NUMERICAL APPROXIMATION OF SINGULAR-DEGENERATE PARABOLIC STOCHASTIC PDES

ĽUBOMÍR BAŇAS, BENJAMIN GESS, AND CHRISTIAN VIETH

Abstract. We study a general class of singular degenerate parabolic stochastic partial differential equations (SPDEs) which include, in particular, the stochastic porous medium equations and the stochastic fast diffusion equation. We propose a fully discrete numerical approximation of the considered SPDEs based on the very weak formulation. By exploiting the monotonicity properties of the proposed formulation we prove the convergence of the numerical approximation towards the unique solution. Furthermore, we construct an implementable finite element scheme for the spatial discretization of the very weak formulation and provide numerical simulations to demonstrate the practicability of the proposed discretization.

1. Introduction

In this paper we study the numerical approximation of a class of singular-degenerate parabolic stochastic partial differential equations

$$\mathrm{d}u = [\Delta(|u|^{p-2}u) + f]\,\mathrm{d}t + \sigma(u)\,\mathrm{d}W \qquad \text{in } (0,T)\times\mathcal{D}\,, \tag{1}$$

where $\mathcal{D}\subset\mathbb{R}^d$, $d\geq 1$ is a bounded, open domain and $\sigma(u)W$ is a multiplicative noise term which will be specified below.

The above equation for $p>2$ is the stochastic *porous medium equation* and for $p\in(1,2)$ the equation corresponds to the stochastic *fast diffusion equation*; the case $p=2$ yields the *stochastic heat equation*.

Stochastic quasilinear diffusion equations of the type (1) appear in several contexts, including, interacting branching diffusion processes [10], self-organized criticality [2, 33], and non-equilibrium fluctuations in non-equilibrium statistical mechanics [27, 20]. We next present three of such instances in more detail.

As a first example, consider the $\mathbb{H}^{-1}$ gradient flow structure of the porous medium equation

$$\partial_t u = -K_u\left(\frac{\delta E}{\delta u}(u)\right) = \Delta(|u|^{p-2}u)$$

with Onsager operator $K_u=-\Delta$ and energy $E(u)=\frac{1}{p}\int|u|^p\,dx$. The corresponding fluctuating system, in accordance with the GENERIC framework of non-equilibrium thermodynamics (see [51]), then reads

$$du = -K_u\left(\frac{\delta E}{\delta u}(u)\right) + B_u\,\mathrm{d}W\,, \tag{2}$$

$$= \Delta(|u|^{p-2}u) + \sqrt{2\kappa_B}\,\operatorname{div}(\,\mathrm{d}W)\,, \tag{3}$$

with $B_uB_u^* = 2\kappa_B K_u$, κ_B the Boltzmann constant and W a vector-valued space-time white noise. Notably, the stochastic PDE (3) is super-critical and, thus, lacks a well-posedness theory. The results of the present paper are applicable to approximate versions of (3), that is, to

$$du = \Delta(|u|^{p-2}u) + \sqrt{2\kappa_B}\,\mathrm{div}(\,\mathrm{d}\tilde{W}), \tag{4}$$

where $\tilde{W}$ is a trace-class Wiener process in $\mathbb{L}^2$; in this case, in one spatial dimension, the stochastic perturbation $\mathrm{div}(\tilde{W})$ still is less-regular than space-time white noise.

The second class of examples arises from fluctuations in non-equilibrium statistical mechanics. This leads to stochastic PDE of the general type

$$du = \Delta\alpha(u)\,\mathrm{d}t + \varepsilon^{\frac{1}{2}}\nabla\cdot(g(u)dW_t), \tag{5}$$

where dW denotes space-time white noise, with the Dean-Kawasaki stochastic PDE

$$du = \Delta u\ \ \mathrm{d}t + \varepsilon^{\frac{1}{2}}\nabla\cdot(\sqrt{u}dW_t),$$

as a model example, see for example [15, 45, 21]. Stochastic PDE of this type serve as continuum models for interacting particle systems, including stochastic corrections reproducing the correct fluctuation behavior on the central limit and large deviations scale, see [19]. Since for large particle number the fluctuations decay, we see the small factor $\varepsilon^{\frac{1}{2}}$ in front of the noise. For example, a concrete example of an interacting particle process is given by the zero range process, see [27, 28], leading to nonlinear, non-degenerate diffusion α in (5) and noise coefficients corresponding to $g(u) = \alpha^{\frac{1}{2}}(u)$. We note that with this choice (5) is in line with the GENERIC framework (2) when considering

$$\partial_t u = \Delta\alpha(u) \tag{6}$$

as a gradient flow on the space of measures with energy given by the Boltzmann entropy. The corresponding stochastic PDE (5) is super-critical and, therefore, lacks a well-posedness theory. Instead, one considers joint scaling limits $\varepsilon \to 0, N \to \infty$ of

$$du = \Delta\alpha(u)\,\mathrm{d}t + \varepsilon^{\frac{1}{2}}\nabla\cdot(g(u)dW^N), \tag{7}$$

where W^N is a regularized noise, see [28, 27]. In the case $\alpha' \geq c > 0$ and g Lipschitz continuous, this class of stochastic PDE is included in the results of the present work.

The third class of equations covered by the present work arises in the continuum scaling limit of the empirical mass of interacting branching diffusions with localized interaction, which, informally, converges to the solution of a stochastic PDE

$$du = \Delta u^2\,\mathrm{d}t + (uc(u))^{\frac{1}{2}}\,\mathrm{d}W, \tag{8}$$

where $\mathrm{d}W$ denotes space-time white noise, see [9, 49]. The results of the present work apply to the particular case of $c(u) = u$ and W being a trace class Wiener process in $\mathbb{H}^{\frac{d+2}{2}}$.

It is common to these stochastic PDE that, due to the irregularity of the random perturbation, solutions are expected to be of low regularity. In fact, in many cases solution take values in spaces of distributions only, causing severe difficulties in even giving meaning to the nonlinear terms appearing in the stochastic PDE.

The lack of regularity of solutions is one of the decisive differences distinguishing the numerical analysis of stochastic PDE from deterministic PDE. While, if the noise and thus the solutions are regular enough, the numerical analysis can proceed similarly to the deterministic case, this ceases to be true in more rough situations. Indeed, if one considers (1) with regular enough noise, the solutions will take values in spaces of functions (L^p spaces), and, therefore, standard finite element basis can be used, such as piecewise constant or piecewise linear functions. The proof of their convergence still requires adaptation from the deterministic arguments, e.g. replacing compactness arguments by a combination of tightness arguments and Skorohod's representation theorem (cf. e.g. [38]), but the numerical method is close to the deterministic case. In contrast, when the noise is not as regular, one cannot expect to close L^p-based estimates, but one has to work in spaces of distributions. Concretely, this means to move from L^p-based estimates for (1) to $\mathbb{H}^{-1}$-based estimates.

While the modification of finite element methods from L^2-based to $\mathbb{H}^{-1}$-based thus is necessary and natural in the context of stochastic PDE, this causes obstacles in their numerical realization: Precisely, while in an L^2-based approach, the choice of piecewise constant (or piecewise linear) finite elements ϕ_i leads to a sparse mass matrix

$$(\tilde{\mathbf{M}}_h)_{i,j} = (\phi_i, \phi_j)_{\mathbb{L}^2},$$

this is not true in the $\mathbb{H}^{-1}$-based approach which leads to a mass matrix

$$(\mathbf{M}_h)_{i,j} = (\phi_i, \phi_j)_{\mathbb{H}^{-1}} = (\phi_i, (-\Delta)^{-1}\phi_j)_{\mathbb{L}^2}. \tag{9}$$

Note that (9) is not a sparse matrix, since $(-\Delta)^{-1}\phi_j$ has global support. Consequently, the resulting numerical scheme is inefficient.

Interestingly, in one spatial dimension this difficulty was addressed in the contribution [24], where an $\mathbb{H}^{-1}$ -based finite element scheme was suggested in the context of a *deterministic* porous medium equation, motivated by the aim to treat irregular initial data and forcing. In [24] it was noticed, that in one spatial dimension a modified finite element basis $\tilde{\phi}_i$ can be constructed, leading to a sparse mass-matrix (9). In view of (9) this requires to choose a basis so that $(-\Delta)^{-1}\phi_j$ has small support. While, in one spatial dimension, this can relatively easily be enforced by choosing ϕ_i of the form

$$-a_{i-1}1_{[x_{i-1},x_i)} + a_i 1_{[x_i,x_{i+1})} - a_{i+1}1_{[x_{i+1},x_{i+2})},$$

for $d \geq 2$ this construction becomes less obvious. In addition, in higher dimension, the proof of the L^p-density of the resulting finite element spaces proves much more challenging.

In the light of this exposition, the contribution of the present work is two-fold: Firstly, motivated by the intrinsic irregularity of stochastic PDE, we provide an $\mathbb{H}^{-1}$ based analysis of a fully discrete finite element scheme for (1) and prove its convergence. Secondly, we construct a finite element basis in dimension $d \geq 2$, which allows for an efficient implementation of the proposed numerical approximation in the $\mathbb{H}^{-1}$-setting, and analyze its approximation properties in L^p. More precisely, motivated by the deterministic numerical approximation [24] we propose a fully discrete finite element based numerical approximation of (1) based on its very weak formulation. We show that the proposed numerical approximation converges for $p \in (1, \infty)$. Furthermore, we generalize the finite element spatial discretization of the very weak formulation, which was restricted to $d = 1$ in [24], to higher dimensions. Moreover, we present numerical simulations to demonstrate the efficiency and convergence behavior of the proposed numerical scheme.

The paper is organized as follows. In Section 2 we state the notation and assumptions along with the definition and basic properties of very weak solutions of (1). We introduce the fully discrete numerical approximation of (1) in Section 3 and show well-posedness of the proposed discrete approximation along with a priori estimates for the numerical solution. The convergence of the numerical approximation towards the very weak solution of (1) is shown in Section 4. In Section 5 we propose and analyze a non-standard finite element scheme for the spatial discretization of the very weak solution which enables an efficient implementation of the resulting fully discrete numerical approximation. Numerical simulations which demonstrate the practicability of the proposed numerical scheme are presented in Section 6.

Comments on the literature. There exists a rich literature on the numerical approximation of deterministic degenerate parabolic equations, i.e. (1) with $\sigma(u) \equiv 0$, where the earlier results include [48], [42]. For more recent results we refer to [23], [24], [18], [22] and the references therein. As far as we are aware, the only result on the numerical approximation of (1) so far is [38], where the convergence of the proposed numerical approximation towards a martingale solution has been shown in dimension $d = 1$ for regular noise and a limited range of the exponent $p \in (2, 3)$, not including the case of the stochastic fast diffusion equation.

In the deterministic setting, the analysis of the equation (1) is well understood, see, e.g. [57]. In the stochastic setting, the well-posedness of (1) in the variational framework goes back to [44, 52] with many details given in [47]; for a generalization of the variational approach to the case of the stochastic fast diffusion equation we refer to [53]. Generalizations to maximal monotone nonlinearities and Cauchy problems can be found in [3], based on monotonicity techniques. Martingale solutions for diffusion coefficients given as Nemytskii operators have been constructed in [37]. In [43] the well-posedness for (1) with additive noise was shown based on a weak convergence approach. An L^1-based alternative approach to well-posedness has been developed based on entropy solutions in [6, 9, 13] and based on kinetic solutions in [35, 17, 34, 26, 28]. Solutions to (1) with space time white multiplicative noise have been constructed in [11].

Besides well-posedness, also the long-time behavior of solutions has been analyzed, see, for example, [26] for the existence of random dynamical systems, [7, 32] for the existence of random attractors, and [3, 14, 58] for ergodicity. For regularity of solutions we refer to [30, 16, 12, 4] and the references therein. Results on finite speed of propagation and waiting times were derived in [31, 3, 29]. Extensions to parabolic-hyperbolic SPDE may be found in [5, 6], and to doubly nonlinear SPDE in [54] and the references therein.

2. Notation and preliminaries

Let $\mathcal{D} \subset \mathbb{R}^d$ be a bounded open domain with $\mathcal{C}^{1,1}$-smooth boundary $\partial\mathcal{D}$ or a rectangular domain. For $1 \leq p \leq \infty$, we denote the conjugate exponent as $p' = \frac{p}{p-1}$. We use the notation $(\mathbb{L}^p, \|\cdot\|_{\mathbb{L}^p})$ for the standard Lebesgue spaces of p-th order integrable functions on $\mathcal{D}$ and $(\mathbb{W}^{k,p}, \|\cdot\|_{\mathbb{W}^{k,p}})$ for the standard Sobolev spaces on $\mathcal{D}$, where $(\mathbb{W}^{k,p}_0, \|\cdot\|_{\mathbb{W}^{k,p}_0})$ stands for the $\mathbb{W}^{k,p}$ space with zero trace on $\partial\mathcal{D}$; for $p = 2$ we denote the corresponding Sobolev spaces as $(\mathbb{H}^k, \|\cdot\|_{\mathbb{H}^k})$ and $(\mathbb{H}^1_0, \|\cdot\|_{\mathbb{H}^1_0})$. We note that the dual space of $\mathbb{H}^1_0$, denoted by $(\mathbb{H}^{-1}, \|\cdot\|_{\mathbb{H}^{-1}})$, is a Hilbert space with the scalar product $(v, w)_{\mathbb{H}^{-1}} := (v, (-\Delta)^{-1}w)_{\mathbb{L}^2} = (\nabla(-\Delta)^{-1}v, \nabla(-\Delta)^{-1}w)_{\mathbb{L}^2}$ where $(-\Delta)^{-1}$ is the inverse Dirichlet Laplace operator $(-\Delta)^{-1} : \mathbb{H}^{-1} \to \mathbb{H}^1_0$.

Throughout the paper we denote $\mathbb{V} := (\mathbb{L}^p \cap \mathbb{H}^{-1})$, and $\mathbb{H} := \mathbb{H}^{-1}$ and note that $\mathbb{V} \hookrightarrow \mathbb{H} \equiv \mathbb{H}' \hookrightarrow \mathbb{V}'$ constitutes a Gelfand triple for the considered range of the exponent p in $d \geq 1$ (for $p \geq 2$ one may take $\mathbb{V} \equiv \mathbb{L}^p$), cf., [46].

For $v \in \mathbb{H}^{-1}$ we define the inverse Laplace operator $\tilde{v} =: (-\Delta)^{-1}v$ as the unique weak solution of the problem

$$
\begin{aligned}
-\Delta\tilde{v} &= v \quad \text{in } \mathcal{D}, \\
\tilde{v} &= 0 \quad \text{on } \partial\mathcal{D}\,.
\end{aligned}
\tag{10}
$$

We note that the above assumption on $\mathcal{D}$ guarantees that $(-\Delta)^{-1}v \in \mathbb{W}^{2,p} \cap \mathbb{W}^{1,p}_0$ for $v \in \mathbb{V} \hookrightarrow \mathbb{H} = \mathbb{H}^{-1}$ and that $\tilde{v}$ depends continuously on v.

We consider W to be a cylindrical Wiener process on a real separable Hilbert space $\mathbb{K}$, that is, for an orthonormal basis $\{\tilde{e}_i\}_{i\in\mathbb{N}}$ of $\mathbb{K}$, we (formally) have $W(t) = \sum_{i\in\mathbb{N}} \tilde{e}_i\beta_i(t)$ with $\{\beta_i(t)\}_{i\in\mathbb{N}}$ independent Brownian motions on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, \mathbb{P})$. Let $L_2(\mathbb{K}, \mathbb{H})$ denote the space of real Hilbert-Schmidt linear operators from $\mathbb{K}$ to $\mathbb{H}$. We note that $(L_2(\mathbb{K}, \mathbb{H}), \|\cdot\|_{L_2(\mathbb{K},\mathbb{H})}, (\cdot\,,\cdot)_{L_2(\mathbb{K},\mathbb{H})})$ is a real separable Hilbert space with inner product

$$(\sigma_1, \sigma_2)_{L_2(\mathbb{K},\mathbb{H})} = \sum_{i=1}^{\infty} (\sigma_1\tilde{e}_i, \sigma_2\tilde{e}_i)_{\mathbb{H}}\,,$$

and the corresponding norm $\|\sigma\|^2_{L_2(\mathbb{K},\mathbb{H})} = \sum_{i=1}^{\infty} \|\sigma\tilde{e}_i\|^2_{\mathbb{H}}$.

We consider a slight generalization of the equation (1):

$$\mathrm{d}u = [\Delta\alpha(u) + f]\,\mathrm{d}t + \sigma(u)\,\mathrm{d}W \quad \text{in } (0,T)\times\mathcal{D}, \tag{11a}$$

$$\alpha(u) = g \quad \text{on } (0,T)\times\partial\mathcal{D}, \tag{11b}$$

$$u(0) = u_0, \quad \text{in } \mathcal{D}, \tag{11c}$$

where $\alpha : \mathbb{R} \to \mathbb{R}$, and $\sigma : \mathbb{V} \to L_2(\mathbb{K}, \mathbb{H})$; the initial condition $u_0 \in L^2(\Omega, \mathbb{H})$ is assumed to be $\mathcal{F}_0$-measurable.

To simplify the presentation we consider (progressively measurable) $f \in L^\infty(\Omega\times(0,T)\times\mathcal{D})$ and $g \in L^\infty(\Omega\times(0,T)\times\partial\mathcal{D})$, a generalization to less regular data is straightforward, cf. [24]. Furthermore, we assume that the function $\alpha : \mathbb{R} \to \mathbb{R}$ is continuous, monotonically increasing, and satisfies a coercivity and growth condition, i.e.,

$$\alpha(z)z \geq \mu|z|^p - \lambda \quad \text{and} \quad |\alpha(z)| \leq c(|z|+1)^{p-1}\,, \quad \forall z \in \mathbb{R}\,, \tag{12}$$

for some $p > 1$ and c, $\mu > 0$, $\lambda \geq 0$, respectively.

Clearly, $\alpha(z) \equiv |z|^{p-2}z$ yields the stochastic porous medium/fast diffusion equation (1) and satisfies the above assumptions for $p > 1$.

We note that $(-\Delta)^{-1}v \in \mathbb{W}^{2,p}$ for $v \in \mathbb{V}$ by standard elliptic regularity theory, cf. [36, Ch. 9], [39, Ch. 4]. Furthermore, for $v \in \mathbb{V}$ the normal trace of $(-\Delta)^{-1}v$ satisfies $\partial_{\vec{n}}\left((-\Delta)^{-1}v\right) \in W^{1/p',p}(\partial\mathcal{D})$ for domains with $\mathcal{C}^{1,1}$-smooth boundary or rectangular domains, cf. [50, Thm. 5.4-5.5 p. 97-99]. Hence, it follows that $b \in L^{p'}(\Omega\times(0,T);\mathbb{V}')$. In the particular case $g \equiv 0$ the following also generalizes to convex domains with piecewise smooth boundary.

Throughout the paper we assume that the following conditions are satisfied.

Assumption 1. *i) Hemi-continuity of A: the function*

$$\epsilon \mapsto \langle A(w+\epsilon z), v\rangle_{\mathbb{V}'\times\mathbb{V}} : [0,1] \to \mathbb{R}$$

is continuous for all $v, w, z \in \mathbb{V}$.

ii) Monotonicity of A: there exists $\lambda_B \geq 0$, such that for all $v, w \in \mathbb{L}^p$

$$2\,\langle Av - Aw, v - w\rangle_{\mathbb{V}'\times\mathbb{V}} + \lambda_B\|v-w\|^2_{\mathbb{H}} \geq \|\sigma(v) - \sigma(w)\|^2_{L_2(\mathbb{K},\mathbb{H})}\,. \tag{13}$$

iii) Coercivity of A: for $\mu > 0$ and $\lambda, \lambda_A, \kappa_\sigma \geq 0$ it holds

$$\langle Av, v\rangle_{\mathbb{V}'\times\mathbb{V}} + \lambda_A\|v\|^2_{\mathbb{H}} \geq \mu\|v\|^p_{\mathbb{V}} - \lambda|\mathcal{D}| + \frac{1}{2}\|\sigma(v)\|^2_{L_2(\mathbb{K},\mathbb{H})} - \kappa_\sigma. \tag{14}$$

iv) Boundedness of A: there exists a $C > 0$ such that

$$\|Av\|_{\mathbb{V}'} \leq C(\|v\|_{\mathbb{V}} + 1)^{p-1} \quad \forall v \in \mathbb{V}.$$

We next generalize the concept of very weak solutions for the deterministic version of (11) with $\sigma(u) \equiv 0$ from [24] to the stochastic problem. We consider the integral form of (11) as

$$u(t) = u_0 + \int_0^t [\Delta\alpha(u(s)) + f(s)]\,\mathrm{d}s + \int_0^t \sigma(u(s))\,\mathrm{d}W(s).$$

We multiply the above equation by $\tilde{v} = (-\Delta)^{-1}v$, integrate over $\mathcal{D}$, and integrate twice by parts in the second order term to obtain, using the boundary condition,

$$\begin{aligned}(u(t), (-\Delta)^{-1}v)_{\mathbb{L}^2} = (u_0, (-\Delta)^{-1}v)_{\mathbb{L}^2} &- \int_0^t (\alpha(u(s)), v)_{\mathbb{L}^2}\,\mathrm{d}s \\ &- \int_0^t (g(s), \partial_{\vec{n}}(-\Delta)^{-1}v)_{L^2(\partial\mathcal{D})}\,\mathrm{d}s \\ &+ \int_0^t (f(s), (-\Delta)^{-1}v)_{\mathbb{L}^2}\,\mathrm{d}s \\ &+ \int_0^t (\sigma(u(s))\,\mathrm{d}W(s), (-\Delta)^{-1}v)_{\mathbb{L}^2}\,.\end{aligned}$$

The above formal construction motivates the following definition of very weak solutions of the stochastic problem (11).

Definition 2.1. *Let $u_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{H})$. Then a $\mathcal{F}_t$-adapted process $u \in L^p(\Omega, \{\mathcal{F}_t\}_t, \mathbb{P}; L^p((0,T); \mathbb{V})) \cap L^2(\Omega, \{\mathcal{F}_t\}_t, \mathbb{P}; C([0,T]; \mathbb{H}))$ is a **very weak solution** of* (11) *if it satisfies $\mathbb{P}$-a.s. for all $v \in \mathbb{V}$ and all $t \in [0,T]$:*

$$\begin{aligned}(u(t), v)_{\mathbb{H}} = (u_0, v)_{\mathbb{H}} &- \int_0^t \langle Au(s), v\rangle_{\mathbb{V}'\times\mathbb{V}}\,\mathrm{d}s \\ &+ \int_0^t \langle b(s), v\rangle_{\mathbb{V}'\times\mathbb{V}}\,\mathrm{d}s + \int_0^t (\sigma(u(s))\,\mathrm{d}W(s), v)_{\mathbb{H}},\end{aligned} \tag{15}$$

with

$$\begin{aligned}\langle Au(s), v\rangle_{\mathbb{V}'\times\mathbb{V}} &= (\alpha(u(s)), v)_{\mathbb{L}^2}\,, \\ \langle b(s), v\rangle_{\mathbb{V}'\times\mathbb{V}} &= (f(s), (-\Delta)^{-1}v)_{\mathbb{L}^2} - (g(s), \partial_{\vec{n}}(-\Delta)^{-1}v)_{L^2(\partial\mathcal{D})}\,.\end{aligned} \tag{16}$$

Remark 2.2. *Owing to the Assumption 1 we may interpret the very weak formulation of* (11) *from Definition 2.1 as a monotone stochastic evolution equation posed on the Gelfand triple $\mathbb{V} \hookrightarrow \mathbb{H} \equiv \mathbb{H}' \hookrightarrow \mathbb{V}'$, cf.* [46, Théorème 3.1], [24]. *Hence, the existence and uniqueness of the very weak solution in Definition 2.1 follows by the standard theory of monotone stochastic evolution equations* [44], [53].

Below we state examples of SPDE problems covered by the framework of Assumption 1; these include all of the problems mentioned in the introduction, in particular. We let $\{e_k\}_{k\in\mathbb{N}}$ be an orthonormal basis of $\mathbb{L}^2$ consisting of eigenvectors of the Laplacian $-\Delta$ with Dirichlet boundary conditions and corresponding eigenvalues $\{\lambda_k\}_{k\in\mathbb{N}}$. We note that

$$\|e_k\|_{\mathbb{L}^\infty} \lesssim {\lambda_k}^{d/4}, \text{ and } \|\nabla e_k\|_{\mathbb{L}^\infty} \lesssim {\lambda_k}^{(d+2)/4}. \tag{17}$$

Example 2.1 (GENERIC framework for the $\mathbb{H}^{-1}$-gradient flow)**.** *We consider* (4) *with $p > 1$ and $W = \sum_{i=1}^\infty \eta_i e_i \beta_i$ a trace-class Wiener process in $\mathbb{L}^2$, i.e., $\sum_{i=1}^\infty \eta_i^2 < \infty$. Then,* $\mathrm{div}(W)$ *is a trace-class Wiener process in $\mathbb{H}^{-1}$ and we choose $\mathbb{V} = \mathbb{L}^p \cap \mathbb{H}^{-1}$, $\mathbb{H} = \mathbb{H}^{-1}$, $\mathbb{K} = \mathbb{L}^2$,*

$A(u) = -\Delta(|u|^{p-2}u)$ extended to $\mathbb{V} \to \mathbb{V}'$ and $\sigma(u)w \equiv \sigma w := \sum_{i=1}^{\infty} \eta_i (e_i, w)_{\mathbb{L}^2} \mathrm{div}(e_i)$. Then, Assumption 1 can be verified analogously to [47].

Example 2.2 (Fluctuations in non-equilibrium systems)**.** *We consider* (7) *so that $\alpha \in \mathcal{C}^1(\mathbb{R})$ satisfies $c^* < \alpha' < C^*$ for some $c^*, C^* > 0$, g is Lipschitz continuous and $W = (\beta_1, \ldots, \beta_N)$ is a $\mathbb{R}^N$-valued Brownian motion, that is,*

$$du = \Delta\alpha(u)dt + \varepsilon^{\frac{1}{2}} \nabla \cdot (g(u)\, \mathrm{d}W),$$

for $\varepsilon \leq \frac{c^*}{2C(N)}$, where $C(N) = \left(\sum_{i=1}^N \|e_i\|_{\mathbb{L}^\infty}^2\right)$. We choose $\mathbb{V} = \mathbb{L}^2$, $\mathbb{H} = \mathbb{H}^{-1}$, $\mathbb{K} = \mathbb{R}^N$, $A(v) = -\Delta\alpha(v)$ extended to $\mathbb{V} \to \mathbb{V}'$, and

$$\sigma(u)w := \varepsilon^{\frac{1}{2}} \sum_{i=1}^{N} \nabla \cdot \big(g(u)e_i(w, \tilde{e}_i)_{\mathbb{R}^N}\big).$$

We then have

$$\begin{aligned}
&- 2\, \langle Av - Aw, v - w\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\sigma(v) - \sigma(w)\|_{L_2(\mathbb{K},\mathbb{H})}^2 \\
&= -2\, \langle Av - Aw, v - w\rangle_{\mathbb{V}'\times\mathbb{V}} + \sum_{j=1}^{N} \|\sigma(v)\tilde{e}_j - \sigma(w)\tilde{e}_j\|_{\mathbb{H}}^2 \\
&= -(\alpha(v) - \alpha(w), v - w)_{\mathbb{L}^2} + \varepsilon \sum_{j=1}^{N} \|\nabla \cdot (g(v)e_i) - \nabla \cdot (g(w)e_i)\|_{\mathbb{H}^{-1}}^2 \\
&\leq -c^*\|v - w\|_{\mathbb{L}^2}^2 + \varepsilon \left(\sum_{i=1}^{N} \|e_i\|_{\mathbb{L}^\infty}^2\right) \|g(v) - g(w)\|_{\mathbb{L}^2}^2 \\
&\leq -c^*\|v - w\|_{\mathbb{L}^2}^2 + C(N)\varepsilon\|g\|_{Lip}\|v - w\|_{\mathbb{L}^2}^2 \leq -\frac{c^*}{2}\|v - w\|_{\mathbb{L}^2}^2.
\end{aligned}$$

The remaining assumptions can be verified similarly. We note that the scaling relation $\varepsilon \leq \frac{c^*}{2C(N)}$ implicitly depends on the dimension d, since the number of frequency modes $\leq N$ depends on the dimension, cf. [19].

Example 2.3 (Branching interacting particle systems)**.** *We consider* (8) *with $c(u) = u$ and $\tilde{W}$ is a trace-class Wiener process in $\mathbb{H}^1$, that is,*

$$du = \Delta u^{[2]}dt + u\, \mathrm{d}\tilde{W}, \tag{18}$$

with $u^{[2]} := |u|u$ and non-negative initial condition u_0. In order to fit this example in the abstract setup of Assumption 1 we choose $\mathbb{V} = \mathbb{L}^3$, $\mathbb{H} = \mathbb{H}^{-1}$, $\mathbb{K} = \ell^2$. Let W be a cylindrical Wiener process on $\mathbb{K}$, $A(v) = -\Delta u^{[2]}$ extended to $\mathbb{V} \to \mathbb{V}'$, and

$$\sigma(u)w := u \sum_{i=1}^{\infty} e_i\eta_i(w, \tilde{e}_i)_{\ell^2},$$

where $\eta_i > 0$, $i \in \mathbb{N}$ *satisfy* $\left(\sum_{i=1}^{\infty} \eta_i^2 \lambda_i^{\frac{d+2}{2}}\right) < \infty$. *Note that then* $\tilde{W} := \sum_{i=1}^{\infty} \eta_i e_i \beta_i$ *defines a trace class Wiener process in* $\mathbb{H}^{\frac{d+2}{2}}$. *We then have, by* (17),

$$
\begin{aligned}
&-2\langle Av - Aw, v - w\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\sigma(v) - \sigma(w)\|^2_{L_2(\mathbb{K},\mathbb{H})} \\
&= -(v^{[2]} - w^{[2]}, v - w)_{L^2} + \sum_{i=1}^{\infty} \|\sigma(v)\tilde{e}_i - \sigma(w)\tilde{e}_i\|^2_{\mathbb{H}^{-1}} \\
&\le \sum_{i=1}^{\infty} \|(v-w)(e_i\eta_i)\|^2_{\mathbb{H}^{-1}} \le \left(\sum_{i=1}^{\infty} \eta_i^2 \|e_i\|^2_{\mathbb{W}^{1,\infty}}\right) \|v - w\|^2_{\mathbb{H}^{-1}} \\
&\le \left(\sum_{i=1}^{\infty} \eta_i^2 \lambda_i^{\frac{d+2}{2}}\right) \|v - w\|^2_{\mathbb{H}^{-1}} \le C\|v - w\|^2_{\mathbb{H}^{-1}}.
\end{aligned}
$$

The remaining assumptions can be verified similarly.

3. Fully discrete Numerical Approximation

We introduce a uniform partition of the time interval $[0, T]$ with a constant time-step size $\tau = T/N$, where $N \in \mathbb{N}$, as $0 = t_0 < t_1 < \ldots < t_N = T$ with $t_n := n\tau$. For a mesh size $h \in (0, 1]$ we consider a family of finite dimensional subspaces $(\mathbb{V}_h)_{h>0} \subset \mathbb{V}$ with the approximation property

$$
\inf_{v_h \in \mathbb{V}_h} \|v - v_h\|_{\mathbb{V}} \to 0 \quad \text{for } h \to 0, \quad \forall v \in \mathbb{V}, \tag{19}
$$

and let $\tilde{J} \equiv \tilde{J}_h = \dim(\mathbb{V}_h)$ for any $h > 0$. We define a family of mappings $R_h : \mathbb{V} \to \mathbb{V}_h$ via the best approximation property, i.e., $R_h v = \underset{v_h \in \mathbb{V}_h}{\arg\inf} \|v - v_h\|_{\mathbb{V}}$ for $v \in \mathbb{V}$. Furthermore, we denote by $P_h : \mathbb{H} \to \mathbb{V}_h$ the family of projection operators which satisfy

$$
\lim_{h\to 0} \|w - P_h w\|_{\mathbb{H}} = 0 \quad \forall w \in \mathbb{H}.
$$

An explicit construction of the discrete finite element spaces $\mathbb{V}_h$ and the operators R_h and P_h will be provided in Section 5 below (see Lemma 5.3, Corollary 5.4 and Remark 5.5).

We define the discrete Brownian increments for $i = 1, 2, \ldots$ as

$$
\Delta_n \beta_i := \begin{cases} 0 & \text{if } n = 1, \\ \beta_i(t_n) - \beta_i(t_{n-1}) & \text{if } n = 2, \ldots, N, \end{cases} \tag{20}
$$

and for $r \in \mathbb{N}$ we define the truncated Hilbert-Schmidt operator $\sigma^r : \mathbb{V} \to L_2(\mathbb{K}, \mathbb{H})$ as

$$
\sigma^r(u)w = \sum_{i=1}^{r} \sigma(u)\tilde{e}_i (w, \tilde{e}_i)_{\mathbb{K}} \qquad \text{for } w \in \mathbb{K},
$$

where $\{\tilde{e}_i\}_{i\in\mathbb{N}}$ is the orthonormal basis of $\mathbb{K}$ and $u \in \mathbb{V}$.

The time-discrete approximation of the right-hand side b (given in Definition 2.1) is obtained as

$$
b^n := \frac{1}{\tau} \int_{t_{n-1}}^{t_n} b(t)\,\mathrm{d}t \approx b(t_n).
$$

Given $N \in \mathbb{N}$, $\tau = \frac{T}{N}$, $h > 0$ and $r \geq 1$, the fully discrete approximation of (11) is obtained as follows: set $u_h^0 = P_h u_0 \in \mathbb{V}_h$, and for $n = 1, \dots, N$ determine $u_h^n \in \mathbb{V}_h$ as the solution of the problem

$$(21) \qquad \left(u_h^n - u_h^{n-1}, v_h\right)_{\mathbb{H}} + \tau \left\langle A u_h^n, v_h \right\rangle_{\mathbb{V}'\times\mathbb{V}} = \tau \left\langle b^n, v_h \right\rangle_{\mathbb{V}'\times\mathbb{V}} + \left(\sigma^r(u_h^{n-1})\Delta_n W, v_h\right)_{\mathbb{H}} .$$

for all $v_h \in \mathbb{V}_h$. We note that the above scheme can be equivalently rewritten as

$$(22) \qquad (u_h^n, v_h)_{\mathbb{H}} + \tau \sum_{k=1}^{n} \left\langle A u_h^k, v_h \right\rangle_{\mathbb{V}'\times\mathbb{V}} = \left(u_h^0, v_h\right)_{\mathbb{H}} \tau \sum_{k=1}^{n} \left\langle b^k, v_h \right\rangle_{\mathbb{V}'\times\mathbb{V}} + \sum_{k=1}^{n} \left(\sigma^r(u_h^{k-1})\Delta_k W, v_h\right)_{\mathbb{H}} .$$

Remark 3.1. *We note that the choice $\Delta_1 \beta_i \equiv 0$, $i \in \mathbb{N}$ in (20) is not strictly required but is convenient since it slightly simplifies the notation and convergence analysis in Section 4 for $u_0 \in \mathbb{H}$. In particular, this choice enables to restate the numerical scheme (22) in the form (26) with the "shifted" interpolant $\overline{u}_\tau^-$ defined in (25) which satisfies the estimate in Corollary 4.1.*

An alternative is to show the convergence by a density argument. For $u_0 \in \mathbb{H}$ one can consider a sufficiently regular sequence $u_0^k \to u_0$, $k \to \infty$, set $\Delta_1 \beta_i \equiv \beta_i(t_1) - \beta_i(t_0)$ and define $\overline{u}_\tau(t) = u_0^k$ for $t \in [0, \tau)$. Then the stochastic integral $\int_\tau^{\theta_\tau^+(t)}$ in (26) is replaced by $\int_0^{\theta_\tau^+(t)}$ and Corollary 4.1 holds for each $k < \infty$.

The measurability of the fully discrete solution is a consequence of the following lemma, c.f. [24, Lemma 3.2], [41, Lemma 3.8].

Lemma 3.2. *Let (S, Σ) be a measure space. Let $\mathbf{f} : S \times \mathbb{V}_h \to \mathbb{V}_h$ be a function that is continuous in its first argument for every (fixed) $\alpha \in S$ and is Σ-measurable in its second argument for every (fixed) $X \in \mathbb{V}_h$. If for every $\alpha \in S$ the equation $\mathbf{f}(\alpha, X) = 0_{\mathbb{V}_h}$ has a unique solution $X = \mathbf{g}(\alpha)$ then $\mathbf{g} : S \to \mathbb{V}_h$ is Σ-measurable.*

The next lemma guarantees the existence, uniqueness and measurability of the fully discrete numerical approximation (21).

Lemma 3.3. *For any $h > 0$, $u_h^0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{H})$, and $\tau \leq \frac{1}{\lambda_B}$ there exists a unique solution $\{u_h^n\}_{n=1}^N$ of the numerical scheme (21). Furthermore, the $\mathbb{V}_h$-valued random variables u_h^n are $\mathcal{F}_{t_n}$-measurable, $n = 1, \dots, N$.*

Proof. We assume that for $u_h^0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{H})$ there exist $\mathbb{V}_h$-valued random variables $\left\{u_h^j\right\}_{j=1}^{n-1}$ that satisfy (21) and that u_h^j are $\mathcal{F}_{t_j}$-measurable for $j = 1, \dots, n-1$. We show the existence of $\mathbb{V}_h$-valued u_h^n, that satisfies (21) and is $\mathcal{F}_{t_n}$-measurable.

For each $\omega \in \Omega$ the scheme (21) defines a canonical mapping $\mathbf{h}_\omega : \mathbb{V}_h \to \mathbb{V}_h$ for which it holds $\mathbf{h}_\omega(u_h^n(\omega)) \equiv 0$. Consequently for $U \in \mathbb{V}_h$ we write

$$\langle \mathbf{h}_\omega(U), U \rangle_{\mathbb{V}_h} := \frac{1}{\tau}(U - u_h^{n-1}(\omega), U)_{\mathbb{H}} + \langle A(U), U \rangle_{\mathbb{V}'\times\mathbb{V}} - \langle b^n(\omega), U \rangle_{\mathbb{V}'\times\mathbb{V}} - \left(\sigma^r(u_h^{n-1}(\omega))\frac{\Delta_n W(\omega)}{\tau}, U\right)_{\mathbb{H}} .$$

We note that

$$(U - u_h^{n-1}(\omega), U)_{\mathbb{H}} \geq \|U\|_{\mathbb{H}}^2 - C\|u_h^{n-1}(\omega)\|_{\mathbb{H}} \|U\|_{\mathbb{V}} .$$

Hence, using the coercivity Assumption 1 iii) along with the embedding $\mathbb{V} \hookrightarrow \mathbb{H}$ we obtain

$$\langle \mathbf{h}_\omega(U), U\rangle_{\mathbb{V}_h} \geq \|U\|_{\mathbb{V}}\left(\mu\|U\|_{\mathbb{V}}^{p-1} - \frac{C}{\tau}\|u_h^{n-1}(\omega)\|_{\mathbb{H}} - C\left\|\sigma(u_h^{n-1}(\omega))\frac{\Delta_n W(\omega)}{\tau}\right\|_{\mathbb{H}}\right) + \left(\frac{1}{\tau} - \lambda_A\right)\|U\|_{\mathbb{H}}^2 + \frac{1}{2}\|\sigma(U)\|_{L_2(\mathbb{K},\mathbb{H})}^2 - C(\lambda_A, \mathcal{D}, b^n)\,.$$

We choose $R_\omega \geq C(\lambda_B, \mathcal{D}, b^n) > 0$ such that

$$\mu R_\omega^{p-1} - \frac{C}{\tau}\|u_h^{n-1}(\omega)\|_{\mathbb{H}} - C\left\|\sigma(u_h^{n-1}(\omega))\frac{\Delta_n W(\omega)}{\tau}\right\|_{\mathbb{H}} \geq 1\,.$$

Since $(1/\tau - \lambda_B) \geq 0$, we get for $\|U\|_{\mathbb{V}} = R_\omega$ that

$$\langle \mathbf{h}_\omega(U), U\rangle_{\mathbb{V}_h} \geq 0\,.$$

Consequently, for each $\omega \in \Omega$ the existence of $u_h^n(\omega) \in \mathbb{V}_h$ that satisfies (21) follows by the Brouwer's fixed point theorem [55, Ch. II, Lemma 1.4].

To show uniqueness we consider U, $\tilde{U} \in \mathbb{V}_h$, such that $\mathbf{h}_\omega(U) = \mathbf{h}_\omega(\tilde{U}) \equiv 0$ and obtain by the monotonicity Assumption 1 ii) that

$$\begin{aligned}0 = \tau\langle \mathbf{h}_\omega(U) - \mathbf{h}_\omega(\tilde{U}), U - \tilde{U}\rangle_{\mathbb{V}_h} &= \|U - \tilde{U}\|_{\mathbb{H}}^2 + \tau\left\langle A(U) - A(\tilde{U}), U - \tilde{U}\right\rangle_{\mathbb{V}'\times\mathbb{V}} \\ &\geq (1 - \lambda_B\tau)\|U - \tilde{U}\|_{\mathbb{H}}^2 \geq 0,\end{aligned}$$

which yields the uniqueness of the discrete solution for $\tau\lambda_B < 1$.

Finally, the $\mathcal{F}_{t_n}$-measurability of the u_h^n follows by Lemma 3.2

□

Under a slightly stronger assumption on τ we obtain the following stability Lemma.

Lemma 3.4. *For $\tau \leq \frac{1}{2(1+\lambda_B)}$ there exist constants $\mu > 0$, $C \geq 0$ such that for $n = 1, \dots, N$ it holds*

$$\mathbb{E}\left[\|u_h^n\|_{\mathbb{H}}^2 + \mu\tau\sum_{j=1}^n \|u_h^j\|_{\mathbb{V}}^p\right] \leq C\,,$$

and

$$\mathbb{E}\left[\sum_{j=1}^n \tau\|Au_h^j\|_{\mathbb{V}'}^{p'}\right] \leq C\,.$$

Proof. i) We set $v_h = u_h^j \in \mathbb{V}_h$ in (21) with $n \equiv j$, use the identity $2(a-b,a)_{\mathbb{H}} = \|a\|_{\mathbb{H}}^2 - \|b\|_{\mathbb{H}}^2 + \|a-b\|_{\mathbb{H}}^2$ and by summing up the resulting equations for $j = 1, \dots, n$ we get, that

$$\begin{aligned}&\|u_h^n\|_{\mathbb{H}}^2 + \sum_{j=1}^n \|u_h^j - u_h^{j-1}\|_{\mathbb{H}}^2 + 2\tau\sum_{j=1}^n \left\langle Au_h^j, u_h^j\right\rangle_{\mathbb{V}'\times\mathbb{V}} \\ &= \|u_h^0\|_{\mathbb{H}}^2 + 2\tau\sum_{j=1}^n \left\langle b^j, u_h^j\right\rangle_{\mathbb{V}'\times\mathbb{V}} + 2\sum_{j=1}^n \left(\sigma^r(u_h^{j-1})\Delta_j W, u_h^j\right)_{\mathbb{H}}\,.\end{aligned} \tag{23}$$

Using the Cauchy-Schwarz and Young's inequalities we estimate the stochastic term as

$$\left(\sigma^r(u_h^{j-1})\Delta_j W, u_h^j\right)_{\mathbb{H}} \leq \left(\sigma^r(u_h^{j-1})\Delta_j W, u_h^{j-1}\right)_{\mathbb{H}} + \frac{1}{2}\left\|\sigma^r(u_h^{j-1})\Delta_j W\right\|_{\mathbb{H}}^2 + \frac{1}{2}\|u_h^j - u_h^{j-1}\|_{\mathbb{H}}^2 .$$

On noting the independence of $\sigma^r(u_h^{j-1})$ and $\Delta_j W$ we estimate

$$\mathbb{E}\left[\left\|\sigma^r(u_h^{j-1})\Delta_j W\right\|_{\mathbb{H}}^2\right] = \tau\mathbb{E}\left[\|\sigma^r(u_h^{j-1})\|_{L_2(\mathbb{K},\mathbb{H})}^2\right] \leq \tau\mathbb{E}\left[\|\sigma(u_h^{j-1})\|_{L_2(\mathbb{K},\mathbb{H})}^2\right].$$

Next, on recalling (16), using the boundedness of f, g we deduce by the Hölder and Young inequalities that

$$\left\langle b^j, u_h^j\right\rangle_{\mathbb{V}'\times\mathbb{V}} \leq C(p,f,g) + \frac{\mu}{2}\|u_h^j\|_{\mathbb{V}}^p .$$

Hence, on recalling the coercivity Assumption 1 *iii)* and using the above inequalities we obtain after taking the expectation in (23) that

$$\begin{aligned}
&\mathbb{E}\left[\|u_h^n\|_{\mathbb{H}}^2 + \mu\tau\sum_{j=1}^{n}\|u_h^j\|_{\mathbb{V}}^p\right] \\
&\quad\leq C + \mathbb{E}\left[\|u_h^0\|_{\mathbb{H}}^2\right] + \tau(1+\lambda_B)\mathbb{E}\left[\sum_{j=1}^{n}\|u_h^j\|_{\mathbb{H}}^2\right].
\end{aligned}$$

The first statement of the Lemma then follows after an application of the discrete Gronwall lemma for $\tau(1+\lambda_B) \leq \frac{1}{2}$.

ii) For the second estimate we use the boundedness Assumption 1 *iv)*, $p' = \frac{p}{p-1}$ and obtain that

$$\|Au_j^n\|_{\mathbb{V}'}^{\frac{p}{p-1}} \leq C_p(\|v\|_{\mathbb{V}}^p + 1).$$

Hence the second estimate follows by part *i)* of the proof. □

Remark 3.5. *The assumption on the step-size τ in the above Lemma (which is required for the application of the discrete Gronwall lemma) is not too restrictive. For instance, for the stochastic porous media equation* (1) *with $\sigma(u) = u$ one may deduce for the constants in Assumption 1,* (13) *that $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 1$, $\lambda_B = 2$. Consequently, we only require a mild condition $\tau \leq \frac{1}{2(1+2)} = \frac{1}{6}$.*

4. Convergence of the numerical approximation

Given the temporal partition $\{t_n\}_{n=0}^N$ with associated discrete random variables $\{u_h^n\}_{n=0}^N$ we define the piecewise constant time-interpolants for $t \in [0,T]$ as follows:

$$\overline{u}_\tau(0) = u_h^1, \quad \overline{u}_\tau(t) = u_h^n \qquad \text{for } t \in (t_{n-1}, t_n] \tag{24}$$

and

$$\begin{aligned}
&\overline{u}_\tau^-(t) = 0 \quad \text{ for } t \in [0,t_1) = [0,\tau), \quad \overline{u}_\tau^-(t) = u_h^{n-1} \quad \text{ for } t \in [t_{n-1}, t_n), \\
&\overline{u}_\tau^-(T) = u_h^N .
\end{aligned} \tag{25}$$

We note that the interpolant $\overline{u}_\tau^-$ is $(\mathcal{F}_t)_{t\in[0,T]}$ adapted by Lemma 3.3.

On recalling (22) we note that the numerical scheme can be restated in terms of the above interpolants, i.e., it holds $\mathbb{P}$-a.s. that

$$
\begin{aligned}
&(\overline{u}_\tau(t), v_h)_{\mathbb{H}} + \int_0^{\theta_\tau^+(t)} \langle A\overline{u}_\tau(s) - b_\tau(s), v_h\rangle_{\mathbb{V}'\times\mathbb{V}}\, \mathrm{d}s \\
&\quad = \left(u_h^0, v_h\right)_{\mathbb{H}} + \int_\tau^{\theta_\tau^+(t)} \left(\sigma^r(\overline{u}_\tau^-(s))\, \mathrm{d}W(s), v_h\right)_{\mathbb{H}} \qquad \text{for all } t \in (0,T),\ \forall v \in \mathbb{V}_h\,,
\end{aligned} \tag{26}
$$

where

$$
\theta_\tau^+(0) := 0, \quad \theta_\tau^+(t) := t_n \qquad \text{for } t \in (t_{n-1}, t_n], \quad n = 1, \dots, N\,. \tag{27}
$$

As a consequence of Lemma 3.4 and Assumption 1 the time interpolants from (24) and (25) satisfy the following a priori estimates.

Corollary 4.1. *For any $h > 0$ and (sufficiently small) $\tau > 0$ it holds that*

$$
\begin{aligned}
&i)\ \sup_{t\in[0,T]} \mathbb{E}\left[\|\overline{u}_\tau^-(t)\|_{\mathbb{H}}^2\right] \le C, &&ii)\ \sup_{t\in[0,T]} \mathbb{E}\left[\|\overline{u}_\tau(t)\|_{\mathbb{H}}^2\right] \le C,\\
&iii)\ \mathbb{E}\left[\int_0^T \|\overline{u}_\tau^-(t)\|_{\mathbb{V}}^p\, \mathrm{d}t\right] \le C, &&iv)\ \mathbb{E}\left[\int_0^T \|\overline{u}_\tau(t)\|_{\mathbb{V}}^p\, \mathrm{d}t\right] \le C,\\
&v)\ \mathbb{E}\left[\int_0^T \|A\overline{u}_\tau^-(t)\|_{\mathbb{V}'}^{p'}\, \mathrm{d}t\right] \le C, &&vi)\ \mathbb{E}\left[\int_0^T \|A\overline{u}_\tau(t)\|_{\mathbb{V}'}^{p'}\, \mathrm{d}t\right] \le C,
\end{aligned}
$$

and

$$
\begin{aligned}
vii)\ \mathbb{E}\left[\int_0^T \|\sigma(\overline{u}_\tau^-(t))\|_{L_2(\mathbb{K},\mathbb{H})}^2\, \mathrm{d}t\right] \le C,\\
viii)\ \mathbb{E}\left[\int_0^T \|\sigma(\overline{u}_\tau(t))\|_{L_2(\mathbb{K},\mathbb{H})}^2\, \mathrm{d}t\right] \le C,
\end{aligned}
$$

where $C > 0$ is a constant that only depends on the data of the problem.

From the a priori bounds in Corollary 4.1 we can directly deduce the following sub-convergence result.

Lemma 4.2. *Let the Assumptions 1 hold and let $u_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{H})$. Then there exists a subsequence h, τ, r (not relabeled) such that for $h, \tau \to 0$, $r \to \infty$ the following holds:*

i) there is a progressively measurable $u \in L^p(\Omega \times (0,T); \mathbb{V})$ such that

$$
\overline{u}_\tau^- \rightharpoonup u \text{ and } \overline{u}_\tau \rightharpoonup u \quad \text{in } L^p(\Omega \times (0,T); \mathbb{V}).
$$

There is a $u_T \in L^2(\Omega; \mathbb{H})$ such that

$$
\overline{u}_\tau^-(T) = \overline{u}_\tau(T) \rightharpoonup u_T \quad \text{in } L^2(\Omega, \mathbb{H})\,.
$$

ii) There exists a progressively measurable $a \in L^{p'}(\Omega \times (0,T); \mathbb{V}')$ such that $A\overline{u}_\tau \rightharpoonup a$ in $L^{p'}(\Omega \times (0,T); \mathbb{V}')$. There is a progressively measurable $\overline{\sigma} \in L^2(\Omega \times (0,T); L_2(\mathbb{K}, \mathbb{H}))$ such that $\sigma^r(\overline{u}_\tau^-)$, $\sigma^r(\overline{u}_\tau)$ and $\sigma(\overline{u}_\tau)$ weakly converge to $\overline{\sigma}$ in $L^2(\Omega \times (0,T); L_2(\mathbb{K}, \mathbb{H}))$.

iii) for $(\mathrm{d}\mathbb{P} \times \mathrm{d}t)$-almost all $(\omega, t) \in \Omega \times (0,T)$ the following equation holds in $\mathbb{V}'$

$$
u(t) = u_0 + \int_0^t b(s) - a(s)\, \mathrm{d}s + \int_0^t \overline{\sigma}(s)\, \mathrm{d}W(s), \tag{28}
$$

iv) there is an $\mathbb{H}$-valued continuous version of u (sill denoted by u) which satisfies (28) *and*

$$\|u(t)\|_{\mathbb{H}}^2 = \|u_0\|_{\mathbb{H}}^2 + \int_0^t \Big(2\,\langle b(s)-a(s), u(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\overline{\sigma}(s)\|_{L_2(\mathbb{K},\mathbb{H})}^2\Big)\,\mathrm{d}s \\ + 2\int_0^t (u(s), \overline{\sigma}(s)\,\mathrm{d}W(s))_{\mathbb{H}}. \tag{29}$$

v) $u_T = u(T)$, i.e. $\overline{u}_\tau(T) \rightharpoonup u(T)$ in $L^2(\Omega;\mathbb{H})$.

Proof. i) We deduce from Corollary 4.1 iii), iv) that $\overline{u}_\tau^- \rightharpoonup u^-$ and $\overline{u}_\tau \rightharpoonup u$ in $L^p(\Omega\times(0,T);\mathbb{V})$. The limit are the same according to [25, Lemma 4.2] see also [40, proof of Prop. 3.3].

Item *ii*) of the Lemma follows from Corollary 4.1 *vii*) and *viii*), the limits again coincide in $L^p(\Omega\times(0,T);\mathbb{V})$ by the arguments from *i*).

To show part *iii*) we consider $v = \psi\phi \in L^\infty(\Omega\times(0,T);\mathbb{V})$ for $\psi \in L^\infty(\Omega\times(0,T);\mathbb{R})$, $\phi\in\mathbb{V}$. We set $v_h = \psi\phi_h \in \mathbb{V}_h$ with $\phi_h = R_h\phi \in \mathbb{V}_h$ in (26), integrate w.r.t. t over $[0,T]$ and take the expectation to get

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T (\overline{u}_\tau(t), v(t))_{\mathbb{H}} + \left\langle \int_0^t A\overline{u}_\tau(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}} \mathrm{d}t\right] \\
&= \mathbb{E}\left[\int_0^T (u_h^0, v(t))_{\mathbb{H}} + \left\langle \int_0^t b_\tau(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}} \right. \\
&\quad \left. + \left(\int_0^t \sigma^r(\overline{u}_\tau^-(s))\,\mathrm{d}W(s), v(t)\right)_{\mathbb{H}} \mathrm{d}t\right] \\
&\quad + \mathcal{R}_{1,\tau,h} + \mathcal{R}_{2,\tau,h} - \mathcal{R}_{3,\tau,h} - \mathcal{R}_{4,\tau,h} - \mathcal{R}_{5,\tau,h} + \mathcal{R}_{6,\tau,h} + \mathcal{R}_{7,\tau,h} + \mathcal{R}_{8,\tau,h}\,,
\end{aligned} \tag{30}$$

where

$$\begin{aligned}
\mathcal{R}_{1,\tau,h} &:= \mathbb{E}\left[\int_0^T \left\langle \int_t^{\theta_\tau^+(t)} b_\tau(s) - A\overline{u}_\tau(s)\,\mathrm{d}s, v_h(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}} \mathrm{d}t\right], \\
\mathcal{R}_{2,\tau,h} &:= \mathbb{E}\left[\int_0^T \left(\int_0^\tau \sigma^r(\overline{u}_\tau^-(s))\,\mathrm{d}W(s), v_h(t)\right)_{\mathbb{H}} \mathrm{d}t\right], \\
\mathcal{R}_{3,\tau,h} &:= \mathbb{E}\left[\int_0^T \left(\int_t^{\theta_\tau^+(t)} \sigma^r(\overline{u}_\tau^-(s))\,\mathrm{d}W(s), v_h(t)\right)_{\mathbb{H}} \mathrm{d}t\right], \\
\mathcal{R}_{4,\tau,h} &:= (\overline{u}_\tau, v_h - v)_{L^2(\Omega\times(0,T);\mathbb{H})}, \\
\mathcal{R}_{5,\tau,h} &:= \left\langle \int_0^\cdot A\overline{u}_\tau(s)\,\mathrm{d}s, v_h - v\right\rangle_{L^{p'}(\Omega\times(0,T);\mathbb{V}')\times L^p(\Omega\times(0,T);\mathbb{V})}, \\
\mathcal{R}_{6,\tau,h} &:= (u_h^0, v_h - v)_{L^2(\Omega\times(0,T);\mathbb{H})}, \\
\mathcal{R}_{7,\tau,h} &:= \left\langle \int_0^\cdot b_\tau(s)\,\mathrm{d}s, v_h - v\right\rangle_{L^{p'}(\Omega\times(0,T);\mathbb{V}')\times L^p(\Omega\times(0,T);\mathbb{V})}, \\
\mathcal{R}_{8,\tau,h} &:= \left(\int_0^\cdot \sigma^r(\overline{u}_\tau^-(s))\,\mathrm{d}W(s), v_h - v\right)_{L^2(\Omega\times(0,T);\mathbb{H})}.
\end{aligned}$$

By the boundedness of b_τ and $A\overline{u}_\tau$ in $L^{p'}(\Omega\times(0,T);\mathbb{V}')$ and $\sigma(\overline{u}_\tau^-)$ in $L^2(\Omega\times(0,T);L_2(\mathbb{K},\mathbb{H},))$ and an application of Itô's isometry we get $\mathcal{R}_{1,\tau,h}$, $\mathcal{R}_{2,\tau,h}$, $\mathcal{R}_{3,\tau,h} \to 0$ for $\tau, h\to 0$.

Further, the boundedness of $\overline{u}_\tau$ in $L^2(\Omega\times(0,T);\mathbb{H})$ and u_h^0 in $L^2(\Omega;\mathbb{H})$ yields for $k=4,\dots,8$ that

$$|\mathcal{R}_{k,\tau,h}| \le C\|v-v_h\|_{L^p(\Omega\times(0,T);\mathbb{V})}\,.$$

On recalling $v=\psi\phi$ and $v_h=\psi\phi_h\in\mathbb{V}_h$, $\phi_h=R_h\phi\in\mathbb{V}_h$ we deduce by (19) that

$$\|v_h-v\|_{L^p(\Omega\times(0,T);\mathbb{V})} = \|\psi\|_{L^p(\Omega\times(0,T),\mathbb{R})}\|\phi-\phi_h\|_{\mathbb{V}}\to 0 \qquad \text{for } h\to 0.$$

Hence, on noting Corollary 4.1 we conclude that $\mathcal{R}_{k,\tau,h}\to 0$, $k=4,\dots,8$ for $h\to 0$.

Next, the weak convergence $A\overline{u}_\tau \rightharpoonup a$, $\sigma(\overline{u}_\tau)\rightharpoonup\overline{\sigma}$ implies for $h,\tau\to 0$, $r\to\infty$

$$\mathbb{E}\left[\int_0^T\left\langle\int_0^t A\overline{u}_\tau(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}}\mathrm{d}t\right]\to\mathbb{E}\left[\int_0^T\left\langle\int_0^t a(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}}\mathrm{d}t\right],$$
$$\mathbb{E}\left[\int_0^T\left(\int_0^t \sigma^r(\overline{u}_\tau^-(s))\,\mathrm{d}W(s), v(t)\right)_{\mathbb{H}}\mathrm{d}t\right]\to\mathbb{E}\left[\int_0^T\left(\int_0^t \overline{\sigma}(s)\,\mathrm{d}W(s), v(t)\right)_{\mathbb{H}}\mathrm{d}t\right].$$

From the weak convergence of $\overline{u}_\tau\rightharpoonup u$ in $L^2(\Omega\times(0,T);\mathbb{H})$ and the strong convergence of $u_h^0\to u_0$ in $L^2(\Omega;\mathbb{H})$ we deduce that

$$\mathbb{E}\left[\int_0^T(\overline{u}_\tau(t),v(t))_{\mathbb{H}}\,\mathrm{d}t\right]\to\mathbb{E}\left[\int_0^T(u(t),v(t))_{\mathbb{H}}\,\mathrm{d}t\right],$$

and

$$\mathbb{E}\left[\int_0^T(u_h^0,v(t))_{\mathbb{H}}\,\mathrm{d}t\right]\to\mathbb{E}\left[\int_0^T(u_0,v(t))_{\mathbb{H}}\,\mathrm{d}t\right].$$

Finally, since $b_\tau\to b$ in $L^{p'}(\Omega\times(0,T);\mathbb{V}')$ it follows that

$$\mathbb{E}\left[\int_0^T\left\langle\int_0^t b_\tau(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}}\mathrm{d}t\right]\to\mathbb{E}\left[\int_0^T\left\langle\int_0^t b(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}}\mathrm{d}t\right].$$

From the above convergence results we conclude, by taking $h,\tau\to 0$, $r\to\infty$ in (30) that

$$\begin{aligned}&\mathbb{E}\left[\int_0^T(u(t),v(t))_{\mathbb{H}}+\left\langle\int_0^t a(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}}\mathrm{d}t\right]\\ &=\mathbb{E}\left[\int_0^T(u_0,v(t))_{\mathbb{H}}+\left\langle\int_0^t b(s)\,\mathrm{d}s, v(t)\right\rangle_{\mathbb{V}'\times\mathbb{V}}+\left(\int_0^t\overline{\sigma}(s)\,\mathrm{d}W(s),v(t)\right)_{\mathbb{H}}\mathrm{d}t\right],\end{aligned}$$

for all $v=\psi\phi$, $\phi\in\mathbb{V}$, which implies (28).

By the standard theory of monotone SPDEs, cf. [44] (or [53]), part *iv*) follows from *iii*) by the Itô formula for the square of the $\mathbb{H}$-norm, which also implies that u has an $\mathbb{H}$-valued continuous modification (which we again denote by u) that satisfies (28).

Finally, to show *v*) we note that $\overline{u}_\tau(T)\rightharpoonup u_T$ by part *i*) which together with *iii*) implies

$$u_T+\int_0^T a(s)\,\mathrm{d}s = u_0+\int_0^T b(s)\,\mathrm{d}s+\int_0^T\overline{\sigma}(s)\,\mathrm{d}W(s)\quad\text{in }\mathbb{L}^{p'}.$$

Since the continuous $\mathbb{H}$-valued modification of u (cf. *iv*)) satisfies (28) we may conclude that $u_T=u(T)$. □

The following variant of the Gronwall lemma, cf. [25, Lemma 5.1], will be useful for the proof of the subsequent theorem.

Lemma 4.3. *Let a and b be real-valued integrable functions such that for all $t \in [0,T]$*

$$a(t) \le a(0) + \int_0^t b(s)\,\mathrm{d}s, \tag{31}$$

then for all $\lambda_B \ge 0$ and for all $t \in [0,T]$

$$e^{-\lambda_B t} a(t) + \lambda_B \int_0^t e^{-\lambda_B s} a(s)\,\mathrm{d}s \le a(0) + \int_0^t e^{-\lambda_B s} b(s)\,\mathrm{d}s. \tag{32}$$

Moreover, if equality holds in (31)*, then equality holds in* (32)*.*

In the next theorem we conclude that the weak limit of the numerical approximation from Lemma 4.2 is the very weak solution of the equation (11).

Theorem 4.4 (Convergence of the numerical approximation)**.** *Let the Assumption 1 hold and let $u_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; \mathbb{H})$. Then, for $h, \tau \to 0$, $r \to \infty$ the fully discrete solution of scheme* (26) *converges to the unique very weak solution $u \in L^p(\Omega \times (0,T); \mathbb{V}) \cap L^2(\Omega; C([0,T]; \mathbb{H}))$ of* (11) *in the sense of Definition 2.1.*

Proof. We have shown in Lemma 4.2 that every weak limit u of the numerical approximation satisfies for $t \in [0,T]$

$$u(t) = u_0 + \int_0^t b(s) - a(s)\,\mathrm{d}s + \int_0^t \overline{\sigma}(s)\,\mathrm{d}W(s)\,.$$

Hence, it remains to show that $a = Au$, $\overline{\sigma} = \sigma(u)$.

Throughout the proof we use the shorthand notation $\ell := (h, \tau, r)$ and $\ell \to \infty$ stands for $h, \tau \to 0$, $r \to \infty$. We define

$$\Xi_\ell(t) := \begin{cases} \|\overline{u}_\tau(t)\|^2_{L^2(\Omega;\mathbb{H})} & \text{if } t \in (0,T], \\ \|u_h^0\|^2_{L^2(\Omega;\mathbb{H})} & \text{if } t = 0\,. \end{cases}$$

Analogously to the proof of Lemma 3.4 we deduce from (23) on noting the definition of the time interpolants (26) that for any $t \in (0,T]$ it holds

$$\begin{aligned} &\Xi_\ell(t) \le \Xi_\ell(0) \\ &\qquad + \mathbb{E}\left[\int_0^t 2\,\langle b_\tau(s) - A\overline{u}_\tau(s), \overline{u}_\tau(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\sigma^r(\overline{u}_\tau(s))\|^2_{L_2(\mathbb{K},\mathbb{H})}\,\mathrm{d}s\right] + \mathcal{R}_\ell(t)\,, \end{aligned}$$

with $\mathcal{R}_\ell(t) = \mathbb{E}\left[\int_t^{\theta_\tau^+(t)} 2\,\langle b_\tau(s) - A\overline{u}_\tau(s), \overline{u}_\tau(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\sigma^r(\overline{u}_\tau(s))\|^2_{L_2(\mathbb{K},\mathbb{H})}\,\mathrm{d}s\right]$.

We use Lemma 4.3 and obtain from the above inequality that

$$\begin{aligned} e^{-\lambda_B T}\Xi_\ell(T) &\le \Xi_\ell(0) - \lambda_B \int_0^T e^{-\lambda_B s}\Xi_\ell(s)\,\mathrm{d}s \\ &\quad + \mathbb{E}\left[\int_0^T e^{-\lambda_B s}\left(2\,\langle b_\tau(s) - A\overline{u}_\tau(s), \overline{u}_\tau(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\sigma^r(\overline{u}_\tau(s))\|^2_{L_2(\mathbb{K},\mathbb{H})}\right)\,\mathrm{d}s\right] \\ &\quad + \lambda_B \int_0^T e^{-\lambda_B s}|\mathcal{R}_\ell(s)|\,\mathrm{d}s. \end{aligned} \tag{33}$$

Note that by the monotonicity property (13) it holds for arbitrary $w \in L^p(\Omega \times (0,T); \mathbb{V})$ that

$$\begin{aligned}
&- 2\mathbb{E}\left[\int_0^T e^{-\lambda_B s} \langle A\overline{u}_\tau(s), \overline{u}_\tau(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \,\mathrm{d}s\right] \\
&\leq \mathbb{E}\bigg[\int_0^T e^{-\lambda_B s}\Big(- \|\sigma(\overline{u}_\tau(s)) - \sigma(w(s))\|^2_{L_2(\mathbb{K},\mathbb{H})} + \lambda_B \|\overline{u}_\tau(s) - w(s)\|^2_{\mathbb{H}}\Big) \,\mathrm{d}s\bigg] \\
&\quad - 2\mathbb{E}\bigg[\int_0^T e^{-\lambda_B s}\Big(\langle Aw(s), \overline{u}_\tau(s) - w(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \langle A\overline{u}_\tau(s), w(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \Big) \,\mathrm{d}s\bigg].
\end{aligned}$$

We substitute the above inequality into (33) and obtain

$$\begin{aligned}
&e^{-\lambda_B T}\|\overline{u}_\tau(T)\|^2_{L^2(\Omega;\mathbb{H})} \\
&\leq \|u_h^0\|^2_{L^2(\Omega;\mathbb{H})} + 2\mathbb{E}\left[\int_0^T e^{-\lambda_B s} \langle b_\tau(s), \overline{u}_\tau(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\bigg[\int_0^T e^{-\lambda_B s}\Big(- \|\sigma(w(s))\|^2_{L_2(\mathbb{K},\mathbb{H})} + 2\left(\sigma(\overline{u}_\tau(s)), \sigma(w(s))\right)_{L_2(\mathbb{K},\mathbb{H})} \\
&\qquad\qquad + \lambda_B \|w(s)\|^2_{\mathbb{H}} - 2\lambda_B \left(\overline{u}_\tau(s), w(s)\right)_{\mathbb{H}} \Big) \,\mathrm{d}s\bigg] \\
&\quad - 2\mathbb{E}\bigg[\int_0^T e^{-\lambda_B s}\Big(\langle Aw(s), \overline{u}_\tau(s) - w(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \langle A\overline{u}_\tau(s), w(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \Big) \,\mathrm{d}s\bigg] \\
&\qquad\qquad + \lambda_B \int_0^T e^{-\lambda_B s} |\mathcal{R}_\ell(s)| \,\mathrm{d}s\,.
\end{aligned} \tag{34}$$

Next, we observe that, by Corollary 4.1,

$$\begin{aligned}
&\lambda_B \int_0^T e^{-\lambda_B} |\mathcal{R}_\ell(t)| \,\mathrm{d}t \\
&\leq \tau\lambda_B \Big(2\left(\|b_\tau\|_{L^{p'}(\Omega\times(0,T);\mathbb{V}')} + \|A\overline{u}_\tau\|_{L^{p'}(\Omega\times(0,T);\mathbb{V}')} \right) \|\overline{u}_\tau\|_{L^p(\Omega\times(0,T);\mathbb{V})} \\
&\qquad\qquad + \|\sigma(\overline{u}_\tau)\|^2_{L^2(\Omega\times(0,T);L_2(\mathbb{K},\mathbb{H}))} \Big) \\
&\leq C\tau \to 0 \quad \text{for } \ell \to \infty\,.
\end{aligned}$$

Hence, using the weak convergence of Lemma 4.2 i), ii) we deduce from (34) by the lower-semicontinuity of norms that

$$
\begin{aligned}
e^{-\lambda_B T}\|u(T)\|^2_{L^2(\Omega;\mathbb{H})} &\leq \liminf_{\ell\to\infty} e^{-\lambda_B T}\|\overline{u}_\tau(T)\|^2_{L^2(\Omega;\mathbb{H})} \\
&\leq \|u_0\|^2_{L^2(\Omega;\mathbb{H})} + 2\mathbb{E}\left[\int_0^T e^{-\lambda_B s}\langle b(s),u(s)\rangle_{\mathbb{V}'\times\mathbb{V}}\,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\Bigg[\int_0^T e^{-\lambda_B s}\Big(-\|\sigma(w(s))\|^2_{L_2(\mathbb{K},\mathbb{H})} + 2(\overline{\sigma}(s),\sigma(w(s)))_{L_2(\mathbb{K},\mathbb{H})} \\
&\qquad + \lambda_B\|w(s)\|^2_{\mathbb{H}} - 2\lambda_B(u(s),w(s))_{\mathbb{H}}\Big)\,\mathrm{d}s\Bigg] \\
&\quad - 2\mathbb{E}\left[\int_0^T e^{-\lambda_B s}\Big(\langle Aw(s),u(s)-w(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \langle a(s),w(s)\rangle_{\mathbb{V}'\times\mathbb{V}}\Big)\,\mathrm{d}s\right].
\end{aligned}
\tag{35}
$$

After a standard stopping argument and taking the expectation in (29) we get for all $t\in[0,T]$

$$
\|u(t)\|^2_{L^2(\Omega;\mathbb{H})} = \|u_0\|^2_{L^2(\Omega;\mathbb{H})} + \mathbb{E}\left[\int_0^t 2\langle b(s)-a(s),u(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\overline{\sigma}(s)\|^2_{L_2(\mathbb{K},\mathbb{H})}\,\mathrm{d}s\right].
$$

Using Lemma 4.3 we obtain from the above equality that

$$
\begin{aligned}
e^{-\lambda_B T}\|u(T)\|^2_{L^2(\Omega;\mathbb{H})} &= \|u_0\|^2_{L^2(\Omega;\mathbb{H})} - \lambda_B\mathbb{E}\left[\int_0^T e^{-\lambda_B s}\|u(s)\|^2_{\mathbb{H}}\,\mathrm{d}s\right] \\
&\quad + \mathbb{E}\left[\int_0^T e^{-\lambda_B s}\left(2\langle b(s)-a(s),u(s)\rangle_{\mathbb{V}'\times\mathbb{V}} + \|\overline{\sigma}(s)\|^2_{L_2(\mathbb{K},\mathbb{H})}\right)\,\mathrm{d}s\right].
\end{aligned}
\tag{36}
$$

Next, we subtract (36) from (35) and get

$$
\begin{aligned}
0 &\leq \mathbb{E}\left[\int_0^T e^{-\lambda_B s}\Big(-\|\sigma(w(s))-\overline{\sigma}(s)\|^2_{L_2(\mathbb{K},\mathbb{H})} + \lambda_B\|w(s)-u(s)\|^2_{\mathbb{H}}\Big)\,\mathrm{d}s\right] \\
&\quad - 2\mathbb{E}\left[\int_0^T e^{-\lambda_B s}\Big(\langle Aw(s),u(s)-w(s)\rangle_{\mathbb{V}'\times\mathbb{V}} - \langle a(s),u(s)-w(s)\rangle_{\mathbb{V}'\times\mathbb{V}}\Big)\,\mathrm{d}s\right].
\end{aligned}
$$

Consequently, it holds that

$$
\begin{aligned}
&2\mathbb{E}\left[\int_0^T e^{-\lambda_B s}\langle Aw(s),u(s)-w(s)\rangle_{\mathbb{V}'\times\mathbb{V}}\,\mathrm{d}s\right] \\
&\leq \mathbb{E}\left[\int_0^T e^{-\lambda_B s}\Big(\lambda_B\|w(s)-u(s)\|^2_{\mathbb{H}} + 2\langle a(s),u(s)-w(s)\rangle_{\mathbb{V}'\times\mathbb{V}}\Big)\,\mathrm{d}s\right].
\end{aligned}
\tag{37}
$$

On taking $w=u$ in (37) we get

$$
0 \leq -\mathbb{E}\left[\int_0^T e^{-\lambda_B s}\|\sigma(u(s))-\overline{\sigma}(s)\|^2_{L_2(\mathbb{K},\mathbb{H})}\,\mathrm{d}s\right] \leq 0,
$$

which implies $\sigma(u(s)) = \overline{\sigma}(s)$ in $L^2(\Omega\times(0,T);L_2(\mathbb{K},\mathbb{H}))$.

Next we choose $w = u - \varepsilon z$ (37) with $z \in L^p(\Omega \times (0,T); \mathbb{V})$, $\varepsilon \in (0,1)$

$$\mathbb{E}\left[\int_0^T e^{-\lambda_B s} \langle A(u(s) - \varepsilon z(s)), z(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \,\mathrm{d}s\right] \leq \mathbb{E}\left[\int_0^T e^{-\lambda_B s}\Big(\frac{1}{2}\varepsilon\lambda_B\|z(s)\|_{\mathbb{H}}^2 + \langle a(s), z(s)\rangle_{\mathbb{V}'\times\mathbb{V}}\Big) \,\mathrm{d}s\right],$$

and obtainusing Assumption 1 i) by the Lebesgue dominated convergence for $\varepsilon \to 0$ that

$$\mathbb{E}\left[\int_0^T e^{-\lambda_B s} \langle Au(s), z(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \,\mathrm{d}s\right] \leq \mathbb{E}\left[\int_0^T e^{-\lambda_B s} \langle a(s), z(s)\rangle_{\mathbb{V}'\times\mathbb{V}} \,\mathrm{d}s\right].$$

This implies that $a = Au$, since $z \in L^p(\Omega \times (0,T); \mathbb{V})$ is arbitrary.

Finally, we conclude by the uniqueness of the very weak solution, that the whole sequence converges to the same limit u.

□

5. Practical finite element approximation in $\mathbb{L}^p$

A natural approach is to construct the numerical solution $u_h^n \in \mathbb{V}_h \subset \mathbb{L}^p$, $n = 0, \ldots, N$ using a finite element space $\mathbb{V}_h$ consisting of piecewise constant functions on a given partition of the domain $\mathcal{D}$ with a given mesh size h. However, the piecewise constant finite element approximation of the very weak formulation is impractical since the resulting finite element matrix associated with the $\mathbb{H}$-scalar product $(\cdot,\cdot)_{\mathbb{H}} = (\cdot, (-\Delta)^{-1}\cdot)_{\mathbb{L}^2} = (\nabla(-\Delta)^{-1}\cdot, \nabla(-\Delta)^{-1}\cdot)_{\mathbb{L}^2}$ in the discrete very weak formulation (21) will be dense. Furthermore, the evaluation of the $\mathbb{H}$-inner product requires the evaluation of the inverse Laplace operator $(-\Delta)^{-1}$, which does not have an explicit formula in general. This is a consequence of the fact that the inverse Laplacian of the characteristic function $\chi_{\mathcal{T}}$ for some subset $\mathcal{T} \subset \mathcal{D}$ does not have compact support in $\mathcal{D}$, i.e., in general $\mathrm{supp}\{(-\Delta)^{-1}\chi_{\mathcal{T}}\} \equiv \mathcal{D}$. A further complication lies in the fact that there is no explicit formula available for $(-\Delta)^{-1}\chi_T$, in general.

Below, we discuss the construction of a finite element basis $\{\phi_i\}_{i=1}^{\tilde{J}}$ of $\mathbb{V}_h$ for $d \geq 1$ on rectangular domains with the property that $\psi_i := (-\Delta)^{-1}\phi_i$ can be computed explicitly and has local support in $\mathcal{D}$ for $i = 1, \ldots, \tilde{J}$.

5.1. **Finite-element basis in $d = 1$.** We summarize the finite element method proposed in [24] for $\mathcal{D} \subset \mathbb{R}^1$. For the domain $\mathcal{D} = (-L, L)$, where $L > 0$ we introduce a partition into disjoint open intervals $\{(\mathbf{x}_{i-1}, \mathbf{x}_i)\}_{i=1}^J$, $\mathbf{x}_0 = -L$, $\mathbf{x}_J = L$ such that $\overline{\mathcal{D}} = \cup_{i=1}^J[\mathbf{x}_{i-1}, \mathbf{x}_i]$ and denote χ_I to be the characteristic function of the interval I. We then set $\mathbb{V}_h = \mathrm{span}\{\phi_i,\ i = 1, \ldots, J\} \subset \mathbb{L}^p$ where $\phi_i : [-L, L] \to \mathbb{R}$ are defined as

$$\phi_1(x) = \frac{3}{2}\chi_{[\mathbf{x}_0,\mathbf{x}_1]}(x) - \frac{1}{2}\chi_{(\mathbf{x}_1,\mathbf{x}_2]}(x)\,, \tag{38}$$

$$\phi_i(x) = -\frac{1}{2}\chi_{(\mathbf{x}_{i-2},\mathbf{x}_{i-1}]}(x) + \chi_{(\mathbf{x}_{i-1},\mathbf{x}_i]}(x) - \frac{1}{2}\chi_{(\mathbf{x}_i,\mathbf{x}_{i+1}]}(x)\,, \tag{39}$$

$$\phi_J(x) = -\frac{1}{2}\chi_{(\mathbf{x}_{J-2},\mathbf{x}_{J-1}]}(x) + \frac{3}{2}\chi_{(\mathbf{x}_{J-1},\mathbf{x}_J]}(x)\,. \tag{40}$$

for any $x \in (-L, L)$.

Note that the proposed approximation is equivalent to a piecewise constant approximation, i.e., $\mathbb{V}_h \equiv \mathrm{span}\{\phi_i\} = \mathrm{span}\{\chi_{(\mathbf{x}_{i-1},\mathbf{x}_i]},\ i=1,\dots,J\}$. The proposed basis has the useful property that $\psi_i := (-\Delta)^{-1}\phi_i$ (with $(-\Delta)^{-1}$ defined on $(-L,L)$) admits an explicit representation for all $i=1,\dots,J$ which has a small support in $\mathcal{D}$. It can be verified by direct calculation that

$$\psi_1(x) = \begin{cases} -\frac{3}{4}(x-\mathbf{x}_0)^2 + h(x-\mathbf{x}_0) & \text{if } x\in[\mathbf{x}_0,\mathbf{x}_1],\\ \frac{1}{4}(x-\mathbf{x}_1)^2 - \frac{h}{2}(x-\mathbf{x}_1) + \frac{h^2}{4} & \text{if } x\in(\mathbf{x}_1,\mathbf{x}_2],\\ 0 & \text{otherwise,}\end{cases} \tag{41}$$

further

$$\psi_i(x) = \begin{cases} \frac{1}{4}(x-\mathbf{x}_{i-2})^2 & \text{if } x\in(\mathbf{x}_{i-2},\mathbf{x}_{i-1}],\\ -\frac{1}{2}(x-\mathbf{x}_{i-1}-\frac{h}{2})^2 + \frac{3h^2}{8} & \text{if } x\in(\mathbf{x}_{i-1},\mathbf{x}_i],\\ \frac{1}{4}(\mathbf{x}_{i+1}-x)^2 & \text{if } x\in(\mathbf{x}_i,\mathbf{x}_{i+1}]\\ 0 & \text{otherwise,}\end{cases} \tag{42}$$

for $i=2,\dots,J-1$, and

$$\psi_J(x) = \begin{cases} \frac{1}{4}(\mathbf{x}_{J-1}-x)^2 - \frac{h}{2}(\mathbf{x}_{J-1}-x) + \frac{h^2}{4} & \text{if } x\in(\mathbf{x}_{J-2},\mathbf{x}_{J-1}]\\ -\frac{3}{4}(\mathbf{x}_J-x)^2 + h(\mathbf{x}_J-x) & \text{if } x\in(\mathbf{x}_{J-1},\mathbf{x}_J]\\ 0 & \text{otherwise,}\end{cases} \tag{43}$$

We note that both basis have a small support in $\mathcal{D}$, i.e., $\mathrm{supp}(\phi_j) = \mathrm{supp}(\psi_j)$, $j=1,\dots,J$ with

$$\mathrm{supp}(\phi_j) = \begin{cases} [\mathbf{x}_0,\mathbf{x}_2] & \text{if } i=1\,,\\ [\mathbf{x}_{j-2},\mathbf{x}_{j+1}] & \text{for } j=2,\dots,J-1\,,\\ [\mathbf{x}_{J-2},\mathbf{x}_J] & \text{if } j=J\,.\end{cases}$$

Consequently, the "mass" matrix

$$\mathbf{M}_h = \{m_{ij}\}_{i,j=1}^J := \{(\phi_j,(-\Delta)^{-1}\phi_i)_{\mathbb{L}^2}\}_{i,j=1}^J \equiv \{(\phi_j,\psi_i)_{\mathbb{L}^2}\}_{i,j=1}^J$$

which corresponds to the $\mathbb{H}$-inner product in the numerical scheme (21) will be sparse.

5.2. **Spatial discretization in higher dimensions.** We consider $\mathcal{D}=(-L,L)^d$ for some $L>0$, $d=1,2,\dots$, and denote $x=(x_1,\dots,x_d)^T\in\mathcal{D}$. Given $m\in\mathbb{N}$ we set $J:=2^m$ and consider a uniform partition of $\mathcal{D}$ with mesh size $h=\frac{2L}{J}$ into $\tilde{J}:=J^d$ rectangular subdomains $\mathcal{D}_{\underline{i}} := (\mathbf{x}_{i_1-1},\mathbf{x}_{i_1}]\times(\mathbf{x}_{i_2-1},\mathbf{x}_{i_2}]\times\dots\times(\mathbf{x}_{i_d-1},\mathbf{x}_{i_d}]$ for a multiindex $\underline{i}\in\{1,\dots,J\}^d$, where $\underline{i}:=(i_1,\dots,i_d)$, $i_k=1,\dots,J$ for $k=1,\dots,d$, and $\mathbf{x}_{i_k} := -L+i_k h$. We denote the above partition of the domain $\mathcal{D}$ as $\mathcal{T}_h=\{\mathcal{D}_{\underline{i}},\ \underline{i}\in\{1,\dots,J\}^d\}$.

We consider ϕ_{i_k}, $\psi_{i_k}=(-\Delta)^{-1}\phi_{i_k}$, $i_k=1,\dots,J$ to be the one dimensional basis functions defined in the previous section and construct the basis functions $\{\boldsymbol{\phi}_{\underline{i}}\}$, $\underline{i}\in\{1,\dots,J\}^d$ of $\mathbb{V}_h$

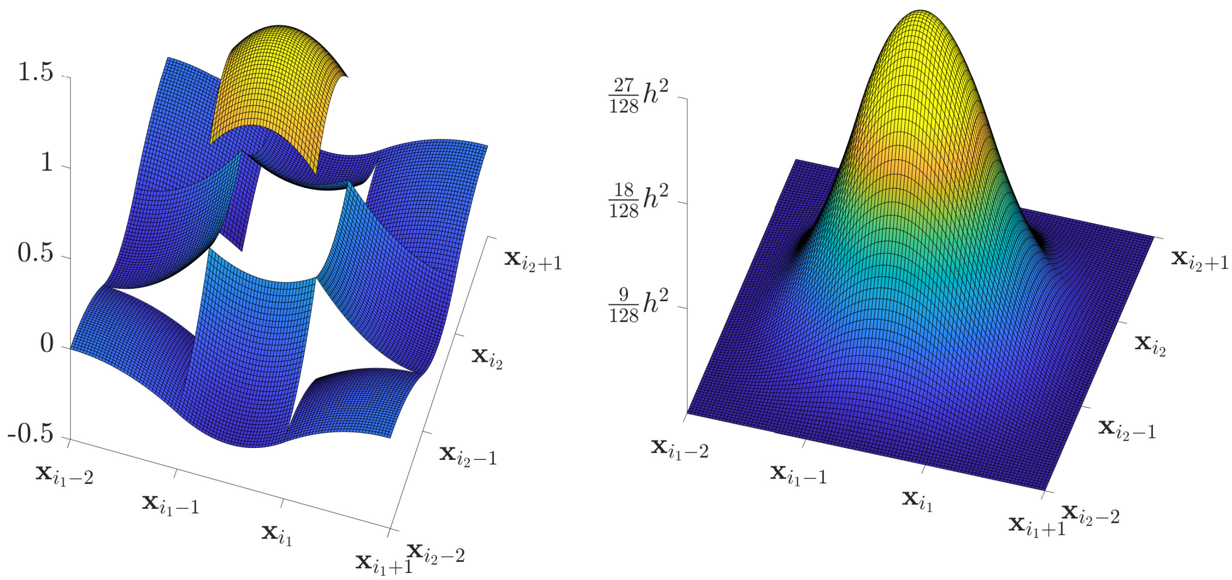


FIGURE 1. $\boldsymbol{\phi}_{(i_1,i_2)}$ and $\boldsymbol{\psi}_{(i_1,i_2)}$ for $d=2$

in $\mathbb{R}^d$ as follows: for $\underline{i} \in \{1,\dots,J\}^d$ we set

$$\begin{aligned}\boldsymbol{\phi}_{\underline{i}}(x) &= \left(\frac{3}{d^{1/(d-1)}}\frac{1}{h^2}\right)^{d-1}\sum_{k=1}^{d}\phi_{i_k}(x_k)\prod_{\substack{l=1\\l\neq k}}^{d}\psi_{i_l}(x_l)\\ &= \sum_{k=1}^{d}\phi_{i_k}(x_k)\prod_{\substack{l=1\\l\neq k}}^{d}\left(\frac{3}{d^{1/(d-1)}}\frac{1}{h^2}\psi_{i_l}(x_l)\right) \qquad x\in\mathcal{D}\,.\end{aligned} \tag{44}$$

On noting $\psi_{i_k} = (-\Delta)^{-1}\phi_{i_k}$ it can be deduced from (44) by a direct calculation that $\boldsymbol{\psi}_i = (-\Delta)^{-1}\boldsymbol{\phi}_{\underline{i}}$ can be expressed explicitly as

$$\begin{aligned}\boldsymbol{\psi}_{\underline{i}}(x) &= \left(\frac{3}{d^{1/(d-1)}}\frac{1}{h^2}\right)^{d-1}\prod_{k=1}^{d}\psi_{i_k}(x_k)\\ &= \left(\frac{d^{1/(d-1)}}{3}h^2\right)\prod_{k=1}^{d}\left(\frac{3}{d^{1/(d-1)}}\frac{1}{h^2}\psi_{i_k}(x_k)\right) \quad x\in\mathcal{D}\,.\end{aligned} \tag{45}$$

Equivalently the basis functions $\boldsymbol{\psi}_{\underline{i}}(x)$, $\underline{i}\in\{1,\dots,J\}^d$ are the solutions of the Poisson problem

$$\begin{aligned}-\Delta\boldsymbol{\psi}_{\underline{i}} = \boldsymbol{\phi}_{\underline{i}} &\quad \text{in } \mathcal{D} = (-L,L)^d,\\ \boldsymbol{\psi}_{\underline{i}} = 0 &\quad \text{on } \partial\mathcal{D}\,.\end{aligned}$$

An example of a basis function for $2 \le i_k \le J-1$ for $d=2$ is given in Figure 1.

Clearly $\boldsymbol{\psi}_{\underline{i}} \in \mathcal{C}^1(\bar{\mathcal{D}})$ since $\psi_{i_k} \in \mathcal{C}^1([-L,L])$ for all $k=1,\dots,d$. In addition, since $\psi_{i_k}, \phi_{i_k} \subset \mathbb{R}$ have a small local support in $[-L,L]$, also $\mathrm{supp}(\boldsymbol{\psi}_{\underline{i}}) = \bigtimes_{k=1}^{d}\mathrm{supp}(\psi_{i_k})$ and $\mathrm{supp}(\boldsymbol{\phi}_{\underline{i}}) =$

$\bigcup_{k=1}^{d} \mathrm{supp}(\phi_{i_k}) \times \left(\bigtimes_{\substack{l=1\\ l\neq k}}^{d} \mathrm{supp}(\psi_{i_l}) \right) \subset \mathbb{R}^d$ remain "small". Consequently, the "mass" matrix for $d \geq 1$

$$\mathbf{M}_h = \{m_{ij}\}_{i,j=1}^{\tilde{J}} := \{(\boldsymbol{\phi}_j, (-\Delta)^{-1}\boldsymbol{\phi}_i)_{\mathbb{L}^2}\}_{i,j=1}^{\tilde{J}} \equiv \{(\boldsymbol{\phi}_j, \boldsymbol{\psi}_i)_{\mathbb{L}^2}\}_{i,j=1}^{\tilde{J}}$$

is sparse; more precisely, there are only 5^d non-zero elements in each row of $\mathbf{M}_h$.

By construction, the finite element space $\mathbb{V}_h$ consists of (discontinuous) piecewise polynomial functions on the rectangular partition $\mathcal{T}_h$ of the domain $\mathcal{D}$. In order to analyze the approximation properties of $\mathbb{V}_h$ in $\mathbb{L}^p$ it is convenient to consider the space of piecewise constant functions on $\mathcal{T}_h$ which is denoted as $\overline{\mathbb{V}}_h = \mathrm{span}\{\boldsymbol{\chi}_{\underline{i}}\}$ where $\boldsymbol{\chi}_{\underline{i}} := \mathbb{1}_{\mathcal{D}_{\underline{i}}}$.

We define the restriction operator $\overline{R}_h : \mathbb{L}^p \to \overline{\mathbb{V}}_h$ as

$$\overline{R}_h v(x) := \sum_{\underline{i}\in\{1,\dots,J\}^d} \overline{v}_{\underline{i}} \boldsymbol{\chi}_{\underline{i}}(x)\,, \tag{46}$$

where $\overline{v}_{\underline{i}} = \left(\frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} v(y)\,\mathrm{d}y \right)$.

Next we analyze the properties of the operator $\overline{R}_h$.

Lemma 5.1. *For any $p \geq 1$ the operator $\overline{R}_h$ is $\mathbb{L}^p$-stable, i.e., $\|\overline{R}_h v\|_{\mathbb{L}^p} \leq \|v\|_{\mathbb{L}^p}$ for all $v \in \mathbb{L}^p$, and for all $v \in \mathbb{W}^{1,p}$ it holds that*

$$\|v - \overline{R}_h v\|_{\mathbb{L}^p} \leq Ch\|\nabla v\|_{\mathbb{L}^p}\,.$$

Proof. The $\mathbb{L}^p$-stability follows from the definition of $\overline{R}_h$ by the Hölder inequality as

$$\|\overline{R}_h v\|_{\mathbb{L}^p}^p \leq \sum_{\underline{i}\in\{1,\dots,J\}^d} |\mathcal{D}_{\underline{i}}| \int_{\mathcal{D}_{\underline{i}}} |v(y)|^p\,\mathrm{d}y \left(\int_{\mathcal{D}_{\underline{i}}} \frac{1}{|\mathcal{D}_{\underline{i}}|^{p/(p-1)}}\,\mathrm{d}y \right)^{p-1} = \|v\|_{\mathbb{L}^p}^p\,.$$

Next, we assume that v is smooth, the result for $v \in \mathbb{W}^{1,p}$ follows by density. By the fundamental theorem of calculus and the Hölder inequality we get that

$$\begin{aligned}\|v - \overline{R}_h v\|_{\mathbb{L}^p}^p &\leq \sum_{\underline{i}\in\{1,\dots,J\}^d} \frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} \int_{\mathcal{D}_{\underline{i}}} |v(x) - v(y)|^p\,\mathrm{d}y\,\mathrm{d}x \\ &\leq d^{p-1} h^p \sum_{k=1}^{d} \sum_{\underline{i}\in\{1,\dots,J\}^d} \int_{\mathcal{D}_{\underline{i}}} |\partial_{x_k} v(x)|^p\,\mathrm{d}x = C(p,d) h^p \|\nabla v\|_{\mathbb{L}^p}^p.\end{aligned}$$

□

Lemma 5.2. *$\{\overline{\mathbb{V}}_h\}_{h>0}$ is a Galerkin scheme for $\mathbb{L}^p$, $p \geq 1$. I.e., for every $v \in \mathbb{L}^p$ it holds that*

$$\inf_{\overline{v}_h \in \overline{\mathbb{V}}_h} \|v - \overline{v}_h\|_{\mathbb{L}^p} \to 0 \text{ for } h \to 0\,.$$

Proof. By density of $\mathbb{W}^{1,p} \hookrightarrow \mathbb{L}^p$ we deduce from Lemma 5.1 that

$$\|v - \overline{R}_h v\|_{\mathbb{L}^p} \to 0 \text{ for } h \to 0 \quad \forall v \in \mathbb{L}^p. \tag{47}$$

Since $\overline{R}_h v \in \overline{\mathbb{V}}_h$ we get from the above that

$$\inf_{\overline{v}_h \in \overline{\mathbb{V}}_h} \|v - \overline{v}_h\|_{\mathbb{L}^p} \leq \|v - \overline{R}_h v\|_{\mathbb{L}^p} \to 0 \text{ for } h \to 0\,.$$

□

For the (piecewise polynomial) basis functions $\boldsymbol{\phi}_{\underline{i}}$ defined in (44) we denote $\overline{\boldsymbol{\phi}}_{\underline{i}} := \overline{R}_h \boldsymbol{\phi}_{\underline{i}} \in \overline{\mathbb{V}}_h$ and observe that $\overline{\mathbb{V}}_h = \mathrm{span}\{\boldsymbol{\chi}_{\underline{i}}\} = \mathrm{span}\{\overline{\boldsymbol{\phi}}_{\underline{i}}\}$.

In order to show the approximation property of the finite element space $\mathbb{V}_h := \mathrm{span}\{\boldsymbol{\phi}_{\underline{i}}\} \subset \mathbb{L}^p$ we define the restriction operator $R_h : \mathbb{L}^p \to \mathbb{V}_h$ as

$$R_h v(x) := \sum_{\underline{i} \in \{1,\dots,J\}^d} v_{\underline{i}} \boldsymbol{\phi}_{\underline{i}}(x)\,, \tag{48}$$

where $v_{\underline{i}} = \dfrac{8}{3h^2} \dfrac{1}{|\mathcal{D}_{\underline{i}}|} \displaystyle\int_{\mathcal{D}_{\underline{i}}} (-\Delta)^{-1} v(y)\,\mathrm{d}y$.

For simplicity we restrict the proof of the convergence of the above restriction operator to $d = 2$ and assume that $\mathcal{D}$ is a rectangle; we expect an analogous proof to hold for $d \geq 3$ and more general domains as well. For $n \in \mathbb{N}$ we denote by $\mathbb{V}_n := \mathrm{span}\{e_k,\ k = 0, \dots, n\}$ the finite-dimensional space spanned by the the first n eigenfunctions of the homogeneous Dirichlet Laplace operator on the rectangular domain $\mathcal{D} = (-L, L) \times (-L, L)$

$$e_k(x_1, x_2) = \sin\left(2\pi k \frac{x_1 + L}{2L}\right) \sin\left(2\pi k \frac{x_2 + L}{2L}\right), \quad k \in \mathbb{N}. \tag{49}$$

By the density of $\cup_{n\in\mathbb{N}} \mathbb{V}_n$ in $\mathbb{L}^p$ it suffices to show the convergence of the restriction operator (48) for $v \in \mathbb{V}_n$.

Lemma 5.3. *Let $n \in \mathbb{N}$ be fixed. For any $p \geq 1$ and $v \in \mathbb{V}_n$ it holds that*

$$\|v - R_h v\|_{\mathbb{L}^p} \longrightarrow 0 \quad \text{for } h \to 0\,.$$

Proof. It is enough to show that the statement holds for $v \equiv e_k$, $k \in \mathbb{N}$.

For $x = (x_1, x_2) \in \mathcal{D}$ we consider the following discrete Laplace operator

$$\begin{aligned}(-\Delta_h^9) u(x_1, x_2) := \frac{8}{3h^2} \Big[& u(x_1, x_2) - \frac{1}{8} u(x_1 + h, x_2 + h) - \frac{1}{8} u(x_1, x_2 + h) \\ & - \frac{1}{8} u(x_1 - h, x_2 + h) - \frac{1}{8} u(x_1 + h, x_2) \\ & - \frac{1}{8} u(x_1 - h, x_2) - \frac{1}{8} u(x_1 + h, x_2 - h) \\ & - \frac{1}{8} u(x_1, x_2 - h) - \frac{1}{8} u(x_1 - h, x_2 - h) \Big]\,.\end{aligned} \tag{50}$$

The discrete Laplace operator $-\Delta_h^9$ corresponds to the 9-point finite difference approximation of the Laplace operator, cf. [8, p. 190, Example 4]; see also Figure 2. We note that for $u \in \mathcal{C}^4(\overline{\mathcal{D}})$ the discrete Laplace operator (50) satisfies the consistency property

$$(-\Delta_h^9) u(x) + \Delta u(x) = \mathcal{O}(h^2) \qquad \forall x \in \mathcal{D}. \tag{51}$$

With each element $\mathcal{D}_{\underline{i}} \in \mathcal{T}_h$ we associate the corresponding basis functions $\boldsymbol{\phi}_{\underline{i}}$, $\boldsymbol{\psi}_{\underline{i}}$. To deal with the complication that the basis functions associated with the elements of the partition $\mathcal{T}_h$ along the boundary of the domain $\mathcal{D}$ have a different shape (c.f., (44) for $i_1, i_2 = 1, J$ and (38), (40)), we introduce a layer of $4(J+1)$ "ghost" cells $\mathcal{D}^*_{(0,i_2)}$, $\mathcal{D}^*_{(J+1,i_2)}$, $\mathcal{D}^*_{(i_1,0)}$, $\mathcal{D}^*_{(i_1,J+1)}$, $i_1, i_2 = 0, \dots, J+1$ (the dimensions of the cells will be specified below) along the outer side of the boundary of $\mathcal{D}$. We then denote the resulting extended partition with $(J+2)^2$ cells as $\mathcal{T}_h^* = \mathcal{T}_h \cup \{\mathcal{D}^*_{(i_1,i_2)}\}$, i.e., $\mathcal{T}_h^*$ includes the elements of $\mathcal{T}_h$ and the "ghost" cells.

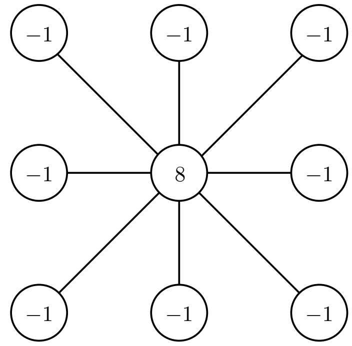


FIGURE 2. Finite difference stencil related to the Discrete Laplace operator $-\Delta_h^9$.

Recall the following trivial symmetry properties of the eigenfunctions e_k from (49) (as well as for $(-\Delta_D^{-1})e_k$, since $(-\Delta_D^{-1})e_k = \lambda_k e_k$) which hold along the boundary of $\mathcal{D}$: $e_k(-L - x_1, x_2) = -e_k(-L + x_1, x_2)$, $e_k(L + x_1, x_2) = -e_k(L - x_1, x_2)$, and $e_k(-L - x_1, -L - x_2) = e_k(-L + x_1, -L + x_2)$, $e_k(L + x_1, -L - x_2) = e_k(L - x_1, -L + x_2)$. We note that (for ghost cells $\mathcal{D}^*_{\underline{i}}$ with dimensions given implicitly via the definition (53)) the symmetry also transfers to the piecewise constant approximation of e_k over $\mathcal{T}_h^*$, i.e., for $\overline{R}_h e_k$ naturally extended on $\mathcal{T}_h^*$. We will use this fact to construct an "extension" of R_h from (48) on $\mathcal{T}_h^*$ (see (55) below).

We consider a (modified) finite element basis associated with the elements of the extended partition $\mathcal{T}_h^*$ with $(J+2)^2$ basis functions which are defined as (44) with the exception that we only use the (suitably shifted) "interior" basis functions (39), (42). Namely, we use (44) where for $i_1 = i$, $i_2 = i$ we set for $i = 0, \dots, J+1$

$$\phi_i^*(x) = -\frac{1}{2}\chi_{(\mathbf{x}_{i-2},\mathbf{x}_{i-1}]}(x) + \chi_{(\mathbf{x}_{i-1},\mathbf{x}_i]}(x) - \frac{1}{2}\chi_{(\mathbf{x}_i,\mathbf{x}_{i+1}]}(x)\,, \tag{52}$$

where we define $\mathbf{x}_{-1} = -L - (\mathbf{x}_1 - \mathbf{x}_0)$, $\mathbf{x}_{J+1} = L + (\mathbf{x}_J - \mathbf{x}_{J-1})$ (i.e., we replace the basis functions (38), (40) and (41), (43) by their "interior" counterparts); we proceed analogously for the basis functions ψ_1, ψ_J, i.e., replace (41), (43) by a suitably shifted analogues ψ_1^*, ψ_J^* of (42).

We note that the "boundary" basis functions satisfy $\phi_1(x)|_{(\mathbf{x}_0,\mathbf{x}_1)} = (\phi_1^*(x) - \phi_0^*(x))|_{(\mathbf{x}_0,\mathbf{x}_1)}$, $\phi_J(x)|_{(\mathbf{x}_{J-1},\mathbf{x}_J)} = (\phi_J^*(x) - \phi_{J+1}^*(x))|_{(\mathbf{x}_{J-1},\mathbf{x}_J)}$ (and similarly for ψ_1, ψ_J). We deduce from (44) that analogous relations also hold for $\boldsymbol{\phi}_{\underline{i}}^*$ and $\boldsymbol{\phi}_{\underline{i}}$ (as well as for $\boldsymbol{\psi}_{\underline{i}}^*$ and $\boldsymbol{\psi}_{\underline{i}}$) for instance it holds at the bottom boundary (analogically for the top, left and right boundaries)

$$\boldsymbol{\phi}_{(i_1,1)}|_{\mathcal{D}_{(i_1,1)}} = \big(\boldsymbol{\phi}^*_{(i_1,1)} - \boldsymbol{\phi}^*_{(i_1,0)}\big)|_{\mathcal{D}_{(i_1,1)}}\,, \tag{53}$$

and similarly for $\boldsymbol{\phi}_{(i_1,1)}|_{\mathcal{D}_{(i_1+1,1)}}$, $\boldsymbol{\phi}_{(i_1,1)}|_{\mathcal{D}_{(i_1-1,1)}}$. Slightly modified relations hold for the basis functions associated with the corner elements $\mathcal{D}_{(1,1)}$, $\mathcal{D}_{(1,J)}$, $\mathcal{D}_{(J,1)}$, $\mathcal{D}_{(J,J)}$ of $\mathcal{T}_h$; for instance for $\mathcal{D}_{(1,1)}$ we deduce

$$\begin{aligned}
\boldsymbol{\phi}_{(1,1)}|_{\mathcal{D}_{(1,1)}} &= \big(\boldsymbol{\phi}^*_{(1,1)} - \boldsymbol{\phi}^*_{(1,0)} - \boldsymbol{\phi}^*_{(0,1)} + \boldsymbol{\phi}^*_{(1,1)}\big)|_{\mathcal{D}_{(1,1)}}\,,\\
\boldsymbol{\phi}_{(1,1)}|_{\mathcal{D}_{(2,1)}} &= \big(\boldsymbol{\phi}^*_{(1,1)} - \boldsymbol{\phi}^*_{(1,0)}\big)|_{\mathcal{D}_{(2,1)}}\,,\\
\boldsymbol{\phi}_{(1,1)}|_{\mathcal{D}_{(1,2)}} &= \big(\boldsymbol{\phi}^*_{(1,1)} - \boldsymbol{\phi}^*_{(0,1)}\big)|_{\mathcal{D}_{(1,2)}}\,,
\end{aligned} \tag{54}$$

and similarly for basis functions at $\mathcal{D}_{(1,J)}$, $\mathcal{D}_{(J,1)}$, $\mathcal{D}_{(J,J)}$.

On noting the aforementioned symmetry properties of eigenfunctions e_k and the relations (53), (54) (along with their counterparts covering the remaining situations) we observe that (48) for $v \equiv e_k$ is equivalent to

$$R_h v(x)|_{\mathcal{D}} \equiv \sum_{\underline{i} \in \{0,1,\dots,J,J+1\}^d} v_{\underline{i}} \boldsymbol{\phi}^*_{\underline{i}}(x)\,, \tag{55}$$

where $\{\boldsymbol{\phi}^*_{\underline{i}}\}$ is the previously constructed extended basis of "interior" basis functions associated with elements of $\mathcal{T}^*_h$.

The equivalent representation (55) of the restriction operator (48) simplifies the subsequent considerations, since it only involves one type of (interior) basis functions. For the rest of the proof we will work with the basis functions $\boldsymbol{\phi}^*_{\underline{i}}$ but drop the superscript "*" to simplify the notation (also note $\boldsymbol{\phi}^*_{(i_1,i_2)} \equiv \boldsymbol{\phi}_{(i_1,i_2)}$ for $1 < i_1, i_2 < J$, i.e., the modification is only required at the boundary).

We consider an element $\mathcal{D}_{\underline{i}} \subset \mathcal{D}$. By a direct calculation of the elementwise mean of the basis functions (44) for $d = 2$ (i.e., evaluating $\overline{\boldsymbol{\phi}}_{\underline{j}} \equiv \overline{R}_h \boldsymbol{\phi}_{\underline{j}}$), we note that for $x \in \mathcal{D}_{\underline{i}}$, fixed $\underline{i} = (i_1, i_2)$ it holds that $\overline{\boldsymbol{\phi}}_{\underline{i}}(x) \equiv 1$ and $\overline{\boldsymbol{\phi}}_{\underline{j}}(x) \equiv -\frac{1}{8}$ for $\underline{j} \in \mathcal{N}(\underline{i}) := \{\underline{j} \in \{1, , \dots, J\}^2;\ \overline{\mathcal{D}}_{\underline{j}} \cap \overline{\mathcal{D}}_{\underline{i}} \neq \emptyset\} \equiv \{\underline{j} = (i_1 + k_1, i_2 + k_2);\ k_1, k_2 = -1, 0, 1\}$, $\underline{j} \neq \underline{i}$, cf. Figure 2; below we denote $\underline{k} = (k_1, k_2) \in \{-1, 0, 1\}^2$ the local index of $\underline{j}$ with respect to $\underline{i}$ and write $\underline{j} \equiv \mathrm{glob}_{\underline{i}}(\underline{k})$. Consequently, we observe that the coefficients in the definition of the discrete Laplace operator (50) for $x \in \mathcal{D}_{\underline{i}}$ correspond to the values $\overline{\boldsymbol{\phi}}_{\underline{j}}|_{\mathcal{D}_{\underline{i}}}$, $\underline{j} \in \mathcal{N}(\underline{i})$, scaled by the factor $\frac{8}{3h^2}$.

Hence, from the above observation, noting the definitions (48), (46) and recalling (50) we deduce for $x \in \mathcal{D}_{\underline{i}}$ that

$$\begin{aligned}
\overline{R}_h[R_h v](x) &= \sum_{\underline{j} \in \mathcal{N}(\underline{i})} v_{\underline{j}} \overline{\boldsymbol{\phi}}_{\underline{j}}(x) = \frac{8}{3h^2} \sum_{\underline{j} \in \mathcal{N}(\underline{i})} \frac{1}{|\mathcal{D}_{\underline{j}}|} \int_{\mathcal{D}_{\underline{j}}} (-\Delta)^{-1} v(y) \overline{\boldsymbol{\phi}}_{\underline{j}}(x) \,\mathrm{d}y \\
&\equiv \frac{8}{3h^2} \sum_{k_1,k_2=-1}^{1} \frac{1}{|\mathcal{D}_{\mathrm{glob}_{\underline{i}}(\underline{k})}|} \int_{\mathcal{D}_{\mathrm{glob}_{\underline{i}}(\underline{k})}} (-\Delta)^{-1} v(y) \overline{\boldsymbol{\phi}}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x) \,\mathrm{d}y \\
&\equiv \frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} \sum_{k_1,k_2=-1}^{1} (-\Delta)^{-1} v(y_1 + k_1 h, y_2 + k_2 h) \frac{8}{3h^2} \overline{\boldsymbol{\phi}}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x) \,\mathrm{d}y \\
&= \frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} (-\Delta_h^9) \left((-\Delta)^{-1} v(y)\right) \,\mathrm{d}y\,,
\end{aligned} \tag{56}$$

where we employed the integral transformation $\mathcal{D}_{\underline{j}} \to \mathcal{D}_{\underline{i}}$ for $\underline{j} \neq \underline{i}$ (i.e., $y = (y_1, y_2) \in \mathcal{D}_{\underline{j}} \to (y_1 + k_1 h, y_2 + k_2 h) \in \mathcal{D}_{\underline{i}}$) along with the fact that $|\mathcal{D}_{\underline{j}}| = |\mathcal{D}_{\underline{i}}|$.

By the consistency property (51) we get from (56) for $x \in \mathcal{D}_{\underline{i}}$ that

$$\begin{aligned}
\overline{R}_h[R_h v](x) &= \frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} -\Delta(-\Delta)^{-1} v(y) \,\mathrm{d}y + \mathcal{O}(h^2) = \frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} v(y) \,\mathrm{d}y + \mathcal{O}(h^2) \\
&\equiv \overline{R}_h v(x) + \mathcal{O}(h^2)\,.
\end{aligned}$$

Consequently, on recalling Lemma 5.1 we conclude for $h \to 0$ that

$$\left\| v - \overline{R}_h[R_h v] \right\|_{\mathbb{L}^p} \leq \| v - \overline{R}_h v \|_{\mathbb{L}^p} + \mathcal{O}(h^2) \to 0\,. \tag{57}$$

Next, we estimate the difference $\overline{R}_h[R_h v] - R_h v$. Due to the local support of the basis functions for $x \in \mathcal{D}_{\underline{i}}$ we may express

$$\left(\overline{R}_h[R_h v] - R_h v\right)(x) = \frac{8}{3h^2} \sum_{\underline{j}\in\mathcal{N}(i)} \frac{1}{|\mathcal{D}_{\underline{j}}|} \int_{\mathcal{D}_{\underline{j}}} (-\Delta)^{-1} v(y)\,\mathrm{d}y \left(\overline{\boldsymbol{\phi}}_{\underline{j}}(x) - \boldsymbol{\phi}_{\underline{j}}(x)\right). \tag{58}$$

As in (56) we employ the transformation $\mathcal{D}_{\underline{j}} \to \mathcal{D}_{\underline{i}}$ for $\underline{j} \neq \underline{i}$ and rewrite the above expression as

$$\begin{aligned}
&\left(\overline{R}_h[R_h v] - R_h v\right)(x) \\
&= \frac{8}{3h^2} \frac{1}{|\mathcal{D}_{\underline{i}}|} \int_{\mathcal{D}_{\underline{i}}} \sum_{k_1,k_2=-1}^{1} \left((-\Delta)^{-1} v(y_1 + k_1 h, y_2 + k_2 h) \left(\overline{\boldsymbol{\phi}}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x) - \boldsymbol{\phi}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x)\right)\right) \mathrm{d}y\,.
\end{aligned}$$

Hence, after expressing the basis functions (44) explicitly (recall $\underline{i} = (i_1, i_2)$, $x = (x_1, x_2) \in \mathcal{D}_{\underline{i}} = (\mathbf{x}_{i_1-1}, \mathbf{x}_{i_1}) \times (\mathbf{x}_{i_2-1}, \mathbf{x}_{i_2})$), for each $y = (y_1, y_2)$ we restate

$$\begin{aligned}
&\sum_{k_1,k_2=-1}^{1} (-\Delta)^{-1} v(y_1 + k_1 h, y_2 + k_2 h) \left(\overline{\boldsymbol{\phi}}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x) - \boldsymbol{\phi}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x)\right) \\
&\quad = \Bigg[(-\Delta)^{-1} v(y_{(1,1)}) \left(-\frac{1}{2} a_{i_1,1}(x_1) - \frac{1}{2} a_{i_2,1}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(0,1)}) \left(-\frac{1}{2} a_{i_1,2}(x_1) + a_{i_2,1}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(-1,1)}) \left(-\frac{1}{2} a_{i_1,3}(x_1) - \frac{1}{2} a_{i_2,1}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(1,0)}) \left(a_{i_1,1}(x_1) - \frac{1}{2} a_{i_2,2}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(0,0)}) \left(a_{i_1,2}(x_1) + a_{i_2,2}(x_2) - 1\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(-1,0)}) \left(a_{i_1,3}(x_1) - \frac{1}{2} a_{i_2,2}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(1,-1)}) \left(-\frac{1}{2} a_{i_1,1}(x_1) - \frac{1}{2} a_{i_2,3}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(0,-1)}) \left(-\frac{1}{2} a_{i_1,2}(x_1) + a_{i_2,3}(x_2) + \frac{1}{8}\right) \\
&\qquad + (-\Delta)^{-1} v(y_{(-1,-1)}) \left(-\frac{1}{2} a_{i_1,3}(x_1) - \frac{1}{2} a_{i_2,3}(x_2) + \frac{1}{8}\right)\Bigg],
\end{aligned} \tag{59}$$

where we employ a shorthand notation $y_{(k_1,k_2)} = (y_1 + k_1 h, y_2 + k_2 h)$ and for $n = 1, 2$ we denote (cf. (44))

$$
\begin{aligned}
a_{i_n,1}(x_n) &:= \frac{3}{2h^2}\psi_{i_n+1}(x_n) = \frac{3}{2h^2}\frac{1}{4}\left(x_n - \mathbf{x}_{i_n-1}\right)^2 && \text{for } x_n \in (\mathbf{x}_{i_n-1}, \mathbf{x}_{i_n}),\\
a_{i_n,2}(x_n) &:= \frac{3}{2h^2}\psi_{i_n}(x_n) = \frac{3}{2h^2}\left[-\frac{1}{2}\left(x_n - \mathbf{x}_{i_n-1} - \frac{h}{2}\right)^2 + \frac{3h^2}{8}\right] && \text{for } x_n \in (\mathbf{x}_{i_n-1}, \mathbf{x}_{i_n}),\\
a_{i_n,3}(x_n) &:= \frac{3}{2h^2}\psi_{i_n-1}(x_n) = \frac{3}{2h^2}\frac{1}{4}\left(\mathbf{x}_{i_n} - x_n\right)^2 && \text{for } x_n \in (\mathbf{x}_{i_n-1}, \mathbf{x}_{i_n}).
\end{aligned}
$$

The following property, which follows from (42) by direct calculation, will be essential in the sequel

$$
a_{i_n,1}(x) + a_{i_n,2}(x) + a_{i_n,3}(x) = \frac{3}{4}\,, \tag{60}
$$

for $i_n = 2, \ldots, J-1$, and $x \in (\mathbf{x}_{i_n-1}, \mathbf{x}_{i_n})$.

Next, we expand the terms $\tilde{v}(y_{(k_1,k_2)}) := (-\Delta)^{-1} v(y_1 + k_1 h, y_2 + k_2 h)$ in (59) at $y \equiv y_{(0,0)}$ using Taylor series as

$$
\sum_{k_1,k_2=-1}^{1} \tilde{v}(y_{(k_1,k_2)})\left(\overline{\phi}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x) - \phi_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x)\right) = I + \cdots + IV\,,
$$

where

$$
\begin{aligned}
I =& \Big[\tilde{v}(y) + \left(\partial_{x_1}\tilde{v}(y) + \partial_{x_2}\tilde{v}(y)\right)h + \left(\frac{1}{2}\partial_{x_1}^2\tilde{v}(y) + \partial_{x_1}\partial_{x_2}\tilde{v}(y) + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)\right)h^2\\
&\quad + \mathcal{O}(h^3)\Big]\left(-\frac{1}{2}a_{i_1,1}(x_1) - \frac{1}{2}a_{i_2,1}(x_2) + \frac{1}{8}\right)\\
&+ \left[\tilde{v}(y) + \partial_{x_2}\tilde{v}(y)h + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)h^2 + \mathcal{O}(h^3)\right]\left(-\frac{1}{2}a_{i_1,2}(x_1) + a_{i_2,1}(x_2) + \frac{1}{8}\right)\,,
\end{aligned}
$$

$$
\begin{aligned}
II =& \Big[\tilde{v}(y) + \left(-\partial_{x_1}\tilde{v}(y) + \partial_{x_2}\tilde{v}(y)\right)h + \left(\frac{1}{2}\partial_{x_1}^2\tilde{v}(y) - \partial_{x_1}\partial_{x_2}\tilde{v}(y) + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)\right)h^2\\
&\quad + \mathcal{O}(h^3)\Big]\left(-\frac{1}{2}a_{i_1,3}(x_1) - \frac{1}{2}a_{i_2,1}(x_2) + \frac{1}{8}\right)\\
&+ \left[\tilde{v}(y) + \partial_{x_1}\tilde{v}(y)h + \frac{1}{2}\partial_{x_1}^2\tilde{v}(y)h^2 + \mathcal{O}(h^3)\right]\left(a_{i_1,1}(x_1) - \frac{1}{2}a_{i_2,2}(x_2) + \frac{1}{8}\right)\\
&+ \tilde{v}(y)\left(a_{i_1,2}(x_1) + a_{i_2,2}(x_2) - 1\right)\,,
\end{aligned}
$$

$$
\begin{aligned}
III =& \left[\tilde{v}(y) - \partial_{x_1}\tilde{v}(y)h + \frac{1}{2}\partial_{x_1}^2\tilde{v}(y)h^2 + \mathcal{O}(h^3)\right]\left(a_{i_1,3}(x_1) - \frac{1}{2}a_{i_2,2}(x_2) + \frac{1}{8}\right)\\
&+ \Big[\tilde{v}(y) + \left(\partial_{x_1}\tilde{v}(y) - \partial_{x_2}\tilde{v}(y)\right)h + \left(\frac{1}{2}\partial_{x_1}^2\tilde{v}(y) - \partial_{x_1}\partial_{x_2}\tilde{v}(y) + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)\right)h^2\\
&\quad + \mathcal{O}(h^3)\Big]\left(-\frac{1}{2}a_{i_1,1}(x_1) - \frac{1}{2}a_{i_2,3}(x_2) + \frac{1}{8}\right)\,,
\end{aligned}
$$

$$
\begin{aligned}
IV =& \left[\tilde{v}(y) - \partial_{x_2}\tilde{v}(y)h + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)h^2 + \mathcal{O}(h^3)\right]\left(-\frac{1}{2}a_{i_1,2}(x_1) + a_{i_2,3}(x_2) + \frac{1}{8}\right) \\
&+ \left[\tilde{v}(y) - \left(\partial_{x_1}\tilde{v}(y) + \partial_{x_2}\tilde{v}(y)\right)h + \left(\frac{1}{2}\partial_{x_1}^2\tilde{v}(y) + \partial_{x_1}\partial_{x_2}\tilde{v}(y) + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)\right)h^2 \right.\\
&\left. + \mathcal{O}(h^3)\right]\left(-\frac{1}{2}a_{i_1,3}(x_1) - \frac{1}{2}a_{i_2,3}(x_2) + \frac{1}{8}\right).
\end{aligned}
$$

We rearrange the above terms $I - IV$, use the identity (60) and obtain

$$
\begin{aligned}
&\sum_{k_1,k_2=-1}^{1}\tilde{v}(y_{(k_1,k_2)})\left(\overline{\phi}_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x) - \phi_{\mathrm{glob}_{\underline{i}}(\underline{k})}(x)\right) \\
&\quad = 0\cdot\left[\tilde{v}(y) + \left(\partial_{x_1}\tilde{v}(y) + \partial_{x_2}\tilde{v}(y)\right)h + \partial_{x_1}\partial_{x_2}\tilde{v}(y)h^2\right] \\
&\qquad + \frac{1}{2}\partial_{x_1}^2\tilde{v}(y)\left[-a_{i_2,1}(x_2) - a_{i_2,2}(x_2) - a_{i_2,3}(x_2) + \frac{3}{4}\right]h^2 \\
&\qquad + \frac{1}{2}\partial_{x_2}^2\tilde{v}(y)\left[-a_{i_1,1}(x_1) - a_{i_1,2}(x_1) - a_{i_1,3}(x_1) + \frac{3}{4}\right]h^2 + \mathcal{O}(h^3) \\
&\quad = \mathcal{O}(h^3).
\end{aligned}
\tag{61}
$$

Hence, we substitute (61) into (58) to conclude that

$$
\begin{aligned}
\|\overline{R}_h[R_h v] - R_h v\|_{\mathbb{L}^p} &= \left\|\frac{8}{3h^2}\sum_{\underline{j}\in\{1,\dots,J\}^d}\frac{1}{|\mathcal{D}_{\underline{j}}|}\int_{\mathcal{D}_{\underline{j}}}(-\Delta)^{-1}v(y)\left(\overline{\phi}_{\underline{j}}(\cdot) - \phi_{\underline{j}}(\cdot)\right)\,\mathrm{d}y\right\|_{\mathbb{L}^p} \\
&= Ch.
\end{aligned}
\tag{62}
$$

Finally, by the triangle inequality we estimate

$$
\|v - R_h v\|_{\mathbb{L}^p} \leq \|v - \overline{R}_h[R_h v]\|_{\mathbb{L}^p} + \|\overline{R}_h[R_h v] - R_h v\|_{\mathbb{L}^p}, \tag{63}
$$

and the statement follows by (57) and (62). □

The above lemma allows us to deduce the density of $\{\mathbb{V}_h\}_{h>0}$ in $\mathbb{L}^p$.

Corollary 5.4 (Approximation property of $\mathbb{V}_h$)**.** *For every $v \in \mathbb{L}^p$, $p \geq 1$ it holds that*

$$
\inf_{v_h\in\mathbb{V}_h}\|v - v_h\|_{\mathbb{L}^p} \to 0 \quad \textit{for } h \to 0.
$$

Proof. Consider $v_\varepsilon \in \mathbb{V}_n$ and note that $\lim_{h\to 0}\|v_\varepsilon - R_h v_\varepsilon\|_{\mathbb{L}^p} = 0$ by Lemma 5.3. Since $R_h v_\varepsilon \in \mathbb{V}_h$ we get

$$
\inf_{v_h\in\mathbb{V}_h}\|v - v_h\|_{\mathbb{L}^p} \leq \|v - R_h v_\varepsilon\|_{\mathbb{L}^p} \leq \|v - v_\varepsilon\|_{\mathbb{L}^p} + \|v_\varepsilon - R_h v_\varepsilon\|_{\mathbb{L}^p}.
$$

The statement then follows by the density of $\cup_{n\in\mathbb{N}}\mathbb{V}_n$ in $\mathbb{L}^p$. □

The restriction operator (48) is not implementable since it requires the evaluation of the function $(-\Delta)^{-1}v$, which is not available in general. For practical purposes (e.g., to compute the discrete approximation of the initial condition) it is convenient to consider the discrete $\mathbb{H}^{-1}$-projection $P_h : \mathbb{H}^{-1} \to \mathbb{V}_h$ which is defined for $v \in \mathbb{H}^{-1}$ as follows

$$
(P_h v, w_h)_{\mathbb{H}^{-1}} = (v, w_h)_{\mathbb{H}^{-1}} \quad \forall w_h \in \mathbb{V}_h. \tag{64}
$$

Remark 5.5. *The $\mathbb{H}^{-1}$-stability of the orthogonal projection, i.e., $\|P_h v\|_{\mathbb{H}^{-1}} \leq C\|v\|_{\mathbb{H}^{-1}}$ follows on taking $w_h = P_h v$ in (64) and using the Cauchy-Schwarz and Young's inequalities. Furthermore, we note that (64) is equivalent to $P_h v = \arg\min_{w_h \in \mathbb{V}_h} \|v - w_h\|^2_{\mathbb{H}^{-1}}$ which in particular implies that $P_h(R_h v) = R_h v$ for $v \in \mathbb{L}^p$.*

Consequently, the $\mathbb{H}^{-1}$-stability of P_h and the continuous embedding $\mathbb{L}^p \hookrightarrow \mathbb{H}^{-1}$, for $p \geq 2$ yield for all $v \in \mathbb{L}^p$, $v_\varepsilon \in \mathbb{V}_n$

$$\begin{aligned}\|v - P_h v\|_{\mathbb{H}^{-1}} &\leq \|v - v_\varepsilon\|_{\mathbb{H}^{-1}} + \|v_\varepsilon - P_h v_\varepsilon\|_{\mathbb{H}^{-1}} + \|P_h(v_\varepsilon - v)\|_{\mathbb{H}^{-1}} \\ &\leq C\|v - v_\varepsilon\|_{\mathbb{H}^{-1}} + \|v_\varepsilon - R_h v_\varepsilon\|_{\mathbb{H}^{-1}} + \|P_h(v_\varepsilon - R_h v_\varepsilon)\|_{\mathbb{H}^{-1}} \\ &\leq C\|v - v_\varepsilon\|_{\mathbb{H}^{-1}} + C\|v_\varepsilon - R_h v_\varepsilon\|_{\mathbb{L}^p}\,.\end{aligned}$$

Hence, by Lemma 5.3, the density of $\mathbb{L}^p$, $p \geq 2$ in $\mathbb{H}^{-1}$ and the density of $\cup_{n\in\mathbb{N}}\mathbb{V}_n$ in $\mathbb{L}^p$ we conclude the approximation property of the $\mathbb{H}^{-1}$-orthogonal projection:

$$\lim_{h\to 0} \|v - P_h v\|_{\mathbb{H}^{-1}} = 0 \qquad \forall v \in \mathbb{H}^{-1}\,.$$

6. Numerical Experiments

6.1. **Convergence of the projection in $d = 2$.** We study the experimental $\mathbb{L}^p$-convergence of the $\mathbb{H}^{-1}$-projection operator (64) as well as of an implementable counterpart $\widetilde{R}_h : \mathbb{L}^p \to \mathbb{V}_h$ of the restriction operator (48) defined as

$$\widetilde{R}_h v(x) := \sum_{\underline{i}\in\{1,\ldots,J\}^d} \tilde{v}_{\underline{i}} \phi_{\underline{i}}(x)\,,$$

where $\tilde{v}_{\underline{i}} = \dfrac{8}{3h^2}[(-\Delta_h^9)^{-1}\overline{R}_h v]_{\underline{i}}$. I.e., the coefficients are the solutions of finite difference scheme

$$-\Delta_h^9\left(\frac{3h^2}{8}\tilde{v}_{\underline{i}}\right) = \overline{R}_h v|_{\mathcal{D}_{\underline{i}}}\,,$$

for $\underline{i} \in \{1,\ldots,J\}^2$; we note that it holds by construction that $\overline{R}_h\widetilde{R}_h v = \overline{R}_h v$ and $\widetilde{R}_h\overline{R}_h v = \widetilde{R}_h v$.

In Figure 3 we display the convergence plot of the $\mathbb{H}^{-1}$-projection of the Barenblatt solution $P_h u_B(t,\cdot)$ at $t = 0.1$ (see (65) below) along with the convergence plot of $P_h\chi_{(-0.5,0.5)^2}$ of the (non-smooth) indicator function of the $(-0.5, 0.5)^2$-square; in both cases $\mathcal{D} = (-1.5, 1.5)^2$. The convergence plot implies convergence of the projection in $\mathbb{L}^p$ of order h for the smooth Barenblatt function and of order of $h^{2/3}$ in the non-smooth case.

In addition we display in Figure 3 the convergence plot of the restriction operator $\widetilde{R}_h$ for the indicator function $\chi_{(-0.5,0.5)^2}$ which is also of order $h^{2/3}$.

6.2. **Barenblatt solution for the deterministic PME.** We consider the equation (1) with $\alpha(u) = |u|^{p-2}u$, $f \equiv 0$, $g \equiv 0$, $\sigma \equiv 0$ which corresponds to the deterministic porous medium equation

$$\partial_t u = \Delta(|u|^{p-2}u).$$

The exact solution of the porous media equation with initial condition $u_0 = \delta_0$ (i.e., the δ-distribution centered at 0) the so-called Barenblatt solution

$$u_B(t,x) = t^{-a}\max\left\{0, C - k|x|^2 t^{-2b}\right\}^{1/(p-2)}\,, \tag{65}$$

where a, b, k, C are suitable constants that depend on p, d, c.f. [57, Ch. 17.5].

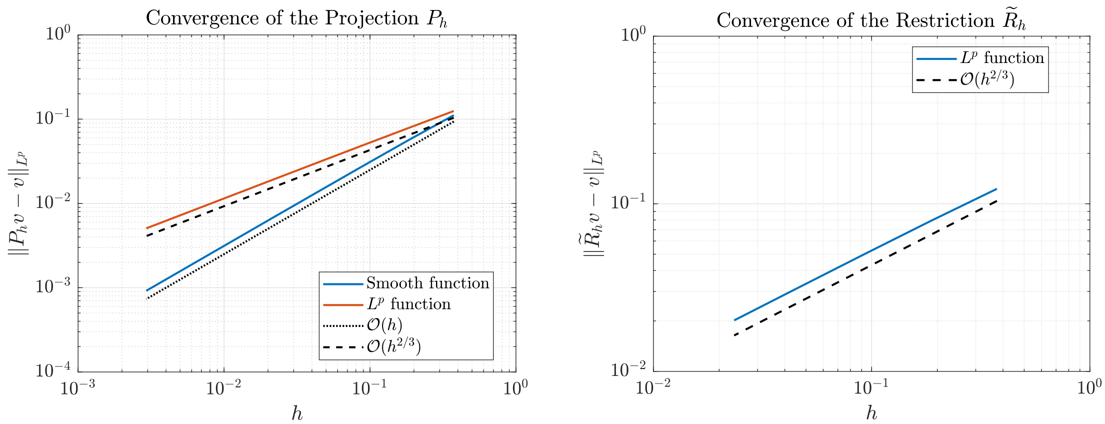


FIGURE 3. Convergence of the $\mathbb{H}^{-1}$-projection (left) and of the restriction operator $\widetilde{R}_h$ (right).

In the experiments below we choose $\mathcal{D} = (-1.5, 1.5)^d$, $d = 1, 2$ and $T = 0.1$, $p = 3$. We consider a regularized initial condition $u_0 = \delta_0 \approx \tilde{u}_{h,0} \in \overline{\mathbb{V}}_h$ with

$$\tilde{u}_{h,0}(x) = \frac{1}{(2h)^d} \begin{cases} 1 & \text{if } x \in \mathcal{D}_{\underline{i}},\ j \in \left\{\frac{J}{2}, \frac{J}{2}+1\right\}^d \\ 0 & \text{else} \end{cases}$$

and set $u_{h,0} = P_h(\tilde{u}_0) \in \mathbb{V}_h$.

We examine the convergence of the numerical approximation with respect to τ, h in the L^p-norm, i.e., we compute the error $\|u_B - \overline{u}_\tau\|_{L^p([\underline{t},T]\times\mathcal{D})}$ with time-interval $[\underline{t}, T] = [0.01, 0.1]$ where we choose $\underline{t} > 0$ to reduce the effect of the approximation of the initial condition.

In Table 1 we display the L^p-error for $\tau = 1/N$, $h = 2L/J$ in $d = 1$. The corresponding convergence plots in Figure 4 indicate that the convergence order of the numerical approximation with respect to τ is slightly less than one and around $\frac{3}{2}$ with respect to h.

$N \setminus J$	8	16	32	64	128	256
8	0.083032	0.02221	0.020881	0.024805	0.025878	0.025931
16	0.07524	0.016254	0.015419	0.015481	0.015893	0.016162
32	0.075711	0.017544	0.0070398	0.0088919	0.0094508	0.0094772
64	0.077172	0.021771	0.0054151	0.0051293	0.0056912	0.0059705
128	0.077702	0.022649	0.0060452	0.0028429	0.0033319	0.0035322
256	0.077591	0.022801	0.006569	0.002342	0.0017761	0.0019579
512	0.077532	0.022934	0.0069462	0.002467	0.0011016	0.0010656
1024	0.077593	0.023061	0.007187	0.0025917	0.00099377	0.00060724

TABLE 1. $L^p((0.01, 0.1) \times \mathcal{D})$-error of the solution, $d = 1$.

To highlight the finite speed of propagation property on the discrete level we display the evolution of the support of the numerical approximation in Figure 5.

Next we examine the convergence behaviour in $d = 2$, we note that in this case $u_0 \notin \mathbb{H}^{-1}$. In Table 2 we display the L^p-error computed for $\tau = 1/N$, $h = 2L/J$.

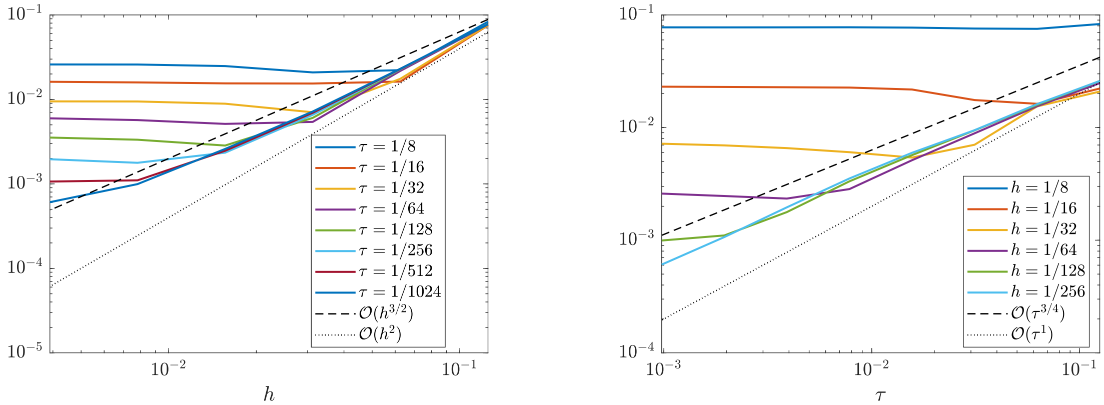


FIGURE 4. Convergence of the $L^p((0.01, 0.1) \times \mathcal{D})$-error of the solution for the deterministic equation in $1d$. Left: convergence of the spacial discretization for different time step-sizes, right: convergence of the time discretization for different spacial step-sizes

The corresponding convergence plots in Figure 6 indicate that the convergence order of the numerical approximation with respect to τ and h are both close to one. As expected, (due to the lower regularity of the initial condition in $d = 2$) the observed convergence order of the spatial discretization is slightly worse than the corresponding convergence order for $d = 1$. We display the time evolution of the numerical solution in Figure 7 and a detail of the numerical solution at $T = 0.1$, $d = 2$ is displayed in Figure 8.

$N \setminus J$	8	16	32	64	128	256
8	0.092154	0.050998	0.045512	0.04534	0.04531	0.045288
16	0.094038	0.047956	0.032504	0.028832	0.027894	0.027644
32	0.095218	0.048104	0.026604	0.019108	0.016976	0.016402
64	0.098787	0.050731	0.026017	0.015466	0.011883	0.010852
128	0.10062	0.052414	0.026237	0.013883	0.0087862	0.0070965
256	0.10077	0.052773	0.026229	0.013247	0.0072316	0.0048007
512	0.10086	0.052981	0.026295	0.013092	0.0067107	0.0037732
1024	0.10109	0.05325	0.026433	0.013108	0.0065808	0.0034176

TABLE 2. $L^p((0.01, 0.1) \times \mathcal{D})$-error of the solution, $d = 2$.

6.3. **Barenblatt solution for the stochastic PME.** Now we consider the stochastic equation with $\phi(u) = u$, set again $f, g = 0$. Then from [56, p. 87,88] and the references cited therein, we get, that for $d = 1$, $\mathcal{D} = \mathbb{R}$ and $p = 3$ the solution at time $t \in (0, T]$ is given by

$$u_B\left(\int_0^t e^{W(s)-\frac{s}{2}}\,\mathrm{d}s, \cdot\right) e^{W(t)-\frac{t}{2}},$$

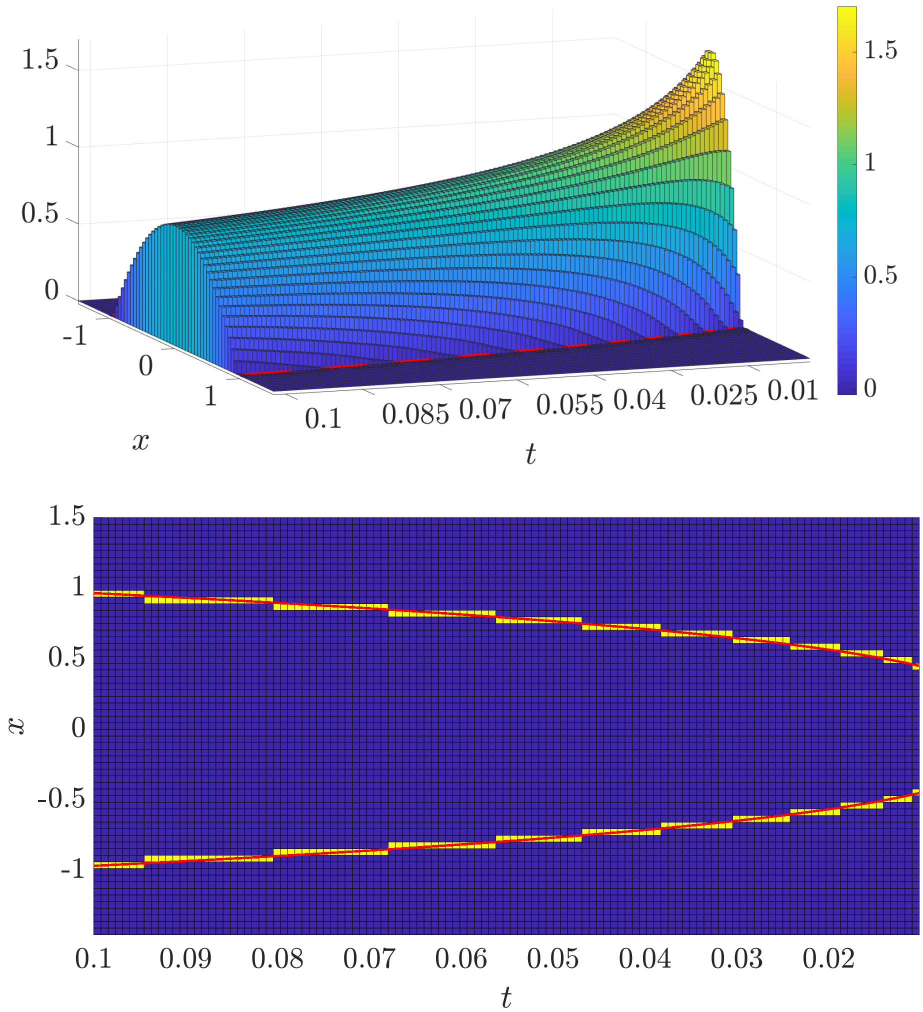


FIGURE 5. (top) Time evolution of the numerical solution for $J = 64$, $N = 128$ in $d = 1$; (bottom) the corresponding support of the numerical solution, in yellow, and the support of the analytical solution, in red.

where u_B is the Barenblatt solution defined above. The support at time $t \in (0, T]$ are all $x \in \mathbb{R}$, such that

$$|x| \leq \sqrt{C\frac{2d(p-1)}{a(p-2)}} \left(\int_0^t e^{W(s)-\frac{s}{2}}\,\mathrm{d}s\right)^{a/d} = \sqrt{12C}\sqrt[3]{\int_0^t e^{W(s)-\frac{s}{2}}\,\mathrm{d}s}, \tag{66}$$

with $C = C(d, p)$ as above. Hence we can cut-off the domain $\mathcal{D}$ at $\pm L$ such that $(-L, L)$ contains the support of the solution in $(0, T]$ for each considered path of W - this is verified for each path during the simulation. For our simulations we take $L = 1.5$.

6.4. **Numerical Results in** $1d$ **for the stochastic equation.** The norm of the error is computed as before, where in addition a Monte-Carlo approximation for the expected value

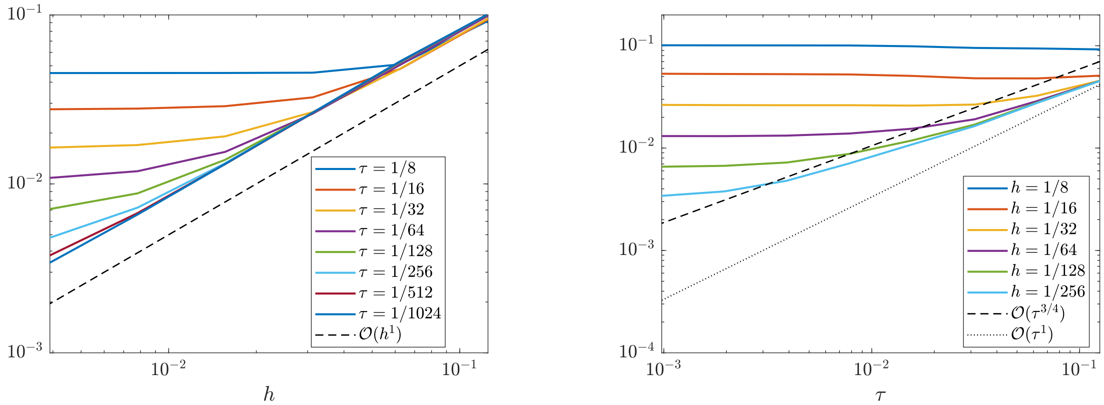


FIGURE 6. Convergence of the $L^p((0.01, 0.1) \times \mathcal{D})$-error of the solution for the deterministic equation in $2d$. Left: convergence of the discretization for different time step-sizes, right: convergence of the discretization for different mesh-sizes.

is used.

In Table 3 and Figure 9 we see, that convergence with respect to (τ, h) also for the approximation stochastic Barenblatt solution holds.

$N \setminus J$	8	16	32	64	128	256
8	0.045434	0.022878	0.036722	0.040406	0.041319	0.041538
16	0.06444	0.01176	0.021544	0.025007	0.025851	0.026057
32	0.069643	0.012408	0.011412	0.014496	0.015296	0.01549
64	0.074465	0.016884	0.0065525	0.0087332	0.0095565	0.0097648
128	0.076666	0.020286	0.0057942	0.00491	0.0056214	0.0058322
256	0.077362	0.02164	0.0063097	0.0029489	0.0031048	0.0032878
512	0.077722	0.022367	0.0067772	0.0025539	0.0017667	0.0018573
1024	0.077983	0.022812	0.0070991	0.002603	0.0012162	0.0010803

TABLE 3. $L^p(\Omega \times (0.01, 0.1) \times \mathcal{D})$-error of the solution in $1d$, a Monte-Carlo Approximation with 10^6 samples was used to approximate the expectation.

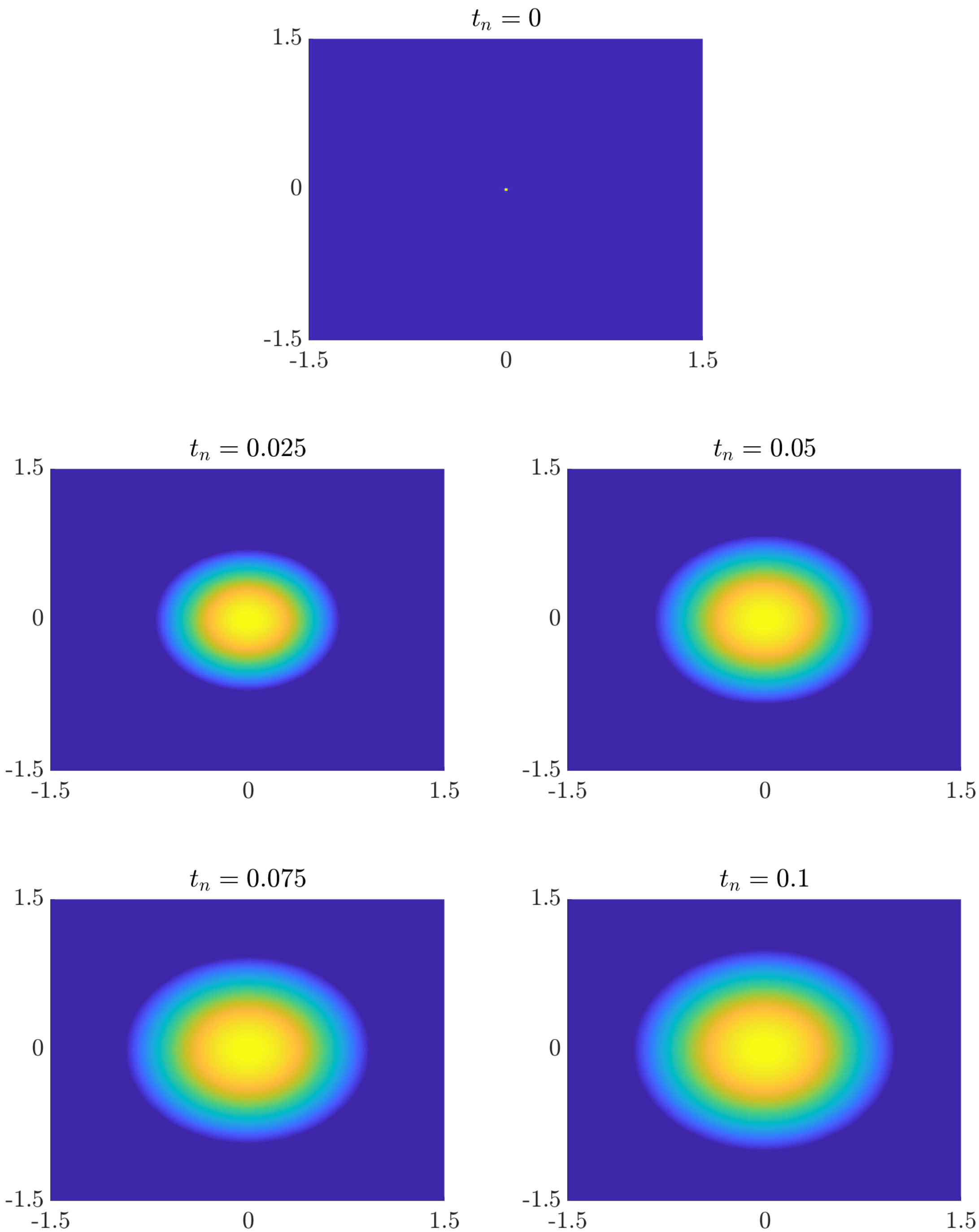


FIGURE 7. Snapshots of the numerical solution computed with $J = N = 256$ at time $t = 0, 0.025, 0.05, 0.075, 0.1$.

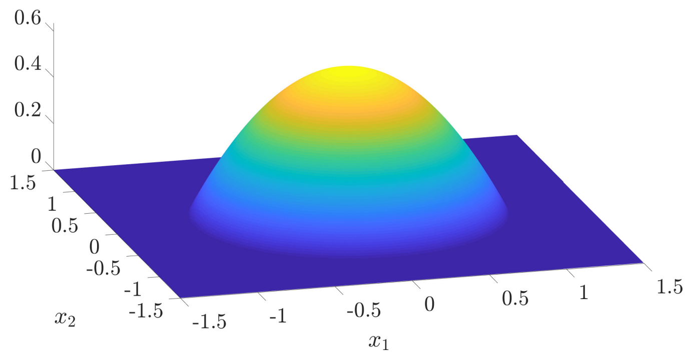


FIGURE 8. Numerical approximation of the Barenblatt solution at time $t = T$ for $J = N = 256$, $d = 2$

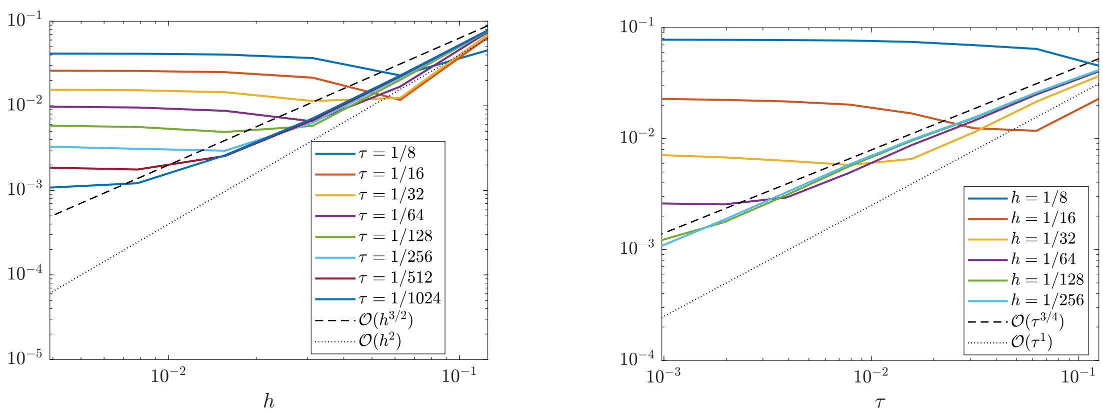
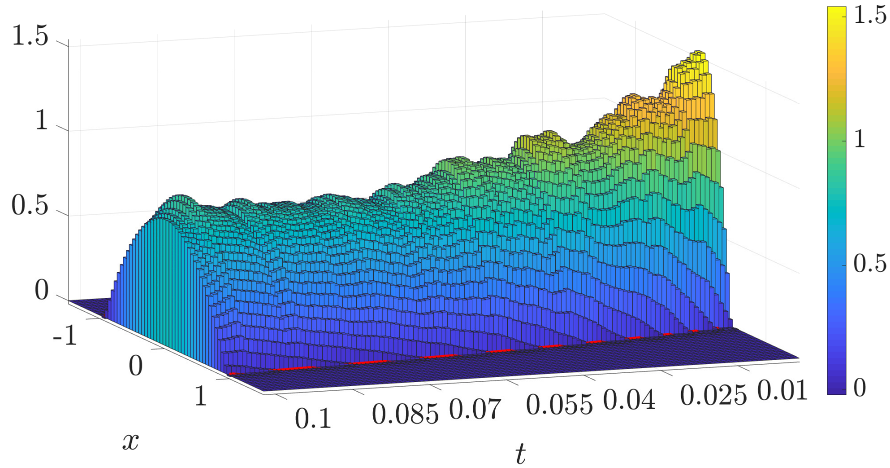
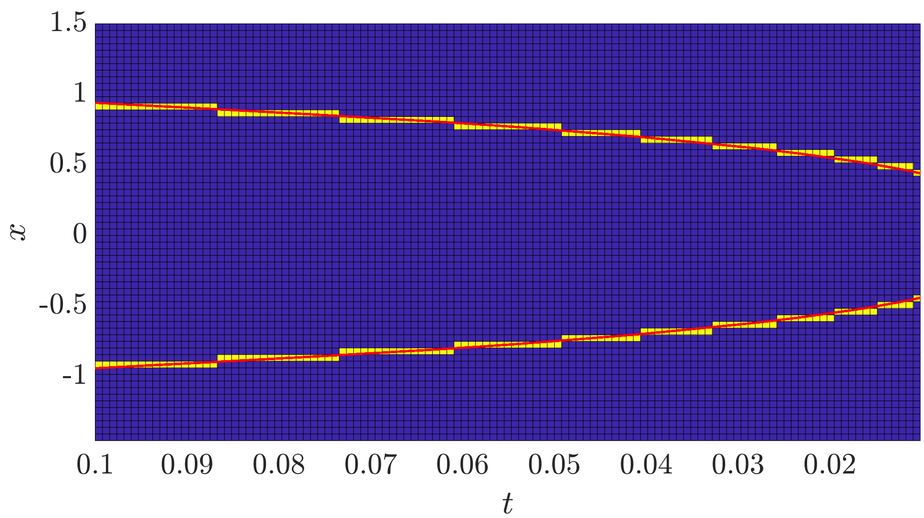


FIGURE 9. Convergence of the $L^p(\Omega \times (0.01, 0.1) \times \mathcal{D})$-error of the numerical solution in $1d$, a Monte-Carlo Approximation with 10^6 samples was used to approximate the expectation. Left: convergence of the spacial discretization for different time step-sizes, right: convergence of the time discretization for different spacial step-sizes.

Figure 10 shows one sample path, the analytical support (for this path) is plotted in red and the support of the approximation in yellow. Finite speed of propagation hold a.s.



(A) $J = 64$, $N = 128$, $1d$

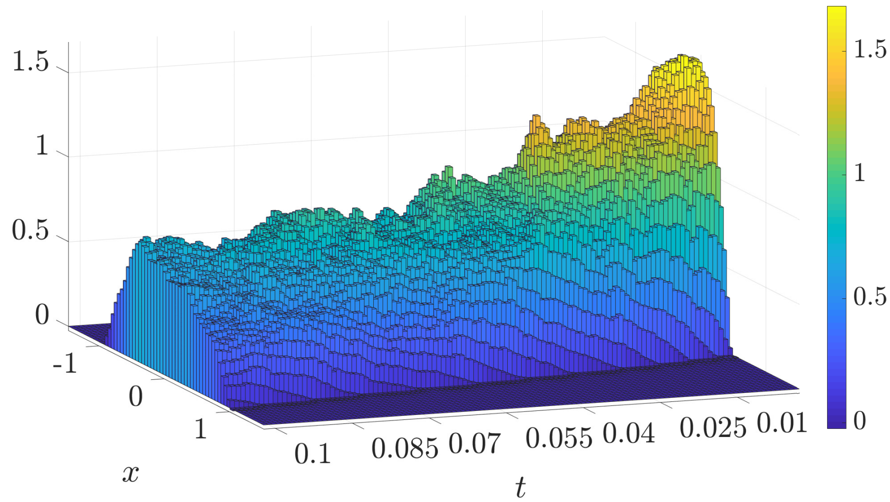


(B) Support of the approximation with $J = 64$, $N = 128$, $1d$. The red line indicates the analytical support from (66).

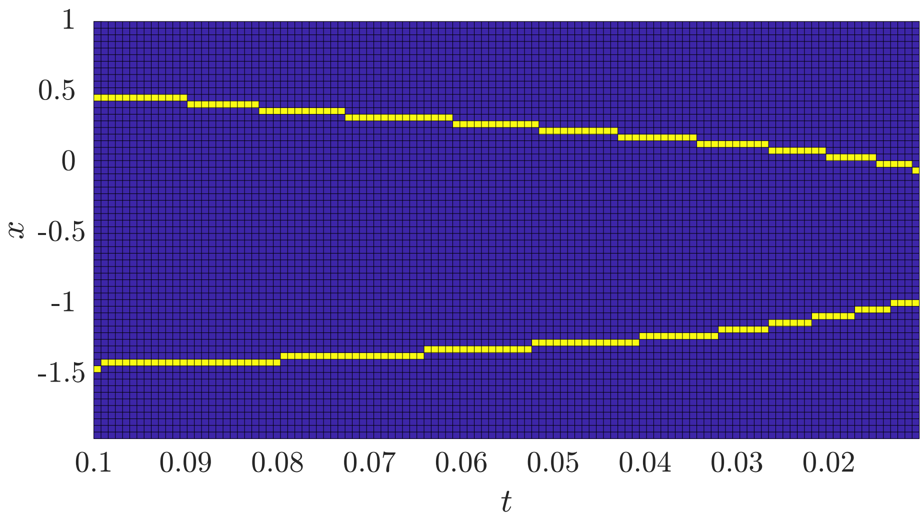
FIGURE 10. Approximation of the stochastic solution in $1d$.

6.5. **Numerical results for space-time noise.** Next, we perform simulation of the stochastic porous media equation with space-time white noise on $\mathcal{D} = [-L, L]$, $L = 1.5$, where no analytical solution is available. Given the the mesh size $h = \frac{2L}{J}$ we take $\sigma(u) \equiv \sigma_h(u) = \sigma_0 \sum_{i=1}^{J} u \frac{\boldsymbol{\chi}_i}{|\mathcal{D}_i|}$ where where $\boldsymbol{\chi}_i = \mathbb{1}_{\mathcal{D}_i}$ are the indicator functions of $\mathcal{D}_i$. We note that the $\bar{\mathbb{V}}_h$-valued noise $\sigma_h(u)W(t,x) = \sigma_0 u \sum_{i=1}^{J} \frac{\boldsymbol{\chi}_i(x)}{|\mathcal{D}_i|}\beta_i(t)$ in an approximation of the multiplicative noise $u\widetilde{W}$ where $\widetilde{W}$ is the space-time white noise, cf. [1].

In Figure 11 we display the numerical solution for one realization of the discrete space-time white noise with $\sigma_0 = \frac{1}{64}$ along with the corresponding support. We observe that the evolution of the support for the space-time white noise does not deviate significantly from the deterministic case. In particular the numerical approximation preserves the finite speed of propagation of the support, see Figure 12.



(A) Evolution of the numerical solution with the space-time-dependent noise, $J = 64$, $N = 128$

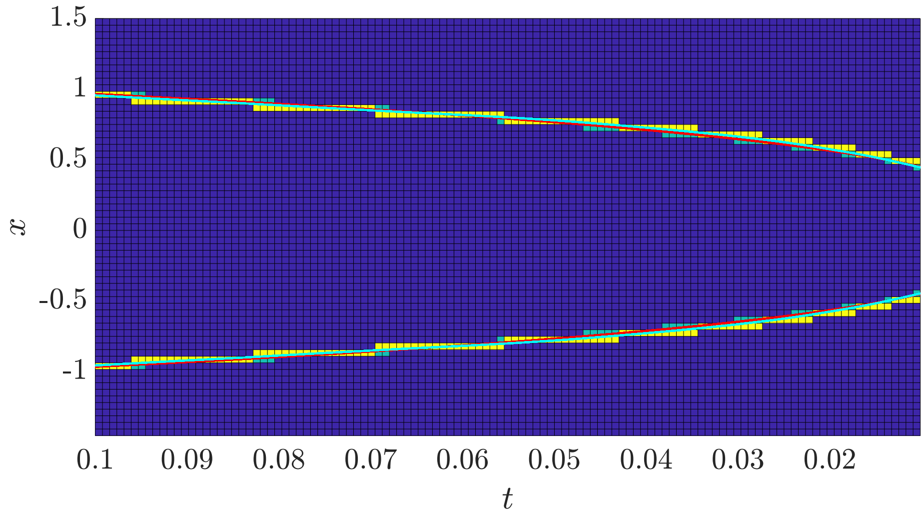


(B) Support of the approximation with $J = 64$, $N = 128$, $1d$ - asymmetric

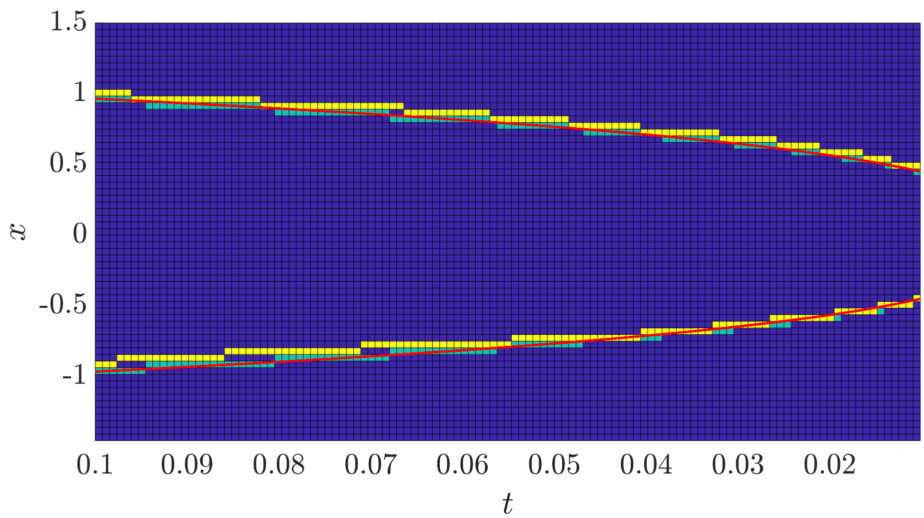
FIGURE 11. Approximation of the stochastic solution with time-space dependent noise in $1d$.

Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft through SFB 1283 "Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications".



(A) Comparison of the spreading of the support $J = 64$, $N = 128$ in $1d$: deterministic in green, stochastic (time dependent noise) in yellow. The lines indicate the analytical support: deterministic in red and stochastic in cyan.



(B) Comparison of the spreading of the support $J = 64$, $N = 128$ in $1d$: deterministic in green, stochastic (space-time dependent) in yellow. The red line indicates the analytical support for the deterministic solution.

FIGURE 12. Support of the different approximations det./stochastic in $1d$.

REFERENCES

[1] E. J. Allen, S. J. Novosel, and Z. Zhang. Finite element and difference approximation of some linear stochastic partial differential equations. *Stochastics Stochastics Rep.*, 64(1-2):117–142, 1998.

[2] Viorel Barbu, Philippe Blanchard, Giuseppe Da Prato, and Michael Röckner. Self-organized criticality via stochastic partial differential equations. In *Potential theory and stochastics in Albac*, volume 11 of *Theta Ser. Adv. Math.*, pages 11–19. Theta, Bucharest, 2009.

[3] Viorel Barbu, Giuseppe da Prato, and Michael Röckner. *Stochastic Porous Media Equations*. Springer International Publishing, 2016.
[4] Viorel Barbu and Michael Röckner. An operatorial approach to stochastic partial differential equations driven by linear multiplicative noise. *J. Eur. Math. Soc. (JEMS)*, 17(7):1789–1815, 2015.
[5] Viorel Barbu and Michael Röckner. Nonlinear Fokker–Planck equations driven by Gaussian linear multiplicative noise. *Journal of Differential Equations*, 265(10):4993–5030, November 2018.
[6] Caroline Bauzet, Guy Vallet, and Petra Wittbold. A degenerate parabolic-hyperbolic Cauchy problem with a stochastic force. *J. Hyperbolic Differ. Equ.*, 12(3):501–533, 2015.
[7] Wolf-Jürgen Beyn, Benjamin Gess, Paul Lescot, and Michael Röckner. The Global Random Attractor for a Class of Stochastic Porous Media Equations. *Comm. Partial Differential Equations*, 36(3):446–469, 2011.
[8] Garrett Birkhoff and Robert E. Lynch. *Numerical solution of elliptic problems*, volume 6 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.
[9] K. Dareiotis, M. Gerencsér, and B. Gess. Entropy solutions for stochastic porous media equations. *Journal of Differential Equations*, 266(6):3732 – 3763, 2019.
[10] Konstantinos Dareiotis, Máté Gerencsér, and Benjamin Gess. Porous media equations with multiplicative space-time white noise. *arXiv preprint arXiv:2002.12924*, 2020.
[11] Konstantinos Dareiotis, Máté Gerencsér, and Benjamin Gess. Porous media equations with multiplicative space-time white noise. *arXiv:2002.12924 [math]*, February 2020.
[12] Konstantinos Dareiotis and Benjamin Gess. Supremum estimates for degenerate, quasilinear stochastic partial differential equations. *arXiv:1712.06655*, 2017.
[13] Konstantinos Dareiotis and Benjamin Gess. Nonlinear diffusion equations with nonlinear gradient noise. *Electron. J. Probab.*, 25:Paper No. 35, 43, 2020.
[14] Konstantinos Dareiotis, Benjamin Gess, and Pavlos Tsatsoulis. Ergodicity of stochastic porous media equations. *arXiv:1907.04605*, 2019.
[15] D. Dean. Langevin equation for the density of a system of interacting Langevin processes. *Journal of Physics A: Mathematical and General*, 29(24):L613, 1996.
[16] Arnaud Debussche, Sylvain de Moor, and Martina Hofmanová. A Regularity Result for Quasilinear Stochastic Partial Differential Equations of Parabolic Type. *SIAM J. Math. Anal.*, 47(2):1590–1614, 2015.
[17] Arnaud Debussche, Martina Hofmanová, and Julien Vovelle. Degenerate parabolic stochastic partial differential equations: Quasilinear case. *The Annals of Probability*, 44(3):1916–1955, May 2016.
[18] Felix Del Teso, Jørgen Endal, and Espen R Jakobsen. Robust numerical methods for nonlocal (and local) equations of porous medium type. part i: Theory. *SIAM Journal on Numerical Analysis*, 57(5):2266–2299, 2019.
[19] Nicolas Dirr, Benjamin Fehrman, and Benjamin Gess. Conservative stochastic pde and fluctuations of the symmetric simple exclusion process. *preprint*, 2020.
[20] Nicolas Dirr, Marios Stamatakis, and Johannes Zimmer. Entropic and gradient flow formulations for nonlinear diffusion. *J. Math. Phys.*, 57(8):081505, 13, 2016.
[21] A. Donev, T. G. Fai, and E. Vanden-Eijnden. A reversible mesoscopic model of diffusion in liquids: From giant fluctuations to Fick's law. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(4):P04004, April 2014.
[22] Jérôme Droniou and Kim-Ngan Le. The gradient discretization method for slow and fast diffusion porous media equations. *SIAM Journal on Numerical Analysis*, 58(3):1965–1992, 2020.
[23] Carsten Ebmeyer and WB Liu. Finite element approximation of the fast diffusion and the porous medium equations. *SIAM journal on numerical analysis*, 46(5):2393–2410, 2008.
[24] Etienne Emmrich and David Šiška. Full discretization of the porous medium/fast diffusion equation based on its very weak formulation. *Commun. Math. Sci.*, 10(4):1055–1080, 2012.
[25] Etienne Emmrich and David Šiška. Nonlinear stochastic evolution equations of second order with damping. *Stoch. Partial Differ. Equ. Anal. Comput.*, 5(1):81–112, 2017.
[26] Benjamin Fehrman and Benjamin Gess. Path-by-path well-posedness of nonlinear diffusion equations with multiplicative noise. *arXiv preprint arXiv:1807.04230*, 2018.
[27] Benjamin Fehrman and Benjamin Gess. Large deviations for conservative stochastic pde and non-equilibrium fluctuations. *arXiv preprint arXiv:1910.11860*, 2019.
[28] Benjamin Fehrman and Benjamin Gess. Well-posedness of nonlinear diffusion equations with nonlinear, conservative noise. *Archive for Rational Mechanics and Analysis*, 233(1):249–322, 2019.

[29] Julian Fischer and Günther Grün. Finite speed of propagation and waiting times for the stochastic porous medium equation: A unifying approach. *SIAM Journal on Mathematical Analysis*, 47(1):825–854, 2015.
[30] Benjamin Gess. Strong solutions for stochastic partial differential equations of gradient type. *J. Funct. Anal.*, 263(8):2355–2383, 2012.
[31] Benjamin Gess. Finite speed of propagation for stochastic porous media equations. *arXiv:1210.2415*, pages 1–26, 2013.
[32] Benjamin Gess. Random attractors for stochastic porous media equations perturbed by space-time linear multiplicative noise. *Ann. Probab.*, 42(2):818–864, 2014.
[33] Benjamin Gess. Finite time extinction for stochastic sign fast diffusion and self-organized criticality. *Comm. Math. Phys.*, 335(1):309–344, 2015.
[34] Benjamin Gess and Martina Hofmanová. Well-posedness and regularity for quasilinear degenerate parabolic-hyperbolic SPDE. *The Annals of Probability*, 46(5):2495–2544, 2018.
[35] Benjamin Gess and Panagiotis E. Souganidis. Stochastic non-isotropic degenerate parabolic–hyperbolic equations. *Stochastic Process. Appl.*, 127(9):2961–3004, 2017.
[36] D. Gilbarg and N.S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.
[37] Benjamin Goldys, Michael Röckner, and Xicheng Zhang. Martingale solutions and Markov selections for stochastic partial differential equations. *Stochastic Processes and their Applications*, 119(5):1725–1764, May 2009.
[38] H. Grillmeier and G. Grün. Nonnegativity preserving convergent schemes for stochastic porous-medium equations. *Math. Comp.*, 88(317):1021–1059, 2019.
[39] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1985.
[40] István Gyöngy. On stochastic squations with respect to semimartingales iii. *Stochastics*, 7(4):231–254, 1982.
[41] I. n. Gyöngy and A. Millet. On discretization schemes for stochastic evolution equations. *POTENTIAL ANALYSIS, 2005, 23, 2, 99*, 2005.
[42] Willi Jäger and Jozef Kačur. Solution of porous medium type systems by linear approximation schemes. *Numerische Mathematik*, 60(1):407–427, 1991.
[43] Jong U. Kim. On the stochastic porous medium equation. *J. Differential Equations*, 220(1):163–194, 2006.
[44] N. V. Krylov and B. L. Rozovskii. Stochastic evolution equations. *Journal of Soviet Mathematics*, 16(4):1233–1277, Jul 1981.
[45] T. Lehmann, V. Konarovskyi, and M. von Renesse. Dean-Kawasaki Dynamics: Ill-posedness vs. Triviality. *arXiv:1806.05018*, June 2018.
[46] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod; Gauthier-Villars, Paris, 1969.
[47] Wei Liu and Michael Röckner. *Stochastic Partial Differential Equations: An Introduction*. Universitext. Springer, Cham, 2015.
[48] E Magenes, RH Nochetto, and C Verdi. Energy error estimates for a linear scheme to approximate non-linear parabolic problems. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 21(4):655–678, 1987.
[49] S. Méléard and S. Roelly. Interacting measure branching processes. Some bounds for the support. *Stochastics Stochastics Rep.*, 44(1-2):103–121, 1993.
[50] J. Nečas. *Les méthodes directes en théorie des équations elliptiques*. Masson, 1967.
[51] Hans Christian Öttinger. *Beyond equilibrium thermodynamics*. John Wiley & Sons, 2005.
[52] Étienne Pardoux. Equations aux dérivées partielles stochastiques non linéaires monotones. *PhD thesis*, 1975.
[53] Jiagang Ren, Michael Röckner, and Feng-Yu Wang. Stochastic generalized porous media and fast diffusion equations. *Journal of Differential Equations*, 238(1):118–152, 2007.
[54] Luca Scarpa and Ulisse Stefanelli. Doubly nonlinear stochastic evolution equations. *arXiv:1905.11294 [math]*, July 2019.
[55] Roger Temam. *Navier-Stokes equations*, volume 2 of *Studies in mathematics and its applications*. North-Holland Publ., Amsterdam [u.a.], rev. ed. edition, 1979.
[56] Michael Röckner (auth.) Viorel Barbu, Giuseppe Da Prato. *Stochastic Porous Media Equations*. Lecture Notes in Mathematics 2163. Springer International Publishing, 1 edition, 2016.

[57] Juan Luis Vázquez. *The porous medium equation.* Oxford mathematical monographs. Clarendon Press, Oxford [u.a.], 2007.

[58] Feng-Yu Wang. Exponential convergence of non-linear monotone SPDEs. *Discrete and Continuous Dynamical Systems. Series A*, 35(11):5239–5253, 2015.

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany
Email address: banas@math.uni-bielefeld.de

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany and Max Planck Institute for Mathematics in the Sciences, Inselstr. 22, 04103 Leipzig, Germany
Email address: bgess@math.uni-bielefeld.de

Department of Mathematics, Bielefeld University, 33501 Bielefeld, Germany
Email address: cvieth@math.uni-bielefeld.de