t \in (0, 1]$,

$$\sup_{t \in [0,T]} \|P_1(\hat{S}_{h,\Delta t}(t) - \hat{S}_{\Delta t}(t))\Theta^{-\frac{\alpha}{2}}\|_{\mathscr{L}(\mathscr{H},H)} \leqslant C\left(\Delta t^{\min\left(\alpha\frac{\rho}{\rho+1},1\right)} + h^{\min\left(\alpha\frac{\kappa}{\kappa+1},\kappa\right)}\right).$$

By the same Gronwall argument, there is a constant $C < \infty$ such that, for all $h, \Delta t \in (0, 1]$ and $t \in [0, T]$,

$$\|\hat{S}_{\Delta t}(t)\|_{\mathscr{L}(\mathscr{H})} \leqslant C \text{ and } \|\hat{S}_{h,\Delta t}(t)\|_{\mathscr{L}(\mathscr{H})} \leqslant C. \tag{3.22}$$

We can now define the fully discrete approximation of (2.8) and prove an error estimate with respect to the first component $\mathrm{Cov}(U(t))$, $t \in [0, T]$. It is denoted by $\tilde{K}_{h,\Delta t}$ and defined by $\tilde{K}_{h,\Delta t}(t_0) = P_h Q_\xi P_h$ and, for $j \geq 1$, $\tilde{K}_{h,\Delta t}(t_j) = \hat{S}_{h,\Delta t}\tilde{K}_{h,\Delta t}(t_{j-1})\hat{S}_{h,\Delta t}^* + \Delta t P_h B(P_h B)^*$. In closed form,

$$\tilde{K}_{h,\Delta t}(t_n) = \hat{S}_{h,\Delta t}^n Q_\xi (\hat{S}_{h,\Delta t}^n)^* + \Delta t \sum_{j=0}^{n-1} \hat{S}_{h,\Delta t}^{n-j-1} B(\hat{S}_{h,\Delta t}^{n-j-1}B)^*$$

$$= \hat{S}_{h,\Delta t}^n Q_\xi (\hat{S}_{h,\Delta t}^n)^* + \Delta t \sum_{j=0}^{n-1} \hat{S}_{h,\Delta t}^{n-j-1} BQB^*(\hat{S}_{h,\Delta t}^{n-j-1})^*,$$

where the adjoint of B is taken with respect to $\mathscr{L}(H, \mathscr{H})$ in the last expression. In Theorem 3.4, the error estimate for $\tilde{K}_{h,\Delta t}$ is given.

THEOREM 3.4 Let Assumptions 3.3 and 3.4 be satisfied for some $r \geq 0$ and $i \in \{1, 2\}$, and let $\tilde{K}_{h,\Delta t}$ be based on the perturbed semigroup approximation $\hat{S}_{h,\Delta t}$ with order $\rho \in \mathbb{N}$ in time and $\kappa \in \{2, 3\}$ in space. Then, there is a constant $C < \infty$ such that for all $h, \Delta t \in (0, 1]$

$$\sup_{n \in \{0,...,N_{\Delta t}\}} \|P_1(K(t_n) - \tilde{K}_{h,\Delta t}(t_n))P_1^*\|_{\mathscr{L}_i(\mathscr{H})} \leqslant C\left(\Delta t^{\min\left(r\frac{\rho}{\rho+1},1\right)} + h^{\min\left(r\frac{\kappa}{\kappa+1},\kappa\right)}\right).$$

Proof. First, we recall that $\hat{S}_{\Delta t} = \tilde{S}_{\Delta t}(I + \Delta t F)$, so that we may rewrite (3.20) as

$$\tilde{K}_{\Delta t}(t_j) = (\tilde{S}_{\Delta t} + \Delta t \tilde{S}_{\Delta t}F)\tilde{K}_{\Delta t}(t_{j-1})(\tilde{S}_{\Delta t} + \Delta t \tilde{S}_{\Delta t}F)^* - \Delta t^2 \tilde{S}_{\Delta t}F\tilde{K}_{\Delta t}(t_{j-1})F^*\tilde{S}_{\Delta t}^* + \Delta t \tilde{S}_{\Delta t}BQB^*\tilde{S}_{\Delta t}^*$$

$$= \hat{S}_{\Delta t}\tilde{K}_{\Delta t}(t_{j-1})\hat{S}_{\Delta t}^* - \Delta t^2 \tilde{S}_{\Delta t}F\tilde{K}_{\Delta t}(t_{j-1})F^*\tilde{S}_{\Delta t}^* + \Delta t \tilde{S}_{\Delta t}BQB^*\tilde{S}_{\Delta t}^*.$$

Iterating this equality and making use of (2.13) yields, for $n = 1, \ldots, N_{\Delta t}$,

$$\tilde{K}_{\Delta t}(t_n) = \hat{S}_{\Delta t}^n Q_\xi (\hat{S}_{\Delta t}^n)^* + \Delta t \sum_{j=0}^{n-1} \hat{S}_{\Delta t}^{n-j-1} BQB^* (\hat{S}_{\Delta t}^{n-j-1})^* + R_{\Delta t}(t_n),$$

where

$$R_{\Delta t}(t_n) = \frac{\Delta t}{2} \sum_{j=0}^{n-1} \hat{S}_{\Delta t}^{n-j-1} (\tilde{S}_{\Delta t} - I)(BQB^* - \Delta t F \tilde{K}_{\Delta t}(t_j) F^*)(\tilde{S}_{\Delta t} + I)^* (\hat{S}_{\Delta t}^{n-j-1})^*$$

$$+ \frac{\Delta t}{2} \sum_{j=0}^{n-1} \hat{S}_{\Delta t}^{n-j-1} (\tilde{S}_{\Delta t} + I)(BQB^* - \Delta t F \tilde{K}_{\Delta t}(t_j) F^*)(\tilde{S}_{\Delta t} - I)^* (\hat{S}_{\Delta t}^{n-j-1})^*$$

$$- \Delta t^2 \sum_{j=0}^{n-1} \hat{S}_{\Delta t}^{n-j-1} F \tilde{K}_{\Delta t}(t_j) F^* (\hat{S}_{\Delta t}^{n-j-1})^* = \mathrm{I} + \mathrm{II} + \mathrm{III}.$$

Adding and subtracting $\tilde{K}_{\Delta t}(t_n)$ in $K(t_n) - \tilde{K}_{h,\Delta t}(t_n)$, we therefore obtain the split

$$\|P_1(K(t_n) - \tilde{K}_{h,\Delta t}(t_n))P_1^*\|_{\mathscr{L}_i(H)}$$

$$\leqslant \|P_1(K(t_n) - \tilde{K}_{\Delta t}(t_n))P_1^*\|_{\mathscr{L}_i(H)} + \|P_1(\hat{S}_{h,\Delta t}^n - \hat{S}_{\Delta t}^n)Q_\xi (\hat{S}_{h,\Delta t}^n + \hat{S}_{\Delta t}^n)^* P_1^*\|_{\mathscr{L}_i(H)}$$

$$+ \left\| \Delta t \sum_{j=0}^{n-1} P_1(\hat{S}_{h,\Delta t}^{n-j-1} - \hat{S}_{\Delta t}^{n-j-1}) BQB^* (\hat{S}_{h,\Delta t}^{n-j-1} + \hat{S}_{\Delta t}^{n-j-1})^* P_1^* \right\|_{\mathscr{L}_i(H)}$$

$$+ \|P_1 \mathrm{I} P_1^*\|_{\mathscr{L}_i(H)} + \|P_1 \mathrm{II} P_1^*\|_{\mathscr{L}_i(H)} + \|P_1 \mathrm{III} P_1^*\|_{\mathscr{L}_i(H)}.$$

The first term of this split is treated by Theorem 3.3, noting that $P_1 \in \mathscr{L}(\mathscr{H}, H)$. For the second term, we use Lemma 3.5, Assumption 3.4.3 and (3.22) to see that

$$\|P_1(\hat{S}_{h,\Delta t}^n - \hat{S}_{\Delta t}^n)Q_\xi (\hat{S}_{h,\Delta t}^n + \hat{S}_{\Delta t}^n)^* P_1^*\|_{\mathscr{L}_i(\mathscr{H})}$$

$$\lesssim \|P_1(\hat{S}_{h,\Delta t}^n - \hat{S}_{\Delta t}^n)\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H},H)} \|\Theta^{\frac{r}{2}} Q_\xi\|_{\mathscr{L}_i(\mathscr{H})} \lesssim \Delta t^{\min(r\frac{\rho}{\rho+1}, 1)} + h^{\min\left(r\frac{\kappa}{\kappa+1}, \kappa\right)}.$$

For the third term, the same estimates combined with Assumption 3.4.2 and (3.17) yield the bound

$$\Delta t \sum_{j=0}^{n-1} \|P_1(\hat{S}_{h,\Delta t}^{n-j-1} - \hat{S}_{\Delta t}^{n-j-1})\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H},H)} \|\Lambda^{\frac{r-1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\mathscr{L}_i(H)} \lesssim \Delta t^{\min\left(r\frac{\rho}{\rho+1}, 1\right)} + h^{\min\left(r\frac{\kappa}{\kappa+1}, \kappa\right)}.$$

For $\|P_1 \mathrm{I} P_1^*\|_{\mathscr{L}_i(H)}$ and $\|P_1 \mathrm{II} P_1^*\|_{\mathscr{L}_i(H)}$, there is a constant $C < \infty$ such that for all $\Delta t \in (0, 1]$

$$\|(\tilde{S}_{\Delta t} - I)\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} \leqslant \|(\tilde{S}_{\Delta t} - S(\Delta t))\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} + \|(S(\Delta t) - S(0))\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} \leqslant C\Delta t^{\min\left(r\frac{\rho}{\rho+1}, 1\right)}.$$

Here, we have used the Hölder regularity of the semigroup S (Kovács *et al.*, 2013, Lemma 4.2) and Lemma 3.2. Using this estimate, we obtain for the term $\|P_1 \mathrm{I} P_1^*\|_{\mathscr{L}_i(H)}$ that

$$\|P_1 \mathrm{I} P_1^*\|_{\mathscr{L}_i(H)} \leqslant \frac{\Delta t}{2} \sum_{j=0}^{n-1} \|P_1 \hat{S}_{\Delta t}^{n-j-1}\|_{\mathscr{L}(\mathscr{H},H)}^2 \|(\tilde{S}_{\Delta t}-I)\Theta^{-\frac{r}{2}}\|_{\mathscr{L}(\mathscr{H})} \|B\Lambda^{\frac{1}{2}}\|_{\mathscr{L}(H,\mathscr{H})}^2 \|\tilde{S}_{\Delta t}+I\|_{\mathscr{L}(\mathscr{H})}$$

$$\times \left(\|\Lambda^{\frac{r-1}{2}}Q\Lambda^{-\frac{1}{2}}\|_{\mathscr{L}_i(H)} + \Delta t \|\Lambda^{\frac{r-1}{2}}G\Lambda^{-\frac{r}{2}}\|_{\mathscr{L}(H)} \|\Theta^{\frac{r}{2}}\tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(H)} \|\Lambda^{-\frac{1}{2}}G\|_{\mathscr{L}(H)}\right)$$

$$\lesssim \Delta t^{\min\left(r\frac{\rho}{\rho+1},1\right)},$$

with an analogous result for $\|P_1 \mathrm{II} P_1^*\|_{\mathscr{L}_i(H)}$. Here, we also made use of Lemma 3.3. This is also used to estimate the final term in the next calculation, completing the proof by Assumption 3.4.3 and (3.17) via

$$\|P_1 \mathrm{III} P_1^*\|_{\mathscr{L}_i(H)} \leqslant \Delta t^2 \sum_{j=0}^{n-1} \|P_1 \hat{S}_{\Delta t}^{n-j-1}\|_{\mathscr{L}(\mathscr{H},H)}^2 \|\Lambda^{-\frac{1}{2}}G\|_{\mathscr{L}(H)}^2 \|\tilde{K}_{\Delta t}(t_j)\|_{\mathscr{L}_i(H)} \lesssim \Delta t.$$

$$\square$$

EXAMPLE 3.5 Finally, we demonstrate the theoretical results obtained above numerically in the case that $\mathscr{D} = (0,1), T = 1$. We again let ξ be deterministic. We set $G = -Q$ in order to obtain the model suggested in Dalang (2009) for the vertical movement of a DNA strand suspended in fluid. We consider the piecewise linear finite element method for our discretization in space and the Crank–Nicolson discretization for our temporal approximation, so that we may take $\rho = \kappa = 2$ in Theorem 3.4. We let Q be an integral operator with kernel q, as in the second part of Example 3.2.

First, we let $q(x,y) = q(x-y)$ be a Matérn covariance function, which, for $z \in \mathbb{R}$, is given by $q(z) = \sigma^2 2^{1-\nu} \Gamma(\nu)^{-1} (\sqrt{2\nu}z\rho^{-1})^\nu K_\nu(\sqrt{2\nu}z\rho^{-1})$ with parameters $\sigma = 10, \nu = 0.01$ and $\rho = 0.1$. Here, K_ν denotes the modified Bessel function of the second kind. Fig. 1(b) shows an approximate realization of the solution to (3.16) for this choice of q and Fig. 2(b) shows its covariance function, corresponding to $P_1 \tilde{K}_{h,\Delta t}(t) P_1^*$, at $t = 0.1$. Since the Fourier transform of q is proportional to $\xi \mapsto (1+\xi^2)^{-\nu-d/2} = (1+\xi^2)^{-0.501}$, the results of Kovács *et al.* (2021, Section 4) implies that $\|\Lambda^{\frac{r-1}{2}}Q\|_{\mathscr{L}_2(H)} < \infty$ for $r < 3/2$. Moreover, since $d = 1$, $\|\Lambda^{-\frac{1}{2}}\|_{\mathscr{L}_2(H)} = \|I\|_{\mathscr{L}_2(\dot{H}^1,H)} < \infty$. By (2.7), therefore, Assumption 3.4 is satisfied with $r < 3/2$ and $i \in \{1,2\}$. From Theorem 3.4, we expect to see a convergence rate of essentially 1 if we plot the errors $\|P_1(K(T) - \tilde{K}_{h,\Delta t}(T))P_1^*\|_{\mathscr{L}_i(H)}, i \in \{1,2\}$, with respect to decreasing values of $h = \Delta t$ in a log-log plot. This is in line with our observations in Fig. 2b that shows the errors for $h = \Delta t = 2^{-1}, \ldots, 2^{-8}$. We have again used a reference solution $\tilde{K}_{h,\Delta t}(T)$ at $h = \Delta t = 2^{-9}$.

Next, we set $q(x,y) = \min(x,y) - xy$; for $x, y \in \mathscr{D}$, i.e., the covariance function of a Brownian bridge on $\mathscr{D}$. By its Karhunen–Loève expansion, one obtains $Q = \Lambda^{-1}$. Since Λ^{-1} has eigenvalues $(j^2\pi^2)_{j=1}^\infty$, Assumption 3.4 is fulfilled for all $r < 7/2$ when $i = 2$ and for all $r < 3$ when $i = 1$. In either case, we expect to see a convergence rate of order 2 if we plot the errors $\|P_1(K(T) - \tilde{K}_{h,\Delta t}(T))P_1^*\|_{\mathscr{L}_i(H)}$, $i \in \{1,2\}$, with respect to decreasing values of $h = \sqrt{\Delta t}$ in a log-log plot. This is in line with Fig. 4(b), where we have plotted the errors with $h = \sqrt{\Delta t} = 2^{-1}, \ldots, 2^{-5}$. In place of $K(T)$, we have used $\tilde{K}_{h,\Delta t}(T)$ at $h = \sqrt{\Delta t} = 2^{-6}$.
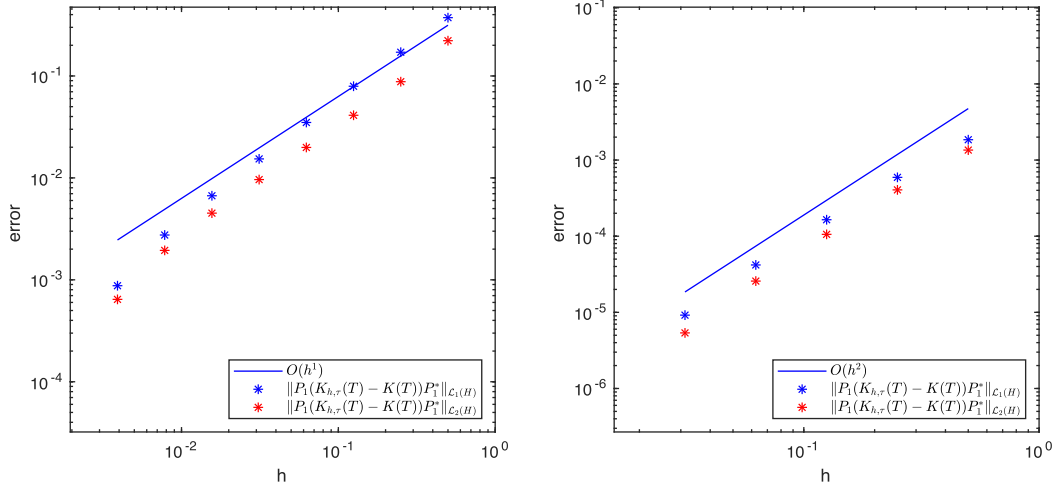
(a) Errors with Matérn noise for mesh sizes $h = \Delta t = 2^{-1}, \ldots, 2^{-8}$.

(b) Errors with Brownian bridge noise for mesh sizes $h = \sqrt{\Delta t} = 2^{-1}, \ldots, 2^{-5}$.

Fig. 4. Approximate errors $\|P_1(K(T) - \tilde{K}_{h,\Delta t}(T))P_1^*\|_{\mathscr{L}_i(H)}$, $i \in \{1, 2\}$, for the equations of Example 3.5.

REFERENCES

ANDERSSON, A., KRUSE, R. & LARSSON, S. (2016) Duality in refined Sobolev–Malliavin spaces and weak approximation of SPDE. *Stoch. Partial Differ. Equ. Anal. Comput.*, **4**, 113–149.

ANDERSSON, A. & LARSSON, S. (2016) Weak convergence for a spatial approximation of the nonlinear stochastic heat equation. *Math. Comp.*, **85**, 1335–1358.

BAKER, G. A. & BRAMBLE, J. H. (1979) Semidiscrete and single step fully discrete approximations for second order hyperbolic equations. *RAIRO Anal. Numér.*, **13**, 75–100.

BOLIN, D., KIRCHNER, K. & KOVÁCS, M. (2020) Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *IMA J. Numer. Anal.*, **40**, 1051–1073.

CHEVET, S. (1983) Compacité dans l'espace des probabilités de radon gaussiennes sur un Banach. *C. R. Acad. Sci. Paris Sér. I Math.*, **296**, 275–278.

DA PRATO, G., JENTZEN, A. & RÖCKNER, M. (2019) A mild Itô formula for SPDEs. *Trans. Amer. Math. Soc.*, **372**, 3755–3807.

DA PRATO, G. & ZABCZYK, J. (2014) *Stochastic Equations in Infinite Dimensions*, 2nd edn. Encyclopedia of Mathematics and Its Applications, vol. 152. Cambridge: Cambridge University Press.

DALANG, R. C. (2009) *The stochastic wave equation. A Minicourse on Stochastic Partial Differential Equations*. Lecture Notes in Mathematics, vol. 1962. Berlin: Springer, pp. 39–71.

FUJITA, H. & SUZUKI, T. (1991) Evolution problems: finite element methods. Part 1. *Handbook of Numerical Analysis*. Amsterdam: North-Holland

GRIGORIEFF, R. D. (1975) Diskrete approximation von Eigenwertproblemen. *Numer. Math.*, **25**, 79–97.

GRISVARD, P. (1985) *Elliptic Problems in Nonsmooth Domains*. Monographs and Studies in Mathematics, vol. 24. Boston, MA: Pitman (Advanced Publishing Program).

HERTERICH, K. & HASSELMANN, K. (1987) Extraction of mixed layer advection velocities, diffusion coefficients, feedback factors and atmospheric forcing parameters from the statistical analysis of North Pacific SST anomaly fields. *J. Phys. Oceanogr.*, **17**, 2145–2156.

KATO, T. (1961) Fractional powers of dissipative operators. *J. Math. Soc. Japan*, **13**, 246–274.

KIRCHNER, K. (2020) Numerical methods for the deterministic second moment equation of parabolic stochastic PDEs. *Math. Comp.*, **89**, 2801–2845.

KIRCHNER, K., LANG, A. & LARSSON, S. (2017) Covariance structure of parabolic stochastic partial differential equations with multiplicative Lévy noise. *J. Differential Equations*, **262**, 5896–5927.

KOVÁCS, M., LANG, A. & PETERSSON, A. (2020) Weak convergence of fully discrete finite element approximations of semilinear hyperbolic SPDE with additive noise. *ESAIM Math. Model. Numer. Anal.*, **54**, 2199–2227.

KOVÁCS, M., LANG, A. & PETERSSON, A. (2021) Hilbert–Schmidt regularity of symmetric integral operators on bounded domains with applications to SPDE approximations. To appear in *Stochastic Anal. Appl.* Available online at https://doi.org/10.1080/07362994.2022.2053541.

KOVÁCS, M., LARSSON, S. & LINDGREN, F. (2013) Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes. *BIT*, **53**, 497–525.

KRUSE, R. (2014) *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*. Lecture Notes in Mathematics, vol. 2093. Cham: Springer.

LANG, A., LARSSON, S. & SCHWAB, C. (2013) Covariance structure of parabolic stochastic partial differential equations. *Stoch. Partial Differ. Equ. Anal. Comput.*, **1**, 351–364.

LINDGREN, F. (2012) On weak and strong convergence of numerical approximations of stochastic partial differential equations. *Ph.D. Thesis*. Gothenburg: Chalmers University of Technology.

LIONS, J.-L. (1962) Espaces d'interpolation et domaines de puissances fractionnaires d'opérateurs. *J. Math. Soc. Japan*, **14**, 233–241.

MANDELBAUM, A. (1984) Linear estimators and measurable linear transformations on a Hilbert space. *Z. Wahrscheinlichkeitstheorie verw Gebiete*, **65**, 385–397.

MIDER, M., SCHAUER, M. & VAN DER MEULEN, F. (2021) Continuous-discrete smoothing of diffusions. *Electron. J. Statist.*, **15**, 4295–4342.

PETERSSON, A. (2020) Rapid covariance-based sampling of linear SPDE approximations in the multilevel Monte Carlo method. *Monte Carlo and Quasi-Monte Carlo Methods* (B. Tuffin & P. L'Ecuyer eds). Springer Proceedings in Mathematics & Statistics, vol. 324. Cham: Springer, pp. 423–443.

TRÈVES, F. (1967) *Topological Vector Spaces, Distributions and Kernels*. New York–London: Academic Press.

YAGI, A. (2010) *Abstract Parabolic Evolution Equations and Their Applications*. Berlin: Springer.

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

A. Appendix A. A derivation of a trace-class error expression

This appendix serves to provide a derivation of the expression

$$\|\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)\|_{\mathscr{L}_1(H)} = \mathrm{Tr}\left(\left|\mathbf{N}_{h,h'}^{\frac{1}{2}}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}^{\frac{1}{2}}\right|\right), \qquad (A.1)$$

in the setting of Example 3.2. First, note that $\|\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)\|_{\mathscr{L}_1(H)} = \mathrm{Tr}(|\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)|)$. Next, let us write $\psi_j = \phi_j^h$ for $j = 1, \ldots, N_h$, $\psi_j = \phi_{j-N_h}^{h'}$ for $j = N_h + 1, \ldots, N_h + N_{h'}$ and

$$\Upsilon = \sum_{i,j=1}^{N_h+N_{h'}} \left(\mathbf{N}_{h,h'}^{-\frac{1}{2}}\left|\mathbf{N}_{h,h'}^{\frac{1}{2}}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}^{\frac{1}{2}}\right|\mathbf{N}_{h,h'}^{-\frac{1}{2}}\right)_{ij}\psi_i \otimes \psi_j.$$

Here, $\mathbf{N}_{h,h'}^{-1/2}$ denotes the pseudoinverse of $\mathbf{N}_{h,h'}^{1/2}$. Note that $(\mathbf{N}_{h,h'})_{ij} = \langle\psi_i,\psi_j\rangle$ so that $\mathbf{N}_{h,h'}, \mathbf{N}_{h,h'}^{1/2}$ and $\mathbf{N}_{h,h'}^{-1/2}$ are in $\Sigma^+(\mathbb{R}^{N_h+N_{h'}})$. We claim that $\Upsilon = |\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)|$. To see this, we first note that $\Upsilon \in \Sigma^+(H)$ as a consequence of the matrix of coefficients in this sum being an element of $\Sigma^+(\mathbb{R}^{N_h+N_{h'}})$. Thus, it suffices to show that $\Upsilon^2 = (\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T))^*(\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T))$. By a direct calculation using the definition of the tensor product $\otimes$ and symmetry of $\mathbf{N}_{h,h'}$, it follows that

$$\Upsilon^2 = \sum_{i,j=1}^{N_h+N_{h'}} \left(\mathbf{N}_{h,h'}^{-\frac{1}{2}}\left|\mathbf{N}_{h,h'}^{\frac{1}{2}}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}^{\frac{1}{2}}\right|\right.$$
$$\left.\times \mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'}^{\frac{1}{2}})}\left|\mathbf{N}_{h,h'}^{\frac{1}{2}}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}^{\frac{1}{2}}\right|\mathbf{N}_{h,h'}^{-\frac{1}{2}}\right)_{ij}\psi_i \otimes \psi_j, \qquad (A.2)$$

where $\mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'}^{1/2})} = \mathbf{N}_{h,h'}^{-1/2}\mathbf{N}_{h,h'}^{1/2} = \mathbf{N}_{h,h'}^{1/2}\mathbf{N}_{h,h'}^{-1/2}$ is the projection onto $\mathrm{im}(\mathbf{N}_{h,h'}^{1/2})$, the range of $\mathbf{N}_{h,h'}^{1/2}$. Since the kernels, and hence the ranges, of a matrix in $\Sigma^+(\mathbb{R}^{N_h+N_{h'}})$ and its symmetric positive semidefinite square root coincide, we have

$$\mathrm{im}\left(\mathbf{N}_{h,h'}\right) = \mathrm{im}\left(\mathbf{N}_{h,h'}^{\frac{1}{2}}\right)$$
$$\supset \mathrm{im}\left(\mathbf{N}_{h,h'}^{\frac{1}{2}}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}^{\frac{1}{2}}\right)$$
$$= \mathrm{im}\left(\left|\mathbf{N}_{h,h'}^{\frac{1}{2}}\left[\begin{matrix}\mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'}\end{matrix}\right]\mathbf{N}_{h,h'}^{\frac{1}{2}}\right|\right). \qquad (A.3)$$

Combining this with (A.2), we find that

$$\Upsilon^2 = \sum_{i,j=1}^{N_h+N_{h'}} \mathbf{U}_{ij}\psi_i \otimes \psi_j,$$

where $\mathbf{U} = \mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}\tilde{\boldsymbol{K}}\mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}$ and

$$\tilde{\boldsymbol{K}} = \left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array}\right]\mathbf{N}_{h,h'}\left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array}\right].$$

Moreover, since

$$\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T) = \sum_{i,j=1}^{N_h+N_{h'}} \left(\left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array}\right]\right)_{ij} \psi_i \otimes \psi_j,$$

it follows from a direct calculation that

$$(\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T))^*(\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)) = \sum_{i,j=1}^{N_h+N_{h'}} \tilde{K}_{ij}\psi_i \otimes \psi_j.$$

Note now that

$$\|(\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T))^*(\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)) - \Upsilon^2\|^2_{\mathscr{L}_2(H)}$$

$$= \sum_{i=1}^{\infty} \left\|\left((\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T))^*(\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)) - \Upsilon^2\right)e_i\right\|^2_H$$

$$= \sum_{i=1}^{\infty} \left\|\sum_{j,k=1}^{N_h+N_{h'}} (\tilde{\boldsymbol{K}}_{jk} - \mathbf{U}_{jk})\langle\psi_k, e_i\rangle_H \psi_j\right\|^2_H$$

$$= \sum_{j,k,\ell,m=1}^{N_h+N_{h'}} \sum_{i=1}^{\infty} (\tilde{\boldsymbol{K}}_{jk} - \mathbf{U}_{jk})\langle\psi_k, e_i\rangle_H (\tilde{\boldsymbol{K}}_{\ell m} - \mathbf{U}_{\ell m})\langle\psi_m, e_i\rangle_H \langle\psi_j, \psi_\ell\rangle_H$$

$$= \sum_{j,k,\ell,m=1}^{N_h+N_{h'}} (\tilde{\boldsymbol{K}}_{jk} - \mathbf{U}_{jk})\langle\psi_k, \psi_m\rangle_H (\tilde{\boldsymbol{K}}_{\ell m} - \mathbf{U}_{\ell m})\langle\psi_j, \psi_\ell\rangle_H$$

$$= \sum_{\ell,m=1}^{N_h+N_{h'}} \left(\mathbf{N}_{h,h'}\left(\tilde{\boldsymbol{K}} - \mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}\tilde{\boldsymbol{K}}\mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}\right)\mathbf{N}_{h,h'}\right)_{\ell m} (\tilde{\boldsymbol{K}}_{\ell m} - \mathbf{U}_{\ell m}) = 0,$$

where the last equality follows from the fact that

$$\mathbf{N}_{h,h'}\left(\tilde{\boldsymbol{K}} - \mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}\tilde{\boldsymbol{K}}\mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}\right)\mathbf{N}_{h,h'} = \mathbf{N}_{h,h'}\left(\tilde{\boldsymbol{K}} - \mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})}\tilde{\boldsymbol{K}}\right)\mathbf{N}_{h,h'}$$

$$= \mathbf{N}_{h,h'}(\mathbf{I} - \mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'})})\tilde{\boldsymbol{K}}\mathbf{N}_{h,h'} = \mathbf{N}_{h,h'}\mathbf{P}_{\mathrm{ker}(\mathbf{N}_{h,h'})}\tilde{\boldsymbol{K}}\mathbf{N}_{h,h'} = 0.$$

With this, we have established that $\Upsilon = |\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)|$. Finally, we obtain (A.1) via

$$
\begin{aligned}
\mathrm{Tr}(|\tilde{K}_{h,\Delta t}(T) - \tilde{K}_{h',\Delta t'}(T)|) &= \sum_{i,j=1}^{N_h+N_{h'}} \left(\mathbf{N}_{h,h'}^{-\frac{1}{2}} \left| \mathbf{N}_{h,h'}^{\frac{1}{2}} \left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array} \right] \mathbf{N}_{h,h'}^{\frac{1}{2}} \left| \mathbf{N}_{h,h'}^{-\frac{1}{2}} \right)_{ij} \langle \psi_j, \psi_i \rangle_H \right. \\
&= \mathrm{Tr}\left(\mathbf{N}_{h,h'}^{-\frac{1}{2}} \left| \mathbf{N}_{h,h'}^{\frac{1}{2}} \left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array} \right] \mathbf{N}_{h,h'}^{\frac{1}{2}} \left| \mathbf{N}_{h,h'}^{\frac{1}{2}} \right) \right. \\
&= \mathrm{Tr}\left(\mathbf{P}_{\mathrm{im}(\mathbf{N}_{h,h'}^{1/2})} \left| \mathbf{N}_{h,h'}^{\frac{1}{2}} \left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array} \right] \mathbf{N}_{h,h'}^{\frac{1}{2}} \right| \right) \right. \\
&= \mathrm{Tr}\left(\left| \mathbf{N}_{h,h'}^{\frac{1}{2}} \left[\begin{array}{cc} \mathbf{K}_{N_{\Delta t},h,\Delta t} & 0 \\ 0 & -\mathbf{K}_{N_{\Delta t'},h',\Delta t'} \end{array} \right] \mathbf{N}_{h,h'}^{\frac{1}{2}} \right| \right).
\end{aligned}
$$

Here, we made use of the cyclic property of the trace and (A.3).