ESAIM: M2AN 56 (2022) 1545–1578
https://doi.org/10.1051/m2an/2022054

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

NUMERICAL ANALYSIS FOR TIME-DEPENDENT ADVECTION-DIFFUSION PROBLEMS WITH RANDOM DISCONTINUOUS COEFFICIENTS

Andrea Barth[1] and Andreas Stein[2,*]

Abstract. As an extension to the well-established stationary elliptic partial differential equation (PDE) with random continuous coefficients we study a time-dependent advection-diffusion problem, where the coefficients may have random spatial discontinuities. In a subsurface flow model, the randomness in a parabolic equation may account for insufficient measurements or uncertain material procurement, while the discontinuities could represent transitions in heterogeneous media. Specifically, a scenario with coupled advection and diffusion coefficients that are modeled as sums of continuous random fields and discontinuous jump components are considered. The respective coefficient functions allow a very flexible modeling, however, they also complicate the analysis and numerical approximation of the corresponding random parabolic PDE. We show that the model problem is indeed well-posed under mild assumptions and show measurability of the pathwise solution. For the numerical approximation we employ a sample-adapted, pathwise discretization scheme based on a finite element approach. This semi-discrete method accounts for the discontinuities in each sample, but leads to stochastic, finite-dimensional approximation spaces. We ensure measurability of the semi-discrete solution, which in turn enables us to derive moments bounds on the mean-squared approximation error. By coupling this semi-discrete approach with suitable coefficient approximation and a stable time stepping, we obtain a fully discrete algorithm to solve the random parabolic PDE. We provide an overall error bound for this scheme and illustrate our results with several numerical experiments.

Mathematics Subject Classification. 65M60, 60H25, 60H30, 60H35, 35R60, 58J65, 35K10.

Received October 4, 2021. Accepted June 6, 2022.

1. Introduction

In this paper we are concerned with the well-posedness of a solution to a time-dependent advection-diffusion equation with discontinuous random coefficients and its numerical discretization. The random coefficient function is modeled by a continuous part and a discontinuous part, inspired by the unique characterization of the Lévy–Khinchine formula for Lévy processes. We adopt this idea to spatial domains and use set-valued random variables to propose coefficients with jumps occurring on lower-dimensional submanifolds. This generalizes the elliptic setting which has drawn great attention over the last decades. While many publications focus on numerical

Keywords and phrases. Flow in heterogeneous media, fractured media, porous media, jump-diffusion coefficient, non-continuous random fields, parabolic equation, finite element method, uncertainty quantification.

[1] Institute of Applied Analysis and Numerical Simulation/SimTech, University of Stuttgart, Stuttgart, Germany.

[2] Seminar for Applied Mathematics, ETH Zürich, Zürich, Switzerland.

*Corresponding author: andreas.stein@sam.math.ethz.ch

methods for continuous stochastic coefficients (see, *e.g.*, [1,4–7,12,16,17,23,29,33,38,39,43,45,46]), the literature on stochastic discontinuous coefficients or stochastic interface problems is sparse (see, *e.g.*, [28, 32, 47]). The reasons are manifold: Gaussian random fields are well-defined mathematical objects and their properties are well studied, simulation methods range from spectral approximations to Fourier methods (see, *e.g.*, [25,31,44]). In contrast, there is no general definition and approximation method for a discontinuous (Lévy) field. Moreover, standard numerical methods for random PDEs, like Monte Carlo-finite element or polynomial chaos approaches, work reasonably for continuous random coefficients, but perform poorly for stochastic interface problems.

This article provides a generalization to the "standard model" in uncertainty quantification, *i.e.* the stationary problem with continuous diffusion coefficient, and provides a thorough analysis of time-dependent advection-diffusion equations with discontinuous random coefficients: Based on the elliptic setting in [10], we introduce a parabolic PDE with advection and diffusion terms given by discontinuous random fields. As our first main result, we derive precise conditions that ensure existence and measurability of pathwise weak solutions. Furthermore, we provide moments bounds on the solution in suitable Lebesgue–Bochner spaces. The random coefficients in our parabolic model problem are in general infinite-dimensional stochastic objects, therefore we also discuss tractable coefficient approximations and derive the corresponding error bounds.

Having ensured well-posedness of the problem, we further address the question of appropriate space-time approximations. Standard finite elements for the spatial discretization are are generally not suitable due to the varying random interfaces in each sample. Hence, we employ a sample-adapted triangulation approach that aligns the finite element grid to each sample of the advection and diffusion coefficient. This means we obtain stochastic, finite-dimensional approximation spaces, which is a crucial difference to common numerical methods in uncertainty quantification, and therefore need special treatment. Our next main contribution is to show the measurability of this semi-discrete sample-adapted solution and to provide the corresponding mean-squared error bounds. As our discretization approach relies on random grids, pathwise convergences rates are random as well, however, we achieve a deterministic control on moments of the error. This paves the way, for instance, to combine the sample-adapted meshing with Monte Carlo methods for further statistical inference. Finally, we introduce a stable time-stepping scheme to obtain a fully discrete algorithm and to avoid numerical oscillations due to the random interfaces. A bound on the overall mean-squared error is provided and our theoretical findings are verified by numerous numerical experiments.

The paper is structured as follows: In Section 2 we state the parabolic model problem in a very general setting and derive existence, uniqueness and measurability for pathwise solutions under mild assumptions on the data. In the following section we define our particular jump-diffusion and jump-advection coefficient and show that the well-posedness theory from Section 2 applies in this setting. Furthermore, we discuss suitable approximations of the discontinuous coefficient functions. These approximations are used in Section 4 to develop a pathwise sample-adapted discretization scheme. This section contains our main results on measurability and convergence for the semi-discrete approximation. Thereafter, we introduce a temporal approximation as the last part of a fully tractable algorithm and we close with several one- and two-dimensional numerical experiments.

2. Parabolic initial-boundary value problems and their solutions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, $\mathbb{T} := [0, T]$ a time interval for some $T > 0$ and $\mathcal{D} \subset \mathbb{R}^d$, $d \in \{1, 2\}$ be a convex, polygonal domain with piecewise linear boundary. We consider the linear, random initial-boundary value problem

$$\begin{aligned} \partial_t u(\omega, x, t) + [Au](\omega, x, t) &= f(\omega, x, t) \quad &&\text{in} \quad \Omega \times \mathcal{D} \times (0, T], \\ u(\omega, x, 0) &= u_0(\omega, x) \quad &&\text{in} \quad \Omega \times \mathcal{D} \times \{0\}, \\ u(\omega, x, t) &= 0 \quad &&\text{on} \quad \Omega \times \partial\mathcal{D} \times \mathbb{T}, \end{aligned} \tag{2.1}$$

where $f : \Omega \times \mathcal{D} \times \mathbb{T} \to \mathbb{R}$ is a random source function and $u_0 : \Omega \times \mathcal{D} :\to \mathbb{R}$ denotes the random initial condition of the partial differential equation (PDE). Furthermore, A is a second order partial differential operator

$$[Au](\omega, x, t) = -\nabla \cdot (a(\omega, x)\nabla u(\omega, x, t)) + b(\omega, x) \cdot \nabla u(\omega, x, t) \tag{2.2}$$

for $(\omega, x, t) \in \Omega \times \mathcal{D} \times \mathbb{T}$ with

- a stochastic jump-diffusion coefficient $a : \Omega \times \mathcal{D} \to \mathbb{R}$ and
- a stochastic jump-advection coefficient $b : \Omega \times \mathcal{D} \to \mathbb{R}^d$.[1]

We base the analysis of Problem (2.1) on the standard Sobolev space $H^k(\mathcal{D})$ with the norm

$$\|v\|_{H^k(\mathcal{D})} := \left(\sum_{|\nu| \le k} \int_{\mathcal{D}} |D^\nu v(x)|^2 \mathrm{d}x \right)^{1/2} \quad \text{for} \quad k \in \mathbb{N},$$

where the $D^\nu = \partial_{x_1}^{\nu_1} \dots \partial_{x_d}^{\nu_d}$ is the mixed partial weak derivative (in space) with respect to the multi-index $\nu \in \mathbb{N}_0^d$. The seminorm corresponding to $H^k(\mathcal{D})$ is denoted by

$$|v|_{H^k(\mathcal{D})} := \left(\sum_{|\nu| = k} \int_{\mathcal{D}} |D^\nu v(x)|^2 \mathrm{d}x \right)^{1/2}.$$

The *fractional order Sobolev spaces* $H^s(\mathcal{D})$ for $s > 0$ are defined by the norm

$$\|v\|_{H^s(\mathcal{D})} := \|v\|_{H^{\lfloor s \rfloor}(\mathcal{D})} + |v|_{H^{s-\lfloor s \rfloor}(\mathcal{D})}, \quad |v|^2_{H^{s-\lfloor s \rfloor}(\mathcal{D})} := \int_{\mathcal{D}} \int_{\mathcal{D}} \frac{|v(x) - v(y)|^2}{|x-y|^{d+2(s-\lfloor s \rfloor)}} \mathrm{d}x \mathrm{d}y,$$

where $|\cdot|_{H^{s-\lfloor s \rfloor}(\mathcal{D})}$ is the so-called *Gagliardo seminorm*, see [19], and $\lfloor \cdot \rfloor : \mathbb{R} \to \mathbb{Z},\ s \mapsto \max(k \in \mathbb{Z}, k \le s)$ is the *floor operator*. Further, we define $H := L^2(\mathcal{D})$ and denote by C a generic positive constant which may change from one line to another. Whenever necessary, the dependence of C on certain parameters is made explicit.

On the domain $\mathcal{D}$, the existence of a bounded, linear operator $\gamma : H^s(\mathcal{D}) \to H^{s-1/2}(\partial\mathcal{D})$ with

$$\gamma : H^s(\mathcal{D}) \cap C^\infty(\overline{\mathcal{D}}) \to H^{s-1/2}(\partial\mathcal{D}), \quad v \mapsto \gamma v = v|_{\partial\mathcal{D}}$$

and

$$\|\gamma v\|_{H^{s-1/2}(\partial\mathcal{D})} \le C \|v\|_{H^s(\mathcal{D})} \tag{2.3}$$

for $s \in (1/2, 3/2)$, $v \in H^s(\mathcal{D})$ is ensured by the trace theorem, see for example [20], where $C = C(s, \mathcal{D}) > 0$ in equation (2.3) depends on the boundary of $\mathcal{D}$. Since we consider homogeneous Dirichlet boundary conditions on $\partial\mathcal{D}$, we may treat γ independently of ω and define the suitable solution space V as

$$V := H_0^1(\mathcal{D}) = \{v \in H^1(\mathcal{D}) |\ \gamma v \equiv 0\},$$

equipped with the $H^1(\mathcal{D})$-norm $\|v\|_V := \|v\|_{H^1(\mathcal{D})}$. Due to the homogeneous Dirichlet boundary conditions, the *Poincaré inequality* $\|v\|_H \le C|v|_{H^1(\mathcal{D})}$ holds with $C = C(\mathcal{D}) > 0$ for all $v \in V$. Hence, $\|\cdot\|_{H^1(\mathcal{D})}$ and $|\cdot|_{H^1(\mathcal{D})}$ are equivalent on V. Furthermore, by Jensen's inequality

$$\left(\sum_{l=1}^d |\partial_{x_l} v(x)| \right)^2 \le 2^{d-1} \sum_{l=1}^d (\partial_{x_l} v(x))^2, \quad x \in \mathcal{D}, \tag{2.4}$$

and hence $\| \sum_{l=1}^d |\partial_{x_l} v| \|_H^2 \le 2^{d-1} |v|^2_{H^1(\mathcal{D})}$ for any $v \in V$.

[1] We could extend the above model problem by including time-dependent diffusion and/or advection coefficients. If a and b are sufficiently smooth with respect to t, *i.e.* continuously differentiable in $\mathbb{T}$, the temporal convergence rates in Section 4.2 are not affected. The focus of this article, however, is on the numerical analysis of Problem (2.1) with coefficients that involve random spatial discontinuities, hence we assume for the sake of simplicity that a and b are time-independent.

We work on the Gelfand triplet $V \subset H \subset V' = H^{-1}(\mathcal{D})$, where $\mathcal{V}'$ denotes the topological dual of any vector space $\mathcal{V}$. As the coefficients a and b are given by random functions, suitable solutions u to Problem (2.1) are in general time-dependent V-valued random variables. To investigate the integrability of u with respect to $\mathbb{T}$ and the underlying probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, we need to introduce the space of *Bochner-integrable* functions.

Definition 2.1. Let (Y, Σ, μ) be a σ-finite and complete measure space, let $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ be a Banach space and define the norm $\|\cdot\|_{L^p(Y;\mathcal{X})}$ for a strongly measurable $\mathcal{X}$-valued function $\varphi : Y \to \mathcal{X}$ by

$$\|\varphi\|_{L^p(Y;\mathcal{X})} := \begin{cases} \left(\int_Y \|\varphi(y)\|_{\mathcal{X}}^p \mu(\mathrm{d}y)\right)^{1/p} & \text{for} \quad 1 \le p < +\infty, \\ \operatorname*{ess\,sup}_{y \in Y} \|\varphi(y)\|_{\mathcal{X}} & \text{for} \quad p = +\infty. \end{cases}$$

The corresponding space of Bochner-integrable random variables is given by

$$L^p(Y; \mathcal{X}) := \{\varphi : Y \to \mathcal{X} \text{ is strongly measurable and } \|\varphi\|_{L^p(Y;\mathcal{X})} < +\infty\}.$$

Furthermore, the space of all continuous functions $\varphi : Y \to \mathcal{X}$ is defined as

$$C(Y; \mathcal{X}) := \{\varphi : Y \to \mathcal{X} \text{ is continuous and } \sup_{y \in Y} \|\varphi(y)\|_{\mathcal{X}} < +\infty\}.$$

We are interested in the two particular cases that

– $(Y, \Sigma, \mu) = (\mathbb{T}, \mathcal{B}(\mathbb{T}), \mu_{\mathbb{T}})$, where $\mathcal{B}(\mathbb{T})$ is the Borel σ-algebra over $\mathbb{T}$ and $\mu_{\mathbb{T}}$ is the Lebesgue-measure on $\mathcal{B}(\mathbb{T})$,
– $(Y, \Sigma, \mu) = (\Omega, \mathcal{F}, \mathbb{P})$.

The space $L^p(\Omega; \mathcal{X})$ is commonly referred to as the *space of Bochner-integrable random variables*. For any $\varphi \in L^1(\mathbb{T}; \mathcal{X})$ we denote by $\partial_t \varphi \in L^1(\mathbb{T}; \mathcal{X})$ the *weak time derivative* of φ if for all $\xi \in C_c^\infty(\mathbb{T}; \mathbb{R})$

$$\int_0^T \partial_t \xi(t) \varphi(t) \mathrm{d}t = -\int_0^T \xi(t) \partial_t \varphi(t) \mathrm{d}t,$$

where $\partial_t \xi$ is the classical (in a strong sense) time derivative of ξ. The set $C_c^\infty(\mathbb{T}; \mathbb{R})$ consists of all functions $\xi \in C^\infty(\mathbb{T}; \mathbb{R})$ with compact support in $(0, T)$. We record the following useful lemma for the calculus in $L^2(\mathbb{T}; H)$ (more precisely in Sect. 4.2).

Lemma 2.2 ([22], Chap. 5.9, Thm. 2)**.** *Let $H = L^2(\mathcal{D})$ and $\varphi, \partial_t \varphi \in L^2(\mathbb{T}; H)$. Then, the mapping $\varphi : \mathbb{T} \to H$ is continuous,*

$$\varphi(t_2) = \varphi(t_1) + \int_{t_1}^{t_2} \partial_t \varphi(t) \mathrm{d}t, \quad \text{for all} \quad 0 \le t_1 \le t_2 \le T,$$

and it holds for $C = C(T) > 0$ that

$$\max_{t \in \mathbb{T}} \|\varphi(t)\|_H^2 \le C \left(\|\varphi\|_{L^2(\mathbb{T};H)}^2 + \|\partial_t \varphi\|_{L^2(\mathbb{T};H)}^2 \right).$$

Remark 2.3. We may as well consider non-homogeneous boundary conditions, that is $u(\omega, x, t) = g_1(\omega, x, t)$ for $g_1 : \Omega \times \partial\mathcal{D} \times \mathbb{T} \to \mathbb{R}$. The corresponding trace operator γ is still well defined provided that $g_1(\omega, \cdot, \cdot)$ can be extended almost surely to a function $\widetilde{g}_1(\omega, \cdot, \cdot) \in L^1(\mathbb{T}; H^1(\mathcal{D}))$ with $\partial_t \widetilde{g}_1(\omega, \cdot, \cdot) \in L^1(\mathbb{T}; H^{-1}(\mathcal{D}))$. Then, $u - \widetilde{g}_1 \in L^1(\mathbb{T}; V)$ may be regarded as a solution to the modified problem

$$\begin{aligned} \partial_t(u - \widetilde{g}_1)(\omega, x, t) + [A(u - \widetilde{g}_1)](\omega, x, t) &= f(\omega, x, t) - [A\widetilde{g}_1](\omega, x, t) - \partial_t \widetilde{g}_1(\omega, x, t) && \text{on} \quad \Omega \times \mathcal{D} \times \mathbb{T}, \\ (u - \widetilde{g}_1)(\omega, x, 0) &= u_0(\omega, x) - \widetilde{g}_1(\omega, x, 0) && \text{on} \quad \Omega \times \mathcal{D} \times \{0\}, \quad \text{and} \\ (u - \widetilde{g}_1)(\omega, x, t) &= 0 && \text{on} \quad \Omega \times \partial\mathcal{D} \times \mathbb{T}. \end{aligned}$$

But this is in fact a version of Problem (2.1) with modified source term and initial value (see also [22], Chap. 6.1).

We introduce the bilinear form associated to A in order to derive a weak formulation of the initial-boundary value Problem (2.1). For fixed $\omega \in \Omega$ and $t \in \mathbb{T}$, multiplying equation (2.1) with a test function $v \in V$ and integrating by parts yields the variational equation

$$\int_{\mathcal{D}} \partial_t u(\omega, x, t) v(x) \mathrm{d}x + B_\omega(u(\omega, \cdot, t), v) = F_{\omega,t}(v). \tag{2.5}$$

The bilinear form $B_\omega : V \times V \to \mathbb{R}$ is given by

$$B_\omega(u, v) = \int_{\mathcal{D}} a(\omega, x) \nabla u(x) \cdot \nabla v(x) + b(\omega, x) \cdot \nabla u(x) v(x) \mathrm{d}x = (a(\omega, \cdot), \sum_{l=1}^{d} \partial_{x_l} u \, \partial_{x_l} v) + (b(\omega, \cdot) \cdot \nabla u, v),$$

where $(\cdot, \cdot)$ denotes the $L^2(\mathcal{D})$-scalar product. The source term is transformed into the right hand side functional

$$F_{\omega,t} : V \to \mathbb{R}, \quad v \mapsto \int_{\mathcal{D}} f(\omega, x, t) v(x) \mathrm{d}x,$$

and the integrals with respect to $\partial_t u$ and f are understood as the duality pairings

$$\int_{\mathcal{D}} \partial_t u(\omega, x, t) v(x) \mathrm{d}x = {}_{V'}\langle \partial_t u(\omega, \cdot, t), v \rangle_V,$$
$$\int_{\mathcal{D}} f(\omega, x, t) v(x) \mathrm{d}x = {}_{V'}\langle f(\omega, \cdot, t), v \rangle_V.$$

Definition 2.4. For fixed $\omega \in \Omega$, the *pathwise weak solution* to Problem (2.1) is a function $u(\omega, \cdot, \cdot) \in L^2(\mathbb{T}; V)$ with $\partial_t u(\omega, \cdot, \cdot) \in L^2(\mathbb{T}; V')$ such that for $t \in \mathbb{T}$ and all $v \in V$,

$${}_{V'}\langle \partial_t u(\omega, \cdot, t), v \rangle_V + B_\omega(u(\omega, \cdot, t), v) = F_{\omega,t}(v), \quad u(\omega, \cdot, 0) = u_0(\omega, \cdot).$$

The following assumptions allow us to show existence and uniqueness of a pathwise weak solution to equation (2.1) and guarantee measurability of the solution map $u : \Omega \to L^2(\mathbb{T}; V)$.

Assumption 2.5. (i) *For each $x \in \mathcal{D}$, the mappings $\omega \mapsto a(\omega, x)$ and $\omega \mapsto b(\omega, x)$ are $\mathcal{F} - \mathcal{B}(\mathbb{R})$-measurable.*

(ii) *For all $\omega \in \Omega$ it holds that $a(\omega, \cdot) \in L^1(\mathcal{D})$ and*

$$a_-(\omega) := \operatorname*{ess\,inf}_{x \in \mathcal{D}} a(\omega, x) > 0, \quad a_+(\omega) := \operatorname*{ess\,sup}_{x \in \mathcal{D}} a(\omega, x) < +\infty.$$

(iii) *It holds that $f \in L^p(\Omega; L^2(\mathbb{T}; V')), u_0 \in L^p(\Omega; H)$ and $1/a_- \in L^q(\Omega; \mathbb{R})$, for some $p, q \in [1, \infty]$, such that $1/p + 1/q \leq 1$.*

(iv) *There are constants $\bar{b}_1, \bar{b}_2 \geq 0$ such that $\|b(\omega, x)\|_\infty \leq \min(\bar{b}_1 a(\omega, x), \bar{b}_2)$ holds for almost all $\omega \in \Omega$ and almost all $x \in \mathcal{D}$. Here $\| \cdot \|_\infty$ denotes the supremum norm in $\mathbb{R}^d$.*

Remark 2.6. Items (i) and (ii) imply measurability of the random variables $a_-, a_+ : \Omega \to \mathbb{R}$. For instance, $a \in L^1(\mathcal{D})$ and $a_+ < +\infty$ imply that $a \in L^n(\mathcal{D})$ for all $n \in \mathbb{N}$. Hence, $a_+(\omega) = \|a(\omega, \cdot)\|_{L^\infty(\mathcal{D})}$ may be written as the point-wise limit of the measurable functions $\|a(\omega, \cdot)\|_{L^n(\mathcal{D})}$ for $n \to \infty$, see *e.g.* Lemma 13.1 of [3].

Theorem 2.7. *For any $w \in L^2(\mathbb{T}; V)$ define the (pathwise) parabolic norm*

$$\|w\|_{*,t} := \left(\|w(\cdot, t)\|_H^2 + \int_0^t |w(\cdot, z)|_{H^1(\mathcal{D})}^2 \mathrm{d}z \right)^{1/2}, \quad t \in \mathbb{T}.$$

Under Assumption 2.5, for any $\omega \in \Omega$, there exists a unique pathwise weak solution $u(\omega,\cdot,\cdot) \in L^2(\mathbb{T};V) \cap C(\mathbb{T};H)$ to Problem (2.1) and $u : \Omega \to L^2(\mathbb{T};V)$, $\omega \mapsto u(\omega,\cdot,\cdot)$ is strongly measurable. Further, for any $r \in [1,(1/p + 1/q)^{-1}]$

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u\|_{*,t}^r\right)^{1/r} \leq C\left(1 + \|1/a_-\|_{L^q(\Omega;\mathbb{R})}\right)\left(\|u_0\|_{L^p(\Omega;H)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};V'))}\right) < +\infty, \tag{2.6}$$

with $C = C(\bar{b}, T, q) > 0$. Moreover, if $f \in L^p(\Omega; L^2(\mathbb{T};H))$, then for any $r \in [1,(1/p + (1/(2q))^{-1}]$

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u\|_{*,t}^r\right)^{1/r} \leq C(1 + \|1/a_-\|_{L^q(\Omega;\mathbb{R})}^{1/2})\left(\|u_0\|_{L^p(\Omega;H)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};H))}\right) < +\infty.$$

Proof. For fixed $\omega \in \Omega$, $a(\omega,\cdot) \in L^\infty(\mathcal{D})$ holds since $a(\omega,\cdot)$ is integrable with bounded supremum. Thus, the bilinear form $B_\omega : V \times V \to \mathbb{R}$ in equation (2.5) is continuous and coercive by Assumption 2.5 and existence and uniqueness of a pathwise weak solution $u(\omega,\cdot,\cdot) \in L^2(\mathbb{T};V) \cap C(\mathbb{T};H)$ to Problem (2.1) follows as for deterministic parabolic problems, see for instance Chapter 7.1 of [22] or Chapter 11 of [41].

Now define the space $\mathcal{X} := L^2(\mathbb{T};V) \times L^2(\mathbb{T};V')$ with norm $\|(y_1,y_2)\|_{\mathcal{X}} := \|y_1\|_{L^2(\mathbb{T};V)} + \|y_2\|_{L^2(\mathbb{T};V')}$ and note that the mapping $\Omega \to \mathcal{X}$, $\omega \mapsto (u(\omega,\cdot,\cdot), \partial_t u(\omega,\cdot,\cdot))$ is well-defined. Let $(v_i, i \in \mathbb{N}) \subset V$ be a basis of V and for fixed $t \in \mathbb{T}$ and $i \in \mathbb{N}$ define the functional

$$J_i : \Omega \times \mathcal{X},\ (\omega, w) \mapsto \int_0^{\mathbb{T}} B_\omega(w(\cdot,t), v_i) - F_{\omega,t}(v_i) + {}_{V'}\langle \partial_t w(\cdot,t), v_i\rangle_V \mathrm{d}t.$$

By Assumption 2.5, it follows that J_i is a *Carathéodory map*, *i.e.* measurable in Ω and continuous in $\mathcal{X}$, and thus $\mathcal{F} \otimes \mathcal{B}(\mathcal{X}) - \mathcal{B}(\mathbb{R})$-measurable. The separability of $L^2(\mathbb{T};V)$ and $L^2(\mathbb{T};V')$ entails separability of $\mathcal{X}$ and, furthermore, $\mathcal{B}(\mathcal{X}) = \mathcal{B}(L^2(\mathbb{T};V)) \otimes \mathcal{B}(L^2(\mathbb{T};V'))$. To show the measurability of u, we define the correspondence

$$\varphi_i(\omega) := \{w \in \mathcal{X} |\, J_i(\omega, w) = 0\}, \quad \omega \in \Omega.$$

By Corollary 18.8 of [3] the graph $\mathrm{Gr}(\varphi_i) := \{(\omega, w) \in \Omega \times \mathcal{X} |\, w \in \varphi_i(\omega)\}$ is measurable, *i.e.* $\mathrm{Gr}(\varphi_i) \in \mathcal{F} \otimes \mathcal{B}(\mathcal{X})$. Since this yields

$$\{(\omega, u(\omega,\cdot,\cdot), \partial_t u(\omega,\cdot,\cdot)) |\, \omega \in \Omega\} = \bigcap_{i\in\mathbb{N}} \mathrm{Gr}(\varphi_i) \in \mathcal{F} \otimes \mathcal{B}(\mathcal{X}),$$

the mapping $\omega \to (u(\omega,\cdot,\cdot), \partial_t u(\omega,\cdot,\cdot))$ is $\mathcal{F} - \mathcal{B}(\mathcal{X})$-measurable (see *e.g.* [3], Thm. 18.25). As $\mathcal{B}(\mathcal{X}) = \mathcal{B}(L^2(\mathbb{T};V)) \otimes \mathcal{B}(L^2(\mathbb{T};V'))$, the marginal mappings $u : \Omega \to L^2(\mathbb{T};V)$ and $\partial_t u : \Omega \to L^2(\mathbb{T};V')$ are strongly $\mathcal{F} - \mathcal{B}(L^2(\mathbb{T};V))$-measurable and $\mathcal{F} - \mathcal{B}(L^2(\mathbb{T};V'))$-measurable, respectively. We note that it is sufficient to test against a basis of V in order to obtain the measurability of the $L^2(\mathbb{T};V')$-valued map $\partial_t u$, since the embeddings $V \hookrightarrow H \hookrightarrow V'$ are dense.

To show the estimate (2.6), we fix $\omega \in \Omega$, $t \in \mathbb{T}$, test against $v = u(\omega,\cdot,t) \in V$ in equation (2.5) and obtain

$${}_{V'}\langle \partial_t u(\omega,\cdot,t), u(\omega,\cdot,t)\rangle_V + B_\omega(u(\omega,\cdot,t), u(\omega,\cdot,t)) = F_{\omega,t}(u(\omega,\cdot,t)).$$

As $u(\omega,\cdot,\cdot) \in L^2(\mathbb{T};V)$ it holds that

$${}_{V'}\langle \partial_t u(\omega,\cdot,t), u(\omega,\cdot,t)\rangle_V = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(\omega,\cdot,t)\|_H^2,$$

see *i.e.* Chapter 5.9 of [22]. Rearranging the terms yields

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(\omega,\cdot,t)\|_H^2 + \left(a(\omega,\cdot)\,, \sum_{l=1}^d (\partial_{x_l} u(\omega,\cdot,t))^2\right) &= -(b(\omega,\cdot)\cdot\nabla u(\omega,\cdot,t), u(\omega,\cdot,t)) + F_{\omega,t}(u(\omega,\cdot,t)) \\ &=: \mathrm{I} + \mathrm{II}.\end{aligned} \tag{2.7}$$

The first term is bounded with Young's inequality, Assumption 2.5 and equation (2.4) *via*

$$\begin{aligned}\mathrm{I} &\leq \frac{2^{1-d}}{4\overline{b}_1}\||\|b(\omega,\cdot)\|_\infty^{1/2}\sum_{l=1}^d|\partial_{x_l}u(\omega,\cdot,t)|\|_H^2+2^{d-1}\overline{b}_1\|\|b(\omega,\cdot)\|_\infty^{1/2}u(\omega,\cdot,t)\|_H^2\\ &\leq \frac{1}{4}\left(a(\omega,\cdot)\,,\,\sum_{l=1}^d(\partial_{x_l}u(\omega,\cdot,t))^2\right)+2^{d-1}\overline{b}_1\overline{b}_2\|u(\omega,\cdot,t)\|_H^2.\end{aligned}$$

By the Poincaré inequality it holds that $\|u\|_V^2=|u|_{H^1(\mathcal{D})}^2+\|u\|_H^2\leq(1+C^2)|u|_{H^1(\mathcal{D})}^2$ and we estimate II by

$$\begin{aligned}\mathrm{II} &\leq (1+C^2)\frac{\|f(\omega,\cdot,t)\|_{V'}^2}{a_-(\omega)}+\frac{a_-(\omega)}{4(1+C^2)}\|u(\omega,\cdot,t)\|_V^2\\ &\leq (1+C^2)\frac{\|f(\omega,\cdot,t)\|_{V'}^2}{a_-(\omega)}+\frac{a_-(\omega)}{4}|u(\omega,\cdot,t)|_{H^1(\mathcal{D})}^2\\ &\leq (1+C^2)\frac{\|f(\omega,\cdot,t)\|_{V'}^2}{a_-(\omega)}+\frac{1}{4}\left(a(\omega,\cdot)\,,\,\sum_{l=1}^d(\partial_{x_l}u(\omega,\cdot,t))^2\right).\end{aligned}$$

Hence, equation (2.7) implies

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(\omega,\cdot,t)\|_H^2+\left(a(\omega,\cdot)\,,\,\sum_{l=1}^d(\partial_{x_l}u(\omega,\cdot,t))^2\right)\leq C\left(\frac{\|f(\omega,\cdot,t)\|_{V'}^2}{a_-(\omega)}+\|u(\omega,\cdot,t)\|_H^2\right).$$

We integrate over $\mathbb{T}$ and use Grönwall's inequality to obtain

$$\begin{aligned}\|u(\omega,\cdot,t)\|_H^2+a_-(\omega)\int_0^t|u(\omega,\cdot,z)|_{H^1(\mathcal{D})}^2\mathrm{d}z &\leq \|u(\omega,\cdot,t)\|_H^2+\int_0^t\left(a(\omega,\cdot)\,,\,\sum_{l=1}^d(\partial_{x_l}u(\omega,\cdot,z))^2\right)\mathrm{d}z\\ &\leq \exp(CT)\left(\|u_0(\omega,\cdot)\|_H^2+\frac{\|f(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};V')}^2}{a_-(\omega)}\right),\end{aligned}$$

where we emphasize that the last estimate is independent of t. If $a_-(\omega)\leq 1$ holds for fixed ω,

$$\begin{aligned}\sup_{t\in\mathbb{T}}\|u(\omega,\cdot,\cdot)\|_{*,t}^2 &= \sup_{t\in\mathbb{T}}\left(\|u(\omega,\cdot,t)\|_H^2+\int_0^t|u(\omega,\cdot,z)|_{H^1(\mathcal{D})}^2\mathrm{d}z\right)\\ &\leq \exp(CT)\left(\frac{\|u_0(\omega,\cdot)\|_H^2+\|f(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};V')}^2}{a_-^2(\omega)}\right).\end{aligned}$$

On the other hand, if $a_-(\omega)>1$, it follows that

$$\sup_{t\in\mathbb{T}}\|u(\omega,\cdot,\cdot)\|_{*,t}^2\leq\exp(CT)\left(\|u_0(\omega,\cdot)\|_H^2+\|f(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};V')}^2\right).$$

With the inequalities $\sqrt{c_1+c_2}\leq\sqrt{c_1}+\sqrt{c_2}$ and $(c_1+c_2)^r\leq 2^{r-1}(c_1^r+c_2^r)$ for $c_1,c_2\geq 0, r\geq 1$, and by taking expectations this yields for any $r\in[1,(1/p+1/q)^{-1}]$

$$\begin{aligned}\mathbb{E}\left(\sup_{t\in\mathbb{T}}\|u\|_{*,t}^r\right)^{1/r} &\leq C\mathbb{E}\left(\frac{\|u_0\|_H^r+\|f\|_{L^2(\mathbb{T};H)}^r}{a_-^r}\mathbb{1}_{\{a_-\leq 1\}}+\left(\|u_0\|_H^r+\|f\|_{L^2(\mathbb{T};V')}^r\right)\mathbb{1}_{\{a_->1\}}\right)^{1/r}\\ &\leq C(1+\|1/a_-\|_{L^q(\Omega;\mathbb{R})})\left(\|u_0\|_{L^p(\Omega;H)}+\|f\|_{L^p(\Omega;L^2(\mathbb{T};V'))}\right),\end{aligned}$$

where we have used Assumption 2.5 and Hölder's inequality for the last estimate.

For the second part of the claim, given that $f \in L^p(\Omega; L^2(\mathbb{T}; H))$, we may bound II *via*

$$\mathrm{II} \leq \frac{1}{2}\|f(\omega,\cdot,t)\|_H^2 + \frac{1}{2}\|u(\omega,\cdot,t)\|_H^2$$

and proceed as for the first term, using Grönwall's inequality, to obtain

$$\|u(\omega,\cdot,t)\|_H^2 + a_-(\omega)\int_0^t |u(\omega,\cdot,z)|_{H^1(\mathcal{D})}^2 \mathrm{d}z \leq C\left(\|u_0(\omega,\cdot)\|_H^2 + \|f(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T},H)}^2\right).$$

Finally, with Hölder's inequality it follows for any $r \in [1, (1/p + 1/(2q))^{-1}]$ that

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u\|_{*,t}^r\right)^{1/r} \leq C(1 + \|1/a_-\|_{L^q(\Omega;\mathbb{R})}^{1/2})\left(\|u_0\|_{L^p(\Omega;H)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};H))}\right).$$

□

To incorporate discontinuities at random submanifolds of $\mathcal{D}$, we introduce the jump-diffusion coefficient a and jump-advection coefficient b in the subsequent section. The introduced coefficients allow us to derive well-posedness and regularity results based on Theorem 2.7 for the solution to the parabolic problem with discontinuous coefficients.

3. Random parabolic problems with discontinuous coefficients

To obtain a stochastic jump-diffusion coefficient representing the permeability in a subsurface flow model, we use the random coefficient a from the elliptic diffusion problem in [11] consisting of a (spatial) Gaussian random field with additive discontinuities on random submanifolds of $\mathcal{D}$. The specific structure of a may be utilized to model the hydraulic conductivity within heterogeneous and/or fractured media and is thus considered time-independent (see also Rem. 2.3). The advection term in this model should then be driven by the same random field and inherit the same discontinuous structure as the diffusion term. Thus, we consider the coefficient b as an essentially linear mapping of a. Since the coefficients usually involve infinite series expansions in the Gaussian field and/or sampling errors in the jump measure, we further describe how to obtain tractable approximations of a and b. Subsequently, existence and stability results for weak solutions of the unapproximated resp. approximated parabolic problems based on Theorem 2.7 are proved. We conclude this section by showing that the approximated solution converges to the solution u of the (unapproximated) advection-diffusion problem in a suitable norm.

3.1. Jump-diffusion coefficients and their approximations

We start by introducing *measurable correspondences*, a concept that is very useful to define a suitable jump-diffusion coefficient in the following.

Definition 3.1 ([3], Def. 18.1)**.** Let $(\Omega_0, \mathcal{F}_0)$ be a measurable space, let $\mathcal{S}$ be a topological space, denote by $2^{\mathcal{S}}$ the power set of $\mathcal{S}$, and let $\mathcal{B}(\mathcal{S})$ be the Borel σ-algebra of $\mathcal{S}$. For any set valued mapping/correspondence $\psi : \Omega_0 \to 2^{\mathcal{S}}$, the *lower inverse* of ψ is defined as

$$\psi^{\ell i} : 2^{\mathcal{S}} \to \Omega_0, \quad S \mapsto \{\omega \in \Omega_0 | \psi(\omega) \cap S \neq \emptyset\}.$$

The correspondence $\psi : \Omega_0 \to 2^{\mathcal{S}}$ is

– *weakly measurable*, if for each open subset $S \subset \mathcal{S}$ it holds that $\psi^{\ell i}(S) \in \mathcal{F}_0$.
– *Borel-measurable*, if for each Borel set $B \subset \mathcal{B}(\mathcal{S})$ it holds that $\psi^{\ell i}(B) \in \mathcal{F}_0$.

Definition 3.2. The *jump-diffusion coefficient* a is defined as

$$a : \Omega \times \mathcal{D} \to \mathbb{R}_{>0}, \quad (\omega, x) \mapsto \overline{a}(x) + \Phi(W(\omega, x)) + P(\omega, x),$$

where

- $\overline{a} \in C^1(\overline{\mathcal{D}}; \mathbb{R}_{\geq 0})$ is non-negative, continuous and bounded.
- $\Phi \in C^1(\mathbb{R}; \mathbb{R}_{>0})$ is a continuously differentiable, positive mapping.
- $W \in L^2(\Omega; H)$ is a zero-mean Gaussian random field. Associated to W is a non-negative, symmetric trace class operator $Q : H \to H$.
- Let $\mathcal{T}_j : \Omega \to \mathcal{B}(\mathcal{D}) \subset 2^{\mathcal{D}}$, $j \in \mathbb{N}$, be a sequence of Borel-measurable correspondences and let $(P_j, j \in \mathbb{N})$ be a sequence of non-negative random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Let the sequence $\mathcal{T} := (\mathcal{T}_j, j \in \mathbb{N})$ be such that $\mathcal{T}(\omega)$ forms P-a.s. a finite partition of $\mathcal{D}$: Let $\mu : \mathcal{B}(\mathcal{D}) \to [0, \infty)$ be the Lebesgue-measure on $\mathcal{D}$ and define

$$\mathcal{I}_{\mathcal{T}}(\omega) := \{j \in \mathbb{N} | \, \mu(\mathcal{T}_j(\omega)) > 0\}, \quad \tau(\omega) := \#|\mathcal{I}_{\mathcal{T}}|.$$

Then, $\tau : \Omega \mapsto \mathbb{N}$, it holds P-a.s. that $\mathcal{T}_i(\omega)$ is an open set for $i \in \mathcal{I}_{\mathcal{T}}(\omega)$, $\mathcal{T}_i(\omega) \cap \mathcal{T}_j(\omega) = \emptyset$ for all $i \neq j$ and

$$\overline{\mathcal{D}} = \bigcup_{j \in \mathbb{N}} \overline{\mathcal{T}_j(\omega)} = \bigcup_{i \in \mathcal{I}_{\mathcal{T}}(\omega)} \overline{\mathcal{T}_i(\omega)}.$$

The jump part of a is then defined by

$$P : \Omega \times \mathcal{D} \to \mathbb{R}_{\geq 0}, \quad (\omega, x) \mapsto \sum_{i \in \mathcal{I}_{\mathcal{T}}(\omega)} \mathbb{1}_{\{\mathcal{T}_i\}}(x) P_i(\omega).$$

Based on a, the *jump-advection coefficient* b is given for vector fields $\widetilde{b}_1, \widetilde{b}_2 \in L^\infty(\mathcal{D})^d$ by

$$b : \Omega \times \mathcal{D} \to \mathbb{R}^d, \quad (\omega, x) \mapsto \min\left(a(\omega, x)\widetilde{b}_1(x), \widetilde{b}_2(x)\right).$$

Remark 3.3. By the finite partition assumption on the sequence $\mathcal{T}$, we have ensured that $P(\omega, x) \in \mathbb{R}_{\geq 0}$ for all $(\omega, x) \in \Omega \times \mathcal{D}$, therefore $a : \Omega \times \mathcal{D} \to \mathbb{R}_{>0}$ as above is well-defined. The measurability and integrability of a and b is shown in Lemma 3.6 below. Therein, we exploit the Borel-measurability of the correspondences $\mathcal{T}_j : \Omega \to \mathcal{B}(\mathcal{D}) \subset 2^{\mathcal{D}}$ and that $\mathcal{T}_j$ is $\mathcal{B}(\mathcal{D})$-valued. Note that the latter is *not implied* by Definition 3.1.

The definition of the jump-advection coefficient immediately implies Assumption 2.5 (iv) since

$$\|b(\omega, x)\|_\infty \leq \min\left(\overline{b}_1 a(\omega, x), \overline{b}_2\right)$$

holds with suitable constants $\overline{b}_1, \overline{b}_2 > 0$ for almost all $\omega \in \Omega$ and almost all $x \in \mathcal{D}$. The upper bound with respect to $\overline{b}_2$ is due to technical reasons and not restrictive in practical applications, as $\overline{b}_2$ may be arbitrary large.

The following assumptions guarantee that a and b are actually measurable mappings as in Assumption 2.5 and we may apply Theorem 2.7 also in the jump-diffusion setting.

Assumption 3.4. (i) *There exists $p > 1$ such that $f \in L^p(\Omega; L^2(\mathbb{T}; V'))$ and $u_0 \in L^p(\Omega; H)$.*

(ii) *The eigenfunctions e_i of Q are continuously differentiable on $\mathcal{D}$ and there exist constants $\alpha, \beta, C_e, C_\eta > 0$ such that for any $i \in \mathbb{N}$*

$$\|e_i\|_{L^\infty(\mathcal{D})} \leq C_e, \quad \max_{l=1,\ldots,d} \|\partial_{x_l} e_i\|_{L^\infty(\mathcal{D})} \leq C_e i^\alpha \quad \text{and} \quad \sum_{i=1}^\infty \eta_i i^\beta \leq C_\eta < +\infty.$$

(iii) *Furthermore, the mapping Φ as in Definition 3.2 and its derivative are bounded for $w \in \mathbb{R}$ by*

$$\phi_1 \exp(\phi_2 w) \geq \Phi(w) \geq \phi_1 \exp(-\phi_2 w), \quad |\frac{\mathrm{d}}{\mathrm{d}x}\Phi(w)| \leq \phi_3 \exp(\phi_4 |w|),$$

where $\phi_1, \dots, \phi_4 > 0$ are arbitrary constants.

(iv) *The sequence $(P_j, j \in \mathbb{N})$ consists of nonnegative and bounded random variables $P_j \in [0, \overline{P}]$ for some $\overline{P} > 0$. In addition, for $s > 1$ such that $1/p + 1/s < 1$ there exists a sequence of approximations $(\widetilde{P}_j, j \in \mathbb{N}) \subset [0, \overline{P}]^{\mathbb{N}}$ so that the sampling error is bounded, for some $\varepsilon > 0$, by*

$$\mathbb{E}(|\widetilde{P}_j - P_j|^s) \leq \varepsilon, \quad i \in \mathbb{N}.$$

Remark 3.5. The exponential bounds on Φ and its derivative imply that $u \in L^r(\Omega; L^2(\mathbb{T}; V))$ for any $r \in [1, p)$. That is, the integrability of u with respect to Ω only depends on the stochastic regularity of f and u_0. In fact, Theorem 2.7 shows that far weaker assumptions on a (resp. Φ) are possible to achieve $u \in L^r(\Omega; L^2(\mathbb{T}; V))$, at the cost that r then also depends on the integrability of a_-. At this point we refer to [11], where the regularity of an elliptic diffusion problem with a as in Definition 3.2, but less restricted functions Φ and P is investigated. However, Assumption 3.4 includes the important case that $\Phi(W)$ is a log-Gaussian random field and the bounds on Φ are merely imposed for a clear and simplified presentation of the results. On a further note, the assumptions on the eigenpairs $((\eta_i, e_i), i \in \mathbb{N})$ are natural and include the case that Q is a Matérn-type or Brownian-motion-type covariance function.

In general, the structure of a as in Definition 3.2 does not allow us to draw samples from the exact distribution of this random function. The Gaussian random field may be approximated by truncated Karhunen–Loève expansions: Let $((\eta_i, e_i), i \in \mathbb{N})$ denote the sequence of eigenpairs of Q, where $Q : H \to H$ is the covariance operator of the Gaussian field W and the eigenvalues are given in decaying order $\eta_1 \geq \eta_2 \geq \cdots \geq 0$. Since Q is trace class, the Gaussian random field W admits the representation

$$W = \sum_{i \in \mathbb{N}} \sqrt{\eta_i} e_i Z_i, \tag{3.1}$$

where $(Z_i, i \in \mathbb{N})$ are independent standard normally distributed random variables. The series above converges in $L^2(\Omega; H)$ and almost surely (see *e.g.* [9]). The truncated Karhunen–Loève expansion W_N of W is given by

$$W_N := \sum_{i=1}^{N} \sqrt{\eta_i} e_i Z_i, \tag{3.2}$$

where we call $N \in \mathbb{N}$ the *cut-off index* of W_N. In addition, it may be possible that the sequence of jumps $(P_i, i \in \mathbb{N})$ cannot be sampled exactly but only with an intrinsic bias (see [11], Rem. 3.4). The biased samples are denoted by $(\widetilde{P}_i, i \in \mathbb{N})$ and the error induced by this approximation is represented by the parameter $\varepsilon > 0$ (see Assumption 3.4). To approximate P using the biased sequence $(\widetilde{P}_i, i \in \mathbb{N})$ instead of $(P_i, i \in \mathbb{N})$ we define

$$P_\varepsilon : \Omega \times \mathcal{D} \to \mathbb{R}, \quad (\omega, x) \mapsto \sum_{i \in \mathcal{I}_{\mathcal{T}}} \mathbb{1}_{\{\mathcal{T}_i\}}(x) \widetilde{P}_i(\omega).$$

The *approximated jump-diffusion coefficient* $a_{N,\varepsilon}$ is then given by

$$a_{N,\varepsilon}(\omega, x) := \overline{a}(x) + \Phi(W_N(\omega, x)) + P_\varepsilon(\omega, x), \tag{3.3}$$

and the *approximated jump-advection coefficient* $b_{N,\varepsilon}$ *via*

$$b_{N,\varepsilon}(\omega, x) := \min\left(a_{N,\varepsilon}(\omega, x)\widetilde{b}_1(x), \widetilde{b}_2(x)\right).$$

Substituting the approximated jump coefficients into the parabolic model Problem (2.1) yields

$$\begin{aligned} \partial_t u_{N,\varepsilon}(\omega,x,t) + [A_{N,\varepsilon} u_{N,\varepsilon}](\omega,x,t) &= f(\omega,x,t) &&\text{in} \quad \Omega\times\mathcal{D}\times(0,T], \\ u_{N,\varepsilon}(\omega,x,0) &= u_0(\omega,x) &&\text{in} \quad \Omega\times\mathcal{D}\times\{0\}, \\ u_{N,\varepsilon}(\omega,x) &= 0 &&\text{on} \quad \Omega\times\partial\mathcal{D}, \end{aligned} \tag{3.4}$$

where the approximated second order differential operator $A_{N,\varepsilon}$ is given by

$$[A_{N,\varepsilon}u](\omega,x,t) = -\nabla\cdot(a_{N,\varepsilon}(\omega,x)\nabla u(\omega,x,t)) + b_{N,\varepsilon}(\omega,x)\cdot\nabla u(\omega,x,t).$$

The pathwise variational formulation of equation (3.4) is then analogous to equation (2.5): For fixed $\omega\in\Omega$ with given $f(\omega,\cdot)$, find $u_{N,\varepsilon}(\omega,\cdot,\cdot)\in L^2(\mathbb{T};V)$ with $\partial_t u_{N,\varepsilon}(\omega,\cdot,\cdot)\in L^2(\mathbb{T};V')$ such that it holds, for $t\in\mathbb{T}$ and for all $v\in V$

$$_{V'}\langle\partial_t u_{N,\varepsilon}(\omega,\cdot,t),v\rangle_V + B_\omega^{N,\varepsilon}(u_{N,\varepsilon}(\omega,\cdot,t),v) = F_{\omega,t}(v). \tag{3.5}$$

The approximated bilinear form is given for $v,w\in V$ by

$$B_\omega^{N,\varepsilon}(v,w) = \int_{\mathcal{D}} a_{N,\varepsilon}(\omega,x)\nabla v(x)\cdot\nabla w(x) + b_{N,\varepsilon}(\omega,x)\cdot\nabla v(x)w(x)\mathrm{d}x.$$

Lemma 3.6. *Let a and b be as in Definition 3.2, let $a_{N,\varepsilon}$ and $b_{N,\varepsilon}$ given by equation (3.3), and let Assumption 3.4 hold. Then, each pair (a,b) and $(a_{N,\varepsilon},b_{N,\varepsilon})$ satisfies Assumption 2.5 (i) and (ii).*

Moreover, define the real-valued random variables

$$a_- := \operatorname*{ess\,inf}_{x\in\mathcal{D}} a(\omega,x),\quad a_{N,\varepsilon,-} := \operatorname*{ess\,inf}_{x\in\mathcal{D}} a_{N,\varepsilon}(\omega,x),\quad a_+ := \operatorname*{ess\,sup}_{x\in\mathcal{D}} a(\omega,x),\quad a_{N,\varepsilon,+} := \operatorname*{ess\,sup}_{x\in\mathcal{D}} a_{N,\varepsilon}(\omega,x).$$

Then, $1/a_-, 1/a_{N,\varepsilon,-}, a_+, a_{N,\varepsilon,+}\in L^q(\Omega;\mathbb{R})$ for any $q\in[1,\infty)$ and there exists $C=C(q,\phi_1,\phi_2)>0$, independent of N and ε, such that

$$\|1/a_-\|_{L^q(\Omega;\mathbb{R})},\quad \|1/a_{N,\varepsilon,-}\|_{L^q(\Omega;\mathbb{R})},\quad \|a_+\|_{L^q(\Omega;\mathbb{R})},\quad \|a_{N,\varepsilon,+}\|_{L^q(\Omega;\mathbb{R})}\le C<+\infty.$$

Proof. Let $x\in\mathcal{D}$ be fixed. We first show that $\omega\mapsto a(\omega,x)$ is $\mathcal{F}-\mathcal{B}(\mathbb{R})$-measurable. By Definition 3.2

$$a(\omega,x) = \overline{a}(x) + \Phi(W(\omega,x)) + P(\omega,x),$$

where $\overline{a}$ is continuous and W a Gaussian random field, hence it suffices to show that $P(\cdot,x):\Omega\mapsto[0,\infty)$ is a random variable. For any $j\in\mathbb{N}$ we consider $\mathbb{1}_{\{\mathcal{T}_j(\cdot)\}}(x):\Omega\mapsto\{0,1\}$, and obtain by Definition 3.1

$$\{\omega\in\Omega|\mathbb{1}_{\{\mathcal{T}_j(\omega)\}}(x)=1\} = \{\omega\in\Omega|x\in\mathcal{T}_j(\omega)\} = \{\omega\in\Omega|\{x\}\cap\mathcal{T}_j(\omega)\neq\emptyset\} = \mathcal{T}_j^{\ell i}(\{x\})\in\mathcal{F}.$$

Repeating this argument for $\mathbb{1}_{\{\mathcal{T}_j(\omega)\}}(x)=0$ shows the measurability of $\mathbb{1}_{\{\mathcal{T}_j(\cdot)\}}(x)$ for any $j\in\mathbb{N}$. Since $\mathcal{T}_j(\omega)\in\mathcal{B}(\mathcal{D})$, it further holds that

$$\mu(\{\mathcal{T}_j(\omega)\}) = \int_{\mathcal{T}_j(\omega)}\mathrm{d}x = \int_{\mathcal{D}}\mathbb{1}_{\{\mathcal{T}_j(\omega)\}}(x)\mathrm{d}x.$$

The measurability of $\mathbb{1}_{\{\mathcal{T}_j(\cdot)\}}(x)$ for any $x\in\mathcal{D}$ now yields the measurability of $\omega\mapsto\mu(\{\mathcal{T}_j(\omega)\})$, and therefore also the measurability of $\mathbb{1}_{\{\mu(\mathcal{T}_j(\cdot))>0\}}:\Omega\to\{0,1\}$. This in turn proves that $P(\cdot,x)$ is $\mathcal{F}-\mathcal{B}(\mathbb{R})$-measurable for fixed $x\in\mathcal{D}$, since the sequence $(P_i,i\in\mathbb{N})$ in Definition 3.2 consists of real-valued random variables and $P(\omega,x)\in[0,\overline{P}]\subset\mathbb{R}$ for all $\omega\in\Omega$ by Assumption 3.4 (iv). Hence, a is measurable, the measurability of b follows immediately. Since $\overline{a}$ and P are nonnegative, and $\Phi\circ W(\cdot,x):\Omega\to(0,+\infty)$ for all $x\in\mathcal{D}$, we have that $a_-:\Omega\to(0,+\infty)$. On the other hand, $\overline{a}$ and P are bounded mappings by Definition 3.2 and therefore $a_+(\omega)<+\infty$ for all $\omega\in\Omega$.

To show that $a(\omega,\cdot) \in L^1(\mathcal{D})$ for fixed ω, note that $(\mathcal{T}_i(\omega), i \in \mathcal{I}_\mathcal{T}(\omega))$ defines a finite partition of $\mathcal{D}$ by Definition 3.2. Moreover, we have $\mathcal{T}_i(\omega) \in \mathcal{B}(\mathcal{D})$ for each partition element, hence a is piecewise continuous, and therefore integrable on $\mathcal{T}_i(\omega)$, *i.e.* $a(\omega,\cdot) \in L^1(\mathcal{T}_i(\omega))$. Thus, $a(\omega,\cdot) \in L^1(\mathcal{D})$ follows as the sets $(\mathcal{T}_i(\omega), i \in \mathcal{I}_\mathcal{T}(\omega))$ are disjoint and form a finite partition of $\mathcal{D}$. We have therefore shown that (a,b) satisfy Assumption 2.5 (i) and (ii). For fixed parameters $N \in \mathbb{N}$ and $\varepsilon > 0$, the assertion follows for $(a_{N,\varepsilon}, b_{N,\varepsilon})$ analogously. By Remark 2.6, we also observe that $a_{N,\varepsilon,-}, a_{N,\varepsilon,+} : \Omega \to \mathbb{R}$ are measurable mappings.

To bound $\|1/a_-\|_{L^q(\Omega;\mathbb{R})}$, we use that W and W_N are centered, almost surely bounded Gaussian random fields ([11], Lem. 3.5) on $\mathcal{D}$ which implies $E := \mathbb{E}(\sup_{x\in\mathcal{D}} W(x)) < +\infty$ as well as

$$\mathbb{P}\left(\sup_{x\in\mathcal{D}} W(\cdot,x) - E \geq c\right) \leq \exp\left(-\frac{c^2}{2\overline{\sigma}^2}\right) \tag{3.6}$$

for all $c > 0$ and $\overline{\sigma}^2 := \sup_{x\in\mathcal{D}} \mathbb{E}(W(\cdot,x)^2) \leq tr(Q)$. Furthermore, by the symmetry of W,

$$\mathbb{P}\left(||W(x)||_{L^\infty(\mathcal{D})} > c\right) \leq 2\mathbb{P}\left(\sup_{x\in\mathcal{D}} W(\cdot,x) > c\right). \tag{3.7}$$

With Assumption 3.4 (iii), and since

$$|| \exp(|W|)||_{L^\infty(\mathcal{D})} \leq \exp\left(||W||_{L^\infty(\mathcal{D})}\right),$$

we then obtain for arbitrary $q \in [1,\infty)$

$$\begin{aligned}
\mathbb{E}(1/a_-^q) &\leq \mathbb{E}\left(\left(\inf_{x\in\mathcal{D}} \Phi(W(\cdot,x)\right)^{-q}\right) \\
&= \mathbb{E}\left(\sup_{x\in\mathcal{D}} \Phi(W(\cdot,x))^{-q}\right) \\
&\leq \frac{1}{\phi_1^q}\mathbb{E}\left(\sup_{x\in\mathcal{D}} \exp(q\phi_2|W(\cdot,x)|)\right) \\
&\leq \frac{1}{\phi_1^q}\mathbb{E}\left(\exp(q\phi_2||W||_{L^\infty(\mathcal{D})})\right).
\end{aligned}$$

By Fubini's Theorem, integration by parts and equations (3.7), (3.6) this yields

$$\begin{aligned}
\mathbb{E}\left(\exp(q\phi_2||W||_{L^\infty(\mathcal{D})})\right) &= \int_0^\infty q\phi_2 \exp(q\phi_2 c)\mathbb{P}(||W||_{L^\infty(\mathcal{D})} > c)\mathrm{d}c \\
&\leq q\phi_2 \exp(q\phi_2 E) + 2\int_E^\infty q\phi_2 \exp(q\phi_2 c))\mathbb{P}\left(\sup_{x\in\mathcal{D}} W(\cdot,x) > c\right)\mathrm{d}c \\
&\leq q\phi_2 \exp(q\phi_2 E) + 2\int_E^\infty q\phi_2 \exp\left(q\phi_2 c - \frac{1}{2\overline{\sigma}^2}c^2\right)\mathrm{d}c.
\end{aligned}$$

The last estimate on the right hand side is finite for each $q \in \mathbb{R}$ which proves the claim for a_-. To bound the expectation of a_+, we may proceed in the same way by noting that

$$\|a_+\|_{L^q(\Omega)} \leq \|\overline{a}\|_{L^\infty(\mathcal{D})} + \mathbb{E}\left(|\sup_{x\in\mathcal{D}} \Phi(W(x))|^q\right)^{1/q} + \overline{P} \leq \|\overline{a}\|_{L^\infty(\mathcal{D})} + \phi_1\mathbb{E}\left(\sup_{x\in\mathcal{D}} \exp(q\phi_2|W(\cdot,x)|)\right)^{1/q} + \overline{P}$$

by Assumption 3.4 (iii) and (iv). Analogously, the claim follows for $a_{N,\varepsilon,-}, a_{N,\varepsilon,+}$ with the same bounds from above as for a_-, a_+ respectively, because

$$\overline{\sigma}_N^2 := \sup_{x\in\mathcal{D}} \mathbb{E}(W_N(x)^2) \leq \sum_{i=1}^N \eta_i \leq tr(Q).$$

□

Theorem 3.7. *Let Assumption 3.4 hold and $N \in \mathbb{N}$ and $\varepsilon > 0$ be fixed. There exist for any $\omega \in \Omega$ unique pathwise weak solutions $u(\omega, \cdot, \cdot) \in L^2(\mathbb{T}; V)$ to Problem (2.1) and $u_{N,\varepsilon}(\omega, \cdot, \cdot) \in L^2(\mathbb{T}; V)$ to Problem (3.4), respectively. Moreover, the mappings $u, u_{N,\varepsilon} : \Omega \to L^2(\mathbb{T}; V)$ are strongly measurable and satisfy for any $r \in [1, p)$ the estimates*

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u\|_{*,t}^r\right)^{1/r}, \quad \mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u_{N,\varepsilon}\|_{*,t}^r\right)^{1/r} \le C\left(\|u_0\|_{L^p(\Omega;H)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};V'))}\right),$$

where $C = C(r, a, b, T) > 0$ is independent of N and ε.

Proof. To apply Theorem 2.7, we need to verify Assumption 2.5. By Definition 3.2, Remark 3.3 and Lemma 3.6, we have already covered Assumption 2.5 (i), (ii) and (iv) for a, b and $a_{N,\varepsilon}, b_{N,\varepsilon}$. From Lemma 3.6 we further obtain $1/a_-, 1/a_{N,\varepsilon,-} \in L^q(\Omega;\mathbb{R})$ for any $q \in [1,\infty)$ and that $\|1/a_{N,\varepsilon,-}\|_{L^q(\Omega;\mathbb{R})}$ is bounded uniformly with respect to N and ε. For given $r \in [1, p)$, we then choose $q = (1/r - 1/p)^{-1} < +\infty$ and the claim follows by Assumption 3.4 (i) and Theorem 2.7. □

Having shown the existence and uniqueness of the weak solutions u and $u_{N,\varepsilon}$, we may bound the difference between both solutions in the (expected) parabolic norm with respect to the parameters N and ε. For this, we record the following estimate on the approximation error $a - a_{N,\varepsilon}$.

Theorem 3.8 ([11], Thm. 3.12)**.** *Under Assumption 3.4, it holds that*

$$\mathbb{E}(\|a - a_{N,\varepsilon}\|_{L^\infty(\mathcal{D})}^s)^{1/s} \le C\left(\Xi_N^{1/2} + \varepsilon^{1/s}\right),$$

where $\Xi_N := \sum_{i>N} \eta_i$ and $C > 0$ is independent of $N \in \mathbb{N}$ and $\varepsilon > 0$.

The final result of this section shows $u_{N,\varepsilon} \to u$ in $L^r(\Omega; L^2(T; V))$ as $N \to +\infty$ and $\varepsilon \to 0$.

Theorem 3.9. *Under Assumption 3.4, for any $r \in [1, (1/s + 1/p)^{-1})$, the approximation error of u is bounded in the parabolic norm by*

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u - u_{N,\varepsilon}\|_{*,t}^r\right)^{1/r} \le C\left(\Xi_N^{1/2} + \varepsilon^{1/s}\right).$$

Proof. By Theorem 3.7, pathwise existence of solutions u and $u_{N,\varepsilon}$ to the variational Problems (2.5), (3.5) is guaranteed, hence for all $\omega \in \Omega, t \in \mathbb{T}$ and $v \in V$

$${}_{V'}\langle \partial_t u(\omega,\cdot,t), v\rangle_V + B_\omega(u(\omega,\cdot,t), v) = {}_{V'}\langle \partial_t u_{N,\varepsilon}(\omega,\cdot,t), v\rangle_V + B_\omega^{N,\varepsilon}(u_{N,\varepsilon}(\omega,\cdot,t), v).$$

This may be reformulated as the problem to find $u - u_{N,\varepsilon} \in L^2(\mathbb{T}; V)$ such that for all $t \in \mathbb{T}$ and $v \in V$

$$\begin{aligned}
{}_{V'}\langle \partial_t(u(\omega,\cdot,t) - u_{N,\varepsilon}(\omega,\cdot,t)), v\rangle_V &+ B_\omega(u(\omega,\cdot,t) - u_{N,\varepsilon}(\omega,\cdot,t), v) \\
&= ((a_{N,\varepsilon} - a)(\omega,\cdot), \nabla u_{N,\varepsilon}(\omega,\cdot,t) \cdot \nabla v) + ((b_{N,\varepsilon} - b)(\omega,\cdot) \cdot \nabla u_{N,\varepsilon}(\omega,\cdot,t), v) \\
&=: {}_{V'}\langle \widehat{f}(\omega,\cdot,t), v\rangle_V,
\end{aligned}$$

with initial value $(u - u_{N,\varepsilon})(\omega,\cdot,0) \equiv 0$. Definition 3.2 and Remark 3.3 imply

$$\|\widehat{f}(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};V')} \le (1 + \overline{b}_1)\|(a - a_{N,\varepsilon})(\omega,\cdot)\|_{L^\infty(\mathcal{D})} \left\|\sum_{l=1}^d |\partial_{x_l} u_{N,\varepsilon}(\omega,\cdot,\cdot)|\right\|_{L^2(\mathbb{T};H)},$$

and by equation (2.4) and Theorem 3.7 we know that for $\overline{r} \in [1, p)$

$$\left\| \sum_{l=1}^{d} |\partial_{x_l} u_{N,\varepsilon}| \right\|_{L^{\overline{r}}(\Omega;L^2(\mathbb{T};H))} \leq 2^{d/2-1/2}\mathbb{E}\left(\left(\int_0^T |u|^2_{H^1(\mathcal{D})}\mathrm{d}t\right)^{\overline{r}/2}\right)^{1/\overline{r}}$$
$$\leq 2^{d/2-1/2}\mathbb{E}\left(\|u_{N,\varepsilon}\|^{\overline{r}}_{*,T}\right)^{1/\overline{r}}$$
$$\leq C\left(\|u_0\|_{L^p(\Omega;H)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};V'))}\right) < +\infty.$$

We may now choose $\overline{p} \in [1, (1/s + 1/\overline{r})^{-1}]$ and obtain by Hölder's inequality and Theorem 3.8

$$\|\widehat{f}(\omega,\cdot,\cdot)\|_{L^{\overline{p}}(\Omega;L^2(\mathbb{T};V'))} \leq C\mathbb{E}(\|a - a_{N,\varepsilon}\|^s_{L^\infty(\mathcal{D})})^{1/s} \left\| \sum_{l=1}^{d} \partial_{x_l} u_{N,\varepsilon} \right\|_{L^{\overline{r}}(\Omega;L^2(\mathbb{T};H))} \leq C\left(\varXi_N^{1/2} + \varepsilon^{1/s}\right)$$

for some $C > 0$ independent of N and ε. The claim now follows with Lemma 3.6 and by applying Theorem 2.7 to $u - u_{N,\varepsilon}$ for $q = (1/r - 1/s - 1/\overline{p})^{-1} < (1/r - 1/s - 1/p)^{-1} < +\infty$. □

To draw samples of $u_{N,\varepsilon}$, we need to employ further numerical techniques since $u_{N,\varepsilon}(\omega,\cdot,\cdot)$ is an element of the infinite-dimensional Hilbert space $L^2(\mathbb{T}; V)$. Hence, we have to find pathwise approximations of $u_{N,\varepsilon}$ in finite-dimensional subspaces of $L^2(\mathbb{T}; V)$ by discretizing the spatial and temporal domain. Next, we construct suitable approximation spaces of V, combine them with a time stepping method and control for the discretization error.

4. Pathwise discretization schemes

In the previous section we demonstrated that u may be approximated by $u_{N,\varepsilon}$ for sufficiently large $N \in \mathbb{N}$ resp. small $\varepsilon > 0$. Nevertheless, even $u_{N,\varepsilon}(\omega,\cdot,\cdot)$ will in general not be accessible analytically for fixed ω, N and ε, thus we need to find pathwise finite-dimensional approximations of $u_{N,\varepsilon}(\omega,\cdot,\cdot)$. In the first part of this section we explain how a semi-discrete solution may be obtained by approximating V with a sequence of *sample-adapted* finite element (FE) spaces. By sample-adaptedness we mean that the FE mesh is aligned *a-priori* with the discontinuities of P in each sample, *i.e.* the grid changes with each $\omega \in \Omega$. This is in contrast to *adaptive* FE schemes based on a-posteriori error estimates that may require several stages of remeshing in each sample, see *e.g.* [18, 21, 30]. To analyze the discretization error for the pathwise sample-adapted strategy, we assume that the random partition $\mathcal{T}$ consists of polygonal elements, which is a particular case in the general setting from the previous section. In the second part we combine the spatial discretization with a backward time stepping scheme in $\mathbb{T}$, with the time step chosen accordingly to the sample-dependent FE basis. Finally, we derive the mean-square error between the unbiased solution u and the fully discrete approximation of $u_{N,\varepsilon}$.

4.1. Sample-adapted spatial discretization

To find approximations of $u_{N,\varepsilon}(\omega,\cdot,t) \in V$ for fixed $\omega \in \Omega$ and $t \in \mathbb{T}$, we use a standard Galerkin approach based on a sequence $\mathcal{V}_\omega = (V_\ell(\omega), \ell \in \mathbb{N}_0)$ of finite-dimensional and sample-dependent subspaces $V_\ell(\omega) \subset V$. An obvious choice for V_ℓ is the space of piecewise linear FE with respect to some triangulation of $\mathcal{D}$. We follow the same approach as in [11] and utilize path-dependent meshes to match the interfaces generated by the jump-diffusion and -advection coefficients: Let $\mathcal{T}(\omega) = (\mathcal{T}_i(\omega), i \in \mathbb{N})$ be a given random partition of $\mathcal{D}$, and recall from Definition 3.2 that $\mathcal{I}_\mathcal{T}(\omega)$ is the (finite) index set of all partition elements with positive measure. We choose a conforming triangulation $\mathcal{K}_\ell(\omega) = \{K_1, \ldots, K_{n_\ell}\}$ of $\mathcal{D}$ such that for all $i \in \mathcal{I}_\mathcal{T}$ there are indices $\mathcal{I}_i \subset \mathbb{N}$ such that

$$\overline{\mathcal{T}}_i(\omega) = \bigcup_{j\in\mathcal{I}_i} K_j, \quad \text{and} \quad h_\ell(\omega) := \max_{K\in\mathcal{K}_\ell(\omega)} \operatorname{diam}(K) \leq \overline{h}_\ell \quad \text{for} \quad \ell \in \mathbb{N}_0. \tag{4.1}$$

Above, $\operatorname{diam}(K)$ is the longest side length of the triangle K and $(\overline{h}_\ell, \ell \in \mathbb{N}_0)$ is a positive sequence of deterministic refinement thresholds, decreasing monotonically to zero. This guarantees that $h_\ell(\omega) \to 0$ almost surely, although the absolute speed of convergence may vary for each ω. Given that $\mathcal{T}$ splits the domain $\mathcal{D}$ into a finite number of piecewise linear polygons (see Assumption 4.1 below), a triangulation $\mathcal{K}_\ell$ satisfying (4.1) for any prescribed refinement $\overline{h}_\ell > 0$ always exists. Consequently, $V_\ell(\omega)$ is chosen as the space of continuous, piecewise linear functions with respect to $K_\ell(\omega)$, *i.e.*

$$V_\ell(\omega) := \{v_{\ell,\omega} \in C^0(\overline{\mathcal{D}}) \big|\, v_{\ell,\omega}|_{\partial\mathcal{D}} = 0 \quad \text{and} \quad v_{\ell,\omega}|_K \in \mathcal{P}_1(K), \quad K \in \mathcal{K}_\ell(\omega)\} \subset V. \tag{4.2}$$

The set $\mathcal{P}_1(K)$ denotes the space of all linear polynomials on the triangle K, and $\{v_{1,\omega}, \dots, v_{d_\ell(\omega),\omega}\}$ is the nodal basis of $V_\ell(\omega)$ that corresponds to the vertices in $\mathcal{K}_\ell(\omega)$. As discussed in Section 4 of [11], the adjustment of $\mathcal{K}_\ell$ to the discontinuities of a and b accelerates convergence of the spatial discretization compared to a fixed, non-adapted FE approach. For a fixed triangle $K \in \mathcal{K}_\ell(\omega)$ let $x_1^K, x_2^K, x_3^K \in \overline{\mathcal{D}}$ denote the corner points of K and let $v_1^K, v_2^K, v_3^K \in \mathcal{P}_1(K)$ be the corresponding linear nodal basis of $\mathcal{P}_1(K)$. We define the *local interpolation operator* on K by

$$\mathcal{I}(K) : C^0(\overline{K}) \to \mathcal{P}_1(K), \quad v \mapsto \sum_{i=1}^{3} v(x_i^K) v_{i,\omega}^K.$$

The *global interpolation operator* $\mathcal{I}_\ell : C^0(\overline{\mathcal{D}}) \to V_\ell(\omega)$ with respect to $\mathcal{K}_\ell(\omega)$ is then given by restrictions to the local operators, that is

$$[\mathcal{I}_\ell v](x) := [\mathcal{I}(K)v](x), \text{ for } x \in \mathcal{D}, \text{ where } K \in \mathcal{K}_\ell \text{ is such that } x \in \overline{K}.$$

For simplicity, we only consider the nodal interpolation of continuous functions $v \in C^0(\overline{\mathcal{D}})$.

The semi-discrete version of (3.5) is to find $u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot) \in L^2(\mathbb{T}; V_\ell(\omega))$ with $\partial_t u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot) \in L^2(\mathbb{T}; (V_\ell(\omega))')$ such that for $t \in \mathbb{T}$ and all $v_{\ell,\omega} \in V_\ell(\omega)$

$$\begin{aligned} {}_{V'}\langle \partial_t u_{N,\varepsilon,\ell}(\omega,\cdot,t), v_{\ell,\omega}\rangle_V + B_\omega^{N,\varepsilon}(u_{N,\varepsilon,\ell}(\omega,\cdot,t), v_{\ell,\omega}) &= F_{t,\omega}(v_{\ell,\omega}), \\ u_{N,\varepsilon,\ell}(\omega,\cdot,0) &= \mathcal{I}_\ell u_0(\omega,\cdot). \end{aligned} \tag{4.3}$$

We have used the nodal interpolation $\mathcal{I}_\ell u_0$ as approximation of the initial value, which is well-defined if $u_0(\omega,\cdot) \in C^0(\overline{\mathcal{D}})$ holds for any ω (see also Assumption 4.1 (i)/Rem. 4.2). The function $u_{N,\varepsilon,\ell}(\omega,\cdot,t)$ may be expanded with respect to the basis $\{v_{1,\omega}, \dots, v_{d_\ell(\omega),\omega}\}$ as

$$u_{N,\varepsilon,\ell}(\omega,x,t) = \sum_{j=1}^{d_\ell(\omega)} c_j(\omega,t) v_{j,\omega}(x), \tag{4.4}$$

where the coefficients $c_1(\omega,t), \dots, c_{d_\ell(\omega)}(\omega,t) \in \mathbb{R}$ depend on $(\omega,t) \in \Omega\times\mathbb{T}$ and the respective coefficient column-vector is defined as $\mathbf{c}(\omega,\mathbf{t}) := (c_1(\omega,t), \dots, c_{d_\ell(\omega)}(\omega,t))^T$. With this, the semi-discrete variational problem in the finite-dimensional space $V_\ell(\omega)$ is equivalent to solving the system of ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{c}(\omega,\mathbf{t}) + \mathbf{A}(\omega)\mathbf{c}(\omega,\mathbf{t}) = \mathbf{F}(\omega,\mathbf{t}), \quad \mathbf{t} \in \mathbb{T},$$

for $\mathbf{c}$ with stochastic stiffness matrix $(\mathbf{A}(\omega))_{jk} = B_\omega^{N,\varepsilon}(v_{j,\omega}, v_{k,\omega})$ and time-dependent load vector $(\mathbf{F}(\omega,t))_j = F_{t,\omega}(v_{j,\omega})$ for $j,k \in \{1,\dots,d_\ell(\omega)\}$. To ensure well-posedness of equation (4.3) and to derive error bounds of the numerical approximation of u in a mean-square sense, we need to modify Assumption 3.4 (items (iii) and (iv) below are unaltered):

Assumption 4.1. (i) *There exists $p > 2$ such that $f, \partial_t f \in L^p(\Omega; L^2(\mathbb{T}; H))$ and $u_0 \in L^p(\Omega; V) \cap L^p(\Omega; H^{1+\epsilon}(\mathcal{D}))$ for some arbitrary $\epsilon > 0$. Furthermore, u_0 and f are stochastically independent of $\mathcal{T}$.*

(ii) *The eigenfunctions e_i of Q are continuously differentiable on $\mathcal{D}$ and there exist constants $\alpha, \beta, C_e, C_\eta > 0$ such that $2\alpha \le \beta$ and for any $i \in \mathbb{N}$*

$$\|e_i\|_{L^\infty(\mathcal{D})} \le C_e, \quad \max_{l=1,\dots,d} \|\partial_{x_l} e_i\|_{L^\infty(\mathcal{D})} \le C_e i^\alpha \quad \text{and} \quad \sum_{i=1}^{\infty} \eta_i i^\beta \le C_\eta < +\infty.$$

(iii) *Furthermore, the mapping Φ as in Definition 3.2 and its derivative are bounded for $w \in \mathbb{R}$ by*

$$\phi_1 \exp(\phi_2 w) \ge \Phi(w) \ge \phi_1 \exp(-\phi_2 w), \quad |\frac{\mathrm{d}}{\mathrm{d}x}\Phi(w)| \le \phi_3 \exp(\phi_4 |w|),$$

where $\phi_1, \dots, \phi_4 > 0$ are arbitrary constants.

(iv) *The sequence $(P_i, i \in \mathbb{N})$ consists of nonnegative and bounded random variables $P_i \in [0, \overline{P}]$ for some $\overline{P} > 0$. In addition, for $s > 2$ such that $1/p + 1/s < 1/2$ there exists a sequence of approximations $(\widetilde{P}_i, i \in \mathbb{N}) \subset [0, \overline{P}]^{\mathbb{N}}$ so that the sampling error is bounded, for some $\varepsilon > 0$, by*

$$\mathbb{E}(|\widetilde{P}_i - P_i|^s) \le \varepsilon, \quad i \in \mathbb{N}.$$

(v) *The (non-trivial) partition elements $(\mathcal{T}_i(\omega), i \in \mathcal{I}_{\mathcal{T}}(\omega))$ are polygons with piecewise linear boundary and a finite number of boundary edges for all $\omega \in \Omega$ and $\mathbb{E}(\tau^n) = \mathbb{E}((\#|\mathcal{I}_{\mathcal{T}}|)^n) < \infty$ for any $n \in \mathbb{N}$.*

(vi) *Let 2^V be the power set of V. For all $\ell \in \mathbb{N}_0$, the correspondence $V_\ell : \Omega \to 2^V$, $\omega \mapsto V_\ell(\omega)$ is weakly measurable (cf. Def. 3.1) and for all $\omega \in \Omega$ it holds that $V_\ell(\omega) \ne \emptyset$.*

(vii) *Conformity: In dimension $d = 2$, let $K_1, K_2 \in \mathcal{K}_\ell(\omega)$ for some fixed $\ell \in \mathbb{N}_0$ and $\omega \in \Omega$. Then, the intersection $\overline{K}_1 \cap \overline{K}_2$ is either empty, a common edge or a common vertex of $\mathcal{K}_\ell(\omega)$.*

(viii) *Shape-regularity: Let $\rho_{K,\mathrm{out}}$ and $\rho_{K,\mathrm{in}}$ denote the radius of the outer respectively inner circle of the triangle K. Then, there is a constant $\overline{\rho} > 0$ such that*

$$\operatorname*{ess\,sup}_{\omega \in \Omega} \sup_{\ell \in \mathbb{N}_0} \sup_{K \in \mathcal{K}_\ell(\omega)} \frac{\rho_{K,\mathrm{out}}}{\rho_{K,\mathrm{in}}} \le \overline{\rho} < +\infty.$$

Remark 4.2. We discuss Assumption 4.1 in the following:

– Item (ii) implies for all $l = 1, \dots, d$ and $x \in \mathcal{D}$

$$\mathbb{E}\left(|\partial_{x_l} W_N(x)|^2\right) = \mathbb{E}\left(\left|\sum_{j=1}^{n} \sqrt{\eta_j} \partial_{x_l} e_j(x) Z_j\right|^2\right) \le C_e \sum_{j=1}^{N} \eta_j j^{2\alpha} \le C_e \sum_{j=1}^{N} \eta_j j^\beta,$$

hence there exist an $L^2(\Omega;\mathbb{R})$-limit $\partial_{x_l} W(\cdot, x) := \lim_{N \to +\infty} \partial_{x_l} W_N(\cdot, x)$. Hence, $2\alpha \le \beta$ entails the mean-square differentiability (or pathwise Lipschitz-continuity) of the Gaussian field W.

– By the fractional Sobolev inequality ([19], Thm. 6.7), $u_0(\omega, \cdot) \in H^{1+\epsilon}(\mathcal{D})$ for $\epsilon > 0$ implies with $d \in \{1, 2\}$ that $u_0(\omega, \cdot) \in C^0(\overline{\mathcal{D}})$ and the nodal interpolation of u_0 is well-defined. The assumptions on f and $\partial_t f$ are necessary to control the error of a temporal discretization scheme. The nodal basis functions $v_{j,\omega}$ are solely determined by $\mathcal{T}(\omega)$ and since f, u_0 are stochastically independent of $\mathcal{T}$, we may expand the sample-adapted semi-discrete solution *via* equation (4.4), *i.e.* obtain a separation of spatial and temporal variables.

– The condition $1/p + 1/s < 1/2$ enables us to derive all errors in a mean-square sense. Furthermore, the partition into piecewise linear polygons enables us to construct triangulations $\mathcal{K}_\ell(\omega)$ resp. approximation spaces $V_\ell(\omega)$ as in equation (4.2).

– The weak measurability of the correspondence $\omega \mapsto V_\ell(\omega)$ ensures the (strong) measurability of the approximated solution $u_{N,\varepsilon,\ell} : \Omega \to L^2(\mathbb{T}; V)$, see Proposition 4.3. This assumption is necessary, since pathological approximation spaces $V_\ell(\omega)$ may still be constructed on a nullset of Ω, even under Assumption 4.1 (v).

– Conformity and shape-regularity of the FE triangulations are necessary to control the FE discretization error.

We show measurability of the semi-discrete approximations and record a bound on the interpolation error.

Proposition 4.3. *Let Assumption 4.1 hold and let $\ell \in \mathbb{N}_0$ be fixed. Then, for any $\omega \in \Omega$ there exists a unique sample-adapted solution $u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot) \in L^2(\mathbb{T};V)$ to the semi-discrete problem (4.3) and the mapping $u_{N,\varepsilon,\ell}:\Omega \to L^2(\mathbb{T};V)$ is strongly measurable.*

Proof. For fixed ω, existence and uniqueness of $u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot)$ follows with Assumption 4.1 as in Theorem 2.7, hence the map $u_{N,\varepsilon,\ell}:\Omega\to L^2(\mathbb{T};V)$ is well-defined. To show measurability, we use again the space $\mathcal{X} := L^2(\mathbb{T};V)\times L^2(\mathbb{T};V')$ with $\|(y_1,y_2)\|_{\mathcal{X}} := \|y_1\|_{L^2(\mathbb{T};V)} + \|y_2\|_{L^2(\mathbb{T};V')}$ as in the proof of Theorem 2.7. Let $\{v_{1,\omega},\dots,v_{d_\ell(\omega),\omega}\}$ be a basis of $V_\ell(\omega)$ and define the sequence

$$\widetilde{v}_{i,\omega} := \begin{cases} v_{i,\omega} & \text{if} \quad i\le d_\ell(\omega) \\ v_{d_\ell(\omega),\omega} & \text{if} \quad i> d_\ell(\omega). \end{cases}$$

By Assumption 4.1 (vi), the correspondence $\omega\mapsto V_\ell(\omega)$ is weakly measurable and has closed, non-empty values, therefore there exists a sequence $(\xi_i, i\in\mathbb{N})$ of measurable functions $\xi_i:\Omega\to V$ such that $\xi_i(\omega)\in V_\ell(\omega)$ and $V_\ell(\omega) = \overline{\{\xi_1(\omega),\xi_2(\omega),\dots\}}$ (see [3], Cor. 18.14). Consequently, each $\widetilde{v}_{i,\cdot}:\Omega\to V$ can be written as the limit of measurable functions and is therefore $\mathcal{F}-\mathcal{B}(V)$-measurable. Now, consider the functional

$$\begin{aligned}\widetilde{J}_i^{N,\varepsilon}:\Omega\times\mathcal{X}\to\mathbb{R},\quad (\omega,w)\mapsto \int_0^T & B_\omega^{N,\varepsilon}(w(\cdot,t),\widetilde{v}_{i,\omega}) - F_{\omega,t}(\widetilde{v}_{i,\omega}) + {}_{V'}\langle \partial_t w(\cdot,t),\widetilde{v}_{i,\omega}\rangle_V \\ &+ \left\| w(\cdot,t) - \sum_{i=1}^{d_\ell(\omega)} (w(\cdot,t),\widetilde{v}_{i,\omega})_V \widetilde{v}_{i,\omega}\right\|_V \mathrm{d}t.\end{aligned}$$

By Theorem 2.7 and the measurability of $\widetilde{v}_{i,\cdot}$ we conclude that $J_i^{N,\varepsilon}$ is a Carathéodory mapping. We define the correspondence

$$\widetilde{\varphi}_i(\omega) := \{w\in\mathcal{X} |\, J_i^{N,\varepsilon}(\omega,w)=0\}.$$

and obtain as in the proof of Theorem 2.7 by Corollary 18.8 of [3] that the graph $\mathrm{Gr}(\widetilde{\varphi}_i) = \{(\omega,w)\in\Omega\times\mathcal{X} |\, w\in \varphi_i(\omega)\}$ is measurable. By construction of $J_i^{N,\varepsilon}$, it then follows that

$$\{(\omega, u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot),\partial_t u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot)) |\, \omega\in\Omega\} = \bigcap_{i\in\mathbb{N}} \mathrm{Gr}(\widetilde{\varphi}_i) \in \mathcal{F}\otimes\mathcal{B}(\mathcal{X}),$$

and the claimed measurability of $u_{N,\varepsilon,\ell}$ follows analogously as in the proof of Theorem 2.7. □

Lemma 4.4. *Under Assumption 4.1, let $\omega\in\Omega$ be fixed, and let $v\in H^\vartheta(\mathcal{T}_i)$ for some $\vartheta\in(1,2]$ and $i\in\mathcal{I}_\mathcal{T}(\omega)$. Then, $\mathcal{I}_\ell v\in C^0(\overline{\mathcal{T}}_i)$ is well-defined on each partition element $\mathcal{T}_i$ and for $m\in\{0,1\}$ there holds*

$$\begin{aligned}\left(\sum_{i\in\mathcal{I}_\mathcal{T}(\omega)} \|(1-\mathcal{I}_\ell)v\|^2_{H^m(\mathcal{T}_i)}\right)^{1/2} &= \left(\sum_{i\in\mathcal{I}_\mathcal{T}(\omega)}\sum_{K\in\mathcal{T}_i} \|(1-\mathcal{I}(K))v\|^2_{H^m(K)}\right)^{1/2} \\ &\le C\overline{h}_\ell^{\vartheta-m}\left(\sum_{i\in\mathcal{I}_\mathcal{T}(\omega)} |v|_{H^\vartheta(\mathcal{T}_i)}\right)^{1/2},\end{aligned} \tag{4.5}$$

where $C=C(\overline{\rho},\vartheta,m,d)>0$ is a deterministic constant.

Proof. By the Sobolev embedding theorem it holds that $\|v\|_{C^0(\overline{\mathcal{T}_i})} \leq C\|v\|_{H^\vartheta(\mathcal{T}_i)}$ and thus $\mathcal{I}_\ell v$ is well-defined on $\mathcal{T}_i$. Moreover, for $m = \{0, 1\}$, we use that $V_\ell(\omega) \neq \emptyset$ and the interpolation estimates from Theorem 4.4.20 of [13] to see that

$$\|\mathcal{I}_\ell v\|_{H^m(\mathcal{T}_i)} \leq C\|v\|_{H^m(\mathcal{T}_i)}$$

for a constant $C = C(\overline{\rho}, m, d) > 0$, independent of $\mathcal{T}_i$ (recall that $\overline{\rho}$ is deterministic and controls the shape regularity of $\mathcal{K}_\ell(\omega)$). Together with $w = \mathcal{I}_\ell w$ for any $w \in V_\ell(\omega)$ we then obtain

$$\begin{aligned}\|(1-\mathcal{I}_\ell)v\|_{H^m(\mathcal{T}_i)} &\leq \inf_{w\in V_\ell(\omega)} \|v-w\|_{H^m(\mathcal{T}_i)} + \|\mathcal{I}_\ell(w-v)\|_{H^m(\mathcal{T}_i)} \\ &\leq C \inf_{w\in V_\ell(\omega)} \|v-w\|_{H^m(\mathcal{T}_i)} \\ &\leq C \inf_{w\in V_\ell(\omega)} \left(\sum_{K\in\mathcal{T}_i} \|v-w\|^2_{H^m(K)}\right)^{1/2}.\end{aligned}$$

Assumption 4.1 guarantees that $K \in \mathcal{T}_i$ holds by construction of the approximation space $V_\ell(\omega)$. The claim now follows, for instance, by the estimates from Chapter 8.5 of [27], summing over $i \in \mathcal{I}_\mathcal{T}$ and by the fact that the constant $C = C(\overline{\rho}, \vartheta, m, d) > 0$ is deterministic, *i.e.*, independent of $\mathcal{T}_i$. □

To bound the pathwise FE discretization error, we now fix $\omega \in \Omega, t \in \mathbb{T}$ and $u_{N,\varepsilon}(\omega, \cdot, t) \in V$ and consider the corresponding pathwise elliptic PDE

$$-\nabla\cdot(a_{N,\varepsilon}(\omega,\cdot)\nabla u_{N,\varepsilon}(\omega,\cdot,t)) = f(\omega,\cdot,t) - b_{N,\varepsilon}(\omega,\cdot)\cdot\nabla u_{N,\varepsilon}(\omega,\cdot,t) - \partial_t u_{N,\varepsilon}(\omega,\cdot,t) =: \widetilde{f}(\omega,\cdot,t) \tag{4.6}$$

on $\mathcal{D}$ with homogeneous Dirichlet boundary conditions. Let $\mathcal{E}$ be the set of all interior edges of $\mathcal{T}(\omega)$ and for every $e \in \mathcal{E}$ let $i_e, i'_e \in \mathcal{I}_\mathcal{T}(\omega)$ with $i_e \neq i'_e$ be the indices such that $e = \overline{\mathcal{T}}_{i_e} \cap \overline{\mathcal{T}}_{i'_e}$. Accordingly, the outward normal vectors on either side of e with respect $\mathcal{T}_{i_e}$ and $\mathcal{T}_{i'_e}$ are denoted by $\vec{n}_{i_e}$ and $\vec{n}_{i'_e}$, respectively. Due to the discontinuities of $a_{N,\varepsilon}(\omega, \cdot)$, this yields the transition condition

$$a_{N,\varepsilon}(\omega,\cdot)\vec{n}_{i_e}\cdot\nabla u_{N,\varepsilon}(\omega,\cdot,t) = a_{N,\varepsilon}(\omega,\cdot)\vec{n}_{i'_e}\cdot\nabla u_{N,\varepsilon}(\omega,\cdot,t) \quad \text{on} \quad e\in\mathcal{E}. \tag{4.7}$$

Therefore, $u_{N,\varepsilon}(\omega, \cdot, t)$ may be regarded as weak solution to an *elliptic interface problem* given by equations (4.6)–(4.7) satisfying for all $v \in V$

$$\int_\mathcal{D} a_{N,\varepsilon}(\omega,x)\nabla u_{N,\varepsilon}(\omega,x,t)\cdot\nabla v(x)\mathrm{d}x = \int_\mathcal{D} \widetilde{f}(\omega,x,t)v(x)\mathrm{d}x. \tag{4.8}$$

Given that $\widetilde{f}(\omega, \cdot, t) \in H$ (which is verified almost surely by Lems. 4.8 and 4.9 below), it is known for dimension $d = 2$, *e.g.* from [35–37,40], that the solution $u_{N,\varepsilon}(\omega, \cdot, t)$ to the elliptic interface problem admits a *decomposition into singular functions* with respect to the corners of $\mathcal{T}(\omega)$. More precisely,

$$u_{N,\varepsilon}(\omega,\cdot,t) = w + \sum_{j\in\mathcal{S}} c_j\chi_j(r^{(j)})\psi_j(r^{(j)},\varphi^{(j)}), \tag{4.9}$$

where $\mathcal{S}$ denotes the set of *singular points* in the partition $\mathcal{T}(\omega)$ (in our case $\mathcal{S}$ is the set of corners in $\mathcal{T}(\omega)$) and $(r^{(j)}, \varphi^{(j)})$ are polar coordinates with respect to the singular point $j \in \mathcal{S}$. For any $i \in \mathcal{I}_\mathcal{T}$, it holds $w \in H^2(\mathcal{T}_i)$ and $\psi \notin H^{1+\kappa_j}(\mathcal{T}_i)$, but $\psi \in H^{1+\kappa_j-\epsilon}(\mathcal{T}_i)$ for some $\kappa_j \in (0, 1]$ and any $\epsilon > 0$. Moreover, $c_j \in \mathbb{R}$ are coefficients and χ_j is a smooth and bounded cutoff function vanishing near the singular point j.

The decomposition in equation (4.9) shows that $u_{N,\varepsilon}(\omega, \cdot, t) \notin H^2(\mathcal{T}_i)$, but we may expect piecewise regularity of $u_{N,\varepsilon}(\omega, \cdot, t) \in H^{1+\underline{\kappa}-\varepsilon}(\mathcal{T}_i)$, where $\underline{\kappa} := \min_{j\in\mathcal{S}} \kappa_j$. The precise values of the exponents $\kappa_j \in (0, 1]$ depend on the shape of the partition elements $\mathcal{T}_i$, *i.e.* their angle at the singular points $\mathcal{S}$, as well as on the magnitude of the jump heights P_i. Furthermore, the results from [35,36] show that the coefficients c_j and w depend continuously on the right hand side $\widetilde{f}$, the gradient of $a_{N,\varepsilon}$ on $\mathcal{T}_i$ and the inverse of $a_{N,\varepsilon,-}$. A detailed analysis on the dependencies of κ_j, c_j and w may be found in the literature (see [36,37]). For the sake of simplicity we assume piecewise regularity of $u_{N,\varepsilon}$ in accordance with the decomposition in equation (4.9).

Assumption 4.5. *Let $\widetilde{f}$ be defined as in equation(4.8). There are deterministic constants $\kappa \in (0,1]$ and $C > 0$, such that for all $N \in \mathbb{N}, \varepsilon > 0$ and $t \in \mathbb{T}$ there holds for almost all $\omega \in \Omega$ and for $i \in \mathcal{I}_{\mathcal{T}}$,*

$$\|u_{N,\varepsilon}(\omega,\cdot,t)\|_{H^{1+\kappa}(\mathcal{T}_i)} \le C \frac{\|\widetilde{f}(\omega,\cdot,t)\|_{L^2(\mathcal{T}_i)} + \|a_{N,\varepsilon}(\omega,\cdot)\|_{\mathcal{W}^{1,\infty}(\mathcal{T}_i)}\|u_{N,\varepsilon}(\omega,\cdot,t)\|_{H^1(\mathcal{T}_i)}}{a_{N,\varepsilon,-}(\omega)}.$$

Remark 4.6. Assumption 4.5 enables the ensuing numerical analysis for $d = 2$, whereas this assumption would not be necessary in $d = 1$. It is based on the decomposition in equation (4.9) as well as the estimates for c_j and w from [35] in terms of the right hand side $\widetilde{f}$ and $a_{N,\varepsilon}$. Although it may seem artificial at a first glance, we recover κ close to one in the numerical examples from Section 5 for $d = 2$. On the other hand, it is actually possible to obtain lower bounds on κ, *i.e.* to ensure a certain minimum of piecewise regularity almost surely. This is for instance the case if:

– the jump heights $\mathcal{P}_i$ and the interior angles of the $\mathcal{T}_i$ are bounded from above and below, or
– if $a_{N,\varepsilon}$ satisfies almost surely a *quasi-monotonicity condition*,

see [40] and the references therein. Since $d \le 2$ and $u_{N,\varepsilon}(\omega,\cdot,t) \in H^{1+\kappa}(\mathcal{T}_i)$ holds for every polygonal subdomain $\mathcal{T}_i$, it follows that $u_{N,\varepsilon}(\omega,\cdot,t) \in H^{\vartheta}(\mathcal{D})$, where $\vartheta = \min(1+\kappa, 3/2-\epsilon)$ for any $\epsilon > 0$ (see [40], Lem. 3.1). This in turn yields that $u_{N,\varepsilon}(\omega,\cdot,t) \in C^0(\overline{\mathcal{D}})$ by the fractional Sobolev inequality. Hence, the nodal interpolation $\mathcal{I}_\ell u_{N,\varepsilon}(\omega,\cdot,t)$ is well-defined.

We are now ready to state our main result on the spatial discretization error:

Theorem 4.7. *Let Assumptions 4.1 and 4.5 hold and let $u_{N,\varepsilon,\ell}$ be the sample-adapted FE approximation of $u_{N,\varepsilon}$ as in equation (4.3), where $\overline{h}_\ell \le 1$. Then, there is a $C > 0$, independent of N, ε and $\overline{h}_\ell$ such that*

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u_{N,\varepsilon} - u_{N,\varepsilon,\ell}\|_{*,t}^2\right)^{1/2} \le C\overline{h}_\ell^{\kappa}.$$

For the proof of Theorem 4.7 we record several technical lemmas as preparation.

Lemma 4.8. *Let $\Theta \in \mathbb{R}^d$ be an open, bounded domain and denote by $\mathcal{W}^{k,\infty}(\Theta)$ the Sobolev space defined by the (semi-)norm*

$$\|v\|_{\mathcal{W}^{k,\infty}(\Theta)} := \sum_{|\nu|\le k} \|D^\nu v\|_{L^\infty(\Theta)}, \quad |v|_{\mathcal{W}^{k,\infty}(\Theta)} := \sum_{|\nu|=k} \|D^\nu v\|_{L^\infty(\Theta)}, \quad k \in \mathbb{N},$$

for any measurable mapping $v : \Theta \to \mathbb{R}$. Under Assumption 4.1, for any $q \in [1,\infty)$

$$\left\|\max_{i\in\mathcal{I}_{\mathcal{T}}} \|a_{N,\varepsilon}\|_{\mathcal{W}^{1,\infty}(\mathcal{T}_i)}\right\|_{L^q(\Omega;\mathbb{R})} \le C < +\infty,$$

where $C = C(q) > 0$ is independent of N and ε.

Proof. As $a_{N,\varepsilon}$ is almost surely continuously differentiable on each partition element $\mathcal{T}_i$, $i \in \mathcal{I}_{\mathcal{T}}$, by Assumption 4.1, we have

$$\|a_{N,\varepsilon}(\omega,\cdot)\|_{\mathcal{W}^{1,\infty}(\mathcal{T}_i)} \le a_{N,\varepsilon,+}(\omega) + \max_{l=1,\dots,d} \|\partial_{x_l}\overline{a}\|_{L^\infty(\mathcal{D})} + \left\|\frac{\mathrm{d}}{\mathrm{d}x}\Phi(W_N(\omega,\cdot))\partial_{x_l}W_N(\omega,\cdot)\right\|_{L^\infty(\mathcal{D})}$$

with $\|\partial_{x_l}\overline{a}\|_{L^\infty(\mathcal{D})} < +\infty$ for all $l = 1,\dots,d$. Moreover, Lemma 3.6 states that $\|a_{N,\varepsilon,+}\|_{L^q(\Omega;\mathbb{R})} < +\infty$ for any

$q \in [1,\infty)$ and the norm is bounded uniformly with respect to N and ε. Thus, we only need to estimate the last term on the right hand side. We use Hölder's inequality and Assumption 4.1 to obtain for any $q \geq 1$

$$\begin{aligned}\left\|\frac{\mathrm{d}}{\mathrm{d}x}\Phi(W_N)\partial_{x_l}W_N\right\|_{L^q(\Omega;L^\infty(\mathcal{D}))} &\leq \left\|\frac{\mathrm{d}}{\mathrm{d}x}\Phi(W_N)\right\|_{L^{2q}(\Omega;L^\infty(\mathcal{D}))}\|\partial_{x_l}W_N\|_{L^{2q}(\Omega;L^\infty(\mathcal{D}))}\\ &\leq \phi_3\mathbb{E}\left(\exp(2q\phi_4||W_N||_{L^\infty(D)})\right)^{1/(2q)}||\partial_{x_l}W_N||_{L^{2q}(\Omega;L^\infty(\mathcal{D}))}.\end{aligned}$$

The random field W_N is centered Gaussian with $\sup_{x\in\mathcal{D}}\mathbb{E}(W_N(x)^2) \leq \sup_{x\in\mathcal{D}}\mathbb{E}(W(x)^2) \leq tr(Q)$ and we proceed as in Lemma 3.6 to conclude that

$$\mathbb{E}\left(\exp(2q\phi_4||W_N||_{L^\infty(D)})\right) \leq \int_0^\infty 2q\phi_4\exp(2q\phi_2 c)\mathbb{P}(||W||_{L^\infty(\mathcal{D})} > c)\mathrm{d}c < +\infty.$$

To estimate $||\partial_{x_l}W_N||_{L^{2q}(\Omega;L^\infty(\mathcal{D}))}$, we note that, for $x \in \mathcal{D}$, $\partial_{x_l}W_N(x)$ is also centered Gaussian with variance $\sum_{j=1}^N \eta_j(\partial_{x_l}e_j(x))^2$. For any $N \in \mathbb{N}$

$$\sup_{x\in\mathcal{D}}|\partial_{x_l}W_N(x)| = \sup_{x\in\mathcal{D}}\left|\sum_{j=1}^N\sqrt{\eta_j}\partial_{x_l}e_j(x)Z_j\right| \leq \sum_{j=1}^N\sqrt{\eta_j}j^\alpha|Z_j|$$

by Assumption 4.1 (ii), hence $\partial_{x_l}W_N$ is almost surely bounded on $\mathcal{D}$. The symmetric distribution of $\partial_{x_l}W_N(x)$ and Theorem 2.1.1 of [2] then imply $\mathbb{E}(\sup_{x\in\mathcal{D}}\partial_{x_l}W_N(x)) \geq 0$,

$$\begin{aligned}\mathbb{E}\left(||\partial_{x_l}W_N||_{L^\infty(\mathcal{D})}\right) \leq 2\mathbb{E}\left(\sup_{x\in\mathcal{D}}\partial_{x_l}W_N(x)\right) =: 2E_{N,l} < +\infty, \quad \text{and}\\ \mathbb{P}\left(\sup_{x\in\mathcal{D}}\partial_{x_l}W_N(x) > c\right) \leq \exp\left(-\frac{(c-E_{N,l})^2}{2\widetilde{\sigma}_{N,i}^2}\right) \leq \exp\left(-\frac{(c-E_{N,l})^2}{2\widetilde{\sigma}^2}\right), \quad c > 0,\end{aligned} \tag{4.10}$$

analogously to Lemma 3.6. The maximal variances in equation (4.10) are given by

$$\widetilde{\sigma}_{N,i}^2 := \sup_{x\in\mathcal{D}}\mathbb{E}((\partial_{x_l}W_N(x))^2) = \sum_{j=1}^N\eta_j(\partial_{x_l}e_j(x)) \leq \widetilde{\sigma}^2 := C_e\sum_{j=1}^\infty\eta_j j^{2\alpha} \leq C_e\sum_{j=1}^\infty\eta_j j^\beta < +\infty.$$

Without loss of generality, we assume $q \in \mathbb{N}$ to obtain $\mathbb{E}(||\partial_{x_l}W_N||_{L^\infty(\mathcal{D})}^{2q}) = \mathbb{E}(\sup_{x\in\mathcal{D}}(\partial_{x_l}W_N(x))^{2q})$. We now have to make sure that $\mathbb{E}(\sup_{x\in\mathcal{D}}(\partial_{x_l}W_N(x))^{2q})$ is bounded uniformly in l and N. Similar to Lemma 3.6, Fubini's Theorem and equation (4.10) yield

$$\begin{aligned}\mathbb{E}\left(\sup_{x\in\mathcal{D}}(\partial_{x_l}W_N(x))^{2q}\right) &= \int_0^\infty\mathbb{P}\left(\sup_{x\in\mathcal{D}}(\partial_{x_l}W_N(x))^{2q} > c\right)\mathrm{d}c\\ &\leq \int_0^\infty\exp\left(-\frac{(c^{1/(2q)}-E_{N,l})^2}{2\widetilde{\sigma}^2}\right)\mathrm{d}c\\ &\leq \int_{\mathbb{R}}\exp\left(-\frac{|c|^{1/q}}{2\widetilde{\sigma}^2}\right)\mathrm{d}c,\end{aligned} \tag{4.11}$$

and the last integral is finite for any $q \in [1,\infty)$ and independent of N and l. □

Lemma 4.9. *Under Assumption 4.1, for any $r \in [1,p)$ it holds that*

$$\|\partial_t u_{N,\varepsilon}\|_{L^r(\Omega;L^2(\mathbb{T};H))} + \left\|\sup_{t\in\mathbb{T}}\|u_{N,\varepsilon}(\cdot,\cdot,t)\|_V\right\|_{L^r(\Omega;\mathbb{R})} \leq C\left(\|u_0\|_{L^p(\Omega;V)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};H))}\right)$$

as well as

$$\|\partial_t u_{N,\varepsilon,\ell}\|_{L^r(\Omega;L^2(\mathbb{T};H))} + \left\|\sup_{t\in\mathbb{T}}\|u_{N,\varepsilon,\ell}(\cdot,\cdot,t)\|_V\right\|_{L^r(\Omega;\mathbb{R})} \le C\left(\|u_0\|_{L^p(\Omega;H^{1+\epsilon}(\mathcal{D}))} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};H))}\right).$$

Proof. We use the first part of the proof from Chapter 7.1 and Theorem 5 of [22] to obtain the pathwise estimate

$$\begin{aligned}
&\|\partial_t u_{N,\varepsilon}(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)} + \sup_{t\in\mathbb{T}}\int_{\mathcal{D}} a_{N,\varepsilon}(\omega,x,t)\nabla u_{N,\varepsilon}(\omega,x,t)\cdot\nabla u_{N,\varepsilon}(\omega,x,t)\mathrm{d}x\\
&\quad\le \int_{\mathcal{D}} a_{N,\varepsilon}(\omega,x,t)\nabla u_{N,\varepsilon}(\omega,x,0)\cdot\nabla u_{N,\varepsilon}(\omega,x,0)\mathrm{d}x\\
&\qquad + \int_0^T \|b_{N,\varepsilon}(\omega,x,t)\cdot\nabla u_{N,\varepsilon}(\omega,\cdot,t)\|^2_H\mathrm{d}t + \|f(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)}\\
&\quad\le a_{N,\varepsilon,+}(\omega)\|u_0(\omega,\cdot)\|^2_V + \bar{b}_2^2 2^{d-1}\|u(\omega,\cdot,\cdot)\|^2_{T,*} + \|f(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)}.
\end{aligned}$$

In the last step, we have used that $\|b_{N,\varepsilon}(\omega,x)\|_\infty \le \bar{b}_2$ (see Rem. 3.3) as well as equation (2.4). On the other hand, we have the lower bound

$$\begin{aligned}
&\|\partial_t u_{N,\varepsilon}(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)} + \sup_{t\in\mathbb{T}}\int_{\mathcal{D}} a_{N,\varepsilon}(\omega,x)\nabla u_{N,\varepsilon}(\omega,x,t)\cdot\nabla u_{N,\varepsilon}(\omega,x,t)\mathrm{d}x\\
&\quad\ge \|\partial_t u_{N,\varepsilon}(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)} + a_{N,\varepsilon,-}(\omega)\sup_{t\in\mathbb{T}}|u_{N,\varepsilon}(\omega,\cdot,t)|^2_{H^1(\mathcal{D})}.
\end{aligned}$$

Since the norms $|\cdot|_{H^1(\mathcal{D})}$ and $\|\cdot\|_{H^1(\mathcal{D})} = \|\cdot\|_V$ are equivalent by the Poincaré inequality, we treat $a_{N,\varepsilon,-}$ once more in the fashion of Theorem 2.7 to arrive at the estimate

$$\begin{aligned}
&\|\partial_t u_{N,\varepsilon}(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)} + \sup_{t\in\mathbb{T}}\int_{\mathcal{D}}\|u_{N,\varepsilon}(\omega,x,t)\|^2_V\mathrm{d}x\\
&\quad\le C(1+1/a_{N,\varepsilon,-}(\omega))\left(a_{N,\varepsilon,+}(\omega)\|u_0(\omega,\cdot)\|^2_V + \|u(\omega,\cdot,\cdot)\|^2_{T,*} + \|f(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)}\right).
\end{aligned}$$

The claim now follows with $1/a_{N,\varepsilon,-}, a_{N,\varepsilon,+} \in L^q(\Omega;\mathbb{R})$ for arbitrary large $q\in[1,\infty)$, Hölder's inequality and Theorem 2.7. The proof for the estimate on $u_{N,\varepsilon,\ell}$ may be carried out analogously with the initial condition $u_{N,\varepsilon,\ell}(\cdot,\cdot,0) = \mathcal{I}_\ell u_0$ and by observing that with Lemma 4.4

$$\|\mathcal{I}_\ell u_0\|_{L^p(\Omega;V)} \le \|\mathcal{I}_\ell u_0 - u_0\|_{L^p(\Omega;V)} + \|u_0\|_{L^p(\Omega;V)} \le C\|u_0\|_{L^p(\Omega;H^{1+\epsilon}(\mathcal{D}))}.$$

□

Lemma 4.10. *Under Assumptions 4.1 and 4.5, for any $r\in[2,p)$ it holds that*

$$\mathbb{E}\left(\left(\int_0^T\sum_{i\in\mathcal{I}_\mathcal{T}}\|u_{N,\varepsilon}\|^2_{H^{1+\kappa}(\mathcal{T}_i)}\mathrm{d}t\right)^{r/2}\right)^{1/r} < +\infty.$$

Proof. Assumptions 4.1 and 4.5 yield for fixed ω and t

$$\begin{aligned}
\sum_{i\in\mathcal{I}_\mathcal{T}(\omega)}\|u_{N,\varepsilon}(\omega,\cdot,t)\|^2_{H^{1+\kappa}(\mathcal{T}_i)} &\le C\frac{\|\widetilde{f}(\omega,\cdot,t)\|^2_H + \|u_{N,\varepsilon}(\omega,\cdot,t)\|^2_V\sum_{i\in\mathcal{I}_\mathcal{T}(\omega)}\|a_{N,\varepsilon}(\omega,\cdot)\|^2_{\mathcal{W}^{1,\infty}(\mathcal{T}_i)}}{a_{N,\varepsilon,-}(\omega)^2}\\
&\le C\frac{\|\widetilde{f}(\omega,\cdot,t)\|^2_H + \|u_{N,\varepsilon}(\omega,\cdot,t)\|^2_V\tau(\omega)\max_{i\in\mathcal{I}_\mathcal{T}(\omega)}\|a_{N,\varepsilon}(\omega,\cdot)\|^2_{\mathcal{W}^{1,\infty}(\mathcal{T}_i)}}{a_{N,\varepsilon,-}(\omega)^2}.
\end{aligned}$$

Now, we integrate with respect to $\mathbb{T}$ and Ω, and use Hölder's inequality to obtain for $r \in [2, p)$

$$\mathbb{E}\left(\left(\int_0^T \sum_{i\in\mathcal{I}_\mathcal{T}} \|u_{N,\varepsilon}\|^2_{H^{1+\kappa}(\mathcal{T}_i)}\mathrm{d}t\right)^{r/2}\right)^{1/r} \le C\Big(\|1/a_{N,\varepsilon,-}\|_{L^q(\Omega;L^2(\mathbb{T};H)}\|\widetilde{f}\|_{L^{r_1}(\Omega;L^2(\mathbb{T};H))}$$
$$+ \|1/a_{N,\varepsilon,-}\|_{L^{4q}(\Omega;L^2(\mathbb{T};H)}\|\tau\|_{L^{2q}(\Omega;\mathbb{R})}$$
$$\times \left\|\max_{i\in\mathcal{I}_\mathcal{T}}\|a_{N,\varepsilon}\|_{\mathcal{W}^{1,\infty}(\mathcal{T}_i)}\right\|_{L^{4q}(\Omega;\mathbb{R})}\|u_{N,\varepsilon}\|_{L^{r_1}(\Omega;L^2(\mathbb{T};V))}\Big)$$
$$\le C\left(\|\widetilde{f}\|_{L^{r_1}(\Omega;L^2(\mathbb{T};H)} + \|u_{N,\varepsilon}\|_{L^{r_1}(\Omega;L^2(\mathbb{T};V))}\right).$$

In the derivation, we have used the Hölder exponents $r_1 \in (r,p)$ and $q := (1/r - 1/r_1)^{-1} < +\infty$. The last estimate holds due to Lemmas 3.6 and 4.8 and Assumption 4.1 (i). By definition

$$\widetilde{f}(\omega,\cdot,t) = f(\omega,\cdot,t) - b_{N,\varepsilon}(\omega,\cdot)\cdot\nabla u_{N,\varepsilon}(\omega,\cdot,t) - \partial_t u_{N,\varepsilon}(\omega,\cdot,t),$$

hence Theorem 3.7, Lemmas 4.8 and 4.9 yield

$$\mathbb{E}\left(\int_0^T \sum_{i\in\mathcal{I}_\mathcal{T}} \|u_{N,\varepsilon}\|^2_{H^{1+\kappa}(\mathcal{T}_i)}\mathrm{d}t\right)^{1/2} \le C\left(\|u_0\|_{L^p(\Omega;V)} + \|f\|_{L^p(\Omega;L^2(\mathbb{T};H))}\right) < +\infty. \tag{4.12}$$

□

We are now ready to prove our main result:

Proof of Theorem 4.7. We define the error $\theta_\ell := u_{N,\varepsilon} - u_{N,\varepsilon,\ell}$ and observe that for fixed $\omega \in \Omega, t \in \mathbb{T}$ equations (4.3) and (3.5) yield

$$_{V'}\langle\partial_t\theta_\ell(\omega,\cdot,t), v_{\ell,\omega}\rangle_V + B^{N,\varepsilon}_\omega(\theta_\ell(\omega,\cdot,t), v_{\ell,\omega}) = 0$$
$$\theta_\ell(\omega,\cdot,0) = (u_0 - \mathcal{I}_\ell u_0)(\omega,\cdot),$$

for all $v_{\ell,\omega} \in V_\ell(\omega)$. We then test against $v_{\ell,\omega} = \mathcal{I}_\ell u_{N,\varepsilon}(\omega,\cdot,t) - u_{N,\varepsilon,\ell}(\omega,\cdot,t)$ and integrate over $[0,t]$ to obtain

$$\frac{1}{2}\|\theta_\ell(\omega,\cdot,t)\|^2_H + \int_0^t \left(a_{N,\varepsilon}(\omega,\cdot), \sum_{l=1}^d(\partial_{x_l}(\theta_\ell(\omega,\cdot,z))^2\right)\mathrm{d}z = \frac{1}{2}\|\theta_\ell(\omega,\cdot,0)\|^2_H$$
$$+ \int_0^t {}_{V'}\langle\partial_t\theta_\ell(\omega,\cdot,z), (1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z)\rangle_V\mathrm{d}z$$
$$+ \int_0^t B^{N,\varepsilon}_\omega\left(\theta_\ell(\omega,\cdot,z),\, (1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z)\right)\mathrm{d}z$$
$$- \int_0^t \left(b_{N,\varepsilon}(\omega,\cdot)\cdot\nabla\theta_\ell(\omega,\cdot,z),\, \theta_\ell(\omega,\cdot,z)\right)\mathrm{d}z$$
$$=: \frac{1}{2}\|\theta_\ell(\omega,\cdot,0)\|^2_H + \mathrm{I} + \mathrm{II} + \mathrm{III}. \tag{4.13}$$

Lemma 4.9 implies that $\partial_t\theta_\ell(\omega,\cdot,\cdot) \in L^2(\mathbb{T};H)$ and we use the Cauchy–Schwarz inequality to bound I:

$$\mathrm{I} = \int_0^t (\partial_t\theta_\ell(\omega,\cdot,z), (1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z))\mathrm{d}z \le \int_0^t \|\partial_t\theta_\ell(\omega,\cdot,z)\|_H\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z))\|_H\mathrm{d}z.$$

We then use the Cauchy–Schwarz inequality and equation (2.4) to bound the second term by

$$\begin{aligned}
\mathrm{II} &= \int_0^t (a_{N,\varepsilon}(\omega,\cdot), \nabla\theta_\ell(\omega,\cdot,z)\cdot\nabla(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z))\mathrm{d}z \\
&\quad + \int_0^t (b_{N,\varepsilon}(\omega,\cdot)\cdot\nabla\theta_\ell(\omega,\cdot,z), (1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z))\mathrm{d}z \\
&\le \int_0^t \left(a_{N,\varepsilon}(\omega,\cdot)\left(\sum_{l=1}^d(\partial_{x_l}\theta_\ell(\omega,\cdot,z))^2\right)^{1/2}, \left(\sum_{l=1}^d(\partial_{x_l}(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z))^2\right)^{1/2}\right)\mathrm{d}z \\
&\quad + \int_0^t 2^{d/2-1/2}\left(\|b_{N,\varepsilon}(\omega,\cdot)\|_\infty\left(\sum_{l=1}^d(\partial_{x_l}\theta_\ell(\omega,\cdot,z))^2\right)^{1/2}, |(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z)|\right)\mathrm{d}z.
\end{aligned}$$

Young's inequality further yields

$$\begin{aligned}
\mathrm{II} &\le \int_0^t \frac{1}{4}\left(a_{N,\varepsilon}(\omega,\cdot), \sum_{l=1}^d(\partial_{x_l}(\theta_\ell(\omega,\cdot,z))^2\right) + a_{N,\varepsilon,+}(\omega)|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z)|^2_{H^1(\mathcal{D})}\mathrm{d}z \\
&\quad + \int_0^t \frac{1}{4}\left(a_{N,\varepsilon}(\omega,\cdot), \sum_{l=1}^d(\partial_{x_l}(\theta_\ell(\omega,\cdot,z))^2\right) + 2^{d-1}\,\overline{b}_1^2 a_{N,\varepsilon,+}(\omega)\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z)\|_H^2\mathrm{d}z \\
&\le \frac{1}{2}\left(a_{N,\varepsilon}(\omega,\cdot), \sum_{l=1}^d(\partial_{x_l}(\theta_\ell(\omega,\cdot,z))^2\right)\mathrm{d}z + Ca_{N,\varepsilon,+}(\omega)\int_0^t\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,z)\|_V^2\mathrm{d}z.
\end{aligned}$$

Similarly, we bound the last term by

$$|\mathrm{III}| \le \frac{1}{4}\int_0^t\left(a_{N,\varepsilon}(\omega,\cdot), \sum_{l=1}^d(\partial_{x_l}(\theta_\ell(\omega,\cdot,z))^2\right)\mathrm{d}z + 2^{d-1}\overline{b}_1\overline{b}_2\int_0^t\|\theta_\ell(\omega,\cdot,z)\|_H^2\mathrm{d}z.$$

We now plug in the estimates for I – III in equation (4.13) and proceed in the fashion of Theorem 2.7 with Grönwalls inequality to arrive at

$$\begin{aligned}
\sup_{t\in\mathbb{T}}\|\theta_\ell\|_{t,*}^2 &\le C(1+1/a_{N,\varepsilon,-}(\omega))\Big(\|\theta_\ell(\omega,\cdot,0)\|_H^2 + \|\partial_t\theta_\ell(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};H)}\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};H)} \\
&\quad + a_{N,\varepsilon,+}(\omega)\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};V)}^2\Big).
\end{aligned}$$

By Lemma 4.9 it holds that $\partial_t\theta_\ell \in L^r(\Omega; L^2(\mathbb{T}; H)$ for any $r \in (2, p)$ with p as in Assumption 4.1 (i). Taking expectations, using Hölder's inequality and $1/a_{N,\varepsilon,-}, a_{N,\varepsilon,+} \in L^q(\Omega;\mathbb{R})$ for all $q \ge 1$ then yields

$$\begin{aligned}
\mathbb{E}(\sup_{t\in\mathbb{T}}\|\theta_\ell\|_{t,*}^2)^{1/2} &\le C\Big(\|\theta_\ell(\omega,\cdot,0)\|_{L^p(\Omega;H)} + \|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,\cdot)\|_{L^r(\Omega;L^2(\mathbb{T};H))}^{1/2} \\
&\quad + \|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,\cdot)\|_{L^r(\Omega;L^2(\mathbb{T};V))}\Big).
\end{aligned}$$

By Assumption 4.1 (i) $u_0 \in L^p(\Omega; H^{1+\epsilon}(\mathcal{D}))$ for some $\epsilon > 0$, which yields with Lemma 4.4

$$\|\theta_\ell(\omega,\cdot,0)\|_{L^p(\Omega;H)} = \|(1-\mathcal{I}_\ell)u_0\|_{L^p(\Omega;H)} \le C\overline{h}_\ell^{1+\epsilon}.$$

Furthermore, Lemmas 4.4 and 4.10 yield

$$\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,\cdot)\|_{L^r(\Omega;L^2(\mathbb{T};H))}^{1/2} \le C\overline{h}_\ell^{(\kappa+1)/2}\mathbb{E}\left(\left(\int_0^T\sum_{i\in\mathcal{I}_\mathcal{T}}\|u_{N,\varepsilon}\|_{H^{1+\kappa}(\mathcal{T}_i)}^2\mathrm{d}t\right)^{r/2}\right)^{1/2r} \le C\overline{h}_\ell^{(\kappa+1)/2}$$

as well as

$$\|(1-\mathcal{I}_\ell)u_{N,\varepsilon}(\omega,\cdot,\cdot)\|_{L^r(\Omega;L^2(\mathbb{T};V))} \le C\overline{h}_\ell^\kappa \mathbb{E}\left(\left(\int_0^T \sum_{i\in\mathcal{I}_\mathcal{T}} \|u_{N,\varepsilon}\|^2_{H^{1+\kappa}(\mathcal{T}_i)}\mathrm{d}t\right)^{r/2}\right)^{1/2r} \le C\overline{h}_\ell^\kappa.$$

The claim now follows since $0<\kappa$, $\overline{h}_\ell \le 1$. □

Remark 4.11. To ensure that the convergence of order $\overline{h}_\ell^\kappa$ in Theorem 4.7 is not affected by the Gaussian field W, Assumption 4.1 (ii) cannot be relaxed. For instance, given that $2\alpha>\beta$, it follows from Proposition 3.4 of [14] that a is piecewise Hölder-continuous with exponent $\varrho<\beta/2\alpha$ and we may only expect a rate of order $\overline{h}_\ell^{\min(\varrho,\kappa)}$, see Section 3 of [15], Section 5 of [45] and Chapter 10.1 of [27]. In fact, we discuss an example with $\varrho = 1/2-\epsilon$ in Section 5 and show that we only achieve a convergence rate of approximately $\overline{h}_\ell^{1/2}$, even for sample-adapted FE.

4.2. Temporal discretization

In the remainder of this section, we introduce a stable temporal discretization for the semi-discrete Problem (4.3) and derive the corresponding mean-square error. To this end, we fix $\omega\in\Omega$ and let $u_{N,\varepsilon,\ell}(\omega,\cdot,\cdot)$ denote the sample-adapted semi-discrete approximation of $u_{N,\varepsilon}(\omega,\cdot,\cdot)$ from equation (4.3). For a fully discrete formulation of Problem (4.3), we consider a time grid $0=t_0<t_1<\cdots<t_n=T$ in $\mathbb{T}$ for some $n\in\mathbb{N}$. The temporal derivative at t_i is approximated by the backward difference

$$\partial_t u_{N,\varepsilon,\ell}(\omega,\cdot,t_i) \approx \frac{u_{N,\varepsilon,\ell}(\omega,\cdot,t_i)-u_{N,\varepsilon,\ell}(\omega,\cdot,t_{i-1})}{t_i-t_{i-1}}, \quad i=1,\dots,n.$$

This yields the fully discrete problem to find $(u^{(i)}_{N,\varepsilon,\ell}(\omega,\cdot),\, i=0,\dots,n)\subset V_\ell(\omega)$ such that for all $i=1,\dots,n$ and $v_{\ell,\omega}\in V_\ell(\omega)$

$$\begin{aligned}\frac{1}{t_i-t_{i-1}}(u^{(i)}_{N,\varepsilon,\ell}(\omega,\cdot)-u^{(i-1)}_{N,\varepsilon,\ell}(\omega,\cdot),v_{\ell,\omega}) + B^{N,\varepsilon}_\omega(u^{(i)}_{N,\varepsilon,\ell}(\omega,\cdot),v_{\ell,\omega}) &= F_{t_i,\omega}(v_{\ell,\omega})\\ u^{(0)}_{N,\varepsilon,\ell}(\omega,\cdot) &= \mathcal{I}_\ell u_0(\omega,\cdot).\end{aligned} \tag{4.14}$$

For convenience, we assume an equidistant temporal grid with fixed time step $\Delta t := t_i - t_{i-1} > 0$. The fully discrete solution is now given by

$$u^{(i)}_{N,\varepsilon,\ell}(\omega,x) = \sum_{j=1}^{d_\ell} c_{i,j}(\omega)v_{j,\omega}(x), \quad i=1,\dots,n,$$

where the coefficient vector $\mathbf{c_i}(\omega) := (c_{i,1}(\omega),\dots,c_{i,d_\ell}(\omega))^T$ solves the linear system of equations

$$(\mathbf{M}(\omega)+\Delta t\mathbf{A}(\omega))\mathbf{c_i}(\omega) = \Delta t\mathbf{F}(\omega,t_i)+\mathbf{M}(\omega)\mathbf{c_{i-1}}(\omega)$$

in every discrete point t_i. The mass matrix consists of the entries $(\mathbf{M}(\omega))_{jk} := (v_{j,\omega},v_{k,\omega})$, the stiffness matrix and load vector are given by $(\mathbf{A}(\omega))_{jk} = B^{N,\varepsilon}_\omega(v_{j,\omega},v_{k,\omega})$ and $(\mathbf{F}(\omega,t_i))_j = F_{t_i,\omega}(v_{j,\omega})$ for $j,k\in\{1,\dots,d_\ell(\omega)\}$, respectively, as in the semi-discrete case. The initial vector c_0 consists of the basis coefficients of $\mathcal{I}_\ell u_0\in V_\ell$ with respect to $\{v_{1,\omega},\dots,v_{d_\ell(\omega),\omega}\}$. To extend the fully discrete solution $(u^{(i)}_{N,\varepsilon,\ell}(\omega,\cdot), i=0,\dots,n)$ to $\mathbb{T}$, we define the linear interpolation

$$\overline{u}_{N,\varepsilon,\ell}(\omega,\cdot,t) := \left(u^{(i)}_{N,\varepsilon,\ell}(\omega,\cdot)-u^{(i-1)}_{N,\varepsilon,\ell}(\omega,\cdot)\right)\frac{(t-t_{i-1})}{\Delta t} + u^{(i-1)}_{N,\varepsilon,\ell}(\omega,\cdot), \quad t\in[t_{i-1},t_i], \quad i=1,\dots,n,$$

and are, therefore, able to estimate the resulting error with respect to the parabolic norm.

Theorem 4.12. *Let Assumption 4.1 hold, let* $\left(u^{(i)}_{N,\varepsilon,\ell}, i=0,\dots,n\right)$ *be the fully discrete sample-adapted approximation of* $u_{N,\varepsilon}$ *as in equation (4.14) and let* $\overline{u}_{N,\varepsilon,\ell}$ *be the linear interpolation in* $\mathbb{T}$. *Then,*

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u_{N,\varepsilon,\ell} - \overline{u}_{N,\varepsilon,\ell}\|^2_{*,t}\right)^{1/2} \le C\Delta t.$$

Proof. We start by investigating the temporal regularity of $u_{N,\varepsilon,\ell}$. For fixed $\omega\in\Omega$ and $0 \le t_{i-1} < t_i \le T$ note that $w_i(\omega,\cdot,t) := u_{N,\varepsilon,\ell}(\omega,\cdot,t) - u_{N,\varepsilon,\ell}(\omega,\cdot,t_{i-1})$ solves the variational problem

$${}_{V'}\langle \partial_t w_i(\omega,\cdot,t), v_{\ell,\omega}\rangle_V + B^{N,\varepsilon}_\omega(w_i(\omega,\cdot,t), v_{\ell,\omega}) = {}_{V'}\langle f(\omega,\cdot,t) - f(\omega,\cdot,t_{i-1}), v_{\ell,\omega}\rangle_V$$

for $t\in[t_{i-1},t_i]$ and $v_{\ell,\omega}\in V_\ell(\omega)$ with initial condition $w(\omega,\cdot,t_{i-1}) = 0$. Therefore, in the fashion of Theorem 2.7, we obtain the pathwise parabolic estimate

$$\begin{aligned}
\sup_{t\in[t_{i-1},t_i]} \|w_i(\omega,\cdot,t)\|^2_H + \int_{t_{i-1}}^t |w_i(\omega,\cdot,z)|^2_{H^1(\mathcal{D})}\mathrm{d}z &\le C(1+1/a_{N,\varepsilon,-}(\omega)) \int_{t_{i-1}}^{t_i} \|f(\omega,\cdot,z) - f(\omega,\cdot,t_{i-1})\|^2_H \mathrm{d}z \\
&= C(1+1/a_{N,\varepsilon,-}(\omega)) \int_{t_{i-1}}^{t_i} \left\| \int_{t_{i-1}}^z \partial_t f(\omega,\cdot,\widetilde{z})d\widetilde{z}\right\|^2_H \mathrm{d}z \\
&= C(1+1/a_{N,\varepsilon,-}(\omega)) \\
&\quad\times \int_{t_{i-1}}^{t_i} \left\| \int_{t_{i-1}}^{t_i} \mathbf{1}_{[t_{i-1},z]}(\widetilde{z})\partial_t f(\omega,\cdot,\widetilde{z})d\widetilde{z}\right\|^2_H \mathrm{d}z \\
&\le C(1+1/a_{N,\varepsilon,-}(\omega)) \\
&\quad\times \int_{t_{i-1}}^{t_i} (z-t_{i-1})\mathrm{d}z \int_{t_{i-1}}^{t_i} \|\partial_t f(\omega,\cdot,z)\|^2_H \mathrm{d}z \\
&= C(1+1/a_{N,\varepsilon,-}(\omega)) \frac{\Delta t^2}{2} \|\partial_t f(\omega,\cdot,\cdot)\|^2_{L^2([t_i,t_{i-1}];H)}. \qquad (4.15)
\end{aligned}$$

For the first identity we have used Lemma 2.2, the second estimate follows with Hölder's inequality. Now let $\overline{\overline{u}}_{N,\varepsilon,\ell}$ be the linear interpolation of the semi-discrete solution $u_{N,\varepsilon,\ell}$ at $t_0,\dots,t_n$ and consider the splitting

$$\mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u_{N,\varepsilon,\ell} - \overline{u}_{N,\varepsilon,\ell}\|^2_{*,t}\right)^{1/2} \le \mathbb{E}\left(\sup_{t\in\mathbb{T}} \|u_{N,\varepsilon,\ell} - \overline{\overline{u}}_{N,\varepsilon,\ell}\|^2_{*,t}\right)^{1/2} + \mathbb{E}\left(\sup_{t\in\mathbb{T}} \|\overline{\overline{u}}_{N,\varepsilon,\ell} - \overline{u}_{N,\varepsilon,\ell}\|^2_{*,t}\right)^{1/2} =: \mathrm{I} + \mathrm{II}.$$

By equation (4.15) it follows that

$$\begin{aligned}
\sup_{t\in\mathbb{T}} \|u_{N,\varepsilon,\ell} - \overline{\overline{u}}_{N,\varepsilon,\ell}\|^2_{*,t} &\le \max_{i=1,\dots,n} \sup_{t\in[t_{i-1},t_i]} \|w_i(\omega,\cdot,t)\|^2_H + 2\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \|w_i(\omega,\cdot,t)\|^2_{H^1(\mathcal{D})}\mathrm{d}t \\
&\le C(1+1/a_{N,\varepsilon,-}(\omega))\frac{\Delta t^2}{2}\|\partial_t f(\omega,\cdot,\cdot)\|^2_{L^2(\mathbb{T};H)}.
\end{aligned}$$

By Assumption 4.1, Hölder's inequality and Lemma 3.6 with $q = (1/2 - 1/p)^{-1}$

$$\mathrm{I} \le C\Delta t(1 + \|1/a_{N,\varepsilon,-}\|_{L^q(\Omega;\mathbb{R})})\|\partial_t f\|_{L^p(\Omega;L^2(\mathbb{T};H))} \le C\Delta t.$$

Now let $\theta^{(i)}(\omega,\cdot) := u_{N,\varepsilon,\ell}(\omega,\cdot,t_i) - u^{(i)}_{N,\varepsilon,\ell}(\omega,\cdot)$ denote the pathwise time discretization error at t_i. For any $t\in[t_{i-1},t_i]$, we observe that $(\overline{u}_{N,\varepsilon,\ell} - \overline{\overline{u}}_{N,\varepsilon,\ell})(\cdot,\cdot,t)$ is a convex combination of θ_i and θ_{i-1}, and it holds that

$$\mathrm{II} \le \mathbb{E}(\max_{i=1,\dots,n} \|\theta^{(i)}\|^2_H + \Delta t\sum_{j=1}^i |\theta^{(j)}|^2_{H^1(\mathcal{D})})^{1/2}. \qquad (4.16)$$

Hence, it is sufficient to control the errors at each t_i. Combining equations (4.14) and (4.3) yields for $i=1,\dots,n$

$$
\begin{aligned}
&{}_{V'}\langle \theta^{(i)}(\omega,\cdot)-\theta^{(i-1)}(\omega,\cdot), v_{\ell,\omega}\rangle_V + \int_{t_{i-1}}^{t_i} B_\omega^{N,\varepsilon}(\theta^{(i)}(\omega,\cdot), v_{\ell,\omega})\mathrm{d}t \\
&\quad = \int_{t_{i-1}}^{t_i} B_\omega^{N,\varepsilon}(u_{N,\varepsilon,\ell}(\omega,\cdot,t_i)-u_{N,\varepsilon,\ell}(\omega,\cdot,t), v_{\ell,\omega}) + {}_{V'}\langle f(\omega,\cdot,t)-f(\omega,\cdot,t_i), v_{\ell,\omega}\rangle_V \mathrm{d}t \\
&\quad =: \int_{t_{i-1}}^{t_i} {}_{V'}\langle \overline{f}_i(\omega,\cdot,t), v_{\ell,\omega}\rangle_V \mathrm{d}t
\end{aligned}
$$

and initial condition $u_{N,\varepsilon,\ell}(\omega,\cdot,0)-u_{N,\varepsilon,\ell}^{(0)}(\omega,\cdot)=0$. We now test against $v_{\ell,\omega}=\theta^{(i)}(\omega,\cdot)$, sum with respect to i and use the discrete Grönwall inequality to obtain (as in Thm. 2.7) the discrete estimate

$$
\begin{aligned}
\max_{i=1,\dots,n}\|\theta^{(i)}(\omega,\cdot)\|_H^2 + \Delta t\sum_{j=1}^{i}|\theta^{(j)}(\omega,\cdot)|_{H^1(\mathcal{D})}^2 &\le C(1+1/a_{N,\varepsilon,-}(\omega))\sum_{j=1}^{n}\|\overline{f}_j(\omega,\cdot,\cdot)\|_{L^2([t_j,t_{j-1}];V')}^2 \\
&\le C(1+1/a_{N,\varepsilon,-}(\omega)) \\
&\quad\times\left(\sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}|u_{N,\varepsilon}(\cdot,\cdot,t)-u_{N,\varepsilon}(\cdot,\cdot,t_j)|_{H^1(\mathcal{D})}^2\mathrm{d}t\right. \\
&\quad\left.+\sum_{j=1}^{n}\|f(\omega,\cdot,t)-f(\omega,\cdot,t_j)\|_H^2\mathrm{d}t\right).
\end{aligned}
$$

Proceeding as for w_i in equation (4.15) to bound each term in both sums, this implies

$$
\max_{i=1,\dots,n}\|\theta^{(i)}(\omega,\cdot)\|_H^2 + \Delta t\sum_{j=1}^{i}|\theta^{(j)}(\omega,\cdot)|_{H^1(\mathcal{D})}^2 \le C(1+1/a_{N,\varepsilon,-}(\omega))a_{N,\varepsilon,+}(\omega)^2\Delta t^2\|\partial_t f(\omega,\cdot,\cdot)\|_{L^2(\mathbb{T};H)}^2.
$$

We use Assumption 4.1, Hölder's inequality and Lemma 3.6 to obtain

$$
\mathbb{E}\left(\max_{i=1,\dots,n}\|\theta^{(i)}\|_H^2 + \Delta t\sum_{j=1}^{i}|\theta^{(j)}|_{H^1(\mathcal{D})}^2\right)^{1/2} \le C\Delta t\|\partial_t f\|_{L^p(\Omega;L^2(\mathbb{T};H))} \le C\Delta t,
$$

and the claim finally follows by equation (4.16). □

To conclude this section, we record a bound on the overall approximation error, which is an immediate consequence of Theorems 3.9, 4.7 and 4.12.

Corollary 4.13. *Let Assumptions 4.1 and 4.5 hold and let $\overline{u}_{N,\varepsilon,\ell}$ be the linear interpolation of the fully discrete approximation of $(u_{N,\varepsilon}^{(i)}, i=0,\dots,n)$. Then,*

$$
\mathbb{E}\left(\sup_{t\in\mathbb{T}}\|u-\overline{u}_{N,\varepsilon,\ell}\|_{*,t}^2\right)^{1/2} \le C\left(\varXi_N^{1/2}+\varepsilon^{1/s}+\overline{h}_\ell^\kappa+\Delta t\right).
$$

5. Numerical experiments

5.1. Experimental setting and error balancing

In all of our numerical experiments we measure the root mean-square error

$$
\mathrm{RMSE} := \mathbb{E}\left(\|u(\cdot,\cdot,T)-\overline{u}_{N,\varepsilon,\ell}(\cdot,\cdot,T)\|_V^2\right)^{1/2}.
$$

For each given FE discretization parameter $\overline{h}_\ell$, we align the error contributions of N, ε and Δt such that $\Xi_N^{1/2} \simeq \varepsilon^{1/s} \simeq \Delta t \simeq \overline{h}_\ell$. Hence, the dominant source of error is the spatial discretization and Corollary 4.13 yields $RMSE \le C\overline{h}_\ell^\kappa$. This allows us to measure the value of κ in Assumption 4.5 by linear regression. While the choices of Δt and ε are usually straightforward for given $\overline{h}_\ell$, we refer to Remark 5.3 of [11], where we describe how to achieve $\Xi_N^{1/2} \simeq \overline{h}_\ell$ for common examples of covariance operators Q. To emphasize the advantage of the sample-adapted FE algorithm introduced in Section 4, we also repeat all experiments with a standard FE approach and compare the resulting errors. For the non-adapted FE algorithm, we use for a given triangulation diameter h_ℓ the same approximation parameters $\Delta t, N$ and ε as for the corresponding sample-adapted method. This ensures that the weaker performance of this non-adapted method is due to the mismatch between FE triangulation and the discontinuities of a and b. We approximate the entries of the stiffness matrix for both FE approaches by the midpoint rule on each interval (in 1D) or triangle (in 2D). If the triangulation is aligned to the discontinuities in a and b, this adds an additional non-dominant term of order $\overline{h}_\ell$ to the error estimate in Corollary 4.13, see for instance Proposition 3.13 of [15]. Thus, the bias stemming from the midpoint rule does not dominate the overall order of convergence in the sample-adapted algorithm. In the other case, we cannot quantify the quadrature error due to the discontinuities on certain triangles but suggest an error of order $\overline{h}_\ell^{1/2}$ based on our experimental observations.

5.2. Numerical examples in 1D

For the first scenario we consider the advection-diffusion problem (2.1) on $\mathcal{D} = (0,1)$, with $T = 1$, $u_0(x) = \sin(\pi x)/10$ and source term $f \equiv 1$. The continuous part of the jump-diffusion coefficient a is given by $\overline{a} \equiv 0$ and $\Phi(w) = \exp(w)$, where W is a Gaussian field characterized by the *Matérn covariance operator*

$$Q_M : H \to H, \quad [Q_M\varphi](y) := \int_{\mathcal{D}} \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\sqrt{2\nu}\frac{|x-y|}{\delta}\right)^\nu K_\nu \left(\sqrt{2\nu}\frac{|x-y|}{\delta}\right) \varphi(x)\mathrm{d}x \quad \text{for} \quad \varphi \in H,$$

with smoothness parameter $\nu > 0$, variance $\sigma^2 > 0$ and correlation length $\delta > 0$. Above, Γ denotes the Gamma function and K_ν is the modified Bessel function of the second kind with ν degrees of freedom. It is known that W is mean square differentiable if $\nu > 1$ and, moreover, the paths of W are almost surely in $C^{\lfloor \nu \rfloor, \varrho}(\overline{\mathcal{D}}; \mathbb{R})$ with $\varrho < \nu - \lfloor \nu \rfloor$ for any $\nu \ge 1/2$, see Section 2.2 of [24]. The spectral basis of Q_M may be efficiently approximated by Nyström's method, see for instance [44]. In our experiments, we use parameters $\nu = 3/2$, $\sigma^2 = 1$ and $\delta = 0.05$.

For each experiment in one dimension, the number of (non-empty) partition elements is given by $\tau = \mathcal{P} + 2$, where $\mathcal{P}$ is Poisson-distributed with intensity parameter 5. On average, this splits the domain in 7 disjoint intervals and the diffusion coefficient has almost surely at least one discontinuity. Given τ, the position of each jump is sampled uniformly on $\mathcal{D}$. More precisely, let $(\widetilde{x}_i, i \in \mathbb{N})$ be an i.i.d. sequence of $\mathcal{U}(\mathcal{D})$-random variables that are independent of τ. We take the first $\tau - 1$ points of this sequence, order them increasingly and denote the ordered subset by $x_0 := 0 < x_1 < \cdots < x_{\tau-1} < x_\tau := 1$. Hence, we obtain the (Borel-measurable) sequence of correspondences

$$\mathcal{T}_i(\omega) := \begin{cases} (x_{i-1}(\omega), x_i(\omega)), & i \le \tau(\omega) \\ \emptyset, & i > \tau(\omega) \end{cases}, \quad i \in \mathbb{N},$$

that generates the random partition $\mathcal{T} = \{(0, x_1), (x_1, x_2), \ldots, (x_{\tau-1}, 1)\}$ for each realization of τ. The conditional distribution of x_i (with respect to $\tau = \mathcal{P} + 2 \ge 2$) is for $i = 1, \ldots, \tau - 1$ given by

$$\mathbb{P}(x_i \le c \,|\tau) = \frac{(\tau-1)!}{(\tau-i)!(i-1)!} c^{\tau-i}(1-c)^{i-1}, \quad c \in \mathcal{D} = (0,1).$$

This can be utilized to derive further statistics, such as the average interval width of $\mathcal{T}$ given by

$$\mathbb{E}(\mathbb{E}(x_1|\tau)) = \mathbb{E}\left(\int_0^1 c^{\tau-1}\mathrm{d}c\right) = \mathbb{E}(1/\tau) = \sum_{k=0}^\infty \frac{5^k e^{-5}}{k!}\frac{1}{k+2} \approx 0.1603$$

with corresponding variance $\mathbb{E}(1/(\tau+1)) - \mathbb{E}(1/\tau)^2 \approx 0.1102$. This also shows that increasing the Poisson parameter in $\mathcal{P}$ resp. τ would yield a longer average computational time, as more and smaller intervals would be sampled. The order of spatial convergence of the sample-adapted FE scheme on the other hand remains unaffected of the distribution of $\mathcal{T}$. We use the jump-advection coefficient given by

$$b(\omega, x) := 2\sin(2\pi x)a(\omega, x), \quad \omega \in \Omega,\ x \in \mathcal{D}.$$

Note that we did not impose an upper deterministic bound $\overline{b}_2$ on b. To obtain pathwise approximations of the samples $u_{N,\varepsilon}(\omega,\cdot,\cdot)$, we use non-adapted and sample-adapted piecewise linear elements and compare both approaches. The FE discretization parameter is given by $\overline{h}_\ell = 2^{-\ell}/4$ and we consider the range $\ell = 1,\dots,6$. We approximate the reference solution u for each sample using sample-adapted FE and set $u_{ref} := \overline{u}_{N_8,\varepsilon_8,8}(\cdot,\cdot,T)$, where we choose $\Delta t_8 \simeq \Xi_{N_8}^{1/2} \simeq \varepsilon_8^{1/2} \simeq 2^{-10}$. The RMSE is estimated by averaging 100 samples of $\|u_{ref} - \overline{u}_{N,\varepsilon,\ell}(\cdot,\cdot,T)\|_V^2$ for $\ell = 1,\dots,6$. To subtract sample-adapted\non-adapted approximations from the reference solution u_{ref}, we use a fixed grid with $2^{10}+1$ equally spaced points in $\mathcal{D}$, thus the error stemming from interpolation\prolongation may be neglected. Given that $RMSE \approx C\overline{h}_\ell^\kappa$, it holds that

$$\log(RMSE) \approx \kappa \log(\overline{h}_\ell) + \log(C)$$

and we estimate the convergence rate κ by a linear regression of the log-RMSE on the log-refinement sizes $\log(\overline{h}_\ell)$. As we consider 1D-problems in this subsection, we expect convergence rates close to one for the sample-adapted method whenever Assumption 4.1 holds.

In our first numerical example, the jump heights P_i follow a *generalized inverse Gaussian* (GIG) distribution with density

$$f_{\mathrm{GIG}}(x) = \frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\psi\chi})} x^{\lambda-1} \exp\left(-\frac{1}{2}(\psi x + \chi x^{-1})\right), \quad x > 0,$$

and parameters $\chi, \psi > 0$, $\lambda \in \mathbb{R}$, see [8]. Unbiased sampling from this distribution may be rather expensive, hence we generate approximations $\widetilde{P}_i$ of P_i by a Fourier inversion technique which guarantees that $\mathbb{E}(|\widetilde{P}_i - P_i|^2) \le \varepsilon$ for any desired $\varepsilon > 0$. This allows us to adjust the sampling bias $\varepsilon > 0$ with $\overline{h}_\ell$ (and the corresponding Δt and Ξ_N) for any $\ell \in \mathbb{N}_0$. Details on the Fourier inversion algorithm, the sampling of GIG distributions and the corresponding $L^2(\Omega;\mathbb{R})$-error may be found in [10]. The GIG parameters are set as $\psi = 0.25, \chi = 9$ and $\overline{\lambda} = -1$, the resulting density and a sample of the coefficients are given in Figure 1. As expected, we see in Figure 1 that the sample-adapted algorithm converges with rate $\kappa = 0.85$. Thus, the sampling error of the GIG jump heights does not dominate the remaining error contributions. Compared to adapted FE, the non-adapted method converges at a significantly lower rate of 0.57. In Remark 4.11, we stated that the condition $2\alpha \le \beta$ on the decay of the eigenvalues of Q entails mean square differentiability of W and thus a convergence rate of order κ in the sample-adapted method. We suggested that this rate will deteriorate if the paths of W are only Hölder continuous with exponent $\varrho < \kappa \le 1$. To illustrate this, we repeat the first experiment with a changed covariance operator. We now consider the *Brownian motion covariance operator*

$$Q_{\mathrm{BM}} : H \to H, \quad [Q_{\mathrm{BM}}\varphi](y) := \int_{\mathcal{D}} \min(x,y)\varphi(x)\mathrm{d}x \quad \text{for} \quad \varphi \in H,$$

with eigenbasis given by $\eta_i = (2\sqrt{2}/((2i+1)\pi))^2$ and $e_i(x) = \sin((2i+1)\pi x/2)$ for $i \in \mathbb{N}_0$. The paths of W generated with Q_{BM} are Hölder-continuous with $\varrho = 1/2 - \epsilon$ for any $\epsilon > 0$ because $\beta = 1 - \epsilon$ and $\alpha = 1$. A sample of the coefficients is given in Figure 2. The sample-adapted RMSE is smaller than the non-adapted curve and decreases slightly faster, but both errors now decay at a lower rate of roughly $1/2$ due to the lack of (piecewise) spatial regularity of a and b.

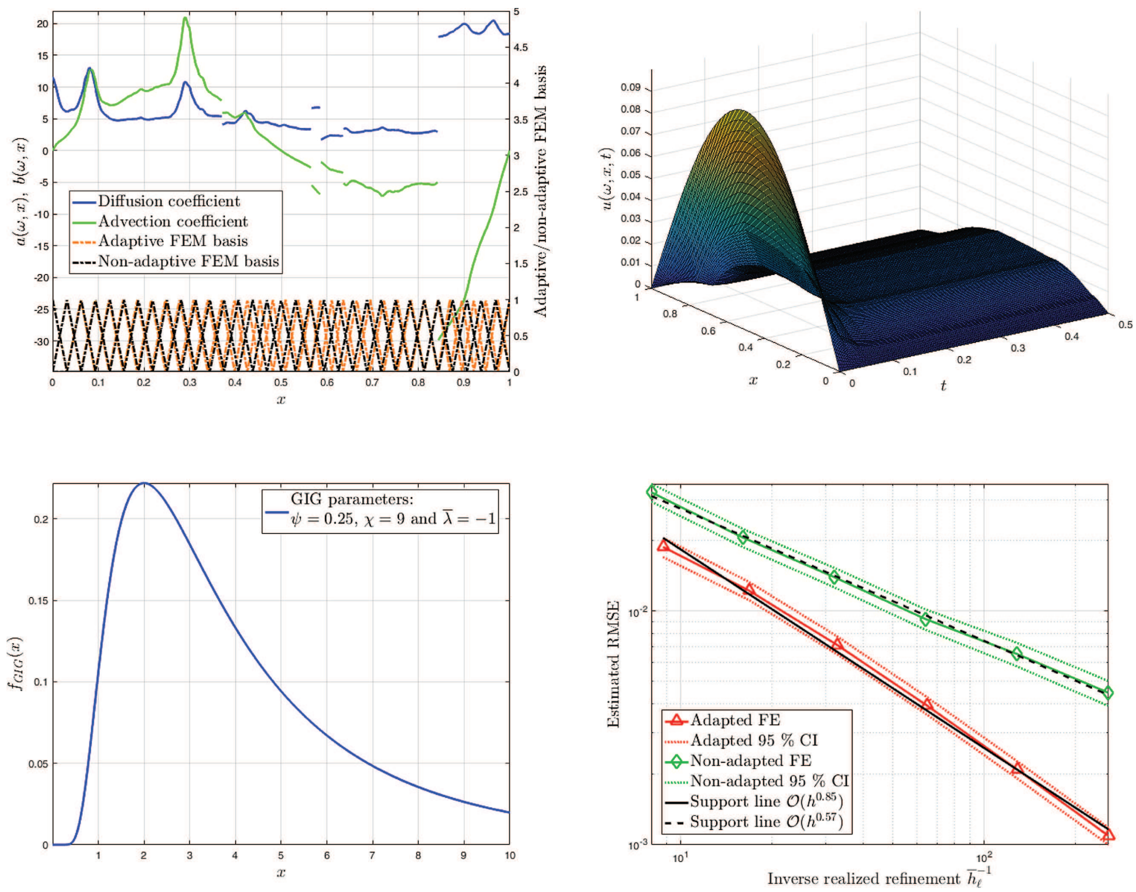

FIGURE 1. *Top left*: Jump-diffusion/advection coefficient and adapted/non-adapted FE basis, *top right*: FE solution corresponding to the sample on the left and the given sample-adapted FE basis, *bottom left*: GIG density function and parameters, *bottom right*: estimated RMSE *vs.* inverse spatial refinement.

5.3. Numerical examples in 2D

In two spatial dimensions, we work on $\mathcal{D} = (0,1)^2$ with $T = 1$, initial data $u_0(x_1, x_2) = \frac{1}{100}\sin(\pi x_1)\sin(\pi x_2)$, source term $f \equiv 1$ and assume that $\bar{a} \equiv 0$. The Gaussian part of a is determined by the Karhunen–Loève expansion

$$W(x) = \sum_{i \in \mathbb{N}} \sqrt{\eta_i} e_i(x) Z_i, \quad x \in \mathcal{D}, \quad Z_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$

with spectral basis given by $\eta_i := \sigma^2 \exp(-\pi^2 i^2 \delta^2)$ and $e_i(x) := \sin(\pi i x_1)\sin(\pi i x_2)$ for $i \in \mathbb{N}$. Again, the parameters $\delta, \sigma^2 > 0$ denote the correlation length and total variance of W, respectively. It can be shown that these eigenpairs solve the integral equation

$$\sigma^2 \int_{\mathcal{D}} \frac{1}{4\pi t} \exp\left(- \frac{-||x-y||_2^2}{2\delta^2} \right) e_i(y) \mathrm{d}y = \eta_i e_i(x), \quad i \in \mathbb{N},$$

with $e_i = 0$ on $\partial\mathcal{D}$, see [26]. This random field vanishes at the boundary and has a very similar regularity properties to a Gaussian field with squared exponential covariance operator. It, further, has the advantage, that

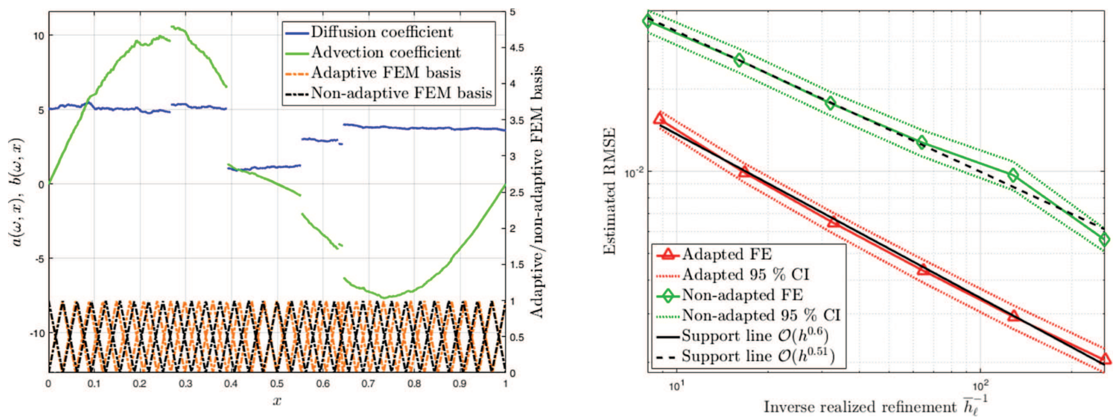

FIGURE 2. Second numerical example in 1D with Brownian motion covariance operator and uniformly distributed jumps. *Left*: Jump-diffusion/advection coefficient and adapted/non-adapted FE basis, *right*: estimated RMSE *vs.* inverse spatial refinement size.

all expressions are available in closed form and we forgo the numerical approximation of the eigenbasis. The eigenvalues decay exponentially fast with respect to i, hence Assumption 4.1 is fulfilled and we use the parameters $\sigma^2 = 0.25$ and $\delta = 0.02$ for all experiments in this section. As before, we consider a log-Gaussian random field, meaning $\Phi(w) = \exp(w)$. To illustrate the flexibility of a jump-diffusion coefficient a as in Definition 3.2, we vary the random partitioning of $\mathcal{D}$ for each example and give a detailed description below. We set the spatial refinement to $\overline{h}_\ell = h_\ell = \frac{2}{5}2^{-\ell}$ and consider the cases $\ell = 1, \dots, 5$. To estimate the RMSE, we sample similar to the one-dimensional case the reference solution $u_{ref} := \overline{u}_{N_7,\varepsilon_7,7}(\cdot,\cdot,T)$ with $\Delta t_7 \simeq \Xi_{N_7}^{1/2} \simeq \varepsilon_7^{1/2} \simeq \frac{2}{5}2^{-7}$ and average again 100 independent samples of $\|u_{ref} - \overline{u}_{N,\varepsilon,\ell}(\cdot,\cdot,T)\|_V^2$. For interpolation/prolongation we use a reference grid with $(2^8+1)\times(2^8+1)$ equally spaced points in $\mathcal{D}$. The convergence rate, *i.e.* the exponent κ from Assumption 4.5, in the sample-adapted method is estimated by linear regression as for the one-dimensional examples. We further use in each scenario the (unbounded) jump-advection coefficient

$$b(\omega,x,y) = 5\sin(\pi x)\sin(\pi y)a(\omega,x,y)\begin{pmatrix}1\\1\end{pmatrix}, \quad \omega \in \Omega,\ (x,y) \in \mathcal{D}.$$

In our first 2D example, we aim to imitate the structure of a heterogeneous medium. For this, we divide the domain by two horizontal and vertical lines. We assume that the horizontal resp. vertical lines do not intersect each other and thus obtain $\tau \equiv 9$ nonempty partition elements. The four intersection points of the lines in $\mathcal{D}$ are independent and uniformly distributed in $(0.2, 0.8)^2$ and thus define a partition of the $\mathcal{D}$ into 9 quadrangles. We assign i.i.d. jump heights $P_i \sim \mathcal{U}(0, 10)$ to each partition element $\mathcal{T}_i$. Figure 3 shows a sample of the advection- and diffusion coefficient for the heterogeneous medium together with the associated (adapted) FE approximation of u. As before, the sample-adapted method is advantageous and the regression suggests that Assumption 4.5 holds with $\kappa = 0.86$. If we use non-adapted FE, we may still recover a convergence rate of 0.66, which is actually slightly better than the expected rate of 0.5.

We now consider an example with lower expected regularity and pure jump field, *i.e.* $\overline{a}$ and Φ are set to zero. Therefore, we need to consider strictly positive jump heights P_i to ensure well-posedness of the problem. We sample one $\mathcal{U}([0.4, 0.6]^2)$-distributed center point $x_c \in \mathcal{D}$ and split the domain by a vertical and horizontal line through x_c. This yields a partition of $\mathcal{D}$ into four squares $\mathcal{T}_1 - \mathcal{T}_4$. We then sample a random variable $P_1 \sim \mathcal{U}([10^{-4}, 10^{-2}])$ and assign the value of P_1 to the lower left and the upper right partition element. The

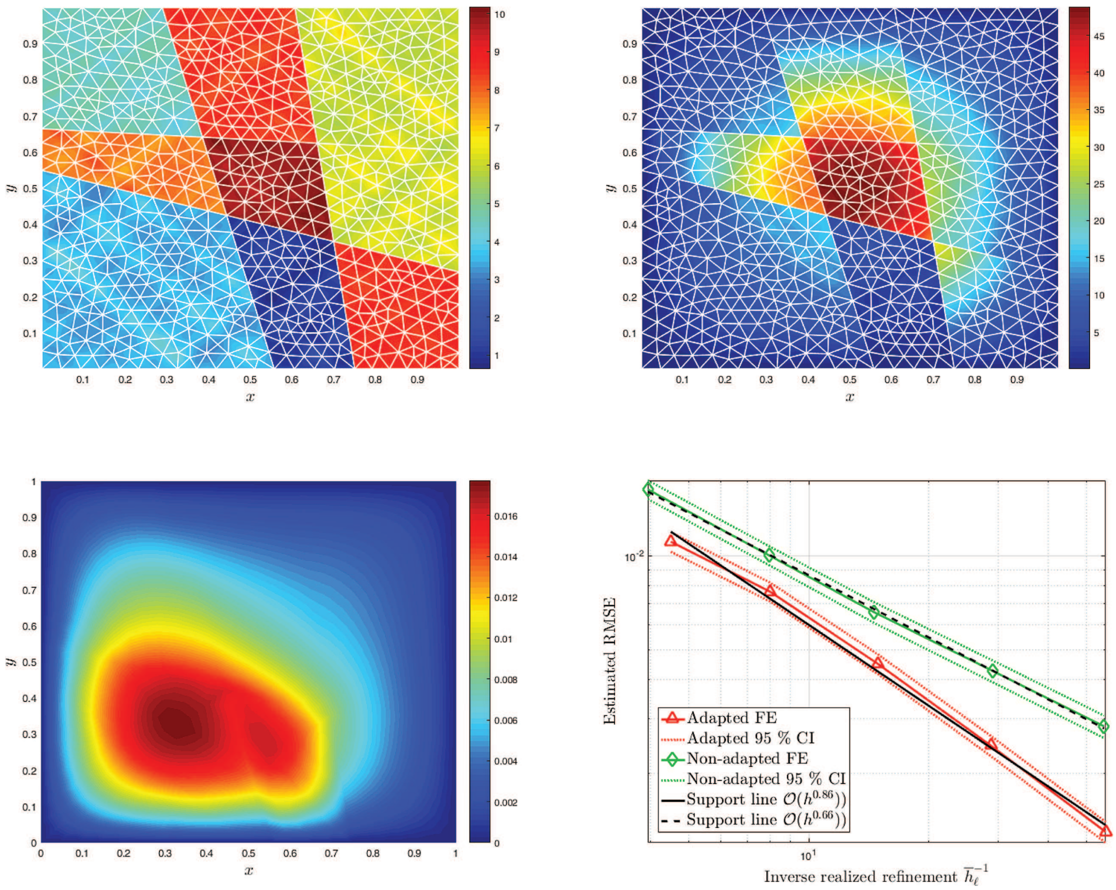

FIGURE 3. First numerical example in 2D (heterogeneous medium). *Top left*: sample of the jump-diffusion coefficient and sample-adapted triangulation, *top right*: sample of the jump-advection coefficient with adapted triangulation, *bottom left*: FE solution at T corresponding to the samples and triangulations on the top, *bottom right*: estimated RMSE *vs.* inverse spatial refinement $\overline{h}_l^{-1}$.

remaining elements are equipped with inverse value $P_2 = P_1^{-1}$, see Figure 4 for a sample of the coefficients. From deterministic regularity theory, it is known that for given P_1 the solution to this problem has only $H^{1+\kappa}$-regularity around x_c, where $\kappa = \mathcal{O}(P_1)$, see *e.g.* [40]. Consequently, we see deteriorated convergence rates compared to the first example. The non-adapted method now performs poorly with an error decay of a rate less than 0.5, whereas the sample-adapted method still recovers a rate of 0.7. A possible explanation for this behavior is that the sample-adapted algorithm generates a mesh with respect to the singularity at x_c. Optimal meshes for this problem refine in the vicinity of x_c and then coarsen on the interior of the partition elements, for instance *graded meshes* or *bisection meshes* as used in [34] and the references therein.

To conclude, we suggest that a more effective refinement in two spatial dimensions may be achieved by *h-finite element methods* (see [42]), *i.e.* by refining the sample-adapted mesh in the reentrant corners. A thorough analysis of this approach for general random geometries is subject to further research.

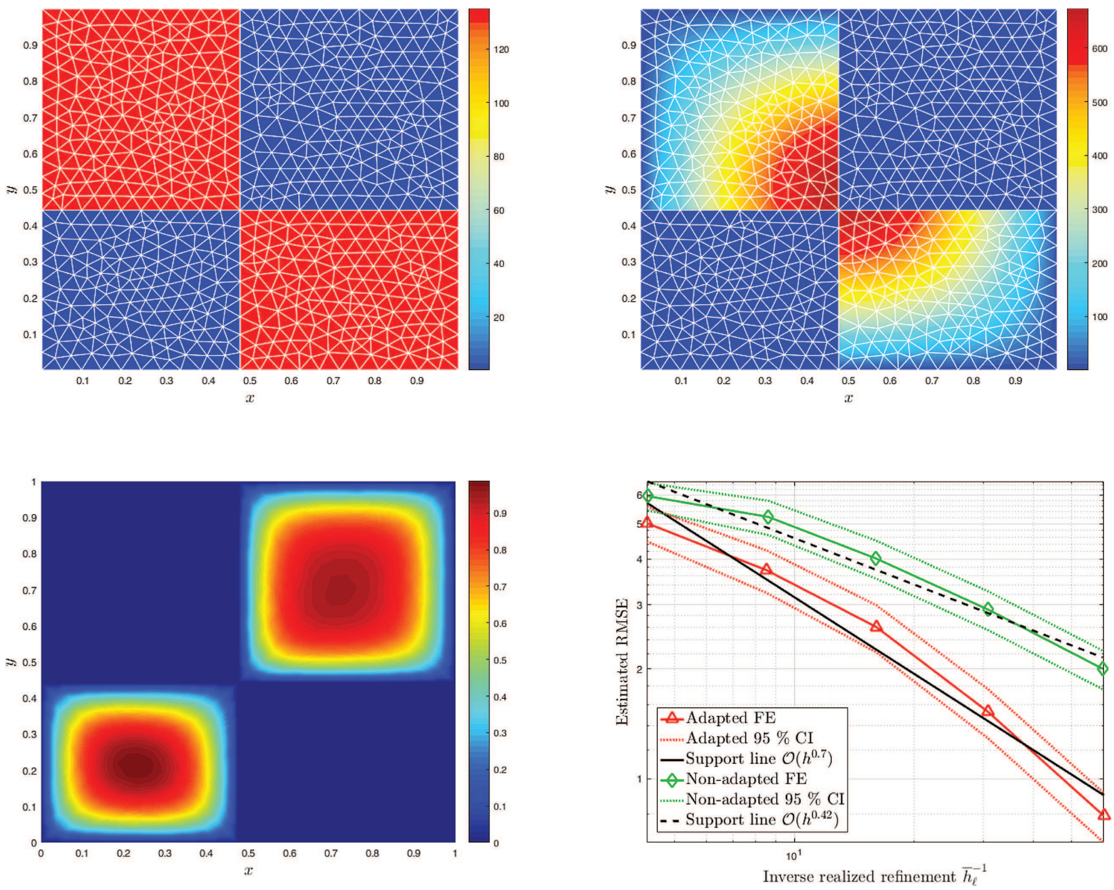

FIGURE 4. Second numerical example in 2D. *Top left*: sample of the jump-diffusion coefficient and sample-adapted triangulation, *top right*: sample of the jump-advection coefficient with adapted triangulation, *bottom left*: FE solution at T corresponding to the samples and triangulations on the top, *bottom right*: estimated RMSE *vs.* inverse spatial refinement parameter $\overline{h}_l^{-1}$.

Acknowledgements. The authors would like to thank Prof. Dr. Christoph Schwab for his valuable suggestions that lead to a significant improvement of this manuscript. The research leading to these results has received funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2075 – 390740016 at the University of Stuttgart and it is gratefully acknowledged.

References

[1] A. Abdulle, A. Barth and C. Schwab, Multilevel Monte Carlo methods for stochastic elliptic multiscale PDEs. *Multiscale Model. Simul.* **11** (2013) 1033–1070.

[2] R.J. Adler and J.E. Taylor, Random Fields and Geometry. Springer Science & Business Media (2009).

[3] C. Aliprantis and K. Border, Infinite Dimensional Analysis: A Hitchhiker's Guide. Springer (2006).

[4] C. Audouze and P.B. Nair, Some a priori error estimates for finite element approximations of elliptic and parabolic linear stochastic partial differential equations. *Int. J. Uncertain. Quantif.* **4** (2014).

[5] I. Babuška, R. Tempone and G.E. Zouraris, Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.* **42** (2004) 800–825 (electronic).

[6] I. Babuška, F. Nobile and R. Tempone, A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.* **45** (2007) 1005–1034 (electronic).

[7] D.A. Barajas-Solano and D.M. Tartakovsky, Stochastic collocation methods for nonlinear parabolic equations with random coefficients. *SIAM/ASA J. Uncertain. Quantif.* **4** (2016) 475–494.

[8] O.E. Barndorff-Nielsen, Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Stat.* **5** (1978) 151–157.

[9] A. Barth and A. Lang, Simulation of stochastic partial differential equations using finite element methods. *Stochastics* **84** (2012, April–June) 217–231.

[10] A. Barth and A. Stein, Approximation and simulation of infinite-dimensional Lévy processes. *Stoch. Part. Differ. Equ.: Anal. Comput.* **6** (2018) 286–334.

[11] A. Barth and A. Stein, A study of elliptic partial differential equations with jump diffusion coefficients. *SIAM/ASA J. Uncertain. Quantif.* **6** (2018) 1707–1743.

[12] A. Barth, C. Schwab and N. Zollinger, Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numer. Math.* **119** (2011) 123–161.

[13] S. Brenner and R. Scott, The Mathematical Theory of Finite Element Methods. Vol. 15. Springer Science & Business Media (2007).

[14] J. Charrier, Strong and weak error estimates for elliptic partial differential equations with random coefficients. *SIAM J. Numer. Anal.* **50** (2012) 216–246.

[15] J. Charrier, R. Scheichl and A.L. Teckentrup, Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods. *SIAM J. Numer. Anal.* **51** (2013) 322–352.

[16] K.A. Cliffe, M.B. Giles, R. Scheichl and A.L. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Comput. Visual. Sci.*

[17] A. Cohen, R. DeVore and C. Schwab, Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDEs. *Anal. App.* **9** (2011) 11–47.

[18] G. Detommaso, T. Dodwell and R. Scheichl, Continuous level Monte Carlo and sample-adaptive model hierarchies. *SIAM/ASA J. Uncertain. Quantif.* **7** (2019) 93–116.

[19] E. Di Nezza, G. Palatucci and E. Valdinoci, Hitchhiker's guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136** (2012) 521–573.

[20] Z. Ding, A proof of the trace theorem of Sobolev spaces on Lipschitz domains. *Proc. Am. Math. Soc.* **124** (1996) 591–600.

[21] M. Eigel, C. Merdon and J. Neumann, An adaptive multilevel Monte Carlo method with stochastic bounds for quantities of interest with uncertain data. *SIAM/ASA J. Uncertain. Quantif.* **4** (2016) 1219–1245.

[22] L. Evans, Partial Differential Equations. American Mathematical Society (2010).

[23] P. Frauenfelder, C. Schwab and R.A. Todor, Finite elements for elliptic problems with stochastic coefficients. *Comput. Methods Appl. Mech. Eng.* **194** (2005) 205–228.

[24] I.G. Graham, F.Y. Kuo, J.A. Nichols, R. Scheichl, C. Schwab and I.H. Sloan, Quasi-Monte Carlo finite element methods for elliptic PDEs with lognormal random coefficients. *Numer. Math.* **131** (2015, October) 329–368.

[25] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl and I.H. Sloan, Circulant embedding with QMC: Analysis for elliptic PDE with lognormal coefficients. *Numer. Math.* **140** (2018) 479–511.

[26] D.S. Grebenkov and B.-T. Nguyen, Geometrical structure of Laplacian eigenfunctions. *SIAM Rev.* **55** (2013) 601–667.

[27] W. Hackbusch, Elliptic Differential Equations: Theory and Numerical Treatment, 2nd softcover print edition. Springer, Berlin, Heidelberg (2017).

[28] H. Harbrecht and J. Li, First order second moment analysis for stochastic interface problems based on low-rank approximation. *ESAIM: M2AN* **47** (2013) 1533–1552.

[29] V.H. Hoang and C. Schwab, Sparse tensor Galerkin discretization of parametric and random parabolic PDEs – Analytic regularity and generalized polynomial chaos approximation. *SIAM J. Math. Anal.* **45** (2013) 3050–3083.

[30] R. Kornhuber and E. Youett, Adaptive multilevel Monte Carlo methods for stochastic variational inequalities. *SIAM J. Numer. Anal.* **56** (2018) 1987–2007.

[31] A. Lang and J. Potthoff, Fast simulation of Gaussian random fields. *Monte Carlo Methods App.* **17** (2011) 195–214.

[32] C. Lang, A. Sharma, A. Doostan and K. Maute, Heaviside enriched extended stochastic FEM for problems with uncertain material interfaces. *Comput. Mech.* **56** (2015) 753–767.

[33] J. Li, X. Wang and K. Zhang, Multi-level Monte Carlo weak Galerkin method for elliptic equations with stochastic jump coefficients. *Appl. Math. Comput.* **275** (2016) 181–194.

[34] F. Müller and C. Schwab, Finite elements with mesh refinement for wave equations in polygons. *J. Comput. Appl. Math.* **283** (2015) 163–181.

[35] S. Nicaise, Polygonal interface problems: Higher regularity results. *Commun. Part. Differ. Equ.* **15** (1990) 1475–1508.

[36] S. Nicaise and A.-M. Sändig, General interface problems – I. *Math. Methods Appl. Sci.* **17** (1994) 395–429.

[37] S. Nicaise and A.-M. Sändig, General interface problems – II. *Math. Methods Appl. Sci.* **17** (1994) 431–450.

[38] F. Nobile and R. Tempone, Analysis and implementation issues for the numerical approximation of parabolic equations with random coefficients. *Int. J. Numer. Methods Eng.* **80** (2009) 979–1006.

[39] F. Nobile, R. Tempone and C.G. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM J. Numer. Anal.* **46** (2008) 2309–2345.

[40] M. Petzoldt, Regularity results for Laplace interface problems in two dimensions. *Z. Anal. Anwend.* **20** (2001) 431–455.

[41] A. Quarteroni and A. Valli, Numerical Approximation of Partial Differential Equations, 2nd edition. Springer Science & Business Media (1997).

[42] C. Schwab, P-and Hp-Finite Element Methods: Theory and Applications in Solid and Fluid Mechanics (Numerical Mathematics and Scientific Computation). Oxford University Press, New York (1999).

[43] C. Schwab and C.J. Gittelson, Sparse tensor discretizations of high-dimensional parametric and stochastic PDEs. *Acta Numer.* **20** (2011) 291–467.

[44] G. Stefanou and M. Papadrakakis, Assessment of spectral representation and Karhunen–Loève expansion methods for the simulation of Gaussian stochastic fields. *Comput. Methods Appl. Mech. Eng.* **196** (2007) 2465–2477.

[45] A.L. Teckentrup, R. Scheichl, M.B. Giles and E. Ullmann, Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients. *Numer. Math.* **125** (2013) 569–600.

[46] G. Zhang and M. Gunzburger, Error analysis of a stochastic collocation method for parabolic partial differential equations with random input data. *SIAM J. Numer. Anal.* **50** (2012) 1922–1940.

[47] T. Zhou, Stochastic Galerkin methods for elliptic interface problems with random input. *J. Comput. Appl. Math.* **236** (2011) 782–792.