Numerische Mathematik (2023) 153:557–609
https://doi.org/10.1007/s00211-022-01343-7

Numerische Mathematik

An averaged space–time discretization of the stochastic p-Laplace system

Lars Diening[1] · **Martina Hofmanová**[1] · **Jörn Wichmann**[1]

Received: 20 April 2022 / Revised: 27 September 2022 / Accepted: 11 December 2022 /
Published online: 31 December 2022

Abstract

We study the stochastic p-Laplace system in a bounded domain. We propose two new space–time discretizations based on the approximation of time-averaged values. We establish linear convergence in space and $1/2$ convergence in time. Additionally, we provide a sampling algorithm to construct the necessary random input in an efficient way. The theoretical error analysis is complemented by numerical experiments.

Mathematics Subject Classification 65N15 · 65N30 · 35K55 · 35K65 · 35K67 · 60H15

Contents

The research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1283/2 2021 – 317210226.

✉ Jörn Wichmann
jwichmann@math.uni-bielefeld.de

Lars Diening
lars.diening@uni-bielefeld.de

Martina Hofmanová
hofmanova@math.uni-bielefeld.de

1 Department of Mathematics, University of Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany

1 Introduction

Let $\mathcal{O} \subset \mathbb{R}^n$ be a polygonal domain, $n \geq 1$, $N \geq 1$, $T > 0$ be finite. We are interested in the approximation of the solution process $u : \Omega \times [0, T] \times \mathcal{O} \to \mathbb{R}^N$ to the stochastic p-Laplace system. Given an initial datum u_0 and a stochastic forcing term $G(u)\,\mathrm{d}W$ (for the precise assumptions see Assumption 1), u is determined by the relations

$$\begin{aligned} \mathrm{d}u - \operatorname{div} S(\nabla u)\,\mathrm{d}t &= G(u)\,\mathrm{d}W && \text{in } \Omega \times (0, T) \times \mathcal{O}, \\ u &= 0 && \text{on } \Omega \times (0, T) \times \partial\mathcal{O}, \\ u(0) &= u_0 && \text{on } \Omega \times \mathcal{O}, \end{aligned} \tag{1.1}$$

where $S(\xi) := (\kappa + |\xi|)^{p-2}\xi \in \mathbb{R}^{N \times n}$, $p \in (1, \infty)$ and $\kappa \geq 0$. Closely related to S is the nonlinear operator $V(\xi) := (\kappa + |\xi|)^{\frac{p-2}{2}}\xi \in \mathbb{R}^{N \times n}$.

1.1 Model

The system (1.1) has many important applications in nature. As one major example, it provides a prototype system towards the modeling of non-Newtonian fluids. More specifically, it is closely related to power law fluids [14, 69]. The case $p = 2$ corresponds to the famous stochastic heat equation and has been studied extensively, both analytically e.g. [12, 47, 48] as well as numerically e.g. [1, 24, 36, 42, 67].

The p-Laplace system arises as a stochastically perturbed gradient flow of the energy $J : W_0^{1,p}(\mathcal{O}) \to [0, \infty)$ defined by

$$\mathcal{J}(u) := \int_{\mathcal{O}} \varphi(|\nabla u|)\,\mathrm{d}x, \tag{1.2}$$

where $\varphi(t) := \int_0^t (\kappa + s)^{p-2} s \, \mathrm{d}s$. In the case $\kappa = 0$ the energy (1.2) corresponds to the classical p-Dirichlet energy.

1.2 Existence and well-posedness

The existence of analytically weak solutions in the space $L^2\left(\Omega; C\left([0,T]; L^2(\mathcal{O})\right)\right) \cap L^p\left(\Omega; L^p\left(0,T; W_0^{1,p}(\mathcal{O})\right)\right)$ can be established by standard monotonicity arguments [61]. It requires a linear growth assumption on the noise coefficient G and $L^2(\mathcal{O})$-integrable initial data.

First results on the existence of strong solutions to stochastically perturbed gradient flows have been obtained by Gess [43]. It includes the degenerate, $p \geq 2$, p-Laplace equation.

In the literature different generalization of (1.1) have been considered. Well-posedness for merely L^1-initial data has been addressed in [68]. More general systems, where p is allowed to depend on (ω, t, x) respectively on (t, x), are considered in [73, 74] respectively in [9]. It was further extended by Breit and Gmeineder [17] to electro-rheological fluids. Electro-rheological fluids are modeled by (1.1) complemented by an additional divergence-free constraint and a free pressure variable. The singular case $p \in [1, 2)$ has been analyzed in [60], [44] and [5].

1.3 Regularity

Regularity properties of a function u become particularly important when it comes to the approximability of u within a discrete function class. Prominently, discrete tensor spaces generated via a time stepping scheme and a finite element discretization in space can approximate smooth functions more easily compared to non-smooth functions.

Historically, many authors addressed Hölder and $C^{1,\alpha}$-regularity of solutions to the deterministic steady p-Laplace equation, e.g. [26, 34, 56, 64, 70–72]. In general $\alpha \in (0, 1)$ is an unknown quantity. While $C^{1,\alpha}$-regularity can be used to measure the approximation quality of finite elements, they usually fail to produce optimal results, since in general $\alpha < 1$.

Sobolev regularity provides an alternative scale of smoothness. In the singular case, $p \in (1, 2)$, $W^{2,2}$-regularity has been proven by Liu and Barrett [59]. The degenerate setting is more delicate and one can not expect $\nabla^2 u$ to be well-defined on $\{\nabla u = 0\}$. In fact, due to the nonlinear structure of the equation, solutions have limited regularity even for smooth data. A sharp result in the 2-dimensional setting about limited regularity for the Hölder as well as Sobolev scale has been obtained by Iwaniec and Manfredi [51].

The nonlinear character of the equation naturally introduces the additional quantities $S(\nabla u)$ and $V(\nabla u)$. In the scalar case both expressions are robust with respect to $p \in (1, \infty)$. Here it is possible to prove $V(\nabla u) \in W^{1,2}$ via a difference quotient technique [13, 58] and $S(\nabla u) \in W^{1,2}$ by a functional inequality [20]. The result $V(\nabla u) \in W^{1,2}$ generalizes to the vectorial case. However, the functional inequality fails in the vectorial case at least point-wise for $p \leq 2(2-\sqrt{2}) \approx 1.1715$. Therefore, it is unclear whether a regularity result $S(\nabla u) \in W^{1,2}$ is achievable for small p. Nevertheless, for $p > 2(2-\sqrt{2})$ it is shown in [2] that $S(\nabla u) \in W^{1,2}$. Regularity for $S(\nabla u)$ on the Besov and Triebel-Lizorkin scale in the plane for $p \geq 2$ has been obtained in [3]. Estimates of $S(\nabla u)$ in terms of Riesz potentials were derived in [53, 54].

Regularity results for the parabolic p-Laplace system were derived in [41, Theorem 6.2.1] (cf. [57] for $p \geq 2$). It was shown that $u \in C^1_{\mathrm{weak}}\big(0,T;L^2(\mathcal{O})\big) \cap C_{\mathrm{weak}}\big(0,T;W^{1,p}_0(\mathcal{O})\big)$. It was extended in [30] to the nonlinear tensor $V(\nabla u)$. The authors showed, by formally testing the equation with $-\Delta u$ respectively $-\partial_t^2 u$, that

$$V(\nabla u) \in L^2\left(0,T;W^{1,2}(\mathcal{O})\right) \cap W^{1,2}\left(0,T;L^2(\mathcal{O})\right), \tag{1.3a}$$

$$u \in L^\infty\left(0,T;W^{1,2}(\mathcal{O})\right) \cap W^{1,\infty}\left(0,T;L^2(\mathcal{O})\right). \tag{1.3b}$$

Additionally, $S(\nabla u) \in L^2(0,T;W^{1,2}(\mathcal{O}))$ was proven in [22] for either $p \in (1,\infty)$ and $N = 1$ or $p \geq 2$ and $N > 1$.

Within the context of the stochastic p-Laplace system it is possible to prove similar spatial regularity as in (1.3). The formal testing needs to be replaced by a suitable application of Itô's formula as done by Breit [15]. However, the time regularity is limited due to the presence of the stochastic forcing. In the super-quadratic regime partial time regularity can be recovered by exploiting the strong formulation of the p-Laplace system as done by Wichmann [76]. Overall, for appropriate data assumptions it is possible to verify (see Sect. 2.5)

$$V(\nabla u) \in L^2\left(\Omega;L^2\left(0,T;W^{1,2}(\mathcal{O})\right) \cap B^{1/2}_{2,\infty}\left(0,T;L^2(\mathcal{O})\right)\right), \tag{1.4a}$$

$$u \in L^2\left(\Omega;L^\infty\left(0,T;W^{1,2}(\mathcal{O})\right) \cap B^{1/2}_{\Phi_2,\infty}\left(0,T;L^2(\mathcal{O})\right)\right). \tag{1.4b}$$

Here $\Phi_2(s) := e^{s^2} - 1$ and B denotes a Besov-Orlicz space (see Sect. 2.1).

1.4 Approximation

In the past many authors have studied the numerical approximation of the deterministic counterpart of (1.1), e.g. [6–8, 10, 11, 16, 19, 25, 30, 33, 35, 37, 75]. The error of the discrete and analytic solution has been expressed in various different quantities. It turned out that the natural quantity to measure the error is given by

$$\max_{0\leq m\leq M} \left\| u(t_m) - u_{m,h}\right\|^2_{L^2(\mathcal{O})} + \tau \sum_{m=1}^{M} \left\| V(\nabla u(t_m)) - V(\nabla u_{m,h})\right\|^2_{L^2(\mathcal{O})}, \tag{1.5}$$

where $V(\xi) := (\kappa + |\xi|)^{\frac{p-2}{2}} \xi$ and $u_{m,h}$ is an approximation of $u(t_m)$. In [33] it has been proved that the expression $\|V(\nabla u) - V(\nabla u_h)\|_{L^2(\mathcal{O})}^2$ is equivalent to the energy error $\mathcal{J}(u_h) - \mathcal{J}(u)$. Here u_h is a minimizer of the energy on a subspace $V_h \subset W_0^{1,p}$.

In the steady case, starting with the seminal work by Barrett and Liu [6] and further improvements in [37] and [29], it has been proved that

$$\|V(\nabla u) - V(\nabla u_h)\|_{L^2(\mathcal{O})} \lesssim h \, \|\nabla V(\nabla u)\|_{L^2(\mathcal{O})} .$$

This settles the question about optimal convergence for piece-wise linear continuous elements. In fact, the paper [30] deals with the parabolic system and optimal convergence under the regularity assumption (1.3) has been achieved, i.e.,

$$\begin{aligned} &\max_{0 \le m \le M} \left\| u(t_m) - u_{m,h} \right\|_{L^2(\mathcal{O})}^2 + \tau \sum_{m=1}^{M} \left\| V(\nabla u(t_m)) - V(\nabla u_{m,h}) \right\|_{L^2(\mathcal{O})}^2 \\ &\lesssim h^2 + \tau^2. \end{aligned} \tag{1.6}$$

The error quantity (1.6) relies on the fact that the mapping $t \mapsto \|V(\nabla u(t))\|_{L^2(\mathcal{O})}$ is continuous. However, if the data is not sufficiently smooth, point-values might not be well-defined. In general, dealing with irregular data and therefore irregular solutions is a delicate task. Different methods have been suggested to recover well-defined error quantities.

In [16] the first and the third author develop a numerical scheme based on time averages to circumvent the usage of point evaluations. A more probabilistic approach has been used in [38]. There the authors replace deterministic evaluation points by random ones.

First results for monotone stochastic equations were derived by Gyöngy and Millet [45, 46]. They developed an abstract discretization theory that also covers the system (1.1) in the superquadratic case. In [45] they gave conditions that guarantee plain convergence of their discretization even for the full degenerate, $\kappa = 0$, system. Contrary, convergence rates have been obtained under a non-degeneracy assumption, $\kappa > 0$, in [46]. In both cases, they require restrictive assumptions on the regularity of the solution.

In the stochastic case, due to the limited time regularity (1.4), robust error quantities need to be used. Breit, Hofmanová and Loisel [18] use randomized time steps to construct an algorithm that achieves almost optimal convergence in time and optimal convergence in space, i.e., for all $\alpha \in (0, 1/2)$,

$$\begin{aligned} &\mathbb{E}_{\mathbf{t}} \otimes \mathbb{E} \Bigg[\max_{0 \le m \le M} \left\| u(\mathbf{t}_m) - u_{m,h} \right\|_{L^2(\mathcal{O})}^2 \\ &\qquad + \sum_{m=1}^{M} \int_{\mathbf{t}_{m-1}}^{\mathbf{t}_m} \left\| V(\nabla u(s)) - V(\nabla u_{m,h}) \right\|_{L^2(\mathcal{O})}^2 \, \mathrm{d}s \Bigg] \\ &\lesssim h^2 + \tau^{2\alpha}. \end{aligned}$$

Here $\mathbb{E}_{\mathbf{t}}$ denotes the expectation with respect to the random time steps $\mathbf{t}_m$ centered around deterministic grid-points t_m. They use Hölder and Sobolev-Slobodeckij spaces to measure the time regularity of the solution. This excludes the limiting case $\alpha = 1/2$.

The classical Euler–Maruyama scheme has been analyzed in [39]. The authors show, for general multi-valued monotone equations, convergence of the algorithm with convergence rate 1/4. In particular, the p-Laplace system can be treated. The authors infer convergence towards the spatially discrete solution $v_h \in V_h$ of

$$\mathrm{d}v_h - \operatorname{div} S(\nabla v_h)\,\mathrm{d}t = G(v_h)\,\mathrm{d}W$$

with rate 1/4 measured in the error

$$\max_{m \le M} \mathbb{E}\left[\| v_h(t_m) - u_{m,h} \|_{L^2_x}^2 \right] \lesssim \tau^{1/2}.$$

The method has the advantage that no regularity on the abstract solution needs to be assumed. However, only a sub-optimal convergence rate can be obtained.

1.5 Main results

Based on deterministic time averages we propose two algorithms (3.14) and (3.15) that are essentially driven by the update rule, $\mathbb{P}$-a.s. for all $m \ge 2$ and $\xi_h \in V_h$

$$(v_m - v_{m-1}, \xi_h) + \tau\,(S(\nabla v_m), \nabla \xi_h) = (G(v_{m-2})\Delta_m \mathbb{W}, \xi_h)\,. \tag{1.7}$$

The randomness enters through the averaged increments $\Delta_m \mathbb{W} := \langle W \rangle_m - \langle W \rangle_{m-1}$, where $\langle W \rangle_m$ is the time averaged value of W on the interval $[t_{m-1}, t_m]$.

Importantly, we manage to achieve optimal convergence in time with rate 1/2 without assuming any time Hölder regularity on the solution process. Instead, we measure time regularity in terms of Nikolskii spaces. Our main results, Theorem 19 and 25, verify under the condition

$$u \in L^2\left(\Omega; B^{1/2}_{2,\infty}\left(0, T; L^2(\mathcal{O})\right) \cap L^\infty\left(0, T; W^{1,2}(\mathcal{O})\right)\right), \tag{1.8a}$$

$$V(\nabla u) \in L^2\left(\Omega; B^{1/2}_{2,\infty}\left(0, T; L^2(\mathcal{O})\right) \cap L^2\left(0, T; W^{1,2}(\mathcal{O})\right)\right), \tag{1.8b}$$

the optimal convergence

$$\mathbb{E}\left[\max_{m=1,\dots,M} \| \langle u \rangle_m - v_m \|_{L^2_x}^2 + \sum_{m=1}^{M} \int_{t_{m-1}}^{t_m} \| V(\nabla u(s)) - V(\nabla v_m) \|_{L^2_x}^2 \,\mathrm{d}s \right]$$
$$\lesssim h^2 + \tau.$$

We want to stress that the regularity assumption (1.8) is even weaker than the provable regularity (1.4).

On the other hand if the solution process u has certain amount of time regularity, e.g. (1.4b), one can recover point evaluations, cf. Lemma 17,

$$\mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|u(t_m)-\langle u\rangle_m\|_{L_x^2}^2\right]\lesssim \tau\ln(1+\tau^{-1})\mathbb{E}\left[[u]^2_{B^{1/2}_{\Phi_2,\infty}L^2_x}\right].$$

For the implementation of (1.7) one needs to sample according to the distribution of the random variable $\Delta_m\mathbb{W}$. We show in Corollary 33 that $\Delta_m\mathbb{W}$ is a Gaussian random variable whose variance is slightly reduced compared to the classical increments $\Delta_m W := W(t_m)-W(t_{m-1})$. Ultimately, we provide a sampling algorithm (4.9) that samples not only the marginal distributions but the joint distribution of the random vector $(\Delta_m W, \Delta_m\mathbb{W})_{m=1}^M$.

1.6 Outline

In Sect. 2 we formulate the functional analytic setup, construct the multiplicative forcing term G and recall known regularity results. Section 3 introduces the discrete setup and contains the main results Theorem 19 and Theorem 25. Next, Sect. 4 clarifies the construction of the discrete random input vectors and provides a sampling algorithm. Lastly, Sect. 5 contains numerical experiments.

2 Mathematical setup

This section contains classical definitions and preliminary results. It is structured as follows: Sect. 2.1 introduces the function analytical framework. Section 2.2 presents the construction of the stochastic forcing. Section 2.3 is about the nonlinear operators S and V. The solution concept is fixed in Sect. 2.4. Lastly, Sect. 2.5 collects regularity results.

Let $\mathcal{O}\subset\mathbb{R}^n$ for $n\geq 1$ be a bounded Lipschitz domain (further assumptions on $\mathcal{O}$ will be needed for the regularity of solutions). For some given $T>0$ we denote by $I:=[0,T]$ the time interval and write $\mathcal{O}_T := I\times\mathcal{O}$ for the time space cylinder. Moreover let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in I},\mathbb{P})$ denote a stochastic basis, i.e. a probability space with a complete and right continuous filtration $(\mathcal{F}_t)_{t\in I}$. We write $f\lesssim g$ for two non-negative quantities f and g if f is bounded by g up to a multiplicative constant. Accordingly we define $\gtrsim$ and $\eqsim$. We denote by c a generic constant which can change its value from line to line.

2.1 Function spaces

As usual, $L^q(\mathcal{O})$ denotes the Lebesgue space and $W^{1,q}(\mathcal{O})$ the Sobolev space, where $1\leq q<\infty$. We denote by $W_0^{1,q}(\mathcal{O})$ the Sobolev spaces with zero boundary values. It is the closure of $C_0^\infty(\mathcal{O})$ (smooth functions with compact support) in the $W^{1,q}(\mathcal{O})$-norm. We denote by $W^{-1,q'}(\mathcal{O})$ the dual of $W_0^{1,q}(\mathcal{O})$. We do not distinguish in the notation between vector- and matrix-valued functions.

For a Banach space $(X, \|\cdot\|_X)$ let $L^q(I; X)$ be the Bochner space of Bochner-measurable functions $u : I \to X$ satisfying $t \mapsto \|u(t)\|_X \in L^q(I)$. Moreover, $C(I; X)$ is the space of continuous functions with respect to the norm-topology. We also use $C^\alpha(I; X)$ for the space of Hölder continuous functions. Given an Orlicz-function $\Phi : [0, \infty] \to [0, \infty]$, i.e. a convex function satisfying $\lim_{t\to 0} \Phi(t)/t = 0$ and $\lim_{t\to\infty} \Phi(t)/t = \infty$ we define the Luxemburg-norm

$$\|u\|_{L^\Phi(I;X)} := \inf \left\{ \lambda > 0 : \int_I \Phi\left(\frac{\|u\|_X}{\lambda} \right) \mathrm{d}s \le 1 \right\}.$$

The Orlicz space $L^\Phi(I; X)$ is the space of all Bochner-measurable functions with finite Luxemburg-norm. For more details on Orlicz-spaces we refer to [32]. Given $h \in I$ and $u : I \to X$ we define the difference operator $\tau_h : \{u : I \to X\} \to \{u : I \cap I - \{h\} \to X\}$ via $\tau_h(u)(s) := u(s + h) - u(s)$. The Besov-Orlicz space $B^\alpha_{\Phi,r}(I; X)$ with differentiability $\alpha \in (0, 1)$, integrability Φ and fine index $r \in (1, \infty]$ is defined as the space of Bochner-measurable functions with finite Besov-Orlicz norm $\|\cdot\|_{B^\alpha_{\Phi,r}(I;X)}$, where

$$\begin{aligned}
\|u\|_{B^\alpha_{\Phi,r}(I;X)} &:= \|u\|_{L^\Phi(I;X)} + [u]_{B^\alpha_{\Phi,r}(I;X)},\\
[u]_{B^\alpha_{\Phi,r}(I;X)} &:= \begin{cases} \left(\int_I h^{-r\alpha} \|\tau_h u\|^r_{L^\Phi(I\cap I-\{h\};X)} \frac{\mathrm{d}h}{h} \right)^{\frac{1}{r}} & \text{if } r \in [1, \infty),\\ \operatorname{ess\,sup}_{h\in I} h^{-\alpha} \|\tau_h u\|_{L^\Phi(I\cap I-\{h\};X)} & \text{if } r = \infty. \end{cases}
\end{aligned}$$

The case $r = \infty$ is commonly called Nikolskii-Orlicz space. When $\Phi(t) = t^p$ we write $B^\alpha_{p,r}(I; X)$ and call the space Besov space.

Similarly, given a Banach space $(Y, \|\cdot\|_Y)$, we define $L^q(\Omega; Y)$ as the Bochner space of Bochner-measurable functions $u : \Omega \to Y$ satisfying $\omega \mapsto \|u(\omega)\|_Y \in L^q(\Omega)$. The space $L^q_{\mathcal{F}}(\Omega \times I; X)$ denotes the subspace of X-valued $(\mathcal{F}_t)_{t\in I}$-progressively measurable processes. Let $(U, \|\cdot\|_U)$ be a separable Hilbert space. $L_2(U; L^2(\mathcal{O}))$ denotes the space of Hilbert-Schmidt operators from U to $L^2(\mathcal{O})$ with the norm $\|z\|^2_{L_2(U;L^2_x)} := \sum_{j\in\mathbb{N}} \left\|z(u_j)\right\|^2_{L^2(\mathcal{O})}$ where $\{u_j\}_{j\in\mathbb{N}}$ is some orthonormal basis of U. We abbreviate the notation $L^q_\omega L^q_t L^q_x := L^q(\Omega; L^q(I; L^q(\mathcal{O})))$ with obvious modifications for Sobolev, Besov and Hölder spaces. Additionally, we write $L^{q-} := \bigcap_{r<q} L^r$.

2.2 Stochastic integrals

In order to construct the stochastic forcing term, we impose the following conditions:

Assumption 1 (a) Let $(U, \|\cdot\|_U)$ be a separable Hilbert space. We assume that W is an U-valued cylindrical Wiener process with respect to $(\mathcal{F}_t)_{t\in I}$. Formally W can be represented as

$$W = \sum_{j\in\mathbb{N}} u_j \beta^j, \tag{2.1}$$

where $\{u_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of U and $\{\beta^j\}_{j\in\mathbb{N}}$ are independent 1-dimensional standard Brownian motions.

(b) Let $v \in L^2_{\mathcal{F}}(\Omega\times I; L^2_x)$. We assume that $G(v)(\cdot) : U \to L^2_{\mathcal{F}}(\Omega\times I; L^2_x)$ is given by

$$u \mapsto G(v)(u) := \sum_{j\in\mathbb{N}} g_j(\cdot, v)(u_j, u)_U,$$

where $\{g_j\}_{j\in\mathbb{N}} \in C(\mathcal{O}\times\mathbb{R}^N; \mathbb{R}^N)$ with

(i) (sublinear growth) for all $x \in \mathcal{O}$ and $\xi \in \mathbb{R}^N$ it holds

$$\sum_{j\in\mathbb{N}} \left|g_j(x,\xi)\right|^2 \le c_{\text{growth}}(1+|\xi|^2), \tag{2.2}$$

(ii) (Lipschitz continuity) for all $x \in \mathcal{O}$ and $\xi_1, \xi_2 \in \mathbb{R}^N$ it holds

$$\sum_{j\in\mathbb{N}} \left|g_j(x,\xi_1) - g_j(x,\xi_2)\right|^2 \le c_{\text{lip}} \left|\xi_1 - \xi_2\right|^2. \tag{2.3}$$

Now it is a classical result, see for example the book of Prévôt and Röckner [66], that we can construct a corresponding stochastic integral.

Proposition 2 *Let Assumption 1 be true. Then the operator $\mathcal{I}$ defined through*

$$\mathcal{I}(G(v)) := \int_0^\bullet G(v)(\,\mathrm{d}W_s) := \sum_{j\in\mathbb{N}} \int_0^\bullet g_j(\cdot, v)\,\mathrm{d}\beta_s^j \tag{2.4}$$

defines a bounded linear operator from $L^2_{\mathcal{F}}(\Omega\times I; L_2(U; L^2_x))$ to $L^2_\omega C_t L^2_x$. Moreover,

(a) *$\mathcal{I}(G(v))$ is an L^2_x-valued martingale with respect to $(\mathcal{F}_t)_{t\in I}$,*

(b) *(Itô isometry) for all $t \in I$ it holds*

$$\mathbb{E}\left[\|\mathcal{I}(G(v))(t)\|^2_{L^2_x}\right] = \mathbb{E}\left[\int_0^t \|G(v)\|^2_{L_2(U;L^2_x)}\,\mathrm{d}s\right]. \tag{2.5}$$

2.3 Perturbed gradient flow

Let $\kappa \ge 0$ and $p \in (1,\infty)$. For $\xi \in \mathbb{R}^{N\times n}$ we define

$$S(\xi) := \varphi'(|\xi|)\frac{\xi}{|\xi|} = (\kappa + |\xi|)^{p-2}\,\xi \tag{2.6}$$

and

$$V(\xi) := \sqrt{\varphi'(|\xi|)\frac{\xi}{|\xi|}} = (\kappa + |\xi|)^{\frac{p-2}{2}}\,\xi, \tag{2.7}$$

where $\varphi(t) := \int_0^t (\kappa + s)^{p-2} s \,\mathrm{d}s$. The nonlinear functions S and V are closely related. In particular the following lemmata are of great importance. The proofs can be found in e.g. [31]. For more details we refer to [10, 27, 29].

Lemma 3 (V-coercivity) *Let $\xi_1, \xi_2 \in \mathbb{R}^{N\times n}$. Then it holds*

$$\begin{aligned}(S(\xi_1) - S(\xi_2)) : (\xi_1 - \xi_2) &\eqsim |V(\xi_1) - V(\xi_2)|^2 \\ &\eqsim (\kappa + |\xi_1| + |\xi_1 - \xi_2|)^{p-2} |\xi_1 - \xi_2|^2 .\end{aligned} \tag{2.8}$$

Lemma 4 (generalized Young's inequality) *Let $\xi_1, \xi_2, \xi_3 \in \mathbb{R}^{N\times n}$ and $\delta > 0$. Then there exists $c_\delta \geq 1$ such that*

$$(S(\xi_1) - S(\xi_2)) : (\xi_2 - \xi_3) \leq \delta \,|V(\xi_1) - V(\xi_2)|^2 + c_\delta \,|V(\xi_2) - V(\xi_3)|^2 . \tag{2.9}$$

Lemma 5 *Let $\xi_1, \xi_2, \xi_3 \in \mathbb{R}^{N\times n}$ and $\delta > 0$. Then there exists $c_\delta \geq 1$ such that*

$$(S(\xi_1) - S(\xi_2)) : \xi_3 \leq \delta \,|V(\xi_1) - V(\xi_2)|^2 + c_\delta \,(\kappa + |\xi_1| + |\xi_1 - \xi_2|)^{p-2} |\xi_3|^2 . \tag{2.10}$$

Remark 6 A continuous, convex and strictly increasing function $\varphi : [0, \infty) \to [0, \infty)$ satisfying

$$\lim_{t\to 0} \frac{\varphi(t)}{t} = \lim_{t\to\infty} \frac{t}{\varphi(t)} = 0$$

is called an N-function. It is called uniformly convex if $\varphi \in C^1[0, \infty) \cap C^2(0, \infty)$ and $\varphi'(t) \eqsim \varphi''(t)t$ uniformly in $t > 0$. Lemmas 3 and 4 are still valid if one replaces φ in (2.6) and (2.7) by any uniformly convex N-function. For more details we refer to e.g. [27].

Given some initial condition $u_0 : \Omega \times \mathcal{O} \to \mathbb{R}^N$ and a stochastic force (G, W) in the sense of Assumption 1 we are interested in the system

$$\mathrm{d}u - \operatorname{div} S(\nabla u)\,\mathrm{d}t = G(u)\,\mathrm{d}W \quad \text{in } \Omega \times \mathcal{O}_T, \tag{2.11a}$$

with boundary and initial conditions given by

$$u = 0 \quad \text{on } \Omega \times I \times \partial\mathcal{O}, \tag{2.11b}$$

$$u(0) = u_0 \quad \text{on } \Omega \times \mathcal{O}. \tag{2.11c}$$

The system (2.11a) is a perturbed version of the gradient flow of the energy $\mathcal{J} : W^{1,p}_{0,x} \to [0, \infty)$ given by

$$\mathcal{J}(u) := \int_{\mathcal{O}} \varphi(|\nabla u|)\,\mathrm{d}x. \tag{2.12}$$

2.4 Weak and strong solutions

We fix the concept of solutions as follows.

Definition 7 Let $u_0 \in L^2_\omega L^2_x$ be $\mathcal{F}_0$-measurable, $p \in (1, \infty)$ and (G, W) be given by Assumption 1. An $(\mathcal{F}_t)$-adapted process $u \in L^2_x$ is called weak solution to (2.11) if

(a) $u \in L^2_\omega C_t L^2_x \cap L^p_\omega L^p_t W^{1,p}_{0,x}$,
(b) for all $t \in I$, $\xi \in C^\infty_{0,x}$ and $\mathbb{P} - a.s.$ it holds

$$\int_{\mathcal{O}} (u(t) - u_0) \cdot \xi \,\mathrm{d}x + \int_0^t \int_{\mathcal{O}} S(\nabla u) : \nabla \xi \,\mathrm{d}x \,\mathrm{d}s = \int_{\mathcal{O}} \int_0^t G(u) \,\mathrm{d}W_s \cdot \xi \,\mathrm{d}x. \tag{2.13}$$

The process u is called strong solution if it is a weak solution and additionally satisfies

(a) $\operatorname{div} S(\nabla u) \in L^2_\omega L^2_t L^2_x$,
(b) for all $t \in I$ and $\mathbb{P} - a.s.$ it holds

$$u(t) - u_0 - \int_0^t \operatorname{div} S(\nabla u) \,\mathrm{d}s = \int_0^t G(u) \,\mathrm{d}W_s \tag{2.14}$$

as an equation in L^2_x.

The next step is to answer the question about existence of weak or even strong solutions. Weak solutions can be constructed through a variational approach that uses the monotonicity of the nonlinear diffusion operator S as presented in [61, Example 4.1.9]. They rely on the Gelfand triple $W^{1,p}_{0,x} \hookrightarrow L^2_x \hookrightarrow W^{-1,p'}_x$ for $p \geq 2$. A quick observation shows that the Gelfand triple remains valid for $p \geq 2n/(n+2)$. Below the threshold we need to modify the energy space. A good choice is the triple $W^{1,p}_{0,x} \cap L^2_x \hookrightarrow L^2_x \hookrightarrow \left(W^{1,p}_{0,x} \cap L^2_x\right)'$. Now, similar arguments as done in [61, Section 4] lead to the existence of a unique weak solution. Proving existence of a strong solution is more delicate and usually requires, not only assumptions on the growth of G, but also on the gradient of G, e.g. for all $x \in \mathcal{O}$, $\xi \in \mathbb{R}^N$

$$\sum_{j \in \mathbb{N}} \left| \nabla_x g_j(x, \xi) \right|^2 \leq c(1 + |\xi|)^2, \tag{2.15a}$$

$$\sum_{j \in \mathbb{N}} \left| \nabla_\xi g_j(x, \xi) \right|^2 \leq c. \tag{2.15b}$$

In [43] a general approach for the construction of strong solutions to gradient flow like equations is presented. In particular, it includes the case of the p-Laplace equations.

Theorem 8 ([43], Theorem 4.12) *Assume $p \geq 2$ and $\mathcal{O}$ is a bounded convex domain. Let (G, W) be given by Assumption 1. Additionally, assume (2.15) and*

$u_0 \in L^{p+\varepsilon}_\omega W^{1,p}_x \cap L^{2+\varepsilon}_\omega L^2_x$ be $\mathcal{F}_0$-measurable for some $\varepsilon > 0$. Then there exists a unique strong solution u to (2.11). Moreover,

$$\mathbb{E}\left[\sup_{t\in I}\|u\|^p_{W^{1,p}_x} + \int_0^T \|\operatorname{div} S(\nabla u)\|^2_{L^2_x}\,\mathrm{d}t\right] \lesssim \mathbb{E}\left[\|u_0\|^p_{W^{1,p}_x}\right] + 1. \tag{2.16}$$

2.5 Regularity of strong solutions

A key ingredient in the error analysis of numerical algorithms are the improved regularity properties of strong solutions in comparison to those of weak solutions. Concerning time regularity, we prove in [76] that the strong solution enjoys $1/2$-time differentiability in an exponential Besov space and even the nonlinear greadient $V(\nabla u)$ obeys $1/2$-time differentiability in a Nikolskii space. The proof uses an assumption on the boundary condition of the noise coefficient G, i.e. for all $x \in \partial\mathcal{O}$,

$$\sum_{j\in\mathbb{N}} \left|g_j(x,0)\right|^2 = 0. \tag{2.17}$$

Theorem 9 ([76] Theorem 3.8 & Theorem 3.11) *Let the assumptions of Theorem 8 be satisfied. Additionally, assume (2.17). Let u be the strong solution to (2.11). Then*

$$u \in L^2_\omega B^{1/2}_{\Phi_2,\infty} L^2_x, \tag{2.18a}$$

$$V(\nabla u) \in L^2_\omega B^{1/2}_{2,\infty} L^2_x, \tag{2.18b}$$

where $\Phi_2(t) = e^{t^2} - 1$.

Spatial regularity is closely connected to the existence of strong solutions. Local regularity has been obtained in [15].

Theorem 10 ([15], Theorem 4) *Let $u_0 \in L^2_\omega W^{1,2}_x$ be $\mathcal{F}_0$-measurable. Let Assumption 1 be satisfied. Additionally, assume (2.15) and denote by u the strong solution of (2.11). Then,*

$$u \in L^2_\omega L^\infty_t W^{1,2}_{x,loc}, \tag{2.19a}$$

$$V(\nabla u) \in L^2_\omega L^2_t W^{1,2}_{x,loc}. \tag{2.19b}$$

On sufficiently regular domains, it is possible to relate the divergence of the nonlinear operator S to the full gradient as presented in [2, 21], i.e.

$$\|\operatorname{div} S(\nabla u)\|_{L^2_\omega L^2_t L^2_x} \eqsim \|\nabla S(\nabla u)\|_{L^2_\omega L^2_t L^2_x}. \tag{2.20}$$

Precisely, given some bounded Lipschitz domain $\mathcal{O}$ such that $\partial\mathcal{O} \in W^{2,1}$, i.e. $\mathcal{O}$ is locally the subgraph of a Lipschitz continuous function of $n-1$ variables, which is also twice weakly differentiable. Denote by $\mathcal{B}$ the second fundamental form on $\partial\mathcal{O}$, by $|\mathcal{B}|$ its norm and define

$$\mathcal{K}_{\mathcal{O}}(r) := \sup_{E\subset\partial\mathcal{O}\cap B_r(x), x\in\partial\mathcal{O}} \frac{\int_E |\mathcal{B}| \,\mathrm{d}\mathcal{H}^{n-1}}{\mathrm{cap}_{B_1(x)}(E)}, \tag{2.21}$$

where $B_r(x)$ denotes the ball of radius r around x, $\mathrm{cap}_{B_1(x)}(E)$ is the capacity of the set E relative to the ball $B_1(x)$ and $\mathcal{H}^{n-1}$ is the $n-1$ dimensional Hausdorff measure. The following result follows from [21, Theorem 2.1].

Lemma 11 *Assume that $\mathcal{O}$ is either*

(a) *bounded and convex,*
(b) *or bounded, Lipschitz and $\partial\mathcal{O} \in W^{2,1}$ with $\lim_{r\to 0}\mathcal{K}_{\mathcal{O}}(r) \le c$.*

Let $v \in L^p_\omega L^p_t W^{1,p}_{0,x}$ with $\operatorname{div} S(\nabla v) \in L^2_\omega L^2_t L^2_x$. Then $\nabla S(\nabla v) \in L^2_\omega L^2_t L^2_x$ and

$$\mathbb{E}\left[\|\nabla S(\nabla v)\|^2_{L^2_t L^2_x}\right] \eqsim \mathbb{E}\left[\|\operatorname{div} S(\nabla u)\|^2_{L^2_t L^2_x}\right].$$

In the non-degenerate setting, $\kappa > 0$, this allows to deduce global spatial regularity.

Corollary 12 ([76], Corollary 2.14) *Let the assumptions of Theorem 8 be satisfied and $\kappa > 0$. Let u be the strong solution to (2.11). Then,*

$$V(\nabla u) \in L^2_\omega L^2_t W^{1,2}_x. \tag{2.22}$$

3 Numerical scheme for the averaged system

In this section we will first present the discrete structures. Afterwards we construct two algorithms that approximate the solution to (2.11). Finally, we prove convergence of the approximation towards the analytic solution with optimal rates.

3.1 Space discretization

Let $\mathcal{O} \subset \mathbb{R}^n$ be a bounded polyhedral domain. By $\mathcal{T}_h$ denote a regular partition (triangulation) of $\mathcal{O}$ (no hanging nodes), which consists of closed n-simplices called *elements*. For each element (n-simplex) $T \in \mathcal{T}_h$ we denote by h_T the diameter of T, and by ρ_T the supremum of the diameters of inscribed balls.

We assume that $\mathcal{T}_h$ is *shape regular*, that is there exists a constant γ (the shape regularity constant) such that

$$\max_{T\in\mathcal{T}_h} \frac{h_T}{\rho_T} \le \gamma. \tag{3.1}$$

We define the maximal mesh-size by

$$h := \max_{T \in \mathcal{T}_h} h_T.$$

We assume further that our triangulation is *quasi-uniform*, i.e.

$$h_T \eqsim h \qquad \text{for all } T \in \mathcal{T}_h. \tag{3.2}$$

For $\ell \in \mathbb{N}_0$ we denote by $\mathscr{P}_\ell(\mathcal{O})$ the polynomials on $\mathcal{O}$ of degree less than or equal to ℓ.

For fixed $r \in \mathbb{N}$ we define the vector valued finite element space V_h as

$$V_h := \{v \in W^{1,1}_{0,x} : v|_T \in \mathscr{P}_r(T) \ \forall T \in \mathcal{T}_h\}. \tag{3.3}$$

Moreover, let $\Pi_2 : L^2_x \to V_h$ be the L^2_x-orthogonal projection defined by

$$\forall \xi_h \in V_h : \quad (\Pi_2 v, \xi_h) = (v, \xi_h)$$

or equivalently

$$\Pi_2 v = \arg\min_{v_h \in V_h} \|v - v_h\|_{L^2_x}.$$

We will need some classical results on the stability properties of the L^2_x-projection for finite elements.

Lemma 13 *Let $r \geq 1$, V_h be given by (3.3) and $\mathcal{T}_h$ be quasi-uniform. Then*

$$\forall v \in W^{1,2}_x : \quad \|v - \Pi_2 v\|_{L^2_x} + h \|\nabla(v - \Pi_2 v)\|_{L^2_x} \lesssim h \|\nabla v\|_{L^2_x},$$
$$\forall v \in W^{2,2}_x : \quad \|v - \Pi_2 v\|_{L^2_x} + h \|\nabla(v - \Pi_2 v)\|_{L^2_x} \lesssim h^2 \left\|\nabla^2 v\right\|_{L^2_x}.$$

Due to the nonlinear structure of the p-Laplace system we also need an adapted stability result.

Proposition 14 ([16], Theorem 7) *Let $r \geq 1$, V_h be given by (3.3) and $\mathcal{T}_h$ be quasi-uniform. Then*

$$\|V(\nabla v) - V(\nabla \Pi_2 v)\|_{L^2_x} \lesssim h \|\nabla V(\nabla v)\|_{L^2_x}.$$

3.2 Time discretization

Let $\{0 = t_0 < \cdots < t_M = T\}$ be a uniform partition of $[0, T]$ with mesh size $\tau = T/M$. For $m \geq 1$ define $I_m := [t_{m-1}, t_m]$. By $|I_m|$ we denote the Lebesgue measure of I_m. By $\fint_{I_m} g \,\mathrm{d}s$ we denote the mean value integral over the set I_m. We also abbreviate $\langle g \rangle_m = \fint_{I_m} g \,\mathrm{d}s$ for the mean value.

Let us discuss the stability properties of piecewise constant approximations generated by the mean values in terms of Nikolskii spaces.

Lemma 15 *Let $\alpha \in (0,1)$ and $r \in [1,\infty)$. Additionally, assume $u \in B^{\alpha}_{r,\infty} L^2_x$. Then*

$$\left(\sum_{m=1}^{M} \int_{I_m} \|u - \langle u \rangle_m\|_{L^2_x}^r \,\mathrm{d}s \right)^{\frac{1}{r}} \lesssim \tau^{\alpha} [u]_{B^{\alpha}_{r,\infty} L^2_x}. \tag{3.4}$$

Proof Due to Jensen's inequality

$$\left(\sum_{m=1}^{M} \int_{I_m} \|u - \langle u \rangle_m\|_{L^2_x}^r \,\mathrm{d}s \right)^{\frac{1}{r}} \leq \left(\sum_{m=1}^{M} \int_{I_m} \fint_{I_m} \|u(s) - u(t)\|_{L^2_x}^r \,\mathrm{d}t \,\mathrm{d}s \right)^{\frac{1}{r}}.$$

A change of variables $t = s + h$ and Fubini's theorem yield

$$\begin{aligned}
&\sum_{m=1}^{M} \tau \fint_{I_m} \fint_{I_m} \|u(s) - u(t)\|_{L^2_x}^r \,\mathrm{d}t \,\mathrm{d}s \\
&= \sum_{m=1}^{M} \tau^{-1} \int_{I_m} \int_{I_m - \{s\}} \|u(s) - u(s+h)\|_{L^2_x}^r \,\mathrm{d}h \,\mathrm{d}s \\
&\leq \tau^{-1} \int_0^{2\tau} \int_{I \cap I - \{h\}} \|u(s) - u(s+h)\|_{L^2_x}^r \,\mathrm{d}s \,\mathrm{d}h \\
&\leq \tau^{-1} \int_0^{2\tau} h^{\alpha r} \,\mathrm{d}h [u]^r_{B^{\alpha}_{r,\infty} L^2_x} = \frac{2^{\alpha r + 1}}{\alpha r + 1} \tau^{\alpha r} [u]^r_{B^{\alpha}_{r,\infty} L^2_x}.
\end{aligned}$$

The proof is finished. □

Remark 16 Lemma 15 is also valid for $r = \infty$. Here we need to substitute the left hand side in (3.4) by

$$\max_{m \in \{1,\dots,M\}} \sup_{s \in I_m} \|u(s) - \langle u \rangle_m\|_{L^2_x}.$$

Note that $B^{\alpha}_{\infty,\infty}(I) = C^{\alpha}(I)$ is the space of α-Hölder continuous functions.

In our application the process u only belongs to $L^2_\omega B^{1/2}_{\Phi_2,\infty} L^2_x \setminus L^2_\omega C^{1/2}_t L^2_x$. Therefore we need to adjust the stability result in terms of exponentially integrable Nikolskii spaces.

Lemma 17 *Let $\alpha \in (0,1)$ and $u \in B^{\alpha}_{\Phi_2,\infty} L^2_x$ where $\Phi_2(t) = e^{t^2} - 1$. Then*

$$\max_{m \in \{1,\dots,M\}} \|u(t_m) - \langle u \rangle_m\|_{L^2_x} \lesssim \tau^{\alpha} \sqrt{\ln(1 + \tau^{-1})} [u]_{B^{\alpha}_{\Phi_2,\infty} L^2_x}. \tag{3.5}$$

Proof Define $I_m^k := [t_m - 2^{-k}\tau, t_m]$. Due to the embedding $B^{\alpha}_{\Phi_2,\infty} \hookrightarrow C_t$ for all $\alpha > 0$ we find u to be continuous as an L^2_x-valued process. Therefore, it holds

$$\lim_{k\to\infty} \langle u \rangle_{I_m^k} = u(t_m).$$

Observe, $\left|I_m^k\right| = 2^{-k}\tau$ and $I_m^{k-1} = [t_m - 2^{-(k-1)}\tau, t_m - 2^{-k}\tau] \cup [t_m - 2^{-k}\tau, t_m]$. A shift in the integral results in

$$\begin{aligned} u(t_m) - \langle u \rangle_m &= \sum_{k=1}^{\infty} \langle u \rangle_{I_m^k} - \langle u \rangle_{I_m^{k-1}} \\ &= \sum_{k=1}^{\infty} 2^{-1} \fint_{I_m^k} u(s) - u(s - 2^{-k}\tau)\, \mathrm{d}s. \end{aligned}$$

Jensen's inequality implies

$$\begin{aligned} &\max_{m\in\{1,\dots,M\}} \|\langle u \rangle_m - u(t_m)\|_{L^2_x} \\ &\leq \max_{m\in\{1,\dots,M\}} \sum_{k=1}^{\infty} 2^{-1} \fint_{I_m^k} \left\| u(s) - u(s - 2^{-k}\tau) \right\|_{L^2_x} \mathrm{d}s \\ &\leq \sum_{k=1}^{\infty} 2^{-1} \max_{m\in\{1,\dots,M\}} \fint_{I_m^k} \left\| u(s) - u(s - 2^{-k}\tau) \right\|_{L^2_x} \mathrm{d}s \\ &\leq \sum_{k=1}^{\infty} 2^{-1} \lambda_k \max_{m\in\{1,\dots,M\}} \Phi_2^{-1} \left(\fint_{I_m^k} \Phi_2 \left(\frac{\left\| u(s) - u(s - 2^{-k}\tau) \right\|_{L^2_x}}{\lambda_k} \right) \mathrm{d}s \right). \end{aligned}$$

Choose λ_k by

$$\lambda_k := \inf \left\{ \mu > 0 : \int_{I \cap I - \{2^{-k}\tau\}} \Phi_2 \left(\frac{\left\| u(s) - u(s - 2^{-k}\tau) \right\|_{L^2_x}}{\mu} \right) \mathrm{d}s \leq 1 \right\}.$$

In particular, since $I_m^k \subset I \cap I - \{2^{-k}\tau\}$,

$$\int_{I_m^k} \Phi_2 \left(\frac{\left\| u(s) - u(s - 2^{-k}\tau) \right\|_{L^2_x}}{\lambda_k} \right) \mathrm{d}s \leq 1.$$

Thus,

$$\sum_{k=1}^{\infty} 2^{-1}\lambda_k \max_{m\in\{1,\dots,M\}} \Phi_2^{-1}\left(\fint_{I_m^k} \Phi_2\left(\frac{\left\| u(s) - u(s-2^{-k}\tau)\right\|_{L_x^2}}{\lambda_k} \right) \mathrm{d}s \right)$$
$$\leq \sum_{k=1}^{\infty} 2^{-1}\lambda_k \max_{m\in\{1,\dots,M\}} \Phi_2^{-1}\left(\left| I_m^k \right|^{-1} \right).$$

Since $u \in B^{\alpha}_{\Phi_2,\infty} L_x^2$ it holds

$$\sup_{k\in\mathbb{N}} (2^{-k}\tau)^{-\alpha}\lambda_k \leq [u]_{B^{\alpha}_{\Phi_2,\infty} L_x^2}.$$

It follows

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{-1}\lambda_k \max_{m\in\{1,\dots,M\}} \Phi_2^{-1}\left(\left| I_m^k \right|^{-1} \right) \\
&\leq \sup_{k\in\mathbb{N}} \left((2^{-k}\tau)^{-\alpha}\lambda_k \right) \sum_{k=1}^{\infty} 2^{-1}(2^{-k}\tau)^{\alpha} \max_{m\in\{1,\dots,M\}} \Phi_2^{-1}\left(\left| I_m^k \right|^{-1} \right) \\
&\leq [u]_{B^{\alpha}_{\Phi_2,\infty} L_x^2} \sum_{k=1}^{\infty} 2^{-1}(2^{-k}\tau)^{\alpha} \sqrt{\ln(1+(2^{-k}\tau)^{-1})}.
\end{aligned} \tag{3.6}$$

Note that

$$\ln\left(1+(2^{-k}\tau)^{-1}\right) = \ln(\tau+2^k) + \ln(\tau^{-1}) \leq \ln(1+2^k) + \ln(1+\tau^{-1}).$$

Therefore,

$$\begin{aligned}
&\sum_{k=1}^{\infty} 2^{-1}(2^{-k}\tau)^{\alpha} \sqrt{\ln(1+(2^{-k}\tau)^{-1})} \\
&\leq \tau^{\alpha} \sum_{k=1}^{\infty} 2^{-1}2^{-\alpha k} \left(\sqrt{\ln(1+2^k)} + \sqrt{\ln(1+\tau^{-1})} \right) \\
&\leq \tau^{\alpha} \left(1 + \sqrt{\ln(1+\tau^{-1})} \right) \sum_{k=1}^{\infty} 2^{-1}2^{-\alpha k} \sqrt{\ln(1+2^k)} \\
&\lesssim \tau^{\alpha} \sqrt{\ln(1+\tau^{-1})}.
\end{aligned} \tag{3.7}$$

The assertion follows by using the estimate (3.7) in (3.6). □

3.3 The averaged algorithm

Let u be the strong solution of (2.11), i.e. for all $t \in I$ and $\mathbb{P}$-a.s.

$$u(t) - u_0 - \int_0^t \operatorname{div} S(\nabla u_\nu)\,\mathrm{d}\nu - \int_0^t G(u_\nu)\,\mathrm{d}W_\nu = 0. \tag{3.8}$$

Take the average over I_1 in (3.8) to get

$$\langle u\rangle_1 - u_0 - \fint_{I_1} \left\{ \int_0^t \operatorname{div} S(\nabla u_\nu)\,\mathrm{d}\nu + \int_0^t G(u_\nu)\,\mathrm{d}W_\nu \right\} \mathrm{d}t = 0. \tag{3.9}$$

To obtain the general evolution of the time averaged values of the solution we subtract (3.8) for t and $t-\tau$ and take the average over I_m

$$\langle u\rangle_m - \langle u\rangle_{m-1} - \fint_{I_m} \left\{ \int_{t-\tau}^t \operatorname{div} S(\nabla u_\nu)\,\mathrm{d}\nu + \int_{t-\tau}^t G(u_\nu)\,\mathrm{d}W_\nu \right\} \mathrm{d}t = 0. \tag{3.10}$$

We denote by $(\cdot,\cdot)$ the L^2_x-inner product. The corresponding weak formulation reads for all $\xi \in L^2_x \cap W^{1,p}_{0,x}$

$$(\langle u\rangle_m - \langle u\rangle_{m-1}, \xi) + \fint_{I_m} \int_{t-\tau}^t (S(\nabla u_\nu), \nabla\xi)\,\mathrm{d}\nu\,\mathrm{d}t = \left(\fint_{I_m} \int_{t-\tau}^t G(u_\nu)\,\mathrm{d}W_\nu\,\mathrm{d}t, \xi \right). \tag{3.11}$$

Due to the (stochastic) Fubini's Theorem we can equivalently write

$$(\langle u\rangle_1 - u_0, \xi) + \int_{\mathbb{R}} a_0(\nu)\,(S(\nabla u_\nu), \nabla\xi)\,\mathrm{d}\nu = \left(\int_{\mathbb{R}} a_0(\nu) G(u_\nu)\,\mathrm{d}W_\nu, \xi \right), \tag{3.12a}$$

$$(\langle u\rangle_m - \langle u\rangle_{m-1}, \xi) + \int_{\mathbb{R}} a_{m-1}(\nu)\,(S(\nabla u_\nu), \nabla\xi)\,\mathrm{d}\nu = \left(\int_{\mathbb{R}} a_{m-1}(\nu) G(u_\nu)\,\mathrm{d}W_\nu, \xi \right), \tag{3.12b}$$

where the weights are given by

$$a_0(\nu) := \frac{\nu}{\tau} 1_{I_1}(\nu), \tag{3.13a}$$

$$a_{m-1}(\nu) := \frac{\nu - t_{m-2}}{\tau} 1_{I_{m-1}}(\nu) + \frac{t_m - \nu}{\tau} 1_{I_m}(\nu). \tag{3.13b}$$

The above considerations motivate the construction of the following numerical scheme:

We initialize the algorithm as

$$v_0 := \Pi_2 u_0. \tag{3.14a}$$

In order to accurately reflect the special character of the first step (3.9), we define v_1 via

$$(v_1 - v_0, \xi_h) + \frac{\tau}{2} \left(S(\nabla v_1), \nabla \xi_h\right) = \left(G(v_0)\langle W\rangle_1, \xi_h\right). \tag{3.14b}$$

Moreover the evolution for $m \in \{2, \ldots, M\}$ is determinded via

$$(v_m - v_{m-1}, \xi_h) + \tau \left(S(\nabla v_m), \nabla \xi_h\right) = \left(G(v_{m-2}) \left(\langle W\rangle_m - \langle W\rangle_{m-1}\right), \xi_h\right) \tag{3.14c}$$

for all $\xi_h \in V_h$ and $\mathbb{P}$-a.s.

The need of a special step size in the first step (3.14b) might be undesirable. It can be overcome by performing a full initial step. We propose the following second algorithm.

Initialize

$$w_0 := \Pi_2 u_0. \tag{3.15a}$$

Full initial step

$$(w_1 - w_0, \xi_h) + \tau \left(S(\nabla w_1), \nabla \xi_h\right) = \left(G(w_0)\langle W\rangle_1, \xi_h\right). \tag{3.15b}$$

Time stepping, for $m \geq 2$,

$$(w_m - w_{m-1}, \xi_h) + \tau \left(S(\nabla w_m), \nabla \xi_h\right) = \left(G(w_{m-2})(\langle W\rangle_m - \langle W\rangle_{m-1}), \xi_h\right). \tag{3.15c}$$

The additional difficulties in the error analysis only occur in the estimate of the initial error.

The next theorem ensures that the numerical schemes are well defined. The arguments are rather standard and we only refer to [40] Section 3 for a detailed analysis of a more general system.

Theorem 18 *Let $u_0 \in L^2_\omega L^2_x$ be $\mathcal{F}_0$-measurable and Assumption 1 be satisfied. Then for all $M \in \mathbb{N}$ there exists a unique $\mathbf{v} = (v_0, \ldots, v_M) \in (V_h)^{M+1}$ such that v_m is $\mathcal{F}_{t_m}$-measurable for all $m \in \{0, \ldots, M\}$ and $\mathbf{v}$ solves* (3.14) *respectively* (3.15).

3.4 Error analysis

We are ready to state and prove the main result.

Theorem 19 (Convergence of Algorithm (3.14)) *Let $u_0 \in L^2_\omega L^2_x$ be $\mathcal{F}_0$-measurable and Assumption 1 be satisfied. Additionally, assume $\mathcal{T}_h$ to be quasi-uniform. Moreover, let u be the weak solution to (2.11) and $\mathbf{v}$ be the numerical solution of (3.14).*

(a) *Assume*

$$u \in L^2_\omega B^{1/2}_{2,\infty} L^2_x \cap L^2_\omega L^\infty_t W^{1,2}_x, \tag{3.16a}$$

$$V(\nabla u) \in L^2_\omega B^{1/2}_{2,\infty} L^2_x \cap L^2_\omega L^2_t W^{1,2}_x. \tag{3.16b}$$

Then it holds

$$\mathbb{E}\left[\max_{m=1,\dots,M} \|\langle u\rangle_m - v_m\|^2_{L^2_x} + \sum_{m=1}^{M} \int_{I_m} \|V(\nabla u_\nu) - V(\nabla v_m)\|^2_{L^2_x} \,\mathrm{d}\nu \right] \lesssim \tau + h^2. \tag{3.17}$$

(b) *If additionally*

$$u \in L^2_\omega B^{1/2}_{\Phi_2,\infty} L^2_x, \tag{3.18}$$

then

$$\mathbb{E}\left[\max_{m=1,\dots,M} \|u(t_m) - v_m\|^2_{L^2_x} \right] \lesssim \tau \ln(1+\tau^{-1}) + h^2. \tag{3.19}$$

Remark 20 The regularity assumption (3.18) is optimal in the sense, that it is the limiting space of the time regularity of the Wiener process W. Hytönen and Veraar prove in [50] that a Banach space valued Brownian motion has full 1/2-differentiability only on the Nikolskii-scale. More precisely, they show for all $q, r \in [1,\infty)$ and $\mathbb{P}$-a.s.

$$W \in B^{1/2}_{\Phi_2,\infty} \quad \text{and} \quad W \notin B^{1/2}_{q,r}.$$

The logarithmic term in (3.19) has already been used in the context of rough stochastic differential equations [23, 55]. It quantifies the distance of L^{Φ_2} and L^∞.

Theorem 19 can be generalized to cover fractional time and space regularity assumptions as done for the deterministic p-Laplace system in [16].

Proof **Part (a)**: Fix $m \in \{2,\dots,M\}$, subtract (3.11) from (3.14c) and choose $\xi_h = \Pi_2 e_m := \Pi_2\langle u\rangle_m - v_m$

$$H_1 + H_2 := (e_m - e_{m-1}, \Pi_2 e_m) + \fint_{I_m} \int_{t-\tau}^{t} (S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 e_m) \,\mathrm{d}\nu \,\mathrm{d}t$$

$$= \left(\fint_{I_m} \int_{t-\tau}^{t} [G(u_\nu) - G(v_{m-2})] \, \mathrm{d}W_\nu \, \mathrm{d}t, \, \Pi_2 e_m \right) =: H_3.$$

Now use the symmetry of the L^2_x-projection and the algebraic identity $2a(a-b) = a^2 - b^2 + (a-b)^2$ to get

$$\begin{aligned} H_1 &= (\Pi_2(e_m - e_{m-1}), \Pi_2 e_m) \\ &= \frac{1}{2}\left(\|\Pi_2 e_m\|_{L^2_x}^2 - \|\Pi_2 e_{m-1}\|_{L^2_x}^2 + \|\Pi_2[e_m - e_{m-1}]\|_{L^2_x}^2 \right). \end{aligned}$$

The second term, due to the V-coercivity (2.8),

$$\begin{aligned} H_2 &= \fint_{I_m} \int_{t-\tau}^{t} (S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 \langle u \rangle_m - \nabla v_m) \, \mathrm{d}\nu \, \mathrm{d}t \\ &\eqsim \fint_{I_m} \int_{t-\tau}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L^2_x}^2 \, \mathrm{d}\nu \, \mathrm{d}t \\ &\quad + \fint_{I_m} \int_{t-\tau}^{t} (S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 \langle u \rangle_m - \nabla u_\nu) \, \mathrm{d}\nu \, \mathrm{d}t. \end{aligned}$$

Summation over $m \in \{2, \dots, m^*\}$ with $m^* \leq M$ results in

$$\begin{aligned} &\|\Pi_2 e_{m^*}\|_{L^2_x}^2 + \sum_{m=2}^{m^*} \|\Pi_2[e_m - e_{m-1}]\|_2^2 + \sum_{m=2}^{m^*} \fint_{I_m} \int_{t-\tau}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L^2_x}^2 \, \mathrm{d}\nu \, \mathrm{d}t \\ &\quad \eqsim \|\Pi_2 e_1\|_{L^2_x}^2 + \sum_{m=2}^{m^*} \left(\fint_{I_m} \int_{t-\tau}^{t} [G(u_\nu) - G(v_{m-2})] \, \mathrm{d}W_\nu \, \mathrm{d}t, \, \Pi_2 e_m \right) \\ &\qquad - \sum_{m=2}^{m^*} \fint_{I_m} \int_{t-\tau}^{t} (S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 \langle u \rangle_m - \nabla u_\nu) \, \mathrm{d}\nu \, \mathrm{d}t. \end{aligned}$$

Take the maximum over $m^* \in \{2, \dots, M\}$ and expectation

$$\begin{aligned} &\mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L^2_x}^2 \right] + \mathbb{E}\left[\sum_{m=2}^{M} \left\|\Pi_2[e_m - e_{m-1}]\right\|_{L^2_x}^2 \right] \\ &\quad + \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \int_{t-\tau}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L^2_x}^2 \, \mathrm{d}\nu \, \mathrm{d}t \right] \end{aligned}$$

$$
\begin{aligned}
&\lesssim \mathbb{E}\left[\|\Pi_2 e_1\|_{L_x^2}^2\right] + \mathbb{E}\left[\max_{m^*\in\{2,\dots,M\}} \sum_{m=2}^{m^*} \left(\fint_{I_m} \int_{t-\tau}^t [G(u_\nu) - G(v_{m-2})] \,\mathrm{d}W_\nu \,\mathrm{d}t, \Pi_2 e_m \right)\right] \\
&\quad + \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \int_{t-\tau}^t |(S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 \langle u\rangle_m - \nabla u_\nu)| \,\mathrm{d}\nu \,\mathrm{d}t\right] \\
&=: J_1 + J_2 + J_3.
\end{aligned} \tag{3.20}
$$

Step 1: We start by estimating the initial error. Similarly as before, we subtract the weak formulation of (3.9) from (3.14b) and choose $\xi_h = \Pi_2 e_1$

$$
\begin{aligned}
(e_1 - e_0, \Pi_2 e_1) + \fint_{I_1} \int_0^t (S(\nabla u_\nu) - S(\nabla v_1), \nabla \Pi_2 \langle u\rangle_1 - \nabla v_1) \,\mathrm{d}\nu \,\mathrm{d}t \\
= \left(\fint_{I_1} \int_0^t G(u_\nu) - G(v_0) \,\mathrm{d}W_\nu \,\mathrm{d}t, \Pi_2 e_1 \right),
\end{aligned}
$$

where $e_0 := u_0 - v_0$. By (3.14a) we have $\Pi_2 e_0 = 0$. Therefore,

$$
\begin{aligned}
&\mathbb{E}\left[\|\Pi_2 e_1\|_{L_x^2}^2\right] + \mathbb{E}\left[\fint_{I_1} \int_0^t \|V(\nabla u_\nu) - V(\nabla v_1)\|_{L_x^2}^2 \,\mathrm{d}\nu \,\mathrm{d}t\right] \\
&\quad \lesssim \mathbb{E}\left[\left(\fint_{I_1} \int_0^t G(u_\nu) - G(v_0) \,\mathrm{d}W_\nu \,\mathrm{d}t, \Pi_2 e_1 \right)\right] \\
&\qquad + \mathbb{E}\left[\fint_{I_1} \int_0^t (S(\nabla u_\nu) - S(\nabla v_1), \nabla u_\nu - \nabla \Pi_2 \langle u\rangle_1) \,\mathrm{d}\nu \,\mathrm{d}t\right] \\
&\quad =: J_{1,a} + J_{1,b}.
\end{aligned} \tag{3.21}
$$

Due to Hölder's and Young's inequalities and Itô isometry

$$
\begin{aligned}
J_{1,a} &\le \frac{1}{2}\mathbb{E}\left[\left\| \fint_{I_1} \int_0^t G(u_\nu) - G(v_0) \,\mathrm{d}W_\nu \,\mathrm{d}t \right\|_{L_x^2}^2\right] + \frac{1}{2}\mathbb{E}\left[\|\Pi_2 e_1\|_{L_x^2}^2\right] \\
&= \frac{1}{2}\mathbb{E}\left[\int_0^\tau a_0(\nu)^2 \|G(u_\nu) - G(v_0)\|_{L_2(U;L_x^2)}^2 \,\mathrm{d}\nu\right] + \frac{1}{2}\mathbb{E}\left[\|\Pi_2 e_1\|_{L_x^2}^2\right].
\end{aligned}
$$

Absorb the second term to the left hand side. For the first term we use $a_0 \le 1$ and the Lipschitz assumption (2.3). Now, since the operator norm of the L_x^2-projection is bounded by one,

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^\tau a_0(\nu)^2 \left\|G(u_\nu)-G(v_0)\right\|_{L_2(U;L_x^2)}^2 \,\mathrm{d}\nu\right] \\
&\quad\lesssim \mathbb{E}\left[\int_0^\tau \|u_\nu - v_0\|_{L_x^2}^2 \,\mathrm{d}\nu\right] \lesssim \tau\left(\mathbb{E}\left[\|u\|_{L_t^\infty L_x^2}^2\right] + \mathbb{E}\left[\|u_0\|_{L_x^2}^2\right]\right). \qquad (3.22)
\end{aligned}
$$

Thanks to the generalized Young's inequality, cf. Lemma 4,

$$
\begin{aligned}
J_{1,b} &\le \delta \mathbb{E}\left[\fint_{I_1}\int_0^t \|V(\nabla u_\nu) - V(\nabla v_1)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right] \\
&\quad + c_\delta \mathbb{E}\left[\fint_{I_1}\fint_{I_1}\int_0^t \|V(\nabla u_\nu) - V(\nabla \Pi_2 u_s)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\,\mathrm{d}s\right].
\end{aligned}
$$

Absorb the first term to the left hand side in (3.21). The second is split up into a space and a time error. Using the nonlinear stability of the L_x^2-projection, cf. Proposition 14,

$$
\begin{aligned}
&\mathbb{E}\left[\fint_{I_1}\fint_{I_1}\int_0^t \|V(\nabla u_\nu) - V(\nabla \Pi_2 u_s)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\,\mathrm{d}s\right] \\
&\quad\lesssim \mathbb{E}\left[\fint_{I_1}\fint_{I_1}\int_0^t \|V(\nabla u_\nu) - V(\nabla u_s)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\,\mathrm{d}s\right] \\
&\qquad + \mathbb{E}\left[\fint_{I_1}\fint_{I_1}\int_0^t \|V(\nabla u_s) - V(\nabla \Pi_2 u_s)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\,\mathrm{d}s\right] \\
&\quad\lesssim \tau \mathbb{E}\left[[V(\nabla u)]_{B_{2,\infty}^{1/2} L_x^2}^2\right] + h^2 \mathbb{E}\left[\|\nabla V(\nabla u)\|_{L_t^2 L_x^2}^2\right].
\end{aligned}
$$

Overall, we arrive at the estimate from (3.21)

$$
\begin{aligned}
&J_1 + (1-\delta)\mathbb{E}\left[\fint_{I_1}\int_0^t \|V(\nabla u_\nu) - V(\nabla v_1)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right] \\
&\quad\lesssim \tau\left(\mathbb{E}\left[\|u\|_{L_t^\infty L_x^2}^2\right] + \mathbb{E}\left[\|u_0\|_{L_x^2}^2\right]\right) \\
&\qquad + c_\delta\left(\tau \mathbb{E}\left[[V(\nabla u)]_{B_{2,\infty}^{1/2} L_x^2}^2\right] + h^2 \mathbb{E}\left[\|\nabla V(\nabla u)\|_{L_t^2 L_x^2}^2\right]\right). \qquad (3.23)
\end{aligned}
$$

Step 2: The stochastic part in (3.20) needs some refined analysis

$$
\begin{aligned}
J_2 &\le \mathbb{E}\left[\max_{m^*\in\{2,\dots,M\}} \sum_{m=2}^{m^*}\left(\fint_{I_m}\int_{t-\tau}^{t}[G(u_\nu)-G(v_{m-2})]\,\mathrm{d}W_\nu\,\mathrm{d}s,\,\Pi_2(e_m-e_{m-2})\right)\right]\\
&\quad+\mathbb{E}\left[\max_{m^*\in\{2,\dots,M\}} \sum_{m=2}^{m^*}\left(\fint_{I_m}\int_{t-\tau}^{t}[G(u_\nu)-G(v_{m-2})]\,\mathrm{d}W_\nu\,\mathrm{d}s,\,\Pi_2 e_{m-2}\right)\right]\\
&=: J_{2,a}+J_{2,b}.
\end{aligned}
$$

The first, due to Hölder's and Young's inequalities and an index shift,

$$
\begin{aligned}
J_{2,a} &\le \mathbb{E}\left[\sum_{m=2}^{M}\left(c_\varepsilon\left\|\fint_{I_m}\int_{t-\tau}^{t}[G(u_\nu)-G(v_{m-2})]\,\mathrm{d}W_\nu\,\mathrm{d}s\right\|_{L^2_x}^2+\varepsilon\left\|\Pi_2(e_m-e_{m-2})\right\|_{L^2_x}^2\right)\right]\\
&\lesssim c_\varepsilon\mathbb{E}\left[\sum_{m=2}^{M}\left\|\fint_{I_m}\int_{t-\tau}^{t}[G(u_\nu)-G(v_{m-2})]\,\mathrm{d}W_\nu\,\mathrm{d}s\right\|_{L^2_x}^2\right]\\
&\quad+\varepsilon\left(\mathbb{E}\left[\sum_{m=2}^{M}\left\|\Pi_2(e_m-e_{m-1})\right\|_{L^2_x}^2\right]+\mathbb{E}\left[\left\|\Pi_2 e_1\right\|_{L^2_x}^2\right]\right).
\end{aligned}
$$

The second term can be absorbed to the left hand side in (3.20). The third term is nothing but the initial error J_1. For the first term we invoke Itô isometry, the Lipschitz condition (2.3) and $a_{m-1}\le 1$

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{m=2}^{M}\left\|\fint_{I_m}\int_{t-\tau}^{t}[G(u_\nu)-G(v_{m-2})]\,\mathrm{d}W_\nu\,\mathrm{d}s\right\|_{L^2_x}^2\right]\\
&=\sum_{m=2}^{M}\int_{t_{m-2}}^{t_m} a_{m-1}^2(\nu)\mathbb{E}\left[\left\|G(u_\nu)-G(v_{m-2})\right\|_{L_2(\mathcal{U};L^2_x)}^2\right]\,\mathrm{d}\nu\\
&\lesssim \sum_{m=2}^{M}\int_{t_{m-2}}^{t_m}\mathbb{E}\left[\left\|u_\nu-v_{m-2}\right\|_{L^2_x}^2\right]\,\mathrm{d}\nu.
\end{aligned}
$$

Decomposition in time and space error, applying Lemmas 13 and 15 and the estimate (3.22)

$$
\begin{aligned}
&\sum_{m=2}^{M} \int_{t_{m-2}}^{t_m} \mathbb{E}\left[\|u_\nu - v_{m-2}\|_{L_x^2}^2 \right] \mathrm{d}\nu \\
&\lesssim \sum_{m=3}^{M} \int_{t_{m-2}}^{t_m} \mathbb{E}\left[\|u_\nu - \langle u \rangle_{m-2}\|_{L_x^2}^2 \right] \mathrm{d}\nu + \tau \sum_{m=3}^{M} \mathbb{E}\left[\|\langle u \rangle_{m-2} - \Pi_2 \langle u \rangle_{m-2}\|_{L_x^2}^2 \right] \\
&\quad + \tau \sum_{m=3}^{M} \mathbb{E}\left[\|\Pi_2 \langle u \rangle_{m-2} - v_{m-2}\|_{L_x^2}^2 \right] + \int_0^{2\tau} \mathbb{E}\left[\|u_\nu - v_0\|_{L_x^2}^2 \right] \mathrm{d}\nu \\
&\lesssim \tau \mathbb{E}\left[[u]_{B_{2,\infty}^{1/2} L_x^2}^2 \right] + h^2 \mathbb{E}\left[\|\nabla u\|_{L_t^2 L_x^2}^2 \right] + \tau \sum_{m=3}^{M} \mathbb{E}\left[\|\Pi_2 e_{m-2}\|_{L_x^2}^2 \right] \\
&\quad + \tau \left(\mathbb{E}\left[\|u\|_{L_t^\infty L_x^2}^2 \right] + \mathbb{E}\left[\|u_0\|_{L_x^2}^2 \right] \right).
\end{aligned} \tag{3.24}
$$

Next, we analyze the second term $J_{2,b}$ in the upper bound for J_2. Define the discrete real-valued stochastic process

$$
K_{m^*} := \sum_{m=2}^{m^*} \left(\fint_{I_m} \int_{t-\tau}^{t} [G(u_\nu) - G(v_{m-2})] \,\mathrm{d}W_\nu \,\mathrm{d}t, \Pi_2 e_{m-2} \right). \tag{3.25}
$$

It is convenient to use stochastic Fubini's theorem to rewrite

$$
K_{m^*} = \sum_{m=2}^{m^*} \left(\int_{t_{m-2}}^{t_m} a_{m-1}(\nu) [G(u_\nu) - G(v_{m-2})] \,\mathrm{d}W_\nu, \Pi_2 e_{m-2} \right).
$$

In the following we abbreviate $\mathcal{F}_m := \mathcal{F}_{t_m}$. Note, (3.25) does not define a martingale with respect to $\mathcal{F}_{t_m}$. In fact, the discrepancy of not being a martingale can be quantified. The general strategy is to split up the sum into a martingale and an error term. The error term is called compensator. To determine how the compensator looks, we first compute the conditional expectations of K_M with respect to $\mathcal{F}_{m^*}$, i.e.,

$$
\begin{aligned}
\mathbb{E}\left[K_M | \mathcal{F}_{m^*} \right] &= \mathbb{E}\left[\sum_{m=m^*+2}^{M} \left(\int_{t_{m-2}}^{t_m} a_{m-1}(\nu) [G(u_\nu) - G(v_{m-2})] \,\mathrm{d}W_\nu, \Pi_2 e_{m-2} \right) \Big| \mathcal{F}_{m^*} \right] \\
&\quad + \mathbb{E}\left[\left(\left(\int_{t_{m^*-1}}^{t_{m^*+1}} a_{m^*}(\nu) [G(u_\nu) - G(v_{m^*-1})] \,\mathrm{d}W_\nu, \Pi_2 e_{m^*-1} \right) \right] + \mathbb{E}\left[K_{m^*} | \mathcal{F}_{m^*} \right] \\
&=: \mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_3.
\end{aligned}
$$

Due to the tower property of conditional expectation, the measurability of e_m with respect to $\mathcal{F}_m$ together with the martingale property of the stochastic integral

$$\begin{aligned} \mathcal{M}_1 &= \sum_{m=m^*+2}^{M} \mathbb{E}\left[\mathbb{E}\left[\left(\int_{t_{m-2}}^{t_m} a_{m-1}(\nu)[G(u_\nu) - G(v_{m-2})]\,\mathrm{d}W_\nu, \Pi_2 e_{m-2}\right)\Big|\mathcal{F}_{m-2}\right]\Big|\mathcal{F}_{m^*}\right] \\ &= \sum_{m=m^*+2}^{M} \mathbb{E}\left[\left(\mathbb{E}\left[\int_{t_{m-2}}^{t_m} a_{m-1}(\nu)[G(u_\nu) - G(v_{m-2})]\,\mathrm{d}W_\nu\Big|\mathcal{F}_{m-2}\right], \Pi_2 e_{m-2}\right)\Big|\mathcal{F}_{m^*}\right] \\ &= 0. \end{aligned}$$

Again using the $\mathcal{F}_m$-measurability of e_m, we conclude that K_{m^*} is $\mathcal{F}_{m^*}$-measurable. Thus,

$$\mathcal{M}_3 = K_{m^*}.$$

It remains to compute the conditional expectation in $\mathcal{M}_2$. Since t_{m^*} is an interior point of $I_{m^*+1} \cup I_{m^*}$ we split up the stochastic integral into a part that only sees values above the threshold t_{m^*} and into a lower part that only sees values below the threshold t_{m^*}, i.e.,

$$\begin{aligned} \mathcal{M}_2 = \mathbb{E}&\left[\left(\int_{t_{m^*}}^{t_{m^*+1}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})]\,\mathrm{d}W_\nu, \Pi_2 e_{m^*-1}\right)\Big|\mathcal{F}_{m^*}\right] \\ &+ \mathbb{E}\left[\left(\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})]\,\mathrm{d}W_\nu, \Pi_2 e_{m^*-1}\right)\Big|\mathcal{F}_{m^*}\right]. \end{aligned}$$

The first vanishes due to the martingale property of the stochastic integral, while the second is measurable with respect to $\mathcal{F}_{m^*}$. Overall,

$$\mathcal{M}_2 = \left(\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})]\,\mathrm{d}W_\nu, \Pi_2 e_{m^*-1}\right).$$

$\mathcal{M}_2$ is called compensator and quantifies the error of not being a martingale, i.e.,

$$\mathbb{E}\left[K_M\big|\mathcal{F}_{m^*}\right] - K_{m^*} = \mathcal{M}_2. \tag{3.26}$$

Furthermore, increments of the discrete stochastic process K satisfy

$$K_{m^*} - K_{m^*-1} = \left(\int_{t_{m^*-2}}^{t_{m^*}} a_{m^*-1}(\nu)[G(u_\nu) - G(v_{m^*-2})]\,\mathrm{d}W_\nu, \Pi_2 e_{m^*-2}\right). \tag{3.27}$$

(3.26) together with (3.27) allow to identify increments of the conditional expectations,

$$\mathbb{E}\left[K_M\big|\mathcal{F}_{m^*}\right] - \mathbb{E}\left[K_M\big|\mathcal{F}_{m^*-1}\right]$$

$$= \left(\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})] \, \mathrm{d}W_\nu, \Pi_2 e_{m^*-1} \right) + \left(\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*-1}(\nu)[G(u_\nu) - G(v_{m^*-2})] \, \mathrm{d}W_\nu, \Pi_2 e_{m^*-2} \right).$$

Observe $\mathbb{E}\left[K_M \middle| \mathcal{F}_1\right] = 0$, since $\Pi_2 e_0 = 0$. The Burkholder-Davis-Gundy's inequality implies

$$\mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left|\mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right]\right|\right] \lesssim \mathbb{E}\left[\left(\sum_{m=2}^{M} \left(\mathbb{E}\left[K_M \middle| \mathcal{F}_m\right] - \mathbb{E}\left[K_M \middle| \mathcal{F}_{m-1}\right]\right)^2\right)^{\frac{1}{2}}\right]$$
$$\lesssim \mathbb{E}\left[\left(\sum_{m=2}^{M} \left\| \int_{t_{m-1}}^{t_m} a_m(\nu)[G(u_\nu) - G(v_{m-1})] \, \mathrm{d}W_\nu \right\|_{L_x^2}^2 \|\Pi_2 e_{m-1}\|_{L_x^2}^2\right)^{\frac{1}{2}}\right] + \mathbb{E}\left[\left(\sum_{m=2}^{M} \left\| \int_{t_{m-1}}^{t_m} a_{m-1}(\nu)[G(u_\nu) - G(v_{m-2})] \, \mathrm{d}W_\nu \right\|_{L_x^2}^2 \|\Pi_2 e_{m-2}\|_{L_x^2}^2\right)^{\frac{1}{2}}\right].$$

Now Young's inequality, Itô isometry and the Lipschitz condition (2.3) imply

$$\mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right]\right]$$
$$\lesssim \varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L_x^2}^2 + \|\Pi_2 e_1\|_{L_x^2}^2\right] + c_\varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \int_{t_{m-1}}^{t_m} \|G(u_\nu) - G(v_{m-2})\|_{L_2(U;L_x^2)}^2 \, \mathrm{d}\nu\right] + c_\varepsilon \mathbb{E}\left[\sum_{m=1}^{M} \int_{t_{m-1}}^{t_m} \|G(u_\nu) - G(v_{m-1})\|_{L_2(U;L_x^2)}^2 \, \mathrm{d}\nu\right]$$
$$\lesssim \varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L_x^2}^2 + \|\Pi_2 e_1\|_{L_x^2}^2\right] + c_\varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \int_{t_{m-1}}^{t_m} \|u_\nu - v_{m-2}\|_{L_x^2}^2 \, \mathrm{d}\nu + \sum_{m=1}^{M} \int_{t_{m-1}}^{t_m} \|u_\nu - v_{m-1}\|_{L_x^2}^2 \, \mathrm{d}\nu\right].$$

The second term is estimated as in (3.24)

$$\mathbb{E}\left[\sum_{m=2}^{M} \int_{t_{m-1}}^{t_m} \|u_\nu - v_{m-2}\|_{L_x^2}^2 \, \mathrm{d}\nu + \sum_{m=1}^{M} \int_{t_{m-1}}^{t_m} \|u_\nu - v_{m-1}\|_{L_x^2}^2 \, \mathrm{d}\nu\right]$$

$$
\begin{aligned}
&\lesssim \tau \mathbb{E}\left[[u]^2_{B^{1/2}_{2,\infty}L^2_x}\right] + h^2 \mathbb{E}\left[\|\nabla u\|^2_{L^2_t L^2_x}\right] + \tau \sum_{m=3}^{M} \mathbb{E}\left[\|\Pi_2 e_{m-2}\|^2_{L^2_x}\right] \\
&\quad + \tau \left(\mathbb{E}\left[\|u\|^2_{L^\infty_t L^2_x}\right] + \mathbb{E}\left[\|u_0\|^2_{L^2_x}\right]\right).
\end{aligned}
$$

Similarly, due to (3.26), Hölder's inequality, Young's inequality and $\ell^1 \hookrightarrow \ell^\infty$

$$
\begin{aligned}
&\mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left(K_{m^*} - \mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right]\right)\right] \\
&= \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left(\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})] \, \mathrm{d}W_\nu, \Pi_2 e_{m^*-1}\right)\right] \\
&\le \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left\|\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})] \, \mathrm{d}W_\nu\right\|_{L^2_x} \|\Pi_2 e_{m^*-1}\|_{L^2_x}\right] \\
&\le \varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*-1}\|^2_{L^2_x}\right] \\
&\quad + c_\varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left\|\int_{t_{m^*-1}}^{t_{m^*}} a_{m^*}(\nu)[G(u_\nu) - G(v_{m^*-1})] \, \mathrm{d}W_\nu\right\|^2_{L^2_x}\right] \\
&\le \varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*}\|^2_{L^2_x} + \|\Pi_2 e_1\|^2_{L^2_x}\right] \\
&\quad + c_\varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \left\|\int_{t_{m-1}}^{t_m} a_m(\nu)[G(u_\nu) - G(v_{m-1})] \, \mathrm{d}W_\nu\right\|^2_{L^2_x}\right].
\end{aligned}
$$

Together we can estimate

$$
\begin{aligned}
J_{2,b} &= \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} K_{m^*}\right] \\
&= \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left(K_{m^*} - \mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right] + \mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right]\right)\right] \\
&\le \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \left(K_{m^*} - \mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right]\right)\right] + \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \mathbb{E}\left[K_M \middle| \mathcal{F}_{m^*}\right]\right] \\
&\lesssim \varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*}\|^2_{L^2_x} + \|\Pi_2 e_1\|^2_{L^2_x}\right] + c_\varepsilon \mathbb{E}\left[\tau \sum_{m=2}^{M} \|\Pi_2 e_{m-1}\|^2_{L^2_x}\right] \\
&\quad + c_\varepsilon \left(\tau \mathbb{E}\left[[u]^2_{B^{1/2}_{2,\infty}L^2_x}\right] + h^2 \mathbb{E}\left[\|\nabla u\|^2_{L^2_t L^2_x}\right] + \tau \mathbb{E}\left[\|u\|^2_{L^\infty_t L^2_x}\right] + \tau \mathbb{E}\left[\|u_0\|^2_{L^2_x}\right]\right).
\end{aligned}
$$

This concludes the bound for J_2 in (3.20)

$$J_2 \lesssim \varepsilon \mathbb{E}\left[\max_{m^* \in \{2,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L_x^2}^2\right] + \varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \|\Pi_2(e_m - e_{m-1})\|_{L_x^2}^2\right] + \varepsilon J_1$$
$$+ c_\varepsilon \left(\tau \mathbb{E}\left[[u]_{B^{1/2}_{2,\infty} L_x^2}^2\right] + h^2 \mathbb{E}\left[\|\nabla u\|_{L_t^2 L_x^2}^2\right] + \tau \mathbb{E}\left[\|u\|_{L_t^\infty L_x^2}^2\right] + \tau \mathbb{E}\left[\|u_0\|_{L_x^2}^2\right]\right)$$
$$+ c_\varepsilon \mathbb{E}\left[\tau \sum_{m=2}^{M} \|\Pi_2 e_m\|_{L_x^2}^2\right]. \tag{3.28}$$

Step 3: In this step we estimate the nonlinear gradient. Jensen's inequality, the generalized Young's inequality (2.9) and the nonlinear stability result Proposition 14 imply

$$J_3 = \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \int_{t-\tau}^{t} |(S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 \langle u \rangle_m - \nabla u_\nu)| \,\mathrm{d}\nu\,\mathrm{d}t\right]$$
$$\leq \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \fint_{I_m} \int_{t-\tau}^{t} \left|\left(S(\nabla u_\nu) - S(\nabla v_m), \nabla \Pi_2 u_{\nu_2} - \nabla u_\nu\right)\right| \,\mathrm{d}\nu_2\,\mathrm{d}\nu\,\mathrm{d}t\right]$$
$$\leq \varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \int_{t-\tau}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right]$$
$$+ c_\varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \int_{t-\tau}^{t} \|V(\nabla u_\nu) - V(\nabla \Pi_2 u_\nu)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right]$$
$$+ c_\varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \fint_{I_m} \int_{t-\tau}^{t} \left\|V(\nabla u_\nu) - V(\nabla u_{\nu_2})\right\|_{L_x^2}^2 \,\mathrm{d}\nu_2\,\mathrm{d}\nu\,\mathrm{d}t\right]$$
$$\leq \varepsilon \mathbb{E}\left[\sum_{m=2}^{M} \fint_{I_m} \int_{t-\tau}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L_x^2}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right]$$
$$+ h^2 c_\varepsilon \mathbb{E}\left[\|\nabla V(\nabla u)\|_{L_t^2 L_x^2}^2\right] + \tau c_\varepsilon \mathbb{E}\left[[V(\nabla u)]_{B^{1/2}_{2,\infty} L_x^2}^2\right].$$

Step 4: We aim at applying a Gronwall type argument. Collecting all estimates we get

$$
\begin{aligned}
&(1-\varepsilon)\mathbb{E}\left[\max_{m^*\in\{1,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L^2_x}^2\right] + (1-\varepsilon)\mathbb{E}\left[\sum_{m=1}^{M} \|\Pi_2[e_m - e_{m-1}]\|_{L^2_x}^2\right] \\
&\quad + (1-\varepsilon)\mathbb{E}\left[\sum_{m=1}^{M} \fint_{I_m}\int_{t-\tau\vee 0}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_2^2 \,\mathrm{d}\nu\,\mathrm{d}t\right] \\
&\lesssim c_\varepsilon \mathbb{E}\left[\sum_{m=1}^{M} \tau \,\|\Pi_2 e_m\|_{L^2_x}^2\right] + c_\varepsilon\left(\tau\mathbb{E}\left[[V(\nabla u)]^2_{B^{1/2}_{2,\infty}L^2_x}\right] + h^2\mathbb{E}\left[\|\nabla V(\nabla u)\|^2_{L^2_t L^2_x}\right]\right) \\
&\quad + c_\varepsilon\left(\tau\mathbb{E}\left[[u]^2_{B^{1/2}_{2,\infty}L^2_x}\right] + h^2\mathbb{E}\left[\|\nabla u\|^2_{L^2_t L^2_x}\right] + \tau\mathbb{E}\left[\|u\|^2_{L^\infty_t L^2_x}\right] + \tau\mathbb{E}\left[\|u_0\|^2_{L^2_x}\right]\right).
\end{aligned}
$$

Choosing ε sufficiently small and applying Gronwall's Lemma ensures

$$
\mathbb{E}\left[\max_{m^*\in\{1,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L^2_x}^2\right] \lesssim e^{cT}\left(\tau + h^2\right).
$$

This implies

$$
\mathbb{E}\left[\max_{m^*\in\{1,\dots,M\}} \|\Pi_2 e_{m^*}\|_{L^2_x}^2 + \sum_{m=1}^{M} \fint_{I_m}\int_{t-\tau\vee 0}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L^2_x}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right] \lesssim \tau + h^2. \tag{3.29}
$$

Step 5: We artificially introduce the desired error quantities and use the stability of mean-value time projections and the L^2-space projection. Let us denote $\langle f\rangle_{a_m} := \fint_{I_m}\fint_{t-\tau\vee 0}^{t} f_\nu \,\mathrm{d}\nu\,\mathrm{d}t$. A slight modification of Lemma 15 implies

$$
\mathbb{E}\left[\sum_{m=1}^{M}\int_{I_m} \left\|V(\nabla u_\nu) - \langle V(\nabla u)\rangle_{a_m}\right\|_{L^2_x}^2 \,\mathrm{d}\nu\right] \lesssim \tau\mathbb{E}\left[[V(\nabla u)]^2_{B^{1/2}_{2,\infty}L^2_x}\right].
$$

Additionally, Jensen's inequality ensures

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{m=1}^{M} \tau \left\|\langle V(\nabla u)\rangle_{a_m} - V(\nabla v_m)\right\|_{L^2_x}^2\right] \\
&\quad \le \mathbb{E}\left[\sum_{m=1}^{M} \fint_{I_m}\int_{t-\tau\vee 0}^{t} \|V(\nabla u_\nu) - V(\nabla v_m)\|_{L^2_x}^2 \,\mathrm{d}\nu\,\mathrm{d}t\right].
\end{aligned}
$$

Therefore, invoking (3.29),

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{m=1}^{M}\int_{I_m}\|V(\nabla u_\nu)-V(\nabla v_m)\|_{L_x^2}^2\,\mathrm{d}\nu\right]\\
&\lesssim \mathbb{E}\left[\sum_{m=1}^{M}\int_{I_m}\left\|V(\nabla u_\nu)-\langle V(\nabla u)\rangle_{a_m}\right\|_{L_x^2}^2\,\mathrm{d}\nu\right]\\
&\quad+\mathbb{E}\left[\sum_{m=1}^{M}\tau\left\|\langle V(\nabla u)\rangle_{a_m}-V(\nabla v_m)\right\|_{L_x^2}^2\right]\\
&\lesssim \tau\mathbb{E}\left[[V(\nabla u)]^2_{B^{1/2}_{2,\infty}L_x^2}\right]+\tau+h^2.
\end{aligned}
$$

The assertion (3.17) follows by an application of Lemma 13

$$
\begin{aligned}
&\mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|\langle u\rangle_m-v_m\|_{L_x^2}^2\right]\\
&\lesssim \mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|\Pi_2 e_m\|_{L_x^2}^2\right]+\mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|\langle u\rangle_m-\Pi_2\langle u\rangle_m\|_{L_x^2}^2\right]\\
&\lesssim \tau+h^2+h^2\mathbb{E}\left[\|\nabla u\|_{L_t^\infty L_x^2}^2\right].
\end{aligned}
$$

Part (b): Now, let us assume $u\in L^2_\omega B^{1/2}_{\Phi_2,\infty}L_x^2$. We apply Lemma 17 to bound

$$
\begin{aligned}
&\mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|u(t_m)-v_m\|_{L_x^2}^2\right]\\
&\lesssim \mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|\langle u\rangle_m-v_m\|_{L_x^2}^2\right]+\mathbb{E}\left[\max_{m\in\{1,\dots,M\}}\|u(t_m)-\langle u\rangle_m\|_{L_x^2}^2\right]\\
&\lesssim \tau+h^2+\tau\ln(1+\tau^{-1}).
\end{aligned}
$$

This verifies (3.19) and the proof is finished. □

One can also use time averages on the nonlinear gradient term to measure the error of the approximation.

Corollary 21 *Let the assumptions of Theorem 19(a) be satisfied. Then*

$$
\mathbb{E}\left[\sum_{m=1}^{M}\tau\|\langle V(\nabla u)\rangle_m-V(\nabla v_m)\|_{L_x^2}^2\right]\lesssim\tau+h^2 \tag{3.30}
$$

and

$$
\mathbb{E}\left[\sum_{m=1}^{M}\tau\|V(\nabla\langle u\rangle_m)-V(\nabla v_m)\|_{L_x^2}^2\right]\lesssim\tau+h^2. \tag{3.31}
$$

Proof The estimate (3.30) immediately follows by an application of Jensen's inequality and the bound (3.17),

$$\mathbb{E}\left[\sum_{m=1}^{M} \tau \left\| \langle V(\nabla u)\rangle_m - V(\nabla v_m) \right\|_{L_x^2}^2 \right] \leq \mathbb{E}\left[\sum_{m=1}^{M} \int_{I_m} \left\| V(\nabla u_\nu) - V(\nabla v_m) \right\|_{L_x^2}^2 \,\mathrm{d}\nu \right] \lesssim \tau + h^2.$$

In order to prove the second estimate (3.31) we use Lemmas 3 and 4

$$\begin{aligned}
&\mathbb{E}\left[\sum_{m=1}^{M} \tau \left\| V(\nabla \langle u\rangle_m) - V(\nabla v_m) \right\|_{L_x^2}^2 \right] \\
&\eqsim \mathbb{E}\left[\sum_{m=1}^{M} \int_{I_m}\int_{\mathcal{O}} (S(\nabla\langle u\rangle_m) - S(\nabla v_m)) : (\nabla u_\nu - \nabla v_m)\,\mathrm{d}x\,\mathrm{d}\nu \right] \\
&\leq c\mathbb{E}\left[\sum_{m=1}^{M} \int_{I_m} \left\| V(\nabla u_\nu) - V(\nabla v_m) \right\|_{L_x^2}^2 \,\mathrm{d}\nu \right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\sum_{m=1}^{M} \tau \left\| V(\nabla \langle u\rangle_m) - V(\nabla v_m) \right\|_{L_x^2}^2 \right].
\end{aligned}$$

Absorbing the second term to the left hand side and applying the bound (3.17) verifies the assertion. ◻

Remark 22 Although both (3.30) and (3.31) enjoy the same convergence rates, it is not clear whether one dominates the other. In the linear case, $p = 2$, the terms coincide.

The averaged error quantities (3.30) and (3.31) are equivalent up to oscillation to the error quantity (3.17).

Lemma 23 *Let $u \in L_t^p W_x^{1,p}$ and $A \in L_x^p$. Then*

$$\begin{aligned}
&\fint_{I_m} \left\| V(\nabla u_\nu) - V(A) \right\|_{L_x^2}^2 \,\mathrm{d}\nu \\
&\quad = \fint_{I_m} \left\| V(\nabla u_\nu) - \langle V(\nabla u)\rangle_m \right\|_{L_x^2}^2 \,\mathrm{d}\nu + \left\| \langle V(\nabla u)\rangle_m - V(A) \right\|_{L_x^2}^2
\end{aligned} \tag{3.32}$$

and

$$\begin{aligned}
&\fint_{I_m} \left\| V(\nabla u_\nu) - V(A) \right\|_{L_x^2}^2 \,\mathrm{d}\nu \\
&\quad \lesssim \fint_{I_m} \left\| V(\nabla u_\nu) - V(\nabla \langle u\rangle_m) \right\|_{L_x^2}^2 \,\mathrm{d}\nu + \left\| V(\nabla \langle u\rangle_m) - V(A) \right\|_{L_x^2}^2 .
\end{aligned} \tag{3.33}$$

Proof The Eq. (3.32) follows by using

$$\fint_{I_m} (V(\nabla u_\nu) - \langle V(\nabla u)\rangle_m, \langle V(\nabla u)\rangle_m - V(A)) \,\mathrm{d}\nu = 0.$$

The estimate (3.33) is obtained trivially,

$$\begin{aligned}
&\fint_{I_m} \|V(\nabla u_\nu) - V(A)\|_{L_x^2}^2 \,\mathrm{d}\nu \\
&\quad = \fint_{I_m} \|V(\nabla u_\nu) - V(\nabla\langle u\rangle_m) - (V(\nabla\langle u\rangle_m) - V(A))\|_{L_x^2}^2 \,\mathrm{d}\nu \\
&\quad \lesssim \fint_{I_m} \|V(\nabla u_\nu) - V(\nabla\langle u\rangle_m)\|_{L_x^2}^2 \,\mathrm{d}\nu + \|V(\nabla\langle u\rangle_m) - V(A)\|_{L_x^2}^2 .
\end{aligned}$$

□

Remark 24 In [28, Lemma 6.2] the authors prove the equivalence (although it is done purely in space but can be extended to time) of

$$\fint_{I_m} \|V(\nabla u_\nu) - V(\nabla\langle u\rangle_m)\|_{L_x^2}^2 \,\mathrm{d}\nu \eqsim \fint_{I_m} \|V(\nabla u_\nu) - \langle V(\nabla u)\rangle_m\|_{L_x^2}^2 \,\mathrm{d}\nu.$$

If $V(\nabla u) \in B^{1/2}_{2,\infty} L_x^2$, then Lemma 15 implies

$$\sum_{m=1}^{M} \fint_{I_m} \|V(\nabla u_\nu) - \langle V(\nabla u)\rangle_m\|_{L_x^2}^2 \,\mathrm{d}\nu \lesssim \tau [V(\nabla u)]^2_{B^{1/2}_{2,\infty} L_x^2}.$$

Theorem 25 (Convergence of Algorithm 3.15) *Let the assumptions of Theorem 19 be satisfied. Denote by* $\mathbf{w} \in (V_h)^{M+1}$ *the solution to* (3.15) *and by* u *the weak solution to* (2.11). *Then*

$$\begin{aligned}
&\mathbb{E}\left[\max_{m=1,\ldots,M} \|\langle u\rangle_m - w_m\|_{L_x^2}^2 + \sum_{m=1}^{M} \int_{I_m} \|V(\nabla u_\nu) - V(\nabla w_m)\|_{L_x^2}^2 \,\mathrm{d}\nu\right] \\
&\lesssim \tau + h^2
\end{aligned} \tag{3.34}$$

and

$$\mathbb{E}\left[\max_{m=1,\ldots,M} \|u(t_m) - w_m\|_{L_x^2}^2\right] \lesssim \tau \ln(1+\tau^{-1}) + h^2. \tag{3.35}$$

Proof The proof proceeds similarly to the proof of Theorem 19. We will only prove the bound for the initial error. Instead of comparing w_1 to $\langle u\rangle_1$, we rather choose $\langle u\rangle_{I_1\cup I_2} := \fint_{I_1\cup I_2} u_\nu \,\mathrm{d}\nu$. The equation for the latter one reads

$$\big(\langle u\rangle_{I_1\cup I_2} - u_0, \xi\big) + \fint_{I_1\cup I_2} \int_0^t (S(\nabla u_\nu), \nabla\xi)\,\mathrm{d}\nu\,\mathrm{d}t = \left(\fint_{I_1\cup I_2} \int_0^t G(u_\nu)\,\mathrm{d}W_\nu\,\mathrm{d}t, \xi\right). \tag{3.36}$$

Subtracting (3.15b) from (3.36) and choosing $\xi_h = \Pi_2\langle u\rangle_{I_1\cup I_2} - w_1$ results in

$$\begin{aligned} A_1 + A_2 &:= \left\|\Pi_2\langle u\rangle_{I_1\cup I_2} - w_1\right\|_{L^2_x}^2 \\ &\quad + \fint_{I_1\cup I_2} \int_0^t \big(S(\nabla u_\nu) - S(\nabla w_1), \nabla(\Pi_2\langle u\rangle_{I_1\cup I_2} - w_1)\big)\,\mathrm{d}\nu\,\mathrm{d}t \\ &= \left(\fint_{I_1\cup I_2} \int_0^t G(u_\nu)\,\mathrm{d}W_\nu\,\mathrm{d}t, \Pi_2\langle u\rangle_{I_1\cup I_2} - w_1\right) \\ &\quad - \big(G(w_0)\langle W\rangle_1, \Pi_2\langle u\rangle_{I_1\cup I_2} - w_1\big) \\ &=: A_3 + A_4. \end{aligned}$$

Due to Hölder's and Young's inequalities, Itô isometry and the growth assumption (2.2)

$$\begin{aligned} \mathbb{E}\left[A_3\right] &\le \frac{1}{4}\mathbb{E}\left[A_1\right] + \mathbb{E}\left[\left\|\fint_{I_1\cup I_2} \int_0^t G(u_\nu)\,\mathrm{d}W_\nu\,\mathrm{d}t\right\|_{L^2_x}^2\right] \\ &= \frac{1}{4}\mathbb{E}\left[A_1\right] + \mathbb{E}\left[\int_{I_1\cup I_2} \left(\frac{\nu}{2\tau}\right)^2 \left\|G(u_\nu)\right\|_{L_2(U;L^2_x)}^2 \,\mathrm{d}\nu\right] \\ &\le \frac{1}{4}\mathbb{E}\left[A_1\right] + c\mathbb{E}\left[\int_{I_1\cup I_2} \left(\frac{\nu}{2\tau}\right)^2 \left\|1 + u_\nu\right\|_{L^2_x}^2 \,\mathrm{d}\nu\right]. \end{aligned}$$

The boundedness of u as an L^2_x valued-process implies

$$\mathbb{E}\left[A_3\right] \le \frac{1}{4}\mathbb{E}\left[A_1\right] + c\frac{2}{3}\tau\left\|1 + u\right\|_{L^2_\omega L^\infty_t L^2_x}^2.$$

The forth term is estimated similarly,

$$\begin{aligned} \mathbb{E}\left[A_4\right] &\le \frac{1}{4}\mathbb{E}\left[A_1\right] + \mathbb{E}\left[\left\|G(w_0)\langle W\rangle_1\right\|_{L^2_x}^2\right] \\ &\le \frac{1}{4}\mathbb{E}\left[A_1\right] + c\frac{1}{3}\tau\mathbb{E}\left[\left\|1 + \Pi_2 u_0\right\|_{L^2_x}^2\right]. \end{aligned}$$

Using the L^2-stability of the L^2-projection we obtain

$$\mathbb{E}\left[A_4\right] \leq \frac{1}{4}\mathbb{E}\left[A_1\right] + c\frac{1}{3}\tau \left\|1 + u_0\right\|^2_{L^2_\omega L^2_x}.$$

It remains to check the second term. Here we use the same arguments as in step 3 of the proof of Theorem 19 to conclude

$$A_2 \geq \int_{I_1} \left\|V(\nabla u_\nu) - V(\nabla w_1)\right\|^2_{L^2_x} \,\mathrm{d}\nu - c\left(\tau[V(\nabla u)]^2_{B^{1/2}_{2,\infty}L^2_x} + h^2 \left\|\nabla V(\nabla u)\right\|^2_{L^2_t L^2_x}\right).$$

Overall, we have established

$$\begin{aligned}
&\mathbb{E}\left[\left\|\Pi_2\langle u\rangle_{I_1\cup I_2} - w_1\right\|^2_{L^2_x}\right] + \mathbb{E}\left[\int_{I_1} \left\|V(\nabla u_\nu) - V(\nabla w_1)\right\|^2_{L^2_x} \,\mathrm{d}\nu\right] \\
&\lesssim \tau\mathbb{E}\left[\left\|1 + u\right\|^2_{L^\infty_t L^2_x} + \left\|1 + u_0\right\|^2_{L^2_x} + [V(\nabla u)]^2_{B^{1/2}_{2,\infty}L^2_x}\right] + h^2\mathbb{E}\left[\left\|\nabla V(\nabla u)\right\|^2_{L^2_t L^2_x}\right].
\end{aligned}$$

Lastly, using Lemmas 13 and 15

$$\begin{aligned}
&\mathbb{E}\left[\left\|\langle u\rangle_1 - w_1\right\|^2_{L^2_x}\right] \\
&\quad\lesssim \mathbb{E}\left[\left\|\langle u\rangle_1 - \langle u\rangle_{I_1\cup I_2}\right\|^2_{L^2_x}\right] + \mathbb{E}\left[\left\|\langle u\rangle_{I_1\cup I_2} - \Pi_2\langle u\rangle_{I_1\cup I_2}\right\|^2_{L^2_x}\right] \\
&\qquad + \mathbb{E}\left[\left\|\Pi_2\langle u\rangle_{I_1\cup I_2} - w_1\right\|^2_{L^2_x}\right] \\
&\quad\lesssim \tau\mathbb{E}\left[\left\|1 + u\right\|^2_{L^\infty_t L^2_x} + \left\|1 + u_0\right\|^2_{L^2_x} + [u]^2_{B^{1/2}_{2,\infty}L^2_x} + [V(\nabla u)]^2_{B^{1/2}_{2,\infty}L^2_x}\right] \\
&\qquad + h^2\mathbb{E}\left[\left\|\nabla u\right\|^2_{L^\infty_t L^2_x} + \left\|\nabla V(\nabla u)\right\|^2_{L^2_t L^2_x}\right].
\end{aligned}$$

The bound for the initial error is complete. □

Remark 26 We want to remark, although the first step (3.15b) is announced to be a full step, it is not a full first step. The deterministic drift is scaled by a full step τ, but the stochastic term is only scaled by $\langle W\rangle_1$. In Corollary 33 we find that $\langle W\rangle_1$ has only half the variance of a full stochastic step.

To overcome this problem we can introduce a stochastic dummy variable $\mathcal{W}_0 \sim \mathcal{N}(0, \tau/3)$[1] that artificially makes up for the missing randomness. Then (3.15b) needs to be substituted by

$$(w_1 - w_0, \xi_h) + \tau\left(S(\nabla w_1), \nabla\xi_h\right) = \left(G(w_0)\left(\langle W\rangle_1 + \mathcal{W}_0\right), \xi_h\right). \tag{3.37}$$

The error analysis of Theorem 25 can be extended to this algorithm.

[1] denotes equality in distribution.

4 Discrete stochastic processes

In this section we investigate the law of averaged Wiener processes and propose an implementable sampling algorithm.

4.1 Wiener process

We introduce the concept of Hilbert space valued Gaussian processes.

Definition 27 A stochastic process Y is called U-valued Gaussian process with mean operator $m : I \times U \to \mathbb{R}$ and variance operator $\Sigma : I \times U \times U \to \mathbb{R}$, if for all $t \in I$ and $u \in U$ it holds

$$\varphi_{Y_t}(u) := \mathbb{E}\left[e^{-i(Y_t,u)_U}\right] = e^{-im_t(u)-\frac{1}{2}\Sigma_t(u,u)}.$$

In short, we write $Y_t \sim \mathcal{N}(m_t, \Sigma_t)$.

It follows that the stochastic forcing W defined by (2.1) is an U-valued Gaussian process with mean operator $m_t(u) = 0$ and variance operator $\Sigma_t(u, v) := t(u, v)_U$ for all $t \in I$ and $u, v \in U$. Moreover, the series (2.1) converges in the weak topology of U, due to the hypercontractivity of normally distributed random variables it holds for any $q > 0$, $u \in U$ and $t, s \in I$

$$\begin{aligned}\left(\mathbb{E}\left[\left|(W_t - W_s, u)_U\right|^q\right]\right)^{\frac{1}{q}} &= \left(\mathbb{E}\left[\left|\sum_{j\in\mathbb{N}} \left(u_j, u\right)_U \left(\beta_t^j - \beta_s^j\right)\right|^q\right]\right)^{\frac{1}{q}} \\ &\eqsim \sqrt{q}\, |t-s|^{\frac{1}{2}} \|u\|_U .\end{aligned}$$

In fact, as soon as the index set is infinite we lose the norm convergence, since

$$\left(\mathbb{E}\left[\|W_t\|_U^q\right]\right)^{\frac{1}{q}} \eqsim \sqrt{q} \left(t \sum_{j\in\mathbb{N}} \|u_j\|_U^2\right)^{\frac{1}{2}} = \infty.$$

4.2 Averaged Wiener process

In this section we compute the distributions of the random variables $(\langle W\rangle_m)_{m=1}^M$ and the joint distribution of the averaged increments.

A key tool in the derivation of the distribution of the averaged Wiener process is the decomposition of the process W adjusted to the equidistant time partition $\{I_m\}_{m=1}^M$. We decompose $W|_{I_m}$ into a Brownian bridge $\mathcal{B}_m$ and its nodal values $W(t_{m-1})$ and $W(t_m)$, i.e., for $t \in I_m$

$$W(t) = W(t_{m-1}) + \mathcal{B}_m(t) + \frac{t - t_{m-1}}{\tau} \Delta_m W, \tag{4.1}$$

where

$$\mathcal{B}_m(t) := W(t) - W(t_{m-1}) - \frac{t - t_{m-1}}{\tau} \Delta_m W. \tag{4.2}$$

Brownian bridges have nice independency properties. They do not look into the past nor future. The following result can be found in [63, Section 1.2].

Proposition 28 *Let $(\mathcal{B}_m)_{m=1}^M$ be given by (4.2). Then for all $m \in \{1, \dots, M\}$*

$$\sigma\left(\mathcal{B}_m(t) \middle| t \in I_m\right) \perp \sigma\left(W(t) \middle| t \in [0, \infty) \backslash (t_{m-1}, t_m)\right), \tag{4.3}$$

i.e. all finite dimensional distributions of the generators of each sigma algebra are independent of each other.

Corollary 29 *Let $(\mathcal{B}_m)_{m=1}^M$ be given by (4.2). Then $\mathcal{B}_1, \dots, \mathcal{B}_M, \Delta_1 W, \dots, \Delta_M W$ are independent.*

Next, we take the time average over the interval I_m in (4.1) and obtain

$$\langle W \rangle_m = W(t_{m-1}) + \langle \mathcal{B}_m \rangle_m + \frac{\Delta_m W}{2}. \tag{4.4}$$

Now, it is our choice whether we want to compute the distribution of $\langle W \rangle_m$ or $\langle \mathcal{B}_m \rangle_m$. The formula (4.4) provides an easy way to compute the remaining one. We choose to compute the distribution of $\langle W \rangle_m$. An application of Itô's formula for $f(s, W_s) = \frac{s - t_m}{\tau} W_s$ implies $\mathbb{P}$-a.s.

$$\langle W \rangle_m = W_{t_{m-1}} + \int_{t_{m-1}}^{t_m} \frac{t_m - s}{\tau} \,\mathrm{d}W_s. \tag{4.5}$$

A stochastic integral that is driven by a Wiener process and a deterministic integrand stays Gaussian.

Lemma 30 ([4] Prop. 7.1) *Let $f \in L^2(I)$. Then*

$$t \mapsto \int_0^t f_s \,\mathrm{d}W_s$$

is an U-valued Gaussian process with zero mean and variance

$$\Sigma_t(u, v) = \int_0^t |f_s|^2 \,\mathrm{d}s \,(u, v)_U\,.$$

Corollary 31 *$\langle W \rangle_m$ is an U-valued Gaussian random variable with zero mean and variance $\Sigma(u, v) = \frac{2t_{m-1} + t_m}{3} (u, v)_U$.*

Proof Let us define

$$\langle W\rangle_m = W_{t_{m-1}} + \int_{t_{m-1}}^{t_m} \frac{t_m - s}{\tau} \,\mathrm{d}W_s =: W_a + W_b.$$

Note, that W_a and W_b are independent. Moreover, W_a is Gaussian with variance $\Sigma_a(u, v) = t_{m-1}\,(u, v)_U$ and due to Lemma 30 W_b is also Gaussian. Therefore $\langle W\rangle_m$ is Gaussian and it suffices to compute the mean and the variance operators.

Let $u, v \in U$. Then $\mathbb{E}\left[(\langle W\rangle_m, u)_U\right] = 0$ and

$$\begin{aligned}
&\mathbb{E}\left[(\langle W\rangle_m, u)_U\,(\langle W\rangle_m, v)_U\right] \\
&\quad = \mathbb{E}\left[(W_a, u)_U\,(W_a, v)_U\right] + \mathbb{E}\left[(W_b, u)_U\,(W_b, v)_U\right] \\
&\quad = (u, v)_U \left(t_{m-1} + \int_{t_{m-1}}^{t_m} \left(\frac{t_m - s}{\tau}\right)^2 \mathrm{d}s\right) \\
&\quad = (u, v)_U\, \frac{2t_{m-1} + t_m}{3}.
\end{aligned}$$

The assertion is proved. □

Corollary 32 *$\langle \mathcal{B}_m\rangle_m$ is an U-valued Gaussian random variable with zero mean and variance $\Sigma(u, v) = \frac{\tau}{12}(u, v)_U$.*

Proof The distribution of a random variable is uniquely determined by its characteristic function. Corollary 31 implies

$$\varphi_{\langle W\rangle_m}(u) = e^{-\frac{1}{2}\frac{2t_{m-1}+t_m}{3}\|u\|_U^2}.$$

Classically, we find

$$\begin{aligned}
\varphi_{W(t_{m-1})}(u) &= e^{-\frac{1}{2}t_{m-1}\|u\|_U^2}, \\
\varphi_{\frac{\Delta_m W}{2}}(u) &= e^{-\frac{1}{2}\frac{\tau}{4}\|u\|_U^2}.
\end{aligned}$$

The characteristic function of $\varphi_{\langle W\rangle_m}$ factors due to the independence of the decomposition (4.4), i.e.,

$$\varphi_{\langle W\rangle_m}(u) = \varphi_{W(t_{m-1})+\langle \mathcal{B}_m\rangle_m+\frac{\Delta_m W}{2}}(u) = \varphi_{W(t_{m-1})}(u)\varphi_{\langle \mathcal{B}_m\rangle_m}(u)\varphi_{\frac{\Delta_m W}{2}}(u).$$

Rearranging implies

$$\begin{aligned}
\varphi_{\langle \mathcal{B}_m\rangle_m}(u) &= \frac{\varphi_{\langle W\rangle_m}(u)}{\varphi_{W(t_{m-1})}(u)\varphi_{\frac{\Delta_m W}{2}}(u)} \\
&= e^{-\frac{1}{2}\left(\frac{2t_{m-1}+t_m}{3} - t_{m-1} - \frac{\tau}{4}\right)\|u\|_U^2} = e^{-\frac{1}{2}\frac{\tau}{12}\|u\|_U^2}.
\end{aligned}$$

Overall, $\langle \mathcal{B}_m \rangle_m$ has the characteristic function of a Gaussian random variable with zero mean and variance $\Sigma(u, v) = \frac{\tau}{12}(u, v)_U$. □

At this point it is a simple task to find the distribution of the increments of the averaged Wiener process. Let us subtract (4.4) for m and $m-1$

$$\Delta_m \mathbb{W} := \langle W \rangle_m - \langle W \rangle_{m-1} = \frac{\Delta_m W + \Delta_{m-1} W}{2} + \langle \mathcal{B}_m \rangle_m - \langle \mathcal{B}_{m-1} \rangle_{m-1}, \quad (4.6)$$

where we define $\langle W \rangle_0 := \Delta_0 W := \langle \mathcal{B}_0 \rangle_0 := 0$.

Corollary 33 *$\Delta_m \mathbb{W}$ is an U-valued Gaussian random variable with zero mean and variance $\Sigma(u, v) = \left(\frac{2\tau}{3} \chi_{\{m \geq 2\}} + \frac{\tau}{3} \chi_{\{m=1\}}\right)(u, v)_U$.*

Proof The right hand side of (4.6) is a sum of independent, centered Gaussian random variables. Thus the left hand side is centered Gaussian. Now, it suffices to compute the variance operator.

Note, in the case $m = 1$ we have $\Delta_1 \mathbb{W} = \langle W \rangle_1$ and the result follows by Corollary 31.

Let $m \geq 2$ and $u, v \in U$. Due to the independence,

$$\begin{aligned}
&\mathbb{E}\left[(\Delta_m \mathbb{W}, u)_U \, (\Delta_m \mathbb{W}, v)_U\right] \\
&\quad = \mathbb{E}\left[\left(\frac{\Delta_m W}{2}, u\right)_U \left(\frac{\Delta_m W}{2}, v\right)_U\right] + \mathbb{E}\left[\left(\frac{\Delta_{m-1} W}{2}, u\right)_U \left(\frac{\Delta_{m-1} W}{2}, v\right)_U\right] \\
&\qquad + \mathbb{E}\left[(\langle \mathcal{B}_m \rangle_m, u)_U \, (\langle \mathcal{B}_m \rangle_m, v)_U\right] + \mathbb{E}\left[(\langle \mathcal{B}_{m-1} \rangle_{m-1}, u)_U \, (\langle \mathcal{B}_{m-1} \rangle_{m-1}, v)_U\right] \\
&\quad = \frac{2\tau}{3} (u, v)_U \, .
\end{aligned}$$

□

So far we have identified how each averaged increment $\Delta_m \mathbb{W}$ is distributed. However, we also need to know what the joint distribution is, i.e., the distribution of a random vector.

Lemma 34 *The random vector $(\Delta_m \mathbb{W})_{m=1}^M$ is an U^M-valued centered Gaussian random variable with variance operator $\Sigma : U^M \times U^M \to \mathbb{R}$ given by*

$$\Sigma(\mathbf{u}, \mathbf{v}) := \sum_{m,l=1}^{M} \sigma_{m,l} \, (u_m, v_l)_U \, ,$$

for $\mathbf{u}, \mathbf{v} \in U^M$ and

$$\sigma_{m,l} = \begin{cases} \frac{1}{3}\tau & \text{if} \quad l = m = 1, \\ \frac{2}{3}\tau & \text{if} \quad l = m > 1, \\ \frac{1}{6}\tau & \text{if} \quad |l - m| = 1, \\ 0 & \text{if} \quad |l - m| > 1. \end{cases} \quad (4.7)$$

Proof The Eq. (4.6) implies, that the random vector $(\Delta_m \mathbb{W})_{m=1}^M$ can be constructed via a linear transformation of independent Gaussian random vectors $(\Delta_m W)_{m=1}^M$ and $(\langle \mathcal{B}_m \rangle_m)_{m=1}^M$, i.e.,

$$(\Delta_m \mathbb{W})_{m=1}^M = \mathbb{K}_1 (\Delta_m W)_{m=1}^M + \mathbb{K}_2 (\langle \mathcal{B}_m \rangle_m)_{m=1}^M,$$

where $\mathbb{K}_1, \mathbb{K}_2 \in \mathbb{R}^{M \times M}$ are given by

$$\mathbb{K}_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 & 1 \end{pmatrix}, \quad \mathbb{K}_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 \\ 0 & -1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & 1 & 0 \\ 0 & 0 & \dots & 0 & -1 & 1 \end{pmatrix}.$$

Therefore, $(\Delta_m \mathbb{W})_{m=1}^M$ is itself a centered Gaussian vector. It remains to compute the covariance matrix. Let $u, v \in U$ and $m, l \in \{1, \dots, M\}$. If $m = l$ Corollary 33 implies

$$\mathbb{E}\left[(\Delta_m \mathbb{W}, u)\,(\Delta_l \mathbb{W}, v)\right] = \left(\frac{2\tau}{3} \chi_{\{m \geq 2\}} + \frac{\tau}{3} \chi_{\{m=1\}} \right) (u, v)_U .$$

If $|m - l| > 1$, then Eq. (4.6) and the independence imply

$$\mathbb{E}\left[(\Delta_m \mathbb{W}, u)\,(\Delta_l \mathbb{W}, v)\right] = 0.$$

It remains to consider the case $|m - l| = 1$. Without loss of generality $l = m + 1$. Now, using (4.6), the independence and Corollary 32

$$\begin{aligned} &\mathbb{E}\left[(\Delta_m \mathbb{W}, u)\,(\Delta_{m+1} \mathbb{W}, v)\right] \\ &\quad = \mathbb{E}\left[\left(\frac{\Delta_m W}{2}, u\right)\left(\frac{\Delta_m W}{2}, v\right)\right] - \mathbb{E}\left[(\langle \mathcal{B}_m \rangle_m, u)\,(\langle \mathcal{B}_m \rangle_m, v)\right] \\ &\quad = \left(\frac{\tau}{4} - \frac{\tau}{12}\right)(u, v)_U = \frac{\tau}{6} (u, v)_U . \end{aligned}$$

The proof is finished. □

4.3 Sampling algorithm

On the computer we are forced to approximate the continuous measure induced by the Wiener process W by an empirical measure. Additionally, if we want to compare different numerical schemes we need to specify how to sample the random input needed for the involved algorithms jointly. More specifically, if we want to compare our algorithms (3.14) and (3.15) and the classical Euler–Maruyama discretization (5.1), we need to sample according to the law of the random vector

$$(\Delta_1 W, \ldots, \Delta_M W, \Delta_1 \mathbb{W}, \ldots, \Delta_M \mathbb{W}) . \tag{4.8}$$

Based on the decomposition (4.6) we propose the following sampling algorithm. The time discretization is achieved by the parameter $M \in \mathbb{N}$ and the series truncation in (2.1) is done by the parameter $J \in \mathbb{N}$.

Given $M, J \in \mathbb{N}$.

(a) (Sampling) Compute i.i.d. random variables $\zeta_m^j, \eta_m^j \sim \mathcal{N}(0,1)$ for $m \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, J\}$.

(b) (Lift to Hilbert space U) For $m \in \{1, \ldots, M\}$ define the random variables

$$Z_m := \sqrt{\tau} \sum_{j=1}^{J} u^j \zeta_m^j, \quad \tilde{\mathbb{Z}}_m := \sqrt{\frac{\tau}{12}} \sum_{j=1}^{J} u^j \eta_m^j, \tag{4.9a}$$

where $\{u^j\}_{j\in\mathbb{N}}$ is an orthonormal system of U.

(c) (Adjusting correlation) For $m \in \{1, \ldots, M\}$ define the random variables

$$\mathbb{Z}_m := \frac{Z_m + Z_{m-1}}{2} + \tilde{\mathbb{Z}}_m - \tilde{\mathbb{Z}}_{m-1}, \tag{4.9b}$$

where $Z_0 := \tilde{\mathbb{Z}}_0 := 0$.

Let Π_J be the U-orthogonal projection onto $U_J := \operatorname{span}(u_1, \ldots, u_J)$. The following proposition guarantees that the sampling algorithm (4.9) approximates the desired random variables.

Proposition 35 *Let* $\mathbf{Z} := (Z_1, \ldots, Z_M, \mathbb{Z}_1, \ldots, \mathbb{Z}_M) \in U^{2M}$ *be generated by* (4.9a) *and* (4.9b). *Then,*

$$\mathbf{Z} \sim (\Pi_J \Delta_1 W, \ldots, \Pi_J \Delta_M W, \Pi_J \Delta_1 \mathbb{W}, \ldots, \Pi_J \Delta_M \mathbb{W}) ,$$

where Π_J *is the* U*-orthogonal projection onto span*$(u^1, \ldots, u^J)$.

Proof First, we need to observe that $Z_m \sim \Pi_J \Delta_m W$ and $\tilde{\mathbb{Z}}_m \sim \Pi_J \langle \mathcal{B}_m \rangle_m$. Then, the statement follows similarly to the proof of Lemma 34. □

Remark 36 Proposition 35 ensures that we can compare our algorithm to the classical Euler–Maruyama discretization (5.1) on an equidistant time grid. It can be adjusted to also match non-equidistant grids.

5 Simulations

In this section we perform numerical simulations to test our algorithms (3.14) and (3.15). We denote the solution of (3.14) as $\mathbf{v}^{\mathrm{Half}}$ and the solution of (3.15) as

$\mathbf{v}^{\text{Full}}$. Additionally, we compare our algorithms to the classical Euler–Maruyama discretization of (2.11). That reads, find $\mathbf{v} \in (V_h)^{M+1}$ such that for all $\xi_h \in V_h$, $m \geq 1$ and $\mathbb{P}$-a.s.

$$(v_m - v_{m-1}, \xi_h) + \tau \left(S(\nabla v_m), \nabla \xi_h\right) = \left(G(v_{m-1})\Delta_m W, \xi_h\right). \tag{5.1}$$

The solution to (5.1) is called $\mathbf{v}^{\text{EM}}$.

A plain convergence result for the Euler–Maruyama scheme without any rate has been obtained in [65]. A more sophisticated analysis has been done in [39]. However, they only obtain convergence for the $L^2_\omega L^\infty_t L^2_x$-error with rate $1/4$.

Originally the Euler–Maruyama scheme has been introduced for stochastic ordinary differential equations. In this context much more is known, see e.g. the book of Kloeden and Platen [52, Section 9.5]. However, when dropping the Lipschitz assumption on the coefficients, divergence with positive probability has been obtained in [49].

We are particularly interested in the experimental study of the following questions:

(a) Do the algorithms (3.14) and (3.15) approximate mean-values or point-values?
(b) How does the Euler–Maruyama scheme (5.1) compares to (3.14) and (3.15) in terms of time and space convergence?
(c) How do the different error quantities (3.17), (3.30) and (3.31) for the gradient relate to each other?
(d) How sensitive are the algorithms with respect to the parameter p?

All simulations are done with the help of the open source tool for solving partial differential equations FEniCS [62].

5.1 An explicit solution

In the linear case, $p = 2$, with a linear right hand side

$$G(v)\Delta W = \lambda v \, \Delta\beta^1, \tag{5.2}$$

for some $\lambda \in \mathbb{R}$, it is possible to find an explicit solution. If we start the evolution defined by (2.11) in an eigenfunction of the Laplace operator, the dynamics become simpler. Let $\mathcal{O} = (0,1)^2$, $T = 1$ and $u_0(x) = \sin(\pi x_1)\sin(\pi x_2)$. Note that u_0 is an eigenfunction of the 2-Laplacian with homogeneous Dirichlet data and corresponding eigenvalue $\mu = 2\pi^2$. The unique solution to (2.11) is given by

$$u(\omega, t, x) = \exp\left\{-\left(\frac{\lambda^2}{2} + \mu\right) t + \lambda\beta^1(\omega, t)\right\} u_0(x). \tag{5.3}$$

Similarly, we can give a closed expression for the solution u_h to the space-discrete equation

$$\mathrm{d}\,(u_h, \xi_h) + (\nabla u_h, \nabla \xi_h)\,\mathrm{d}t = \lambda\,(u_h, \xi_h)\,\mathrm{d}\beta^1(t). \tag{5.4}$$

This is equivalent to the system of linear stochastic differential equations

$$\mathrm{d}M\mathbf{u} + S\mathbf{u}\,\mathrm{d}t = \lambda M\mathbf{u}\,\mathrm{d}\beta^1(t), \tag{5.5}$$

where $M_{i,j} = \left(\xi_h^j, \xi_h^i\right)$ and $S_{i,j} = \left(\nabla\xi_h^j, \nabla\xi_h^i\right)$ is the mass respectively the stiffness matrix and $\{\xi_h^i\}$ form a basis of V_h. The eigenpairs of the discretized Laplacian are related to the linear system

$$S\mathbf{u} = \mu M\mathbf{u} \Leftrightarrow M^{-1}S\mathbf{u} = \mu\mathbf{u}. \tag{5.6}$$

Let $(\mu_h, \mathbf{u}_h) \in (0,\infty) \times \mathbb{R}^{|V_h|}$ be a solution to (5.6), then the solution to (5.4) started in $u_h(0) = \mathbf{u}_h \cdot \boldsymbol{\xi}_h$ is given by

$$u_h(\omega, t, x) = \exp\left\{-\left(\frac{\lambda^2}{2} + \mu_h\right)t + \lambda\beta^1(\omega, t)\right\}u_h(0). \tag{5.7}$$

The main advantage of having an analytic solution of the space discrete equation is, that it rules out any space discretization errors.

To accurately compare continuous processes and discrete vectors, one needs to either lift the vector to a process or project the process to a vector. We do the latter approach and evaluate the continuous process u_h on the equidistant partition of I. This leads to the definition

$$\mathbf{u}^{\mathrm{Point}} := \left(u_h(t_m)\right)_{m=1}^{M}. \tag{5.8}$$

Since the algorithms (3.14) and (3.15) approximate the mean value of the analytic solution, we define

$$\mathbf{u}^{\mathrm{Aver}}_{\mathrm{exact}} := \left(\langle u_h\rangle_{I_m}\right)_{m=1}^{M}. \tag{5.9}$$

Although we know the exact solution, the time averages of the exact solution are non-treatable without knowledge of the full trajectory of the Brownian motion β^1. Numerically, we substitute $\langle u_h\rangle_{I_m}$ by a Riemann sum approximation, i.e. we fix an equidistant partition $\{[t_{m,k-1}, t_{m,k}]\}_{k=1}^{r}$ of I_m with resolution $r \in \mathbb{N}$ and define

$$\mathbf{u}^{\mathrm{Aver}} := \left(\frac{1}{r}\sum_{k=1}^{r} u_h(t_{m,k})\right)_{m=1}^{M}. \tag{5.10}$$

The approximation quality is measured in the error quantities

$$\mathcal{E}(\mathbf{u}, \mathbf{v}) := \mathbb{E}\left[\max_{m=1,\dots,M} \|u_m - v_m\|_{L^2_x}^2\right], \tag{5.11a}$$

$$\mathcal{V}(\mathbf{u}, \mathbf{v}) := \mathbb{E}\left[\sum_{m=1}^{M} \tau\,\|\nabla(u_m - v_m)\|_{L^2_x}^2\right], \tag{5.11b}$$

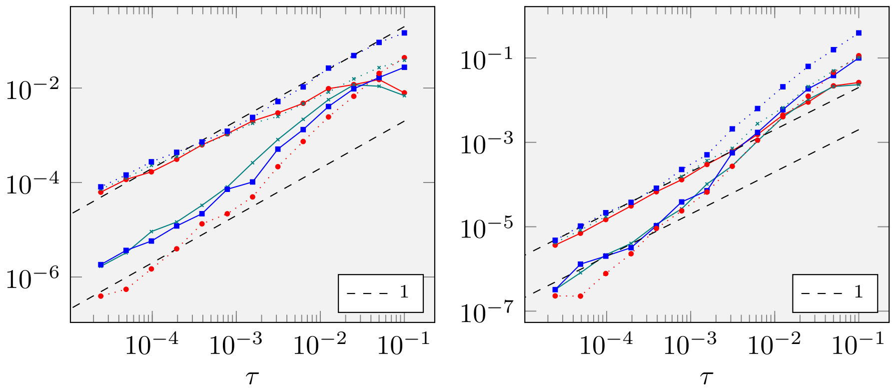

Fig. 1 Time convergence of $\mathbf{v}^{\mathrm{EM}}$ (red), $\mathbf{v}^{\mathrm{Half}}$ (blue) and $\mathbf{v}^{\mathrm{Full}}$ (green) towards $\mathbf{u}^{\mathrm{Point}}$ (dashed) respectively $\mathbf{u}^{\mathrm{Aver}}$ (solid) measured in the error terms $\mathcal{E}$ (left) and $\mathcal{V}$ (right) (color figure online)

where $\mathbf{u} \in \{\mathbf{u}^{\mathrm{Point}}, \mathbf{u}^{\mathrm{Aver}}\}$ and $\mathbf{v} \in \{\mathbf{v}^{\mathrm{EM}}, \mathbf{v}^{\mathrm{Full}}, \mathbf{v}^{\mathrm{Half}}\}$.

In Fig. 1 we plot the time convergence of the error quantities (5.11a) and (5.11b). We approximate the expectation by the Monte-Carlo method with 20 samples. Additionally, we let $\lambda = 1$ and V_h to be the space of piecewise linear, continuous elements on a uniform mesh with $|V_h| = 121$. The average values (5.10) are approximated by $r = 10$.

The numerical results support that $\mathbf{v}^{\mathrm{EM}}$ approximates the solution on the grid points, while $\mathbf{v}^{\mathrm{Full}}$ and $\mathbf{v}^{\mathrm{Half}}$ approximate the average values of the solution. The gap is due to the difference

$$\mathbb{E}\left[\left\| \langle u \rangle_{I_m} - u(t_m) \right\|_{L^2_x}^2 \right] \geq c\tau \left\| u_0 \right\|_{L^2_x}^2 .$$

Initially, we observe a preasymptotic effect. It stabilizes at the time scale $\tau \approx 10^{-3}$. Afterwards the predicted convergence speed of order 1 is achieved.

5.2 Beyond known solutions

In general, a major obstacle is the absence of an analytic solution to the Eq. (2.11). In particular, the distance of the numerical solution and the analytic solution as presented in (3.17), (3.19) respectively (3.34) and (3.35) are non-computable.

To overcome this difficulty we measure the error of a fine reference approximation $\mathbf{v}_f$ and a coarse approximation $\mathbf{v}_c$. Both, $\mathbf{v}_f$ and $\mathbf{v}_c$, are generated via the same algorithm (either (3.14), (3.15) or (5.1)) on a fine respectively coarse scale. Let $h_c \geq h_f > 0$ be coarse respectively fine space mesh sizes. Similarly, let $M_c, M_f \in \mathbb{N}$ be coarse respectively fine time discretization parameters with corresponding timestep sizes τ_c, τ_f. For simplicity, we assume $M_c/M_f = r \in \mathbb{N}$. Now, the coarse intervals are generated by the fine ones, i.e.

$$I_m^c := [(m-1)\tau_c, m\tau_c] = \bigcup_{k=1}^{r} I^f_{(m-1)r+k}.$$

The coarse averaging operator can be decomposed into the fine averaging operator

$$\langle u\rangle_{I_m^c} = \fint_{I_m^c} u_\nu \,\mathrm{d}\nu = \frac{1}{r\tau_f}\sum_{k=1}^{r}\int_{I^f_{(m-1)r+k}} u_\nu \,\mathrm{d}\nu = \frac{1}{r}\sum_{k=1}^{r}\langle u\rangle_{I^f_{(m-1)r+k}}.$$

To accurately substitute the analytic averaging operator, we define the discrete time averaging operator

$$\langle \mathbf{v}_f\rangle_m^r := \frac{1}{r}\sum_{k=1}^{r} v^f_{(m-1)r+k}.$$

Additionally, we define the error quantities

$$d^{\text{aver}}_{L^\infty L^2}(\mathbf{v}_f, \mathbf{v}_c) := \mathbb{E}\left[\max_{m=1,\ldots,M_c} \left\| \langle \mathbf{v}_f\rangle_m^r - v_m^c \right\|^2_{L^2_x}\right], \tag{5.12a}$$

$$d^{\text{point}}_{L^\infty L^2}(\mathbf{v}_f, \mathbf{v}_c) := \mathbb{E}\left[\max_{m=1,\ldots,M_c} \left\| v^f_{mr} - v_m^c \right\|^2_{L^2_x}\right], \tag{5.12b}$$

and

$$d^{\text{classic}}_{L^2 V}(\mathbf{v}_f, \mathbf{v}_c) := \mathbb{E}\left[\sum_{m=1}^{M_c}\frac{\tau_c}{r}\sum_{k=1}^{r}\left\| V(\nabla v^f_{(m-1)r+k}) - V(\nabla v_m^c)\right\|^2_{L^2_x}\right], \tag{5.13a}$$

$$d^{\text{inner}}_{L^2 V}(\mathbf{v}_f, \mathbf{v}_c) := \mathbb{E}\left[\sum_{m=1}^{M_c}\tau_c \left\| V(\nabla\langle \mathbf{v}_f\rangle_m^r) - V(\nabla v_m^c)\right\|^2_{L^2_x}\right], \tag{5.13b}$$

$$d^{\text{outer}}_{L^2 V}(\mathbf{v}_f, \mathbf{v}_c) := \mathbb{E}\left[\sum_{m=1}^{M_c}\tau_c \left\| \langle V(\nabla \mathbf{v}_f)\rangle_m^r - V(\nabla v_m^c)\right\|^2_{L^2_x}\right]. \tag{5.13c}$$

5.3 Joint sampling on fine and coarse scales

It is crucial to use the same stochastic input when computing $\mathbf{v}_f$ and $\mathbf{v}_c$. This can be done in two different ways. Either one first samples coarse stochastic data and a posteriori samples the fine data based on the conditional probabilities of the coarse one, or we can sample the fine stochastic input and try to reconstruct the coarse stochastic data. The latter approach is more suitable for the averaged increments.

Lemma 37 *Let $M_f, M_c \in \mathbb{N}$. Assume $M_f = rM_c$ for some $r \in \mathbb{N}$. Then*

$$\Delta_1^c \mathbb{W} = \sum_{l=1}^{r} \left(1 - \frac{l-1}{r}\right) \Delta_l^f \mathbb{W}, \tag{5.14a}$$

$$\Delta_j^c \mathbb{W} = \sum_{l=0}^{r-1} \frac{l+1}{r} \Delta_{rj-l}^f \mathbb{W} + \sum_{l=0}^{r-2} \left(1 - \frac{l+1}{r}\right) \Delta_{r(j-1)-l}^f \mathbb{W}, \tag{5.14b}$$

for $j \in \{2, \ldots, M_c\}$.

Proof Since $M_f = rM_c$, we have $\tau_c = r\tau_f$. Thus,

$$\begin{aligned}
\Delta_1^c \mathbb{W} &= \frac{1}{\tau_c} \int_0^{\tau_c} W_s \, \mathrm{d}s = \frac{1}{r\tau_f} \int_{0\tau_f}^{r\tau_f} W_s \, \mathrm{d}s = \frac{1}{r\tau_f} \sum_{l=1}^{r} \int_{(l-1)\tau_f}^{l\tau_f} W_s \, \mathrm{d}s \\
&= \frac{1}{r\tau_f} \sum_{l=1}^{r} \int_{(l-1)\tau_f}^{l\tau_f} W_s - W_{s-(l-1)\tau_f} \, \mathrm{d}s + \frac{1}{r\tau_f} \sum_{l=1}^{r} \int_{(l-1)\tau_f}^{l\tau_f} W_{s-(l-1)\tau_f} \, \mathrm{d}s \\
&=: \mathrm{I} + \mathrm{II}.
\end{aligned}$$

Due to the discrete Fubini's theorem

$$\begin{aligned}
\mathrm{I} &= \frac{1}{r\tau_f} \sum_{l=1}^{r} \int_{(l-1)\tau_f}^{l\tau_f} \sum_{l'=0}^{l-2} W_{s-l'\tau_f} - W_{s-(l'+1)\tau_f} \, \mathrm{d}s \\
&= \frac{1}{r\tau_f} \sum_{l=1}^{r} \sum_{l'=0}^{l-2} \int_{(l-l'-1)\tau_f}^{(l-l')\tau_f} W_s - W_{s-\tau_f} \, \mathrm{d}s \\
&= \frac{1}{r} \sum_{l=1}^{r} \sum_{l'=0}^{l-2} \Delta_{l-l'}^f \mathbb{W} = \sum_{l=2}^{r} \left(1 + \frac{1-l}{r}\right) \Delta_l^f \mathbb{W}.
\end{aligned}$$

The second term is easily computed

$$\mathrm{II} = \frac{1}{r\tau_f} \sum_{l=1}^{r} \int_0^{\tau_f} W_s \, \mathrm{d}s = \Delta_1^f \mathbb{W}.$$

Overall,

$$\Delta_1^c \mathbb{W} = \sum_{l=1}^{r} \left(1 + \frac{1-l}{r}\right) \Delta_l^f \mathbb{W}.$$

Let $j \in \{2, \dots, M_c\}$. Since $t_j^c = t_{rj}^f$, it holds

$$\begin{aligned}
\Delta_j^c \mathbb{W} &= \frac{1}{\tau_c} \int_{t_{j-1}^c}^{t_j^c} W_s - W_{s-\tau_c} \,\mathrm{d}s = \frac{1}{r\tau_f} \int_{t_{r(j-1)}^f}^{t_{rj}^f} W_s - W_{s-r\tau_f} \,\mathrm{d}s \\
&= \frac{1}{r\tau_f} \sum_{l=0}^{r-1} \int_{t_{rj-(l+1)}^f}^{t_{rj-l}^f} \sum_{l'=0}^{r-1} W_{s-l'\tau_f} - W_{s-(l'+1)\tau_f} \,\mathrm{d}s \\
&= \frac{1}{r\tau_f} \sum_{l=0}^{r-1} \sum_{l'=0}^{r-1} \int_{t_{rj-(l+l'+1)}^f}^{t_{rj-(l+l')}^f} W_s - W_{s-\tau_f} \,\mathrm{d}s = \frac{1}{r} \sum_{l=0}^{r-1} \sum_{l'=0}^{r-1} \Delta_{rj-(l+l')}^f \mathbb{W}.
\end{aligned}$$

The discrete Fubini's theorem implies

$$\sum_{l=0}^{r-1} \sum_{l'=0}^{r-1} \Delta_{rj-(l+l')}^f \mathbb{W} = \sum_{l=0}^{r-1} (l+1) \Delta_{rj-l}^f \mathbb{W} + \sum_{l=0}^{r-2} (r-(l+1)) \Delta_{r(j-1)-l}^f \mathbb{W}.$$

Therefore,

$$\Delta_j^c \mathbb{W} = \sum_{l=0}^{r-1} \frac{l+1}{r} \Delta_{rj-l}^f \mathbb{W} + \sum_{l=0}^{r-2} \left(1 - \frac{l+1}{r}\right) \Delta_{r(j-1)-l}^f \mathbb{W}.$$

The claim is proved. □

Remark 38 The reconstruction formula (5.14) is the key ingredient, why it is possible to compare fine and coarse numerical solutions in an efficient way. If one tries to establish a corresponding formula for randomized algorithms as proposed in [18], this task becomes more challenging.

5.4 Simulation: unknown solution

Let $\mathcal{O} = (0, 1)^2$ and $T = 1$. We choose $p \in \{1.5, 3\}$, $u_0(x) = \sin(\pi x_1)\sin(\pi x_2)$ and

$$G(u)\Delta W := \underbrace{\sin(\pi x_1) x_2 u}_{=: g_1(x,u)} \Delta\beta^1 + \underbrace{\sin(\pi x_2) x_1 u}_{=: g_2(x,u)} \Delta\beta^2. \tag{5.15}$$

V_h denotes the space of piece wise linear elements with zero boundary values on a triangulation $\mathcal{T}_h$ of $\mathcal{O}$. We initialize the coarse triangulation $\mathcal{T}_{h_c}$ as a uniform triangulation of $\mathcal{O}$ such that $\left|V_{h_c}\right| = 121$ and $h_c \approx 1.4 * 10^{-1}$. $\mathcal{T}_{h_f}$ is generated by three uniform refinements of $\mathcal{T}_{h_c}$. Then $\left|V_{h_f}\right| = 6561$ and $h_f \approx 1.7 * 10^{-2}$. For the time discretization we use $M_f = 1280$ and $M_c = 40$. Therefore $\tau_f \approx 7.8 * 10^{-4}$ and $\tau_c \approx 2.5 * 10^{-2}$. We measure the error of the fine numerical solution $\mathbf{v}_f \in \{\mathbf{v}_f^{\mathrm{EM}}, \mathbf{v}_f^{\mathrm{Half}}, \mathbf{v}_f^{\mathrm{Full}}\}$ versus the coarse numerical solution $\mathbf{v}_c \in \{\mathbf{v}_c^{\mathrm{EM}}, \mathbf{v}_c^{\mathrm{Half}}, \mathbf{v}_c^{\mathrm{Full}}\}$ of the same algorithm in the error

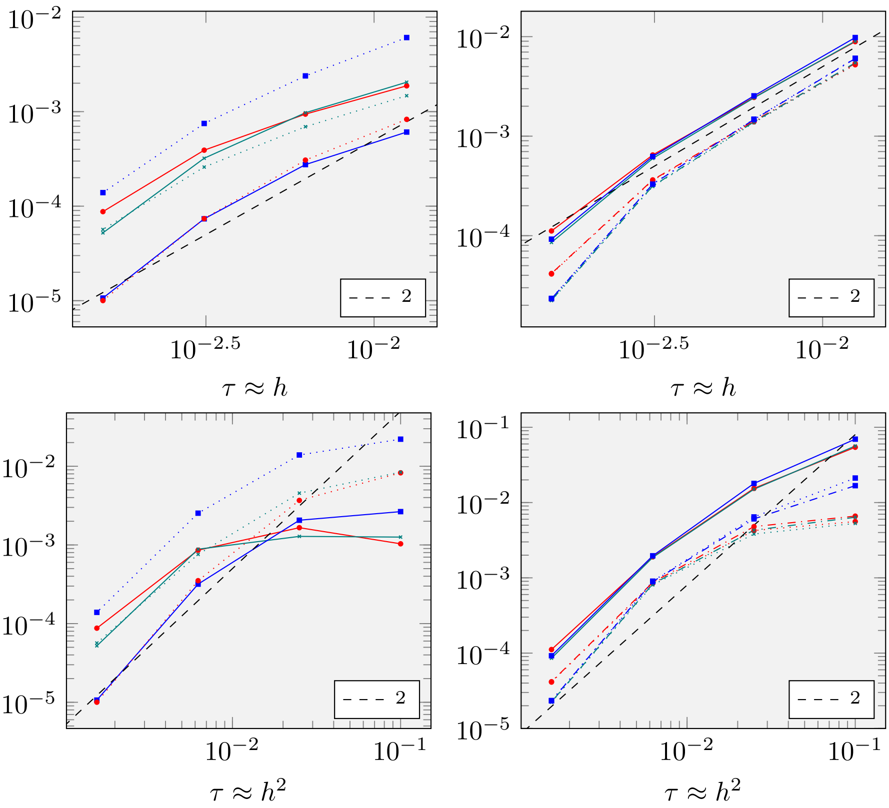

Fig. 2 Convergence for $p = 3$ of $\mathbf{v}_c^{\mathrm{EM}}$ (red), $\mathbf{v}_c^{\mathrm{Half}}$ (blue) and $\mathbf{v}_c^{\mathrm{Full}}$ (green) towards $\mathbf{v}_f$ measured in $d_{L^\infty L^2}^{\mathrm{aver}}$ (solid, left), $d_{L^\infty L^2}^{\mathrm{point}}$ (dashed, left), $d_{L^2 V}^{\mathrm{classic}}$ (solid, right), $d_{L^2 V}^{\mathrm{inner}}$ (dashed, right) and $d_{L^2 V}^{\mathrm{outer}}$ (dash dotted, right). In the top row we use $\tau \approx h$ and in the bottom row $\tau \approx h^2$ (color figure online)

quantities (5.12) and (5.13). The expectation is approximated by the Monte-Carlo method with 20 samples.

In Fig. 2 respectively Fig. 3 we plot the evolution of the error for $p = 3$ respectively $p = 1.5$. In both cases we observe linear convergence. This indicates that on the used discretization scale the space error dominates the time error. We do not see a substantial difference in the gradient error quantities. Additionally, $\mathbf{v}^{\mathrm{EM}}$ measured in the point distance $d_{L^\infty L^2}^{\mathrm{point}}$ and $\mathbf{v}^{\mathrm{Half}}$ measured in the averaged distance $d_{L^\infty L^2}^{\mathrm{aver}}$ perform equally well.

If $p = 3$ then $\mathbf{v}^{\mathrm{Full}}$ behaves similarly in both error terms and performs slightly worse than $\mathbf{v}^{\mathrm{EM}}$ and $\mathbf{v}^{\mathrm{Half}}$. Contrary in the case $p = 1.5$, while still performing slightly worse compared to $\mathbf{v}^{\mathrm{EM}}$ and $\mathbf{v}^{\mathrm{Half}}$, $\mathbf{v}^{\mathrm{Full}}$ is better approximated in $d_{L^\infty L^2}^{\mathrm{aver}}$ than in $d_{L^\infty L^2}^{\mathrm{point}}$. This indicates that at least in the singular setting $\mathbf{v}^{\mathrm{Full}}$ also approximates average values.

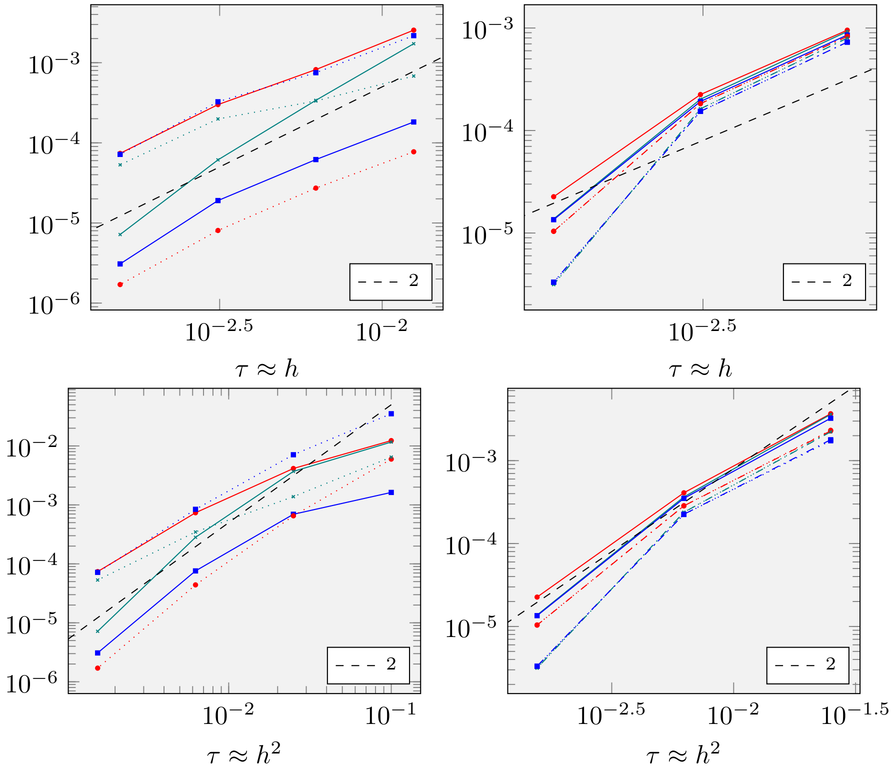

Fig. 3 Convergence for $p = 1.5$ of $\mathbf{v}_c^{\mathrm{EM}}$ (red), $\mathbf{v}_c^{\mathrm{Half}}$ (blue) and $\mathbf{v}_c^{\mathrm{Full}}$ (green) towards $\mathbf{v}_f$ measured in $d^{\mathrm{aver}}_{L^\infty L^2}$ (solid, left), $d^{\mathrm{point}}_{L^\infty L^2}$ (dashed, left), $d^{\mathrm{classic}}_{L^2 V}$ (solid, right), $d^{\mathrm{inner}}_{L^2 V}$ (dashed, right) and $d^{\mathrm{outer}}_{L^2 V}$ (dash dotted, right). In the top row we use $\tau \approx h$ and in the bottom row $\tau \approx h^2$ (color figure online)

5.5 Conclusion

Our algorithms achieve optimal linear convergence in space and optimal 1/2-convergence in time with minimal regularity assumptions. Experimentally, the Euler–Maruyama scheme (5.1) and our algorithms (3.14) and (3.15) share the same computational complexity and accuracy. However, the convergence of the Euler–Maruyama scheme is only proven with rate 1/4, cf. [39]. The construction of the random inputs is fast and can be done with the simple sampling algorithm (4.9).

Funding Open Access funding enabled and organized by Projekt DEAL.

References

1. Andersson, A., Larsson, S.: Weak convergence for a spatial approximation of the nonlinear stochastic heat equation. Math. Comput. **85**(299), 1335–1358 (2016). https://doi.org/10.1090/mcom/3016
2. Balci, A.K., Cianchi, A., Diening, L., Maz'ya, V.: A pointwise differential inequality and second-order regularity for nonlinear elliptic systems. In: Mathematische Annalen (Aug. 2021). https://doi.org/10.1007/s00208-021-02249-9
3. Balci, A.K., Diening, L., Weimar, M.: Higher order Calderón–Zygmund estimates for the p-Laplace equation. J. Differ. Equ. **268**(2), 590–635 (2020). https://doi.org/10.1016/j.jde.2019.08.009
4. Baldi, P.: Stochastic Calculus. Universitext. An Introduction Through Theory and Exercises, pp. xiv+627. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-62226-2
5. Baňas, L'., Röckner, M., Wilke, A.: Convergent numerical approximation of the stochastic total variation flow. Stoch. Partial Differ. Equ. Anal. Comput. **9**(2), 437–471 (2021). https://doi.org/10.1007/s40072-020-00169-4
6. Barrett, J.W., Liu, W.B.: Finite element approximation of the p-Laplacian. Math. Comput. **61**(204), 523–537 (1993). https://doi.org/10.2307/2153239
7. Barrett, J.W., Liu, W.B.: Finite element approximation of the parabolic p-Laplacian. SIAM J. Numer. Anal. **31**(2), 413–428 (1994). https://doi.org/10.1137/0731022
8. Bartels, S., Diening, L., Nochetto, R.H.: Unconditional stability of semiimplicit discretizations of singular flows. SIAM J. Numer. Anal. **56**(3), 1896–1914 (2018). https://doi.org/10.1137/17M1159166
9. Bauzet, C., Vallet, G., Wittbold, P., Zimmermann, A.: On a $p(t, x)$-Laplace evolution equation with a stochastic force. Stoch. Partial Differ. Equ. Anal. Comput. **1**(3), 552–570 (2013). https://doi.org/10.1007/s40072-013-0017-z
10. Belenki, L., Diening, L., Kreuzer, C.: Optimality of an adaptive finite element method for the p-Laplacian equation. IMA J. Numer. Anal. **32**(2), 484–510 (2012). https://doi.org/10.1093/imanum/drr016
11. Berselli, L.C., Růžička, M.: Space–time discretization for nonlinear parabolic systems with p-structure. IMA J. Numer. Anal. **42**(1), 260–299 (2022) https://doi.org/10.1093/imanum/draa079
12. Bertini, L., Cancrini, N.: The stochastic heat equation: Feynman–Kac formula and intermittence. J. Stat. Phys. **78**(5–6), 1377–1401 (1995). https://doi.org/10.1007/BF02180136
13. Bojarski, B., Iwaniec, T.: p-Harmonic equation and quasiregular mappings. Banach Center Publications. **1**(19), 25–38 (1987)
14. Breit, D.: Existence theory for stochastic power law fluids. J. Math. Fluid Mech. **17**(2), 295–326 (2015). https://doi.org/10.1007/s00021-015-0203-z
15. Breit, D.: Regularity theory for nonlinear systems of SPDEs. Manuscr. Math. **146**(3–4), 329–349 (2015). https://doi.org/10.1007/s00229-014-0704-8
16. Breit, D., Diening, L., Storn, J., Wichmann, J.: The parabolic p-Laplacian with fractional differentiability. IMA J. Numer. Anal. **41**(3), 2110–2138 (2021). https://doi.org/10.1093/imanum/draa081
17. Breit, D., Gmeineder, F.: Electro-rheological fluids under random influences: martingale and strong solutions. Stoch. Partial Differ. Equ. Anal. Comput. **7**(4), 699–745 (2019). https://doi.org/10.1007/s40072-019-00138-6
18. Breit, D., Hofmanová, M., Loisel, S.: Space–time approximation of stochastic p-Laplace-type systems. SIAM J. Numer. Anal. **59**(4), 2218–2236 (2021). https://doi.org/10.1137/20M1334310
19. Breit, D., Mensah, P.R.: Space–time approximation of parabolic systems with variable growth. IMA J. Numer. Anal. (2019). https://doi.org/10.1093/imanum/drz039
20. Cianchi, A., Maz'ya, V.G.: Second-order two-sided estimates in nonlinear elliptic problems. Arch. Ration. Mech. Anal. **229**(2), 569–599 (2018). https://doi.org/10.1007/s00205-018-1223-7
21. Cianchi, A., Maz'ya, V.G.: Optimal second-order regularity for the p-Laplace system. J. Math. Pures Appl. **9**(132), 41–78 (2019). https://doi.org/10.1016/j.matpur.2019.02.015
22. Cianchi, A., Maz'ya, V.G.: Second-order regularity for parabolic p-Laplace problems. J. Geom. Anal. (2020). https://doi.org/10.1007/s12220-019-00213-3

23. Dareiotis, K., Gerencsér, M.: On the regularisation of the noise for the Euler–Maruyama scheme with irregular drift. Electron. J. Probab. **25**, Paper No. 82, 18 (2020). https://doi.org/10.1214/20-ejp479
24. Debussche, A., Printems, J.: Weak order for the discretization of the stochastic heat equation. Math. Comput. **78**(266), 845–863 (2009). https://doi.org/10.1090/S0025-5718-08-02184-4
25. Di Pietro, D.A., Droniou, J., Harnist, A.: Improved error estimates for hybrid high-order discretizations of Leray–Lions problems. Calcolo **58**(2), Paper No. 19, 24 (2021). https://doi.org/10.1007/s10092-021-00410-z
26. DiBenedetto, E.: $C^{1+\alpha}$ local regularity of weak solutions of degenerate elliptic equations. Nonlinear Anal. **7**(8), 827–850 (1983). https://doi.org/10.1016/0362-546X(83)90061-5
27. Diening, L., Fornasier, M., Tomasi, R., Wank, M.: A relaxed Kacanov iteration for the p-Poisson problem. Numer. Math. **145**(1), 1–34 (2020). https://doi.org/10.1007/s00211-020-01107-1
28. Diening, L., Kaplický, P., Schwarzacher, S.: BMO estimates for the p-Laplacian. Nonlinear Anal. **75**(2), 637–650 (2012). https://doi.org/10.1016/j.na.2011.08.065
29. Diening, L., Růžička, M.: Interpolation operators in Orlicz–Sobolev spaces. Numer. Math. **107**(1), 107–129 (2007). https://doi.org/10.1007/s00211-007-0079-9
30. Diening, L., Ebmeyer, C., Růžička, M.: Optimal convergence for the implicit space–time discretization of parabolic systems with p-structure. SIAM J. Numer. Anal. **45**(2), 457–472 (2007). https://doi.org/10.1137/05064120X
31. Diening, L., Ettwein, F.: Fractional estimates for non-differentiable elliptic systems with general growth. Forum Math. **20**(3), 523–556 (2008). https://doi.org/10.1515/FORUM.2008.027
32. Diening, L., Harjulehto, P., Hästö, P., Růžička, M.: Lebesgue and Sobolev Spaces with Variable Exponents. Vol. 2017. Lecture Notes in Mathematics, pp. x+509. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-18363-8
33. Diening, L., Kreuzer, C.: Linear convergence of an adaptive finite element method for the p-Laplacian equation. SIAM J. Numer. Anal. **46**(2), 614–638 (2008). https://doi.org/10.1137/070681508
34. Diening, L., Stroffolini, B., Verde, A.: Everywhere regularity of functionals with φ-growth. Manuscr. Math. **129**(4), 449–481 (2009). https://doi.org/10.1007/s00229-009-0277-0
35. Droniou, J., Goldys, B., Le, K.-N.: Design and convergence analysis of numerical methods for stochastic evolution equations with Leray–Lions operator. IMA J. Numer. Anal. **42**(2), 1143–1179 (2022). https://doi.org/10.1093/imanum/draa105
36. Dunst, T., Prohl, A.: The forward-backward stochastic heat equation: numerical analysis and simulation. SIAM J. Sci. Comput. **38**(5), A2725–A2755 (2016). https://doi.org/10.1137/15M1022951
37. Ebmeyer, C., Liu, W.: Quasi-norm interpolation error estimates for the piecewise linear finite element approximation of p-Laplacian problems. Numer. Math. **100**(2), 233–258 (2005). https://doi.org/10.1007/s00211-005-0594-5
38. Eisenmann, M., Kovács, M., Kruse, R., Larsson, S.: On a randomized backward Euler method for nonlinear evolution equations with time-irregular coefficients. Found. Comput. Math. **19**(6), 1387–1430 (2019). https://doi.org/10.1007/s10208-018-09412-w
39. Eisenmann, M., Kovács, M., Kruse, R., Larsson, S.: Error estimates of the backward Euler–Maruyama method for multi-valued stochastic differential equations. BIT Numer. Math. **62**(3), 803–848 (2022). https://doi.org/10.1007/s10543-021-00893-w
40. Emmrich, E., Šiška, D.: Nonlinear stochastic evolution equations of second order with damping. Stoch. Partial Differ. Equ. Anal. Comput. **5**(1), 81–112 (2017). https://doi.org/10.1007/s40072-016-0082-1
41. Gajewski, H., Gröger, K., Zacharias, K.: Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen. Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38. Akademie, Berlin, ix+281 pp. (loose errata) (1974)
42. Geissert, M., Kovács, M., Larsson, S.: Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise. BIT **49**(2), 343–356 (2009). https://doi.org/10.1007/s10543-009-0227-y
43. Gess, B.: Strong solutions for stochastic partial differential equations of gradient type. J. Funct. Anal. **263**(8), 2355–2383 (2012). https://doi.org/10.1016/j.jfa.2012.07.001
44. Gess, B., Tölle, J.M.: Ergodicity and local limits for stochastic local and nonlocal p-Laplace equations. SIAM J. Math. Anal. **48**(6), 4094–4125 (2016). https://doi.org/10.1137/15M1049774
45. Gyöngy, I., Millet, A.: On discretization schemes for stochastic evolution equations. Potential Anal. **23**(2), 99–134 (2005). https://doi.org/10.1007/s11118-004-5393-6
46. Gyöngy, I., Millet, A.: Rate of convergence of space time approximations for stochastic evolution equations. Potential Anal. **30**(1), 29–64 (2009). https://doi.org/10.1007/s11118-008-9105-5

47. Hairer, M., Labbé, C.: Multiplicative stochastic heat equations on the whole space. J. Eur. Math. Soc.: JEMS **20**(4), 1005–1054 (2018). https://doi.org/10.4171/JEMS/781
48. Hu, Y., Huang, J., Nualart, D., Tindel, S.: Stochastic heat equations with general multiplicative Gaussian noises: Hölder continuity and intermittency. Electron. J. Probab. **20**(55), 50 (2015). https://doi.org/10.1214/EJP.v20-3316
49. Hutzenthaler, M., Jentzen, A., Kloeden, P.E.: Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **467**(2130), 1563–1576 (2011). https://doi.org/10.1098/rspa.2010.0348
50. Hytönen, T.P., Veraar, M.C.: On Besov regularity of Brownian motions in infinite dimensions. Probab. Math. Stat. **28**(1), 143–162 (2008)
51. Iwaniec, T., Manfredi, J.J.: Regularity of p-harmonic functions on the plane. Rev. Mat. Iberoam. **5**(1–2), 1–19 (1989). https://doi.org/10.4171/RMI/82
52. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Vol. 23. Applications of Mathematics (New York). Springer, Berlin, pp. xxxvi+632 (1992). https://doi.org/10.1007/978-3-662-12616-5
53. Kuusi, T., Mingione, G.: Linear potentials in nonlinear potential theory. Arch. Ration. Mech. Anal. **207**(1), 215–246 (2013). https://doi.org/10.1007/s00205-012-0562-z
54. Kuusi, T., Mingione, G.: Vectorial nonlinear potential theory. J. Eur. Math. Soc.: JEMS **20**(4), 929–1004 (2018). https://doi.org/10.4171/JEMS/780
55. Lê, K., Ling, C.: Taming singular stochastic differential equations: a numerical method. (2021). https://doi.org/10.48550/ARXIV.2110.01343
56. Lieberman, G.M.: Boundary regularity for solutions of degenerate elliptic equations. Nonlinear Anal. **12**(11), 1203–1219 (1988). https://doi.org/10.1016/0362-546X(88)90053-3
57. Lions, J.-L.: Quelques méthodes de résolution des problèmes aux limites non linéaires, pp. xx+554. Dunod, Paris (1969)
58. Liu, W.B., Barrett, J.W.: A further remark on the regularity of the solutions of the p-Laplacian and its applications to their finite element approximation. Nonlinear Anal. **21**(5), 379–387 (1993). https://doi.org/10.1016/0362-546X(93)90081-3
59. Liu, W.B., Barrett, J.W.: A remark on the regularity of the solutions of the p-Laplacian and its application to their finite element approximation. J. Math. Anal. Appl. **178**(2), 470–487 (1993). https://doi.org/10.1006/jmaa.1993.1319
60. Liu, W.: On the stochastic p-Laplace equation. J. Math. Anal. Appl. **360**(2), 737–751 (2009). https://doi.org/10.1016/j.jmaa.2009.07.020
61. Liu, W., Röckner, M.: SPDE in Hilbert space with locally monotone coefficients. J. Funct. Anal. **259**(11), 2902–2922 (2010). https://doi.org/10.1016/j.jfa.2010.05.012
62. Logg, A., Mardal, K.-A., Wells, G.N. (eds.): Automated Solution of Differential Equations by the Finite Element Method. Vol. 84. Lecture Notes in Computational Science and Engineering. The FEniCS book, pp. xiv+723. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-23099-8
63. Mansuy, R., Yor, M.: Aspects of Brownian Motion. Universitext. Springer, Berlin, pp. xiv+195 (2008). https://doi.org/10.1007/978-3-540-49966-4
64. Marcellini, P., Papi, G.: Nonlinear elliptic systems with general growth. J. Differ. Equ. **221**(2), 412–443 (2006). https://doi.org/10.1016/j.jde.2004.11.011
65. Ondreját, M., Prohl, A., Walkington, N.J.: Numerical approximation of nonlinear SPDE's. In: Stochastics and Partial Differential Equations: Analysis and Computations (2022). https://doi.org/10.1007/s40072-022-00271-9
66. Prévôt, C., Röckner, M.: A Concise Course on Stochastic Partial Differential Equations. Vol. 1905. Lecture Notes in Mathematics, pp. vi+144. Springer, Berlin (2007)
67. Prohl, A., Wang, Y.: Strong rates of convergence for a space–time discretization of the backward stochastic heat equation, and of a linear-quadratic control problem for the stochastic heat equation. ESAIM Control Optim. Calc. Var. **27**, Paper No. 54, 30 (2021). https://doi.org/10.1051/cocv/2021052
68. Sapountzoglou, N., Zimmermann, A.: Well-posedness of renormalized solutions for a stochastic p-Laplace equation with L^1-initial data. Discrete Contin. Dyn. Syst. **41**(5), 2341–2376 (2021). https://doi.org/10.3934/dcds.2020367
69. Terasawa, Y., Yoshida, N.: Stochastic power law fluids: existence and uniqueness of weak solutions. Ann. Appl. Probab. **21**(5), 1827–1859 (2011). https://doi.org/10.1214/10-AAP741

70. Tolksdorf, P.: Regularity for a more general class of quasilinear elliptic equations. J. Differ. Equ. **51**(1), 126–150 (1984). https://doi.org/10.1016/0022-0396(84)90105-0
71. Uhlenbeck, K.: Regularity for a class of non-linear elliptic systems. Acta Math. **138**(3–4), 219–240 (1977). https://doi.org/10.1007/BF02392316
72. Ural'ceva, N.N.: Degenerate quasilinear elliptic systems. Zap. Naucn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **7**, 184–222 (1968)
73. Vallet, G., Wittbold, P., Zimmermann, A.: On a stochastic $p(w, t, x)$-Laplace equation. In: Thirteenth International Conference Zaragoza-Pau on Mathematics and its Applications. Vol. 40. Monogr. Mat. García Galdeano. Prensas Univ. Zaragoza, Zaragoza, pp. 125–134 (2016)
74. Vallet, G., Zimmermann, A.: The stochastic $p(w, t, x)$-Laplace equation with cylindrical Wiener process. J. Math. Anal. Appl. **444**(2), 1359–1371 (2016). https://doi.org/10.1016/j.jmaa.2016.07.018
75. Wei, D.: Existence, uniqueness, and numerical analysis of solutions of a quasilinear parabolic problem. SIAM J. Numer. Anal. **29**(2), 484–497 (1992). https://doi.org/10.1137/0729029
76. Wichmann, J.: On temporal regularity for strong solutions to stochastic *p*-Laplace systems (2021). arXiv: 2111.09601 [math.AP]