DESIGN AND CONVERGENCE ANALYSIS OF NUMERICAL METHODS FOR STOCHASTIC EVOLUTION EQUATIONS WITH LERAY–LIONS OPERATOR

JÉRÔME DRONIOU, BENIAMIN GOLDYS, AND KIM-NGAN LE

ABSTRACT. The gradient discretisation method (GDM) is a generic framework, covering many classical methods (Finite Elements, Finite Volumes, Discontinuous Galerkin, etc.), for designing and analysing numerical schemes for diffusion models. In this paper, we study the GDM for a general stochastic evolution problem based on a Leray–Lions type operator. The problem contains the stochastic p-Laplace equation as a particular case. The convergence of the Gradient Scheme (GS) solutions is proved by using Discrete Functional Analysis techniques, Skorohod theorem and the Kolmogorov test. In particular, we provide an independent proof of the existence of weak martingale solutions for the problem. In this way, we lay foundations and provide techniques for proving convergence of the GS approximating stochastic partial differential equations.

KEYWORDS: p-Laplace equation, stochastic PDE, numerical methods, gradient discretisation method, convergence analysis

1. INTRODUCTION

The parabolic p-Laplacian problem occurs in many mathematical models of physical processes, such as nonlinear diffusion [1] and non-Newtonian flows [32]. However, in practical situations with large scales, rapid velocity and pressure fluctuations, the motion of flow becomes unsteady and it is described as being turbulent [34]. Turbulence is a combination of slow oscillating (deterministic) component and fast oscillating component that can be modelled as a white noise perturbation of regular fluid velocity field. Therefore, in order to investigate turbulence in the parabolic p-Laplacian problem, the first step is to develop the theory and numerical algorithms for the stochastic parabolic p-Laplacian problem. Motivated by this problem, we study in this paper a more general stochastic partial differential equation based on a Leray–Lions type operator with homogeneous Dirichlet boundary condition. The model reads

$$
\begin{aligned}
du - \operatorname{div}(a(u,\nabla u))dt &= f(u)dW_t \quad \text{in } (0,T)\times\Theta,\\
u(0,\cdot) &= u_0 \quad \text{in } \Theta,\\
u &= 0 \quad \text{on } (0,T)\times\partial\Theta,
\end{aligned} \tag{1}
$$

where $T > 0$, Θ is an open bounded domain in $\mathbb{R}^d$, $d = 1,2,3$, and the initial data $u_0 \in L^2(\Theta)$. Here, f is a continuous function of linear growth acting between appropriate Banach spaces, see Section 2 for details. We assume that $W = \{W(t), t \in [0,T]\}$ is a $\mathcal{K}$-valued Wiener process with a trace class covariance

Date: April 25, 2020.

operator $\mathcal{Q}$, for a certain Hilbert space $\mathcal{K}$. Particular choices of a include the p-Laplace operator corresponding to $a(u, \boldsymbol{v}) = |\boldsymbol{v}|^{p-1}\boldsymbol{v}$ for some $p \in (1, +\infty)$, and some nonlinear and nonlocal diffusion operators corresponding to $a(u, \boldsymbol{v}) = \Lambda(u)\boldsymbol{v}$ for u in a given functional space. If $f = 0$ then the noise term vanishes, hence our stochastic model (1) includes deterministic equation as a special case.

The existence and uniqueness theory for equation (1) is well developed at present, see [8, 15, 25, 26, 33], but our assumption that $a(u, \nabla u)$ may be a nonlinear function of both variables is more general than in the aformentioned papers. The methods of proof used in those works do not provide convenient numerical algorithms, while we provide a proof of the existence of weak martingale solutions via a converging sequence of numerical approximations.

Numerical methods of the deterministic version of model (1) (i.e. $f = 0$) and their proofs of convergence are studied in [6, 12, 20, 29] and the references cited therein. However, there is no numerical approximation of the stochastic model (1) due to difficulties arising in the nonlinear term and the infinite dimensional nature of the driving noise processes.

There are an increasing number of numerical methods for the solution of stochastic evolution equations mentioned in the literature [30, 31, 35], where unique mild solutions are required and the approximate schemes are treated in terms of the semi-group approach. However, these assumptions are not applicable for a class of stochastic equations involving strongly nonlinear terms, such as Navier–Stokes, magnetohydrodynamics (MHD), Strödinger, Landau–Lifshitz–Gilbert, Landau–Lifshitz–Bloch and nonlinear porous media equations. In recent works, the stochastic Navier-Stokes equation [10, 11] and the stochastic Landau–Lifshitz–Gilbert equation [4, 5, 23, 24] are investigated by using the conforming finite element method to approximate their solutions. Furthermore, the convergence of the approximate solutions is also proved which implies the existence of weak martingale solutions.

All these previous work, however, only deal with conforming approximations, which use for the spatial discretisation a subspace of the Sobolev space appearing in the weak formulation of the continuous problem. This usually imposes restrictions on the types of mesh that can be considered – typically, triangular/tetrahedral or quadrangular/hexahedral meshes. Moreover, conforming methods are know to be ill-suited in some applications, e.g. when mesh locking appears, when inf-sup stability is sought, or when some physical properties of the model must be respected (such as balance and conservativity of approximate fluxes). In such circumstances, non-conforming methods might be better suited; such methods include non-conforming finite elements and finite volume methods, and also recent high-order methods for polytopal meshes with cell and face unknowns – such as Hybrid-High Order schemes and Virtual Element Methods. We refer the reader to [2, 7, 13, 16, 18] and reference therein for detailed presentations of these methods.

In this work, we approximate (1) by using the Gradient Discretisation Method (GDM) [19]. The GDM is a generic convergence analysis framework for a wide variety of methods (conforming or nonconforming) written in discrete variational formulation, and based on independent approximations of functions and gradients using the same degrees of freedom. Several well-known methods fall in the GDM framework, in particular:

- Galerkin methods, including the conforming Finite Element methods,

- Nonconforming Finite Element methods, including the nonconforming $\mathbb{P}_1$ scheme,
- Symmetric Interior Penalty Galerkin (SIPG) methods,
- Mixed Finite Element methods,
- Hybrid Mimetic Mixed methods and Mimetic Finite Difference methods,
- Hybrid High-Order and Virtual Elements Methods.

By writing numerical schemes for (1) and performing their analysis in the GDM framework, we provide a unified convergence result for all these methods. We refer to [17, 19, 21, 22] for details of the GDM and the methods it covers.
Our convergence analysis is based on Discrete Functional Analysis techniques, Skorohod theorem and the Kolmogorov test; we show the convergence of the Gradient Scheme (GS) solutions to a weak martingale solution of (1). In particular, we provide an independent proof of the existence of weak martingale solutions for the problem. In this way, we lay foundations and provide techniques for proving convergence of the GS approximating stochastic partial differential equations.
The paper is organised as follows. In Section 2 we recall the notations of the gradient discretisation method and propose the GS for approximating the stochastic model (1). Weak martingale solutions to (1) are defined and our main result is stated in this section. Section 3 provides priori estimates of approximated solutions and the noise term added at each step of the scheme in various norms. In Section 4, we first show the tightness of the sequence including the GS solutions and then prove the almost sure convergence to a certain limit, up to a change of probability space. The continuity of the limit and the martingale part are also proved in this section. Section 5 is devoted to the proof of the main theorem. Finally, in the Appendix we prove necessary results that are used in the course of the proof.

2. Gradient scheme and main resutls

Before introducing the GS for approximation of (1), we introduce notations and assumptions used in the rest of the paper.
Notations: We let $p' = \frac{p}{p-1}$ be the conjugate exponent of p. To alleviate the formulas, when written without specifying the space, the Lebesgue spaces we consider are those on Θ; so, most of the time, we write L^q instead of $L^q(\Theta)$. Correspondingly, $\|\cdot\|_{L^q}$ is the norm in $L^q(\Theta)$, $\langle\cdot,\cdot\rangle_{L^{p'},L^p}$ is the duality product between $L^{p'}(\Theta)$ and $L^p(\Theta)$, and $\langle\cdot,\cdot\rangle_{L^2}$ the inner product in $L^2(\Theta)$; we use the same notations in vector-valued Lebesgues spaces $L^q(\Theta)^e$ for $e \geqslant 2$. We will use the notation $\Theta_T := (0,T) \times \Theta$. In proofs of theorems and lemmas, C will stand for a generic constant that depends only on the data above, and on any constant appearing in the statement of the corresponding theorem or lemma.

Assumptions: The following Leray–Lions type standing assumptions will not be enunciated again:

- the initial condition u_0 belongs to L^2,
- the function $a : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ is continuous,
- for $p \in (1, +\infty)$ there exist constants c_1, c_2 such that for all $(x, \boldsymbol{y}) \in \mathbb{R} \times \mathbb{R}^d$ and $\boldsymbol{z} \in \mathbb{R}^d$

$$a(x, \boldsymbol{y}) \cdot \boldsymbol{y} \geqslant c_1 |\boldsymbol{y}|^p \tag{2}$$

$$|a(x, \boldsymbol{y})| \leqslant c_2(1 + |\boldsymbol{y}|^{p-1}) \tag{3}$$

$$(a(x, \boldsymbol{y}) - a(x, \boldsymbol{z})) \cdot (\boldsymbol{y} - \boldsymbol{z}) \geqslant 0, \tag{4}$$

- the function $f : L^p \cap L^2 \to \mathcal{L}(\mathcal{K}, L^2)$ is continuous with linear growth, i.e., there exist constants $F_1, F_2 > 0$ such that for any $v \in L^p \cap L^2$ and any sequence $\{w_n\}_{n\geqslant 0}$ which converges to w in L^p

$$\begin{aligned} &\|f(v)\|^2_{\mathcal{L}(\mathcal{K},L^2)} \leqslant F_1\|v\|^2_{L^2} + F_2 \quad \text{ and} \\ &f(w_n) \to f(w) \ \text{ in } \mathcal{L}(\mathcal{K}, L^2). \end{aligned} \tag{5}$$

- $T > 0$ is a given constant and $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ is a stochastic basis, that is $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $\mathbb{F}$ is a filtration.

2.1. **Gradient scheme.** We recall here the notions of the gradient discretisation method. The idea of this general analysis framework is to replace, in the weak formulation of the problem, the continuous space and operators by discrete ones; the set of discrete space and operators is called a gradient discretisation (GD), and the scheme obtained after substituting these elements into the weak formulation is called a gradient scheme (GS). The convergence of the obtained GS can be established based on only a few general concepts on the underlying GD. Moreover, different GDs correspond to different classical schemes (finite elements, finite volumes, etc.). Hence, the analysis carried out in the GDM directly applies to all these schemes, and does not rely on the specificity of each particular method.

Definition 2.1. $\mathcal{D} = \left(X_{\mathcal{D},0}, \Pi_{\mathcal{D}}, \nabla_{\mathcal{D}}, \mathcal{I}_{\mathcal{D}}, \left(t^{(n)}\right)_{n=0,\cdots,N}\right)$ *is a space-time gradient discretisation for homogeneous Dirichlet boundary conditions, if its elements satisfy the following properties*

(i) $X_{\mathcal{D},0}$ *is a finite dimensional vector space of functions of discrete argument and* $X_{\mathcal{D},0}$ *encodes homogeneous Dirichlet boundary conditions.*
(ii) the function reconstruction $\Pi_{\mathcal{D}} : X_{\mathcal{D},0} \to L^\infty$ *is a linear mapping that reconstructs, from an element of* $X_{\mathcal{D},0}$*, a function over* Θ*,*
(iii) the linear mapping $\nabla_{\mathcal{D}} : X_{\mathcal{D},0} \to (L^p)^d$ *gives a reconstructed discrete gradient. It must be chosen in such a way that* $\|\nabla_{\mathcal{D}} \cdot\|_{L^p}$ *is a norm on* $X_{\mathcal{D},0}$*,*
(iv) $\mathcal{I}_{\mathcal{D}} : L^2 \to X_{\mathcal{D},0}$ *is an interpolation operator,*
(v) $t^{(0)} = 0 < t^{(1)} < \cdots < t^{(N)} = T$ *is a uniform time discretisation in the sense that* $\delta t_{\mathcal{D}} := t^{(n+1)} - t^{(n)}$ *is a constant time step.*

For any $\left(v^{(n)}\right)_{n=0,\cdots,N} \subset X_{\mathcal{D},0}$, we define piecewise-constant-in-time functions $\Pi_{\mathcal{D}}v : [0,T] \to L^\infty$, $\nabla_{\mathcal{D}}v : (0,T] \to (L^p)^d$ and $d_{\mathcal{D}}v : (0,T] \to L^\infty$ by: For $n = 0, \cdots, N-1$, for any $t \in (t^{(n)}, t^{(n+1)}]$, for a.e. $\boldsymbol{x} \in \Theta$

$$\Pi_{\mathcal{D}}v(0,\boldsymbol{x}) := \Pi_{\mathcal{D}}v^{(0)}(\boldsymbol{x}), \qquad \Pi_{\mathcal{D}}v(t,\boldsymbol{x}) := \Pi_{\mathcal{D}}v^{(n+1)}(\boldsymbol{x}),$$
$$\nabla_{\mathcal{D}}v(t,\boldsymbol{x}) := \nabla_{\mathcal{D}}v^{(n+1)}(\boldsymbol{x}), \qquad d_{\mathcal{D}}v(t) = d_{\mathcal{D}}^{(n+\frac{1}{2})}v := \Pi_{\mathcal{D}}v^{(n+1)} - \Pi_{\mathcal{D}}v^{(n)}.$$

We now describe the scheme.

Algorithm 2.2 (Gradient scheme for (1))**.** *Set* $u^{(0)} := \mathcal{I}_{\mathcal{D}}u_0$ *and take random variables* $u(\cdot) = \left(u^{(n)}(\omega,\cdot)\right)_{n=0,\cdots,N} \subset X_{\mathcal{D},0}$ *such that:*

- *u is adapted to the filtration* $(\mathcal{F}_N^n)_{0\leqslant n\leqslant N}$ *defined by*

$$\mathcal{F}_N^n := \sigma\{W(t^{(k)}), 0 \leqslant k \leqslant n\}.$$

- *for any function* $\phi \in X_{\mathcal{D},0}$ *and any* $\omega \in \Omega$,

$$\left\langle d_{\mathcal{D}}^{(n+\frac{1}{2})}u, \Pi_{\mathcal{D}}\phi \right\rangle_{L^2} + \delta t_{\mathcal{D}}\langle a(\Pi_{\mathcal{D}}u^{(n+1)}, \nabla_{\mathcal{D}}u^{(n+1)}), \nabla_{\mathcal{D}}\phi\rangle_{L^{p'},L^p} = \left\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W, \Pi_{\mathcal{D}}\phi \right\rangle_{L^2}. \quad (6)$$

Here $\Delta^{(n+1)}W := W(t^{(n+1)}) - W(t^{(n)})$.

In order to establish the stability and convergence of GS (6), sequences of space-time gradient discretisations $(\mathcal{D}_m)_{m\in\mathbb{N}}$ are required to satisfy *consistency, limit-conformity* and *compactness* properties [19]. The consistency is slightly adapted here to account for the non-linearity we consider. In the following, we let $\widehat{p} = \max\{2, p'\}$.

Definition 2.3 (Consistency). *A sequence* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations in the sense of Definition 2.1 is said to be consistent if*

- *for all* $\phi \in L^{\widehat{p}}(\Theta) \cap W_0^{1,p}(\Theta)$, *letting*

$$S_{\mathcal{D}_m}(\phi) := \min_{w\in X_{\mathcal{D}_m}} \left(\|\Pi_{\mathcal{D}_m}w - \phi\|_{L^{\widehat{p}}} + \|\nabla_{\mathcal{D}_m}w - \nabla\phi\|_{L^p} \right),$$

we have $S_{\mathcal{D}_m}(\phi) \to 0$ *as* $m \to \infty$,
- *for all* $\phi \in L^2$, $\Pi_{\mathcal{D}_m}\mathcal{I}_{\mathcal{D}_m}\phi \to \phi$ *in* L^2 *as* $m \to \infty$
- $\delta t_{\mathcal{D}_m} \to 0$ *as* $m \to \infty$.

It follows from the consistency property that there exists a constant $C_{u_0} > 0$ not depending on m such that

$$\|\Pi_{\mathcal{D}_m}u^{(0)}\|_{L^2} \leqslant C_{u_0}. \quad (7)$$

Definition 2.4 (Limit-conformity). *A sequence* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations in the sense of Definition 2.1 is said to be limit-conforming if, for all* $\boldsymbol{\phi} \in W^{\mathrm{div},p'}(\Theta) := \{\boldsymbol{\phi} \in L^{p'}(\Theta)^d \,:\, \mathrm{div}\boldsymbol{\phi} \in L^{p'}(\Theta)\}$ *letting*

$$W_{\mathcal{D}_m}(\boldsymbol{\phi}) := \max_{v\in X_{\mathcal{D}_m}\setminus\{0\}} \frac{\left| \int_\Omega \left(\nabla_{\mathcal{D}_m}v(\boldsymbol{x}) \cdot \boldsymbol{\phi}(\boldsymbol{x}) + \Pi_{\mathcal{D}_m}v(\boldsymbol{x})\mathrm{div}\boldsymbol{\phi}(\boldsymbol{x}) \right) d\boldsymbol{x} \right|}{\|\nabla_{\mathcal{D}_m}v\|_{L^p}},$$

we have $W_{\mathcal{D}_m}(\boldsymbol{\phi}) \to 0$ *as* $m \to \infty$.

Definition 2.5 (Compactness). *A sequence* $(\mathcal{D}_m)_{m\in\mathbb{N}}$ *of space-time gradient discretisations in the sense of Definition 2.1 is said to be compact if*

$$\lim_{\boldsymbol{\xi}\to 0} \sup_{m\in\mathbb{N}} T_{\mathcal{D}_m}(\boldsymbol{\xi}) = 0,$$

where

$$T_{\mathcal{D}_m}(\boldsymbol{\xi}) := \max_{v\in X_{\mathcal{D}_m}\setminus\{0\}} \frac{\|\Pi_{\mathcal{D}_m}v(\cdot + \boldsymbol{\xi}) - \Pi_{\mathcal{D}_m}v\|_{L^p(\mathbb{R}^d)}}{\|\nabla_{\mathcal{D}_m}v\|_{L^p}}, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d,$$

with $\Pi_{\mathcal{D}_m}v$ *extended by* 0 *outside* Θ.

Remark 2.6. *The definition we use here is often considered as a characterisation of the compactness of GDs, see* [19, Lemma 2.21].

A sequence of GDs that is compact also satisfies another important property: the coercivity [19, Lemma 2.10].

Lemma 2.7 (Coercivity of sequences of GDs). *If a sequence $(\mathcal{D}_m)_{m\in\mathbb{N}}$ of space-time gradient discretisations in the sense of Definition 2.1 is compact, then it is coercive: there exists a constant C_p such that*

$$\max_{v\in X_{\mathcal{D}_m}\backslash\{0\}} \frac{\|\Pi_{\mathcal{D}_m} v\|_{L^p}}{\|\nabla_{\mathcal{D}_m} v\|_{L^p}} \leqslant C_p, \quad \forall m \in \mathbb{N}.$$

Finally, we will need sequences of GDs that satisfy the following discrete Sobolev embeddings. As shown in [19], and especially in Appendix B therein, such embeddings are known for all classical gradient discretisations.

Definition 2.8 (Discrete Sobolev embeddings). *A sequence of gradient discretisations $(\mathcal{D}_m)_{m\in\mathbb{N}}$ satisfies the discrete Sobolev embeddings if there exists $p^* > p$ and $C \geqslant 0$ such that, for all $m \in \mathbb{N}$ and all $v_m \in X_{\mathcal{D}_m,0}$, it holds $\|\Pi_{\mathcal{D}_m} v\|_{L^{p^*}} \leqslant C\|\nabla_{\mathcal{D}_m} v\|_{L^p}$.*

2.2. **Main results.** We first define a weak martingale solution to (1).

Definition 2.9. *Given $T \in (0,\infty)$, a weak martingale solution $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}, W, u)$ to* (1) *consists of*

- *(a) a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with the filtration satisfying the usual (normal) conditions* [14, page 71],
- *(b) a $\mathcal{K}$-valued $\mathbb{F}$-adapted Wiener process with the covariance operator $\mathcal{Q}$,*
- *(c) a progressively measurable process $u : [0,T] \times \Omega \to L^p$*

such that

- *(1) There is a ball $B_{\rm w}$ of L^2, endowed with the weak topology, such that, $\mathbb{P}$-a.s. $\omega \in \Omega$, $u(\cdot,\omega) \in C([0,T]; B_{\rm w})$.*
- *(2) $\mathbb{E}\left(\sup_{t\in[0,T]} \|u(t)\|^2_{L^2}\right) < \infty$;*
- *(3) $\mathbb{E}\left(\|u\|^p_{L^p(0,T;W^{1,p}_0(\Theta))}\right) < \infty$;*
- *(4) for every $t \in [0,T]$, for all $\psi \in W^{1,p}_0(\Theta) \cap L^{\widehat{p}}(\Theta)$, $\mathbb{P}$-a.s.:*

$$\left\langle u(t), \psi\right\rangle_{L^2} - \left\langle u_0, \psi\right\rangle_{L^2} + \int_0^t \left\langle a(u, \nabla u(s)), \nabla\psi\right\rangle_{L^{p'},L^p} ds = \left\langle \int_0^t f(u)(s,\cdot) dW(s), \psi\right\rangle_{L^2}.$$

Remark 2.10 (Continuity of the solution). *The weak continuity of $u(\omega,\cdot) : [0,T] \to B_{\rm w}$ implies its continuity $[0,T] \to H^{-1}(\Theta)$ for the standard norm topology on $H^{-1}(\Theta)$.*

The main result of this paper is the following theorem, which states the existence of a solution to the GS and its convergence, up to a subsequence, towards a weak martingale solution of the continuous problem.

Theorem 2.11. *Assume that we are given an initial data $u_0 \in L^2(\Theta)$ and $T > 0$. Let $(\mathcal{D}_m)_{m\in\mathbb{N}}$ be a sequence of gradient discretisations that is consistent, limit-conforming, compact, and satisfies the discrete Sobolev embeddings. For every $m \geqslant 1$, there exists random process u_m solution to the gradient scheme (Algorithm 2.2 with $\mathcal{D} := \mathcal{D}_m$).*
Moreover, there exists a weak martingale solution $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, (\widetilde{\mathcal{F}}_t)_{t\in[0,T]}, \widetilde{\mathbb{P}}, \widetilde{W}, \widetilde{u})$ to (1) *in the sense of Definition 2.9, and a sequence $\{\widetilde{u}_m\}$ of random processes defined on*

$\widetilde{\Omega}$ with the same law as u_m, so that up to a subsequence, the following convergences hold

$$\Pi_{\mathcal{D}_m}\widetilde{u}_m \to \widetilde{u}, \quad \widetilde{\mathbb{P}} - a.s. \text{ in } L^p(\Theta_T)$$
$$\nabla_{\mathcal{D}_m}\widetilde{u}_m \to \nabla\widetilde{u}, \quad \widetilde{\mathbb{P}} - a.s. \text{ in } \big(L^p(\Theta_T)\big)_{\mathrm{w}}.$$

Remark 2.12. *The existence of a weak solution to* (1) *is obtained as a by-product of the convergence analysis. This existence is not assumed* a priori*, and no regularity property is required on the continuous solution to get the convergence of the GDM.*

3. A priori estimates

We first provide a priori estimates for the solution u to (6) and then deduce its existence in the following lemma. For legibility, we drop the index m in sequences of gradient discretisations, and we simply write $\mathcal{D}$ instead of $\mathcal{D}_m$.

Lemma 3.1. *There exists at least one $u_{\mathcal{D}}$ solution to the Algorithm 2.2 and there exists a constant $C_{f,a,T,\mathcal{Q},u_0} > 0$ depending only on $f, a, T, \mathcal{Q}$ and u_0 such that*

$$\mathbb{E}\left[\max_{1\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2} + \|\nabla_{\mathcal{D}}u\|^p_{L^p(\Theta_T)} + \sum_{n=0}^{N-1}\|\Pi_{\mathcal{D}}u^{(n+1)} - \Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\right] \leqslant C_{f,a,T,\mathcal{Q},u_0}. \quad (8)$$

We also have for any integer number $q \geqslant 1$

$$\mathbb{E}\left[\max_{1\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^q}_{L^2} + \|\nabla_{\mathcal{D}}u\|^{p2^{q-1}}_{L^p(\Theta_T)}\right] \leqslant C_{f,a,T,\mathcal{Q},u_0,q}. \quad (9)$$

Proof.

<u>A priori estimates on $\Pi_{\mathcal{D}}u$ in (8).</u>

We first prove a priori energy estimates of solution u. We choose the test function $\phi = u^{(n+1)} \in X_{\mathcal{D},0}$ in (6) and use the following fundamental identity

$$(a-b)a = \frac{1}{2}(a^2 - b^2) + \frac{1}{2}(a-b)^2, \quad \forall a, b \in \mathbb{R}, \quad (10)$$

to write

$$\begin{aligned}\frac{1}{2}\|\Pi_{\mathcal{D}}u^{(n+1)}\|^2_{L^2} &+ \frac{1}{2}\|\Pi_{\mathcal{D}}u^{(n+1)} - \Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2} \\ &+ \delta t_{\mathcal{D}}\big\langle a(\Pi_{\mathcal{D}}u^{(n+1)}, \nabla_{\mathcal{D}}u^{(n+1)}),\, \nabla_{\mathcal{D}}u^{(n+1)}\big\rangle_{L^{p'},L^p} \\ &= \frac{1}{2}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2} + \big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}}(u^{(n+1)} - u^{(n)})\big\rangle_{L^2} \\ &\quad + \big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2}. \end{aligned} \quad (11)$$

By taking the sum in the above equation from $n = 0$ to $n = k$, for an arbitrary $k \in \{0, \ldots, N-1\}$, and using (2), Cauchy–Schwarz inequality and the Young inequality $ab \leqslant a^2 + \frac{b^2}{4}$ for the second term in the right hand side, we obtain

$$\begin{aligned}\frac{1}{2}\|\Pi_{\mathcal{D}}u^{(k+1)}\|^2_{L^2} &+ \frac{1}{4}\sum_{n=0}^{k}\|\Pi_{\mathcal{D}}u^{(n+1)} - \Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2} + c_1\sum_{n=0}^{k}\delta t_{\mathcal{D}}\|\nabla_{\mathcal{D}}u^{(n+1)}\|^p_{L^p} \\ &\leqslant \frac{1}{2}\|\Pi_{\mathcal{D}}u^{(0)}\|^2_{L^2} + \sum_{n=0}^{k}\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}\end{aligned}$$

$$+\sum_{n=0}^{k}\big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W,\,\Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2}. \tag{12}$$

Note that the last term on the right hand side of (12) vanishes when taking its expectation since $\Pi_{\mathcal{D}}u^{(n)}$ is $\mathcal{F}_{t^{(n)}}$ measurable, and thus independent with $\Delta^{(n+1)}W$ which has a zero expectation. We obtain from (12)

$$\begin{aligned}\frac{1}{2}\mathbb{E}\Big[\|\Pi_{\mathcal{D}}u^{(k+1)}\|^2_{L^2}+\frac{1}{4}\sum_{n=0}^{k}\|\Pi_{\mathcal{D}}u^{(n+1)}-\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\Big]&\\ +\mathbb{E}\Big[c_1\sum_{n=0}^{k}\delta t_{\mathcal{D}}\|\nabla_{\mathcal{D}}u^{(n+1)}\|^p_{L^p}\Big]&\\ \leqslant\frac{1}{2}\|\Pi_{\mathcal{D}}u^{(0)}\|^2_{L^2}+\sum_{n=0}^{k}\mathbb{E}\big[\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}\big].&\end{aligned} \tag{13}$$

By the tower property of the conditional expectation, the independence of the increments of the Wiener process, and assumptin on f we find for the last term

$$\begin{aligned}\mathbb{E}\big[\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}\big]&\\ &=\mathbb{E}\Big[\mathbb{E}\big[\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}|\mathcal{F}_{t^{(n)}}\big]\Big]\\ &=\mathbb{E}\Big[\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\mathbb{E}\big[\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}|\mathcal{F}_{t^{(n)}}\big]\Big]\\ &=\delta t_{\mathcal{D}}(\mathrm{Tr}\mathcal{Q})\mathbb{E}[\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}]\\ &\leqslant\delta t_{\mathcal{D}}(\mathrm{Tr}\mathcal{Q})\big(F_1\mathbb{E}[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}]+F_2\big).\end{aligned} \tag{14}$$

Together with (13), this implies

$$\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(k+1)}\|^2_{L^2}\big]\leqslant\|\Pi_{\mathcal{D}}u^{(0)}\|^2_{L^2}+2(\mathrm{Tr}\mathcal{Q})F_2T+2(\mathrm{Tr}\mathcal{Q})F_1\sum_{n=0}^{k}\delta t_{\mathcal{D}}\mathbb{E}[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}].$$

By applying the discrete version of Gronwall's lemma to the above inequality and using (7), we obtain

$$\max_{1\leqslant n\leqslant N}\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]\leqslant C_{f,a,T,\mathcal{Q},u_0}. \tag{15}$$

It follows from (13)–(15) that

$$\mathbb{E}\big[\|\nabla_{\mathcal{D}}u\|^p_{L^p(\Theta_T)}+\sum_{n=0}^{N-1}\|\Pi_{\mathcal{D}}u^{(n+1)}-\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]\leqslant C_{f,T,\mathcal{Q},u_0}.$$

By taking the maximum of (12) over $0\leqslant k\leqslant N-1$ and appying the expectations, we get

$$\begin{aligned}\mathbb{E}\big[\max_{1\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]\leqslant\|\Pi_{\mathcal{D}}u^{(0)}\|^2_{L^2}+2\mathbb{E}\big[\sum_{n=0}^{N-1}\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}\big]&\\ +2\mathbb{E}\big[\max_{0\leqslant k\leqslant N-1}\sum_{n=0}^{k}\big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W,\,\Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2}\big].&\end{aligned} \tag{16}$$

To bound the last term in the right hand side, we treat the sum as the stochastic integral of a piecewise constant integrand and use the Burkholder–Davis–Gundy inequality: [9, Theorem 2.4]

$$
\begin{aligned}
&\mathbb{E}\Big[\max_{0\leqslant k\leqslant N-1}\sum_{n=0}^{k}\big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W,\,\Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2}\Big]\\
&\quad\leqslant C\mathbb{E}\Big[\big(\sum_{n=0}^{N-1}\delta t_{\mathcal{D}}\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big)^{1/2}\Big]\\
&\quad\leqslant C\mathbb{E}\Big[\max_{0\leqslant n\leqslant N-1}\|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}\big(\sum_{n=0}^{N-1}\delta t_{\mathcal{D}}(F_1\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}+F_2)\big)^{1/2}\Big]\\
&\quad\leqslant \frac{1}{4}\mathbb{E}\big[\max_{0\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]+C^2F_1\sum_{n=0}^{N-1}\delta t_{\mathcal{D}}\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]+C^2F_2T\\
&\quad\leqslant \frac{1}{4}\mathbb{E}\big[\max_{0\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]+C^2F_1T\max_{0\leqslant n\leqslant N}\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]+C^2F_2T. \qquad (17)
\end{aligned}
$$

We use (14) to bound the second term in the right hand side of (16).

$$
\begin{aligned}
&\mathbb{E}\Big[\sum_{n=0}^{N-1}\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}\Big]\\
&\qquad\leqslant (\mathrm{Tr}\mathcal{Q})F_1\sum_{n=0}^{N-1}\delta t_{\mathcal{D}}\mathbb{E}[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}]+(\mathrm{Tr}\mathcal{Q})F_2T\\
&\qquad\leqslant (\mathrm{Tr}\mathcal{Q})F_1T\max_{1\leqslant n\leqslant N}\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]+(\mathrm{Tr}\mathcal{Q})F_2T. \qquad (18)
\end{aligned}
$$

By using (15), (17) and (18), we deduce from (16) that

$$
\mathbb{E}\Big[\max_{1\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\Big]\leqslant C_{f,a,T,\mathcal{Q},u_0},
$$

which completes the proof of the a priori estimates (8).
The existence of at least one solution to the Algorithm 2.2 is then done as in the proof of [19, Theorem 2.44].

Higher moments bound (9).
We adapt the ideas from [10], where different type of difficulties had to be dealt with.
We will use induction to proof this result. First, from (8) we have the assertion for $q = 1$. We assume therefore that (9) holds for any integer number $\bar{q} \in [1, q-1]$, that is,

$$
\mathbb{E}\big[\max_{1\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^{\bar{q}}}_{L^2}\big]\leqslant C_{f,a,T,\mathcal{Q},u_0,\bar{q}}. \qquad (19)
$$

In what follow, we will prove that (19) holds for $\bar{q} = q$. We begin by multiplying identity (11) by $\|\Pi_{\mathcal{D}}u^{(n+1)}\|^2_{L^2}$ and use the positive-definiteness (2) of a to obtain

$$
\begin{aligned}
&\frac{1}{2}\|\Pi_{\mathcal{D}}u^{(n+1)}\|^2_{L^2}\big(\|\Pi_{\mathcal{D}}u^{(n+1)}\|^2_{L^2}-\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big)\\
&\qquad+\frac{1}{4}\|\Pi_{\mathcal{D}}u^{(n+1)}\|^2_{L^2}\|\Pi_{\mathcal{D}}u^{(n+1)}-\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\leqslant I_1+I_2, \qquad (20)
\end{aligned}
$$

where

$$I_1 := \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 \big\langle f(\Pi_{\mathcal{D}} u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}}(u^{(n+1)} - u^{(n)})\big\rangle_{L^2},$$
$$I_2 := \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 \big\langle f(\Pi_{\mathcal{D}} u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}} u^{(n)}\big\rangle_{L^2}.$$

By using the Cauchy–Schwarz and Young inequalities, we estimate I_1 and I_2 as follow

$$\begin{aligned}
I_1 &\leqslant \|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 \\
&\quad + \frac{1}{4}\|\Pi_{\mathcal{D}} u^{(n+1)} - \Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 \\
&= \|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \big[\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 + \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 - \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2\big] \\
&\quad + \frac{1}{4}\|\Pi_{\mathcal{D}} u^{(n+1)} - \Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 \\
&\leqslant \|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \\
&\quad + 4\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^4 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^4 + \frac{1}{16}\left(\|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 - \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2\right)^2 \\
&\quad + \frac{1}{4}\|\Pi_{\mathcal{D}} u^{(n+1)} - \Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2,
\end{aligned}$$

and

$$\begin{aligned}
I_2 &= \big\langle f(\Pi_{\mathcal{D}} u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}} u^{(n)}\big\rangle_{L^2} \big[\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 + \|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 - \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2\big] \\
&\leqslant \big\langle f(\Pi_{\mathcal{D}} u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}} u^{(n)}\big\rangle_{L^2} \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \\
&\quad + 4\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \\
&\quad + \frac{1}{16}\left(\|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^2 - \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2\right)^2.
\end{aligned}$$

By using the above estimates together with (10), we infer from (20) that

$$\begin{aligned}
\frac{1}{4}\|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^4 - \frac{1}{4}\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^4 &\leqslant 5\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \\
&\quad + 4\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^4 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^4 \\
&\quad + \big\langle f(\Pi_{\mathcal{D}} u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}} u^{(n)}\big\rangle_{L^2} \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2. \qquad (21)
\end{aligned}$$

Using (10) and (21), it is easily proved by induction on q (the inductive step from q to $q+1$ consisting in multiplying this estimate by $\|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^{2^q}$) that

$$\begin{aligned}
&\frac{1}{2^q}\|\Pi_{\mathcal{D}} u^{(n+1)}\|_{L^2}^{2^q} - \frac{1}{2^q}\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q} \\
&\qquad \leqslant 5\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q-2} \\
&\qquad\quad + 4\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^4 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^4 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q-4} \\
&\qquad\quad + \big\langle f(\Pi_{\mathcal{D}} u^{(n)})\Delta^{(n+1)}W,\, \Pi_{\mathcal{D}} u^{(n)}\big\rangle_{L^2} \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q-2}. \qquad (22)
\end{aligned}$$

Then, proceeding as in (14), the first two terms in the right hand side of (21) are estimated as follow

$$\begin{aligned}
&\mathbb{E}\big[\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q-2}\big] \\
&\qquad \leqslant \delta t_{\mathcal{D}}(\mathrm{Tr}\mathcal{Q})\mathbb{E}\big[(F_1\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 + F_2)\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q-2}\big], \qquad (23)
\end{aligned}$$

$$\mathbb{E}\big[\|f(\Pi_{\mathcal{D}} u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^4 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^4 \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^{2^q-4}\big]$$

$$\leqslant \delta_{\mathcal{D}}^2 (\mathrm{Tr}\mathcal{Q})^2 \mathbb{E}\Big[(F_1\|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^2 + F_2)^2 \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q-4}\Big]. \quad (24)$$

We note that last term on the right hand side of (22) vanishes when taking expectation. Hence, summing (22) from $n = 0$ to $n = k$ (for an arbitrary $k = 0, \dots, N-1$), taking the expectations and using (15), the above estimates, and the discrete version of Gronwall lemma, we obtain

$$\max_{1\leqslant n\leqslant N} \mathbb{E}\Big[\|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q}\Big] \leqslant C_{f,a,T,\mathcal{Q},u_0,q}. \quad (25)$$

By summing (22) from $n = 0$ to $n = k$ (for an arbitrary $k = 0, \dots, N-1$), and taking the maximum over k and then applying $\mathbb{E}$, we get

$$\begin{aligned}
\mathbb{E}\Big[\max_{1\leqslant n\leqslant N} \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q}\Big] &\leqslant \|\Pi_{\mathcal{D}}u^{(0)}\|_{L^2}^{2^q} \\
&+ 20\mathbb{E}\Big[\sum_{n=0}^{N-1} \|f(\Pi_{\mathcal{D}}u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2 \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q-2}\Big] \\
&+ 16\mathbb{E}\Big[\sum_{n=0}^{N-1} \|f(\Pi_{\mathcal{D}}u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^4 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^4 \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q-4}\Big] \\
&+ 4\mathbb{E}\Big[\max_{0\leqslant k\leqslant N-1} \sum_{n=0}^{k} \big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W, \Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2} \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q-2}\Big]. \quad (26)
\end{aligned}$$

Proceeding as in (17), the last term of the right hand side is estimated as follow

$$\begin{aligned}
&\mathbb{E}\Big[\max_{0\leqslant k\leqslant N-1} \sum_{n=0}^{k} \big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W, \Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2} \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q-2}\Big] \\
&\qquad \leqslant \frac{1}{2^{q+1}}\mathbb{E}\Big[\max_{0\leqslant n\leqslant N} \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q}\Big] + CF_1T \max_{0\leqslant n\leqslant N} \mathbb{E}\Big[\|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q}\Big] \\
&\qquad\quad + CF_2T \max_{0\leqslant n\leqslant N} \mathbb{E}\Big[\|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q-2}\Big].
\end{aligned}$$

By using the above inequality, (23)–(25) and (15), we obtain from (26) that

$$\mathbb{E}\Big[\max_{1\leqslant n\leqslant N} \|\Pi_{\mathcal{D}}u^{(n)}\|_{L^2}^{2^q}\Big] \leqslant C_{f,a,T,\mathcal{Q},u_0,q}, \quad (27)$$

which completes the proof of the inductive step.

A priori estimates on $\nabla_{\mathcal{D}}u$ in (9).

By using Jensen's inequality, we obtain from (12) with $k = N-1$ that

$$\begin{aligned}
\|\nabla_{\mathcal{D}}u\|_{L^p(\Theta_T)}^{p2^{q-1}} &\leqslant C_q \|\Pi_{\mathcal{D}}u^{(0)}\|_{L^2}^{2^q} \\
&+ C_q \left(\sum_{n=0}^{N-1} \|f(\Pi_{\mathcal{D}}u^{(n)})\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \|\Delta^{(n+1)}W\|_{\mathcal{K}}^2\right)^{2^{q-1}} \\
&+ C_q \left(\sum_{n=0}^{N-1} \big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W, \Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2}\right)^{2^{q-1}}. \quad (28)
\end{aligned}$$

We estimate the second term in the right hand side of (28) by using, for $\gamma \geqslant 1$,

$$\Big(\sum_{n=0}^{N-1} a_n\Big)^{\gamma} \leqslant N^{\gamma-1} \sum_{n=0}^{N-1} a_n^{\gamma}, \quad (29)$$

which can be proved using Jensen's inequality on the sum. Applying the above inequality, arguments used in the proof of (14) and invoking (27), we have

$$
\begin{aligned}
&\mathbb{E}\Big[\Big(\sum_{n=0}^{N-1}\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Delta^{(n+1)}W\|^2_{\mathcal{K}}\Big)^{2^{q-1}}\Big]\\
&\leqslant C_{f,\mathcal{Q}}N^{2^{q-1}-1}\delta\!t_{\mathcal{D}}^{2^{q-1}}\sum_{n=0}^{N-1}\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^{q-1}}_{L^2}+\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^q}_{L^2}\big]\\
&\leqslant C_{f,a,T,\mathcal{Q},u_0,q},
\end{aligned}
\tag{30}
$$

where we have used the inequality $\mathbb{E}[(\|\Delta^{(n+1)}W\|^2_{\mathcal{K}})^r]\leqslant C_{\mathcal{Q},r}(\delta\!t_{\mathcal{D}})^r$ for all integer $r\geqslant 1$, see [27, Corollary 1.1]. Proceeding as (17) and using (29), (27), we estimate the third term in the right hand side of (28):

$$
\begin{aligned}
&\mathbb{E}\big[\big(\sum_{n=0}^{N-1}\big\langle f(\Pi_{\mathcal{D}}u^{(n)})\Delta^{(n+1)}W,\,\Pi_{\mathcal{D}}u^{(n)}\big\rangle_{L^2}\big)^{2^{q-1}}\big]\\
&\leqslant C\mathbb{E}\big[\big(\sum_{n=0}^{N-1}\delta\!t_{\mathcal{D}}\|f(\Pi_{\mathcal{D}}u^{(n)})\|^2_{\mathcal{L}(\mathcal{K},L^2)}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big)^{2^{q-2}}\big]\\
&\leqslant C\mathbb{E}\big[\max_{0\leqslant n\leqslant N-1}\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^{q-1}}_{L^2}\big(\sum_{n=0}^{N-1}\delta\!t_{\mathcal{D}}(F_1\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}+F_2)\big)^{2^{q-2}}\big]\\
&\leqslant \frac14\mathbb{E}\big[\max_{0\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^q}_{L^2}\big]+C^2\delta\!t_{\mathcal{D}}^{2^{q-1}}\mathbb{E}\big[\big(\sum_{n=0}^{N-1}F_1\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}+F_2\big)^{2^{q-1}}\big]\\
&\leqslant \frac14\mathbb{E}\big[\max_{0\leqslant n\leqslant N}\|\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]+C^2F_1T\max_{1\leqslant n\leqslant N}\mathbb{E}\big[\|\Pi_{\mathcal{D}}u^{(n)}\|^{2^q}_{L^2}\big]+C^2F_2^{2^{q-1}}T\\
&\leqslant C_{f,a,T,\mathcal{Q},u_0,q},
\end{aligned}
$$

where we have used the Burkholder–Davis–Gundy inequality in the second line, a Young inequality in the fourth line, and (29) in the fifth line. Together with (28) and (30), this implies

$$
\mathbb{E}\big[\|\nabla_{\mathcal{D}}u\|^{p2^{q-1}}_{L^p(\Theta_T)}\big]\leqslant C_{f,a,T,\mathcal{Q},u_0,q},
$$

which completes the proof of this lemma. ∎

In order to estimate the time-translate of $\Pi_{\mathcal{D}}u$, we will need the following relation.

Lemma 3.2. *Let u be a solution of the Algorithm 2.2. Then, for all $\ell\in\{1,\dots,N-1\}$,*

$$
\mathbb{E}\big[\delta\!t_{\mathcal{D}}\sum_{n=1}^{N-\ell}\|\Pi_{\mathcal{D}}u^{(n+\ell)}-\Pi_{\mathcal{D}}u^{(n)}\|^2_{L^2}\big]\leqslant C_{f,T,\mathcal{Q},p,\|\Pi_{\mathcal{D}}u^{(0)}\|_{L^2}}\,t^{(\ell)}.
$$

Proof. For any function $\phi\in X_{\mathcal{D},0}$, we deduce from (6) that

$$
\begin{aligned}
\big\langle\Pi_{\mathcal{D}}u^{(n+\ell)}-\Pi_{\mathcal{D}}u^{(n)},\Pi_{\mathcal{D}}\phi\big\rangle_{L^2}&=-\delta\!t_{\mathcal{D}}\sum_{i=0}^{\ell-1}\big\langle a(\Pi_{\mathcal{D}}u^{(n+i+1)},\nabla_{\mathcal{D}}u^{(n+i+1)}),\nabla_{\mathcal{D}}\phi\big\rangle_{L^{p'},L^p}\\
&\quad+\sum_{i=0}^{\ell-1}\big\langle f(\Pi_{\mathcal{D}}u^{(n+i)})\Delta^{(n+i+1)}W,\Pi_{\mathcal{D}}\phi\big\rangle_{L^2}.
\end{aligned}
\tag{31}
$$

Choosing $\phi = \delta t_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})$ and taking the sum over n from 1 to $N-\ell$, we have

$$
\begin{aligned}
&\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|\Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 \\
&\qquad = -\delta t_{\mathcal{D}}^2 \sum_{n=1}^{N-\ell}\sum_{i=0}^{\ell-1} \big\langle a(\Pi_{\mathcal{D}} u^{(n+i+1)}, \nabla_{\mathcal{D}} u^{(n+i+1)}), \nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\big\rangle_{L^{p'},L^p} \\
&\qquad\quad + \delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell}\sum_{i=0}^{\ell-1} \big\langle f(\Pi_{\mathcal{D}} u^{(n+i)})\Delta^{(n+i+1)} W, \Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)}\big\rangle_{L^2} \\
&\qquad =: I_1 + I_2.
\end{aligned} \tag{32}
$$

We now estimate the expectation of I_1 by using (3), Hölder inequality, and Lemma 3.1.

$$
\begin{aligned}
&\mathbb{E}[I_1] \leqslant c_2 \mathbb{E}\Big[\delta t_{\mathcal{D}}^2 \sum_{n=1}^{N-\ell}\sum_{i=0}^{\ell-1} \big\langle 1 + |\nabla_{\mathcal{D}} u^{(n+i+1)}|^{p-1}, |\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})|\big\rangle_{L^2}\Big] \\
&\leqslant C t^{(\ell)} \mathbb{E}\big[\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\|_{L^1}\big] \\
&\quad + C\mathbb{E}\Big[\delta t_{\mathcal{D}}^2 \sum_{n=1}^{N-\ell} \|\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\|_{L^p} \sum_{i=0}^{\ell-1} \|\nabla_{\mathcal{D}} u^{(n+i+1)}\|_{L^p}^{p-1}\Big] \\
&\leqslant C t^{(\ell)} \mathbb{E}\big[\int_0^T \|\nabla_{\mathcal{D}} u(t)\|_{L^1}\,dt\big] \\
&\quad + C\mathbb{E}\Big[\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\|_{L^p} \int_{t^{(n)}}^{t^{(n+\ell)}} \|\nabla_{\mathcal{D}} u(t)\|_{L^p}^{p-1}\,dt\Big] \\
&\leqslant C t^{(\ell)} \mathbb{E}\Big[\|\nabla_{\mathcal{D}} u\|_{L^p(\Theta_T)}^p\Big]^{1/p} \\
&\quad + C(t^{(\ell)})^{1/p} \mathbb{E}\Big[\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\|_{L^p} \Big(\int_{t^{(n)}}^{t^{(n+\ell)}} \|\nabla_{\mathcal{D}} u(t)\|_{L^p}^p\,dt\Big)^{1/p'}\Big].
\end{aligned} \tag{33}
$$

The second term in the right hand side is estimated as follows:

$$
\begin{aligned}
&C(t^{(\ell)})^{1/p} \mathbb{E}\Big[\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\|_{L^p} \Big(\int_{t^{(n)}}^{t^{(n+\ell)}} \|\nabla_{\mathcal{D}} u(t)\|_{L^p}^p\,dt\Big)^{1/p'}\Big] \\
&\qquad \leqslant C(t^{(\ell)})^{1/p} \mathbb{E}\Big[\delta t_{\mathcal{D}} \Big(\sum_{n=1}^{N-\ell} \|\nabla_{\mathcal{D}}(u^{(n+\ell)} - u^{(n)})\|_{L^p}^p\Big)^{1/p} \\
&\qquad\qquad \times \Big(\sum_{n=1}^{N-\ell} \int_{t^{(n)}}^{t^{(n+\ell)}} \|\nabla_{\mathcal{D}} u(t)\|_{L^p}^p\,dt\Big)^{1/p'}\Big] \\
&\qquad \leqslant C(t^{(\ell)})^{1/p} (\delta t_{\mathcal{D}} \ell)^{1/p'} \mathbb{E}\big[\|\nabla_{\mathcal{D}} u\|_{L^p(\Theta_T)}^p\big],
\end{aligned} \tag{34}
$$

where the conclusion follows by noticing that, in the last sum of integrals term in the second line, each interval $[t^{(n)}, t^{(n+1)}]$ appears at most ℓ times.

To estimate the expectation of I_2, we use the Young inequality and write

$$\begin{aligned}
\mathbb{E}[I_2] &= \delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \mathbb{E}\Big[\Big\langle \int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t) f(\Pi_{\mathcal{D}} u(t)) dW(t), \Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)} \Big\rangle_{L^2}\Big] \\
&\leqslant \frac{1}{4} \delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \mathbb{E}\big[\|\Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2\big] \\
&\quad + \delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \mathbb{E}\Big[\Big\| \int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t) f(\Pi_{\mathcal{D}} u(t)) dW(t) \Big\|_{L^2}^2\Big]
\end{aligned} \tag{35}$$

By using the Itô isometry, (5) and Lemma 3.1, we bound the last term in the right hand side:

$$\begin{aligned}
&\mathbb{E}\Big[\Big\| \int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t) f(\Pi_{\mathcal{D}} u(t)) dW(t) \Big\|_{L^2}^2\Big] \\
&\qquad \leqslant (\mathrm{Tr}\mathcal{Q}) \mathbb{E}\Big[\int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t) \|f(\Pi_{\mathcal{D}} u(t))\|_{\mathcal{L}(\mathcal{K},L^2)}^2 \, dt\Big] \\
&\qquad \leqslant (\mathrm{Tr}\mathcal{Q}) \mathbb{E}\Big[\int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t) (F_1 \|\Pi_{\mathcal{D}} u(t)\|_{L^2}^2 + F_2) \, dt\Big] \\
&\qquad \leqslant (\mathrm{Tr}\mathcal{Q}) t^{(\ell)} \mathbb{E}\big[F_1 \max_{1\leqslant n\leqslant N} \|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2 + F_2\big] \\
&\qquad \leqslant C_{f,T,\mathcal{Q},p,\|\Pi_{\mathcal{D}} u^{(0)}\|_{L^2}} t^{(\ell)}.
\end{aligned}$$

Together with (35), (34), (33) and (32), this implies

$$\mathbb{E}\Big[\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|\Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^2\Big] \leqslant C_{f,T,\mathcal{Q},p,\|\Pi_{\mathcal{D}} u^{(0)}\|_{L^2}} t^{(\ell)},$$

which completes the proof of the lemma. ∎

Remark 3.3. *The result of Lemma 3.2 will be used to obtain compactness-in-time of the approximate functions. The approach used here based on this estimate fills an apparent gap in* [3, 10] *where the result of* [3, Lemma 4.4] *(*[10, Lemma 3.2]*) is not sufficient for proving* [3, Theorem 4.6] *(*[10, Lemma 4.1], *respectively).*

We can now estimate the time-translate of $\Pi_{\mathcal{D}} u$. It follows from Lemmas 6.2 and 3.2, and estimate (76) that, for any $\rho \in (0, T)$,

$$\mathbb{E}\Big[\int_0^{T-\rho} \|\Pi_{\mathcal{D}} u(t+\rho) - \Pi_{\mathcal{D}} u(t)\|_{L^2}^2 \, dt\Big] \leqslant C\rho, \tag{36}$$

and

$$\mathbb{E}\big[\|\Pi_{\mathcal{D}} u\|_{H^\beta(0,T;L^2)}^2\big] \leqslant C, \quad \text{for any } \beta \in (0, 1/2). \tag{37}$$

In the following lemma, we estimate the dual norm of the time variation of the iterates $\{\Pi_{\mathcal{D}} u^{(n)}\}_{n=0}^N$. The dual norm $|\cdot|_{*,\mathcal{D}}$ on $\Pi_{\mathcal{D}}(X_{\mathcal{D},0}) \subset L^2$ is defined by: for all $v \in \Pi_{\mathcal{D}}(X_{\mathcal{D},0})$,

$$|v|_{*,\mathcal{D}} := \sup\Big\{ \int_\Omega v(\boldsymbol{x}) \, \Pi_{\mathcal{D}} \phi(\boldsymbol{x}) d\boldsymbol{x} \ : \ \phi \in X_{\mathcal{D},0}, \|\Pi_{\mathcal{D}} \phi\|_{L^2} + \|\nabla_{\mathcal{D}} \phi\|_{L^p} \leqslant 1 \Big\}.$$

Lemma 3.4. *For any $q \in \mathbb{N}$ let $r = 2^q$ and $\alpha = \min\{1/2, 1/p\}$. Then, for all $\ell = 1, \ldots, N-1$,*

$$\mathbb{E}\big[|\Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)}|_{*,\mathcal{D}}^r\big] \leqslant C_{f,T,\mathcal{Q},p,q,\|\Pi_{\mathcal{D}} u^{(0)}\|_{L^2}} (t^{(\ell)})^{\alpha r}. \tag{38}$$

As a consequence, for any $t, s \in [0,T]$

$$\mathbb{E}\big[|\Pi_{\mathcal{D}} u(t) - \Pi_{\mathcal{D}} u(s)|_{*,\mathcal{D}}^r\big] \leqslant C_{f,T,\mathcal{Q},p,q,\|\Pi_{\mathcal{D}} u^{(0)}\|_{L^2}} \big(|t-s| + \delta t_{\mathcal{D}}\big)^{\alpha r}. \tag{39}$$

Proof. It follows from (31) that

$$\begin{aligned}
&\mathbb{E}\big[|\Pi_{\mathcal{D}} u^{(n+\ell)} - \Pi_{\mathcal{D}} u^{(n)}|_{*,\mathcal{D}}^r\big] \\
&\quad\leqslant 2^{r-1} \delta t_{\mathcal{D}}^r \mathbb{E}\Big[\Big(\sup_{\phi\in\mathcal{A}} \sum_{i=0}^{\ell-1} \big\langle a(\Pi_{\mathcal{D}} u^{(n+i+1)}, \nabla_{\mathcal{D}} u^{(n+i+1)}), \nabla_{\mathcal{D}}\phi\big\rangle_{L^{p'},L^p}\Big)^r\Big] \\
&\qquad + 2^{r-1}\mathbb{E}\Big[\Big(\sup_{\phi\in\mathcal{A}} \sum_{i=0}^{\ell-1} \big\langle f(\Pi_{\mathcal{D}} u^{(n+i)})\Delta^{(n+i+1)}W, \Pi_{\mathcal{D}}\phi\big\rangle_{L^2}\Big)^r\Big] \\
&\quad =: I_1 + I_2,
\end{aligned} \tag{40}$$

where we have set $\mathcal{A} := \{\phi \in X_{\mathcal{D},0}, \|\Pi_{\mathcal{D}}\phi\|_{L^2} + \|\nabla_{\mathcal{D}}\phi\|_{L^p} \leqslant 1\}$. We estimate the first term I_1 by using (3) and Lemma 3.1.

$$\begin{aligned}
I_1 &\leqslant C\delta t_{\mathcal{D}}^r \mathbb{E}\Big[\sup_{\phi\in\mathcal{A}}\Big(\sum_{i=0}^{\ell-1}\big\langle 1 + |\nabla_{\mathcal{D}} u^{(n+i+1)}|^{p-1}, |\nabla_{\mathcal{D}}\phi|\big\rangle_{L^2}\Big)^r\Big] \\
&\leqslant C(t^{(\ell)})^r \mathbb{E}\big[\sup_{\phi\in\mathcal{A}}\|\nabla_{\mathcal{D}}\phi\|_{L^1}^r\big] + C\delta t_{\mathcal{D}}^r \mathbb{E}\Big[\sup_{\phi\in\mathcal{A}}\|\nabla_{\mathcal{D}}\phi\|_{L^p}^r \Big(\sum_{i=0}^{\ell-1}\|\nabla_{\mathcal{D}} u^{(n+i+1)}\|_{L^p}^{p-1}\Big)^r\Big] \\
&\leqslant C(t^{(\ell)})^r + C(t^{(\ell)})^{r/p}\mathbb{E}\Big[\Big(\int_{t^{(n)}}^{t^{(n+\ell)}} \|\nabla_{\mathcal{D}} u(t)\|_{L^p}^p\,dt\Big)^{r/p'}\Big] \\
&\leqslant C(t^{(\ell)})^r + C(t^{(\ell)})^{r/p}\mathbb{E}\big[\|\nabla_{\mathcal{D}} u\|_{L^p(\Theta_T)}^{(p-1)r}\big] \\
&\leqslant C(t^{(\ell)})^r + C(t^{(\ell)})^{r/p}\big(\mathbb{E}\big[\|\nabla_{\mathcal{D}} u\|_{L^p(\Theta_T)}^{pr}\big]\big)^{1/p'} \leqslant C(t^{(\ell)})^{r/p}.
\end{aligned} \tag{41}$$

The last term I_2 is estimated by using the Burkholder–Davis–Gundy inequality, (5) and Lemma 3.1.

$$\begin{aligned}
I_2 &\leqslant C\mathbb{E}\big[\|\int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t) f(\Pi_{\mathcal{D}} u(t))dW(t)\|_{L^2}^r\big] \\
&\leqslant C\mathbb{E}\Big[\Big(\int_0^T \mathbb{1}_{[t^{(n)},t^{(n+\ell)}]}(t)(F_1\|\Pi_{\mathcal{D}} u(t)\|_{L^2}^2 + F_2)\,dt\Big)^{r/2}\Big] \\
&\leqslant C(t^{(\ell)})^{r/2}\mathbb{E}\big[F_1^{r/2}\max_{1\leqslant n\leqslant N}\|\Pi_{\mathcal{D}} u^{(n)}\|_{L^2}^r + F_2^{r/2}\big] \leqslant C(t^{(\ell)})^{r/2}.
\end{aligned} \tag{42}$$

The estimate (38) follows from (40)–(42). The bound (39) follows by noticing that, if $t < s \in [0,T]$ and $n \leqslant r$ are such that $t \in (t^{(n)}, t^{(n+1)}]$ and $s \in (t^{(r)}, t^{(r+1)}]$, then $t^{(r-n)} \leqslant |s-t| + \delta t_{\mathcal{D}}$. ■

For any $t \in [0,T]$, there exists $n \in \{0, \cdots, N-1\}$ such that $t \in (t^{(n)}, t^{(n+1)}]$. Using this notation, we define

$$M_{\mathcal{D}}(t) := M_{\mathcal{D}}^{(n)} := \sum_{i=0}^{n} f(\Pi_{\mathcal{D}} u^{(i)})\Delta^{(i+1)}W.$$

The term $f(\Pi_{\mathcal{D}} u^{(i)})\Delta^{(i+1)}W$ corresponds to the noise term added at each time step of the GS. The following lemma shows that $M_{\mathcal{D}}$ is bounded in various norms.

Lemma 3.5. *For any $\beta \in (0, 1/2)$, for any $r = 2^q$ with $q \in \mathbb{N}$, there exists $C \geqslant 0$ such that*

$$\mathbb{E}\big[\|M_{\mathcal{D}}\|^2_{H^\beta(0,T;L^2)}\big] \leqslant C \quad \text{and} \quad \mathbb{E}\big[\|M_{\mathcal{D}}\|^r_{L^\infty(0,T;L^2)}\big] \leqslant C. \tag{43}$$

Proof. It follows, in a similar way as (42), that

$$\mathbb{E}\big[\|M_{\mathcal{D}}^{(n+\ell)} - M_{\mathcal{D}}^{(n)}\|^r_{L^2}\big] \leqslant C(t^{(\ell)})^{r/2}. \tag{44}$$

Together with Lemma 6.2, this implies the first estimate. The second estimate follows from the uniform bound of $\mathbb{E}\big[\|\Pi_{\mathcal{D}} u\|^r_{L^\infty(0,T;L^2)}\big]$ and the Burkholder–Davis–Gundy. ∎

4. Tightness and construction of new probability space and processes

In this section, we show that the sequence $\{(\Pi_{\mathcal{D}_m} u_m, \nabla_{\mathcal{D}_m} u_m, M_{\mathcal{D}_m}, W)\}_{m\in\mathbb{N}}$ is tight. To prove the tightness of $M_{\mathcal{D}_m}$, we introduce the following space. For any $r \geqslant 2$, let us consider

$$\begin{aligned} L^r(0,T;L^2_{\rm w}) := {}& \text{the space of } r\text{-integrable functions } v : [0,T] \to L^2, \text{ endowed} \\ & \text{with the weakest topology such that, for all } \phi \in L^2, \text{ the mapping} \\ & v \in L^r(0,T;L^2_{\rm w}) \mapsto L^r(0,T;\mathbb{R}) \ni \langle v(\cdot), \phi\rangle_{L^2} \text{ is continuous.} \end{aligned}$$

In particular, $v_n \to v$ in $L^r(0,T;L^2_{\rm w})$ if and only if for all $\phi \in L^2$:

$$\langle v_n(\cdot), \phi\rangle_{L^2} \to \langle v(\cdot), \phi\rangle_{L^2} \quad \text{in } L^r(0,T;\mathbb{R}).$$

Let $(\phi_i)_{i\in\mathbb{N}} \subset C_c^\infty(\Theta)$ be a dense sequence in L^2 and equip the ball B of radius C_B in L^2 with the following metric

$$d_{L^2_{\rm w}}(v,w) = \sum_{i\in\mathbb{N}} \frac{\min(1, |\langle v-w, \phi_i\rangle_{L^2}|)}{2^i} \quad \text{for } v,w \in B.$$

It is easily checked that bounded sets in $L^\infty(0,T;L^2)$ are metrisable for the topology of $L^r(0,T;L^2_{\rm w})$, with metric

$$d_{L^r(L^2_{\rm w})}(v,w) := \left(\int_0^T d_{L^2_{\rm w}}(v(s),w(s))^r\,ds\right)^{1/r}.$$

To prove the tightness of $\Pi_{\mathcal{D}_m} u_m$, we define the following norm on $X_{\mathcal{D}_m}^{N_m+1}$: for any $v_m \in X_{\mathcal{D}_m}^{N_m+1}$

$$\|v_m\|_{\mathcal{D}_m} := \|\nabla_{\mathcal{D}_m} v_m\|_{L^p(\Theta_T)} + \|\Pi_{\mathcal{D}_m} v_m\|_{H^\beta(0,T,L^2)}.$$

By Lemma 3.1 and Estimate (37), we have

$$\mathbb{E}\big[\|u_m\|^q_{\mathcal{D}_m}\big] \leqslant C, \quad \text{with } q = \min(2,p).$$

Since the norm $\|\cdot\|_{\mathcal{D}_m}$ changes with m, we need to use Lemma 6.4 to establish the tightness of $\{\Pi_{\mathcal{D}_m} u_m\}_{m\in\mathbb{N}}$.
We now define the space $\mathcal{E}$

$$\mathcal{E} := L^p(0,T;L^p) \times \big(L^p(0,T;L^p)^d\big)_{\rm w} \times L^r(0,T;L^2_{\rm w}) \times C([0,T];L^2),$$

where $\big(L^p(0,T;L^p)\big)_{\rm w}$ is the space $L^p(0,T;L^p)$ endowed with the weak topology. The sequence $\big\{\big(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m, M_{\mathcal{D}_m}, W\big)\big\}_{m\in\mathbb{N}}$ is proved to be tight in the following lemma.

Lemma 4.1. *The measures of law of* $\big\{\big(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m, M_{\mathcal{D}_m}, W\big)\big\}_{n\in\mathbb{N}}$ *on* $\mathcal{E}$ *are tight.*

Proof. Let us first establish a (deterministic) compactness result. Consider, for a fixed constant C, the sets

$$K_m(C) := \Big\{v \in \Pi_{\mathcal{D}_m}X_{\mathcal{D}_m,0} \,:\, \exists w_m \in X_{\mathcal{D}_m,0} \text{ satisfying } \Pi_{\mathcal{D}_m}w_m = v,\ \|w_m\|_{\mathcal{D}_m} \leqslant C \\ \text{and} \quad \int_0^{T-\rho} \|v(t+\rho)-v(t)\|_{L^2}^2\,dt \leqslant C\rho, \quad \forall \rho \in (0,T)\Big\}$$

and define

$$\mathcal{K}(C) = \left(\bigcup_{m\in\mathbb{N}} K_m(C)\right) \cap \{v \in L^\infty(0,T;L^2) \,:\, \|v\|_{L^\infty(0,T;L^2)} \leqslant C\}.$$

Each $K_m(C)$ is relatively compact in $L^1(0,T;L^1)$ since it is bounded in the finite-dimensional space $\Pi_{\mathcal{D}_m}X_{\mathcal{D}_m,0}$. Moreover, by the compactness of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ (Definition 2.5), [19, Proposition C.5] shows that any sequence $\{v_m\}_{m\in\mathbb{N}}$ satisfying $v_m \in K_m(C)$ for any $m \in \mathbb{N}$ is relatively compact in $L^1(0,T;L^1)$. Hence, Lemma 6.4 shows that $\bigcup_{m\in\mathbb{N}} K_m(C)$, and thus $\mathcal{K}(C)$ is relatively compact in $L^1(0,T;L^1)$. The bound on $\|w_m\|_{\mathcal{D}_m}$ stated in $K_m(C)$ and the discrete Sobolev embeddings (Definition 2.8) ensure that $\mathcal{K}(C)$ is bounded in $L^p(0,T;L^{p^*})$ for $p^* > p$. Together with the bound in $L^\infty(0,T;L^2)$ and standard interpolation results, this proves that $\mathcal{K}(C)$ is bounded in $L^{\bar p}(0,T;L^{\bar p})$ for some $\bar p > p$. Using again interpolation inequality, this proves that the relative compactness of $\mathcal{K}(C)$ not only holds in $L^1(0,T;L^1)$, but also in $L^p(0,T;L^p)$.
This compactness of $\mathcal{K}(C)$, Lemma 6.3 and the bounds on $\{\Pi_{\mathcal{D}_m}u_m\}_{m\in\mathbb{N}}$, $\{\nabla_{\mathcal{D}_m}u_m\}_{m\in\mathbb{N}}$ and $\{M_{\mathcal{D}_m}\}_{m\in\mathbb{N}}$ stated in Lemma 3.1, (36), (37) and Lemma 3.5 imply the tightness law of $\big\{\big(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m, M_{\mathcal{D}_m}, W\big)\big\}_{m\in\mathbb{N}}$ in $\mathcal{E}$. ∎

By using Jakubowski's version of the Skorohod theorem [28, Theorem 2], we show the almost sure convergence of $\big\{\big(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m, M_{\mathcal{D}_m}, W\big)\big\}_{m\in\mathbb{N}}$, up to a change of probability space, in the following lemma.

Lemma 4.2. *There exists a new probability space* $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}}, \overline{\mathbb{P}})$*, a sequence of random variables* $\big(\widetilde{u}_m, \overline{M}_m, \overline{W}_m\big)_{m\in\mathbb{N}}$ *and random variables* $(\overline{u}, \overline{M}, \overline{W})$ *on this space such that*

- $\widetilde{u}_m \in X_{\mathcal{D}_m,0}$ *for each* $m \in \mathbb{N}$,
- $\big(\Pi_{\mathcal{D}_m}\widetilde{u}_m, \nabla_{\mathcal{D}_m}\widetilde{u}_m, \overline{M}_m, \overline{W}_m\big)$ *takes its values in space* $\mathcal{E}$ *with the same laws, for each* $m \in \mathbb{N}$*, as* $\big(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m, M_{\mathcal{D}_m}, W\big)$,
- $(\overline{u}, \overline{M}, \overline{W})$ *takes its values in* $L^p(0,T;W_0^{1,p}(\Theta)) \times L^r(0,T;L^2_{\rm w}) \times C([0,T];L^2)$,
- *up to a subsequence as* $m \to \infty$,

$$\Pi_{\mathcal{D}_m}\widetilde{u}_m \to \overline{u} \quad a.s. \ in \ L^p(0,T;L^p), \tag{45}$$

$$\nabla_{\mathcal{D}_m}\widetilde{u}_m \to \nabla\overline{u} \quad a.s. \ in \ \big(L^p(0,T;L^p)^d\big)_{\rm w}, \tag{46}$$

$$\overline{M}_m \to \overline{M} \quad a.s. \ in \ L^r(0,T;L^2_{\rm w}), \tag{47}$$

$$\overline{W}_m \to \overline{W} \quad a.s. \ in \ C([0,T];L^2), \tag{48}$$

- *$\widetilde{u}_m$ is a solution to the gradient scheme (Algorithm 2.2 with $\mathcal{D} = \mathcal{D}_m$) in which W is replaced by $\overline{W}_m$.*

Furthermore, up to a subsequence as $m \to \infty$, for almost all $t, s \in (0,T)$, for all $r \geqslant 1$,

$$\Pi_{\mathcal{D}_m}\widetilde{u}_m(t) - \Pi_{\mathcal{D}_m}\widetilde{u}_m(s) \to \overline{u}(t) - \overline{u}(s) \quad \text{in } L^p(\overline{\Omega} \times \Theta), \tag{49}$$

$$\overline{M}_m(t) - \overline{M}_m(s) \to \overline{M}(t) - \overline{M}(s) \quad \text{in } L^r(\overline{\Omega}; L^2_{\mathrm{w}}). \tag{50}$$

Proof. By using Jakubowski's version of the Skorohod theorem [28, Theorem 2], we find a new probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}}, \overline{\mathbb{P}})$, a sequence of random variables on this space $\big(\overline{u}_m, \overline{z}_m, \overline{M}_m, \overline{W}_m\big)$ taking its values in space $\mathcal{E}$ with the same laws, for each $m \in \mathbb{N}$, as $\big(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m, M_{\mathcal{D}_m}, W\big)$, and random variables $(\overline{u}, \overline{z}, \overline{M}, \overline{W})$ in $\mathcal{E}$, so that up to a subsequence as $m \to \infty$,

$$\overline{u}_m \to \overline{u} \quad \text{a.s. in } L^p(0,T;L^p), \tag{51}$$

$$\overline{z}_m \to \overline{z} \quad \text{a.s. in } \big(L^p(0,T;L^p)^d\big)_{\mathrm{w}}, \tag{52}$$

and the convergences (47), (48) hold.

Since $(\overline{u}_m, \overline{z}_m)$ has the same law as $(\Pi_{\mathcal{D}_m}u_m, \nabla_{\mathcal{D}_m}u_m)$, there exists $\widetilde{u}_m \in X_{\mathcal{D}_m,0}$ such that

$$\overline{u}_m = \Pi_{\mathcal{D}_m}\widetilde{u}_m, \quad \overline{z}_m = \nabla_{\mathcal{D}_m}\widetilde{u}_m$$

and $\widetilde{u}_m$ is a solution to the gradient scheme (Algorithm 2.2 with $\mathcal{D} = \mathcal{D}_m$) in which W is replaced by $\overline{W}_m$. More precisely, for any $n \in \{0, \cdots, N_m - 1\}$ and $\phi \in X_{\mathcal{D}_m,0}$, $\widetilde{u}_m$ satisfies, $\overline{\mathbb{P}}$ a.s.,

$$\begin{aligned}\Big\langle d^{(n+\frac{1}{2})}_{\mathcal{D}_m}\widetilde{u}_m, \Pi_{\mathcal{D}_m}\phi\Big\rangle_{L^2} + \delta t_{\mathcal{D}}\Big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}^{(n+1)}_m, \nabla_{\mathcal{D}_m}\widetilde{u}^{(n+1)}_m), \nabla_{\mathcal{D}_m}\phi\Big\rangle_{L^{p'},L^p}& \\ = \Big\langle f(\Pi_{\mathcal{D}_m}\widetilde{u}^{(n)}_m)\Delta^{(n+1)}\overline{W}_m, \Pi_{\mathcal{D}_m}\phi\Big\rangle_{L^2}.&\end{aligned} \tag{53}$$

Furthermore, applying [19, Lemma 4.8] and the a.s. convergences (51) and (52), the limit-conformity of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ ensures that

$$\overline{z} = \nabla\overline{u}, \quad \nabla_{\mathcal{D}_m}\widetilde{u}_m \to \nabla\overline{u} \quad \text{a.s. in } \big(L^p(\Theta_T)^d\big)_{\mathrm{w}}, \text{ and } \overline{u} \in L^p(0,T;W^{1,p}_0(\Theta)). \tag{54}$$

From (51)–(54) we obtain the first part of the lemma including (45) and (46).

We now prove (49) and (50) as the second part of the lemma. We obtain, from (8)–(9), the coercivity of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ and (43), for any $q \geqslant 1$

$$\begin{aligned}\sup_{m\in\mathbb{N}}\mathbb{E}\Big[\|\Pi_{\mathcal{D}_m}\widetilde{u}_m\|^q_{L^p(\Theta_T)} + \|\Pi_{\mathcal{D}_m}\widetilde{u}_m\|^2_{L^\infty(0,T;L^2)} + \|\nabla_{\mathcal{D}_m}\widetilde{u}_m\|^p_{L^p(\Theta_T)}\Big]& \\ + \sup_{m\in\mathbb{N}}\mathbb{E}\|\overline{M}_m\|^q_{L^\infty(0,T;L^2)} \leqslant C.&\end{aligned} \tag{55}$$

From (45), (47) and (55), we obtain the following result by applying the Vitali theorem

$$\Pi_{\mathcal{D}_m}\widetilde{u}_m \to \overline{u} \quad \text{in } L^p(\overline{\Omega} \times (0,T) \times \Theta) \text{ as } m \to \infty, \tag{56}$$

$$\overline{M}_m \to \overline{M} \quad \text{in } L^r(\overline{\Omega} \times (0,T); L^2_{\mathrm{w}}) \text{ as } m \to \infty. \tag{57}$$

Hence, up to a subsequence, one has (49) for almost all $t, s \in (0,T)$. The convergence (50) can be obtained from (57) using the classical a.e. extraction in $L^r(0,T)$ on the function $t \mapsto \int_{\overline{\Omega}} d_{L^2_{\mathrm{w}}}(\overline{M}_m(t) - \overline{M}(t), 0)^r d\overline{\mathbb{P}}$. ∎

The continuity of the stochastic processes $\overline{u}$ and $\overline{M}$ is showed in the following lemma.

Lemma 4.3. *The stochastic processes* $\overline{u}$ *and* $\overline{M}$ *have continuous versions in* $C([0,T],L^2_{\rm w})$ *and* $C([0,T],L^2)$*, respectively.*

Proof. The continuity of $\overline{u}$ will be proved using Kolmogorov's test [14, Theorem 3.3]. Let $(\psi_i)_{i\in\mathbb{N}}\subset C^\infty_c(\Theta)\backslash\{0\}$ be a dense sequence in L^2 and define the metric

$$\widehat{d}_{L^2_{\rm w}}(v,w)=\sum_{i\in\mathbb{N}}\frac{|\langle v-w,\phi_i\rangle_{L^2}|}{2^i}\quad\text{for } v,w\in L^2,$$

with $\phi_i:=\psi_i/(\|\psi_i\|_{L^{\widehat{p}}}+\|\nabla\psi_i\|_{L^p})$, where we recall that $\widehat{p}=\max\{2,p'\}$. This metric defines the weak topology of L^2 on its closed balls, which are compact and thus complete for this topology. To estimate the continuity of u, we start by estimating $\widehat{d}_{L^2_{\rm w}}\big(\Pi_{\mathcal{D}_m}u_m(s),\Pi_{\mathcal{D}_m}u_m(s')\big)$ for $0\leqslant s\leqslant s'\leqslant T$.
We first define the interpolator $P_{\mathcal{D}_m}:W^{1,p}_0(\Theta)\cap L^{\widehat{p}}\to X_{\mathcal{D}_m,0}$ by

$$P_{\mathcal{D}_m}\phi:=\operatorname{argmin}_{w\in X_{\mathcal{D}_m,0}}\big(\|\Pi_{\mathcal{D}_m}w-\phi\|_{L^{\widehat{p}}}+\|\nabla_{\mathcal{D}_m}w-\nabla\phi\|_{L^p}\big).\tag{58}$$

We have, for $r\geqslant 1$,

$$\begin{aligned}
&\mathbb{E}\Bigg[\bigg|\int_\Theta\Big(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s',\boldsymbol{x})-\Pi_{\mathcal{D}_m}\widetilde{u}_m(s,\boldsymbol{x})\Big)\phi_i(\boldsymbol{x})d\boldsymbol{x}\bigg|^r\Bigg]\\
&\leqslant 2^{r-1}\mathbb{E}\Bigg[\bigg|\int_\Theta\Big(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s',\boldsymbol{x})-\Pi_{\mathcal{D}_m}\widetilde{u}_m(s,\boldsymbol{x})\Big)\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i(\boldsymbol{x})d\boldsymbol{x}\bigg|^r\Bigg]\\
&\quad+2^{r-1}\mathbb{E}\Bigg[\bigg|\int_\Theta\Big(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s',\boldsymbol{x})-\Pi_{\mathcal{D}_m}\widetilde{u}_m(s,\boldsymbol{x})\Big)\Big(\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i(\boldsymbol{x})-\phi_i(\boldsymbol{x})\Big)d\boldsymbol{x}\bigg|^r\Bigg]\\
&\leqslant 2^{r-1}\mathbb{E}\Bigg[\bigg|\Pi_{\mathcal{D}_m}\widetilde{u}_m(s',\boldsymbol{x})-\Pi_{\mathcal{D}_m}\widetilde{u}_m(s,\boldsymbol{x})\bigg|^r_{*,\mathcal{D}}\Bigg]\big(\|\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i\|_{L^2}+\|\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i\|_{L^p}\big)^r\\
&\quad+2^{r-1}\mathbb{E}\Big[\|\Pi_{\mathcal{D}_m}\widetilde{u}_m\|^r_{L^\infty(0,T;L^2)}\Big]\|\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i-\phi_i\|^r_{L^2}
\end{aligned}\tag{59}$$

It follows from (58) and $\|\phi_i\|_{L^{\widehat{p}}}+\|\nabla\phi_i\|_{L^p}\leqslant C$ that

$$\begin{aligned}
\|\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i-\phi_i\|_{L^2}\leqslant CS_{\mathcal{D}_m}(\phi_i)\leqslant C,\ \text{and}\\
\|\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i\|_{L^2}+\|\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\phi_i\|_{L^p}\leqslant C.
\end{aligned}$$

Note that the bound $S_{\mathcal{D}_m}(\phi_i)\leqslant 1$ is obtained selecting $w=0$ in the definition of this quantity. We then estimate the right hand side of (59) using Lemmas 3.1 and 3.4 to obtain

$$\begin{aligned}
\mathbb{E}\Bigg[\bigg|\int_\Omega\Big(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s',\boldsymbol{x})-\Pi_{\mathcal{D}_m}\widetilde{u}_m(s,\boldsymbol{x})\Big)\phi_i(\boldsymbol{x})d\boldsymbol{x}\bigg|^r\Bigg]&\\
\leqslant C\big(|s'-s|+\delta t_{\mathcal{D}_m}\big)^{\alpha r}+CS_{\mathcal{D}_m}(\phi_i)&\\
\leqslant C|s'-s|^{\alpha r}+C\delta t^{\alpha r}_{\mathcal{D}_m}+CS_{\mathcal{D}_m}(\phi_i).&
\end{aligned}$$

Recalling the definition of $\widehat{d}_{L^2_{\rm w}}$ and using Jensen's inequality to write

$$\widehat{d}_{L^2_{\rm w}}(u,v)^r=\left(\sum_{i\in\mathbb{N}}\frac{|\langle v-w,\phi_i\rangle_{L^2}|}{2^i}\right)^r\leqslant\sum_{i\in\mathbb{N}}\frac{|\langle v-w,\phi_i\rangle_{L^2}|^r}{2^i},$$

we infer

$$\mathbb{E}\Big[\widehat{d}_{L^2_{\rm w}}\big(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s),\Pi_{\mathcal{D}_m}\widetilde{u}_m(s')\big)^r\Big]\leqslant C|s'-s|^{\alpha r}+C\sum_{i\in\mathbb{N}}\frac{C\delta t^{\alpha r}_{\mathcal{D}_m}+S_{\mathcal{D}_m}(\phi_i)}{2^i}.$$

Since $\delta t_{\mathcal{D}_m} \to 0$ and $S_{\mathcal{D}_m}(\phi_i) \to 0$ for all $i \in \mathbb{N}$, while being uniformly bounded as seen above, we can apply the dominated convergence theorem ont the last sum to see that it tends to 0 as $m \to \infty$. Together with (49) and Fatou's lemma, this implies, for a.e. s, s',

$$\mathbb{E}\left[\widehat{d}_{L^2_{\rm w}}\left(\overline{u}(s), \overline{u}(s')\right)^r\right] \leqslant C|s'-s|^{\alpha r}.$$

By choosing r such that $\alpha r > 1$, we obtain the desired continuity of $\overline{u}$ by applying the Kolmogorov test.

We now prove the continuity of $\overline{M}$. It follows from (44) and the fact that $\overline{M}_m$ has the same law as M_m that

$$\mathbb{E}\big[\|\overline{M}_m(s') - \overline{M}_m(s)\|^r_{L^2}\big] \leqslant C(|s'-s| + \delta t_{\mathcal{D}_m})^{r/2}, \tag{60}$$

and $\mathbb{E}[\|\overline{M}_m\|^r_{L^\infty(0,T;L^2)}] \leqslant C$, which implies $\|\overline{M}_m\|^r_{L^\infty(0,T;L^r(\overline{\Omega};L^2))} \leqslant C$. Estimate (60) and the discontinuous Ascoli-Arzelà theorem [19, Theorem C.11] imply

$$\overline{M}_m \to \overline{M} \quad \text{ uniformly on } [0,T] \text{ in } (L^r(\overline{\Omega};L^2))_{\rm w}, \text{ as } m \to \infty,$$

and $\overline{M} \in C\big([0,T];(L^r(\overline{\Omega};L^2))_{\rm w}\big)$. It follows from this convergence, (60), the weak lower semicontinuity of norms and Fatou's lemma that

$$\mathbb{E}\big[\|\overline{M}(s') - \overline{M}(s)\|^r_{L^2}\big] \leqslant C|s'-s|^{r/2}.$$

The continuity of $\overline{M}$ follows immediately by choosing $r > 3$ and applying the Kolmogorov test. ∎

5. Identification of the limit

In this section, we first find a representation of the martingale part $\overline{M}$. Since $\overline{M}$ is continuous from $[0,T]$ to L^2, the representation theorem in [14, Theorem 8.2] can be used. We will check conditions of [14, Theorem 8.2] in the following lemma.

Lemma 5.1. *The process $t \in [0,T] \mapsto \overline{M}(t,\omega) \in L^2$ is a square integrable continnuous martingale, with quadratic variaton defined for all $a, b \in L^2$ by*

$$\langle\langle M(t)\rangle\rangle(a,b) = \int_0^t \big\langle\big(f(\overline{u})\mathcal{Q}^{1/2}\big)^*(a), \big(f(\overline{u})\mathcal{Q}^{1/2}\big)^*(b)\big\rangle_{\mathcal{K}}\, ds, \tag{61}$$

for any $t \geqslant 0$.

Proof. It follows from the fact that $M_{\mathcal{D}_m}$ is piecewise constant and the same laws that $\overline{M}_m$ is piecewise constant for any $m \in \mathbb{N}$. Furthermore, for all $t \in [0,T]$ and $\overline{\mathbb{P}}$ a.e., $\overline{M}_m$ satisfies

$$\overline{M}_m(t) = \sum_{0 \leqslant i\,\delta t_{\mathcal{D}_m} < t} f(\Pi_{\mathcal{D}_m}\widetilde{u}^{(i)}_m)\Delta^{(i+1)}\overline{W}_m.$$

Since $\widetilde{u}_m$ is a solution to the gradient scheme (Algorithm 2.2 with $\mathcal{D} = \mathcal{D}_m$), $\widetilde{u}_m$ is adapted to

$$\mathcal{F}_{i\delta t_{\mathcal{D}_m}} := \sigma\{\overline{W}_m(k\,\delta t_{\mathcal{D}_m}),\, k = 1,\cdots,i\},$$

and the process $\overline{M}^{(i)}_m := \overline{M}_m(i\,\delta t_{\mathcal{D}_m})$ defines a martingale with respect to this filtration. In particular, we have the following identity

$$\mathbb{E}\big[\big(\langle a, \overline{M}^{(j)}_m\rangle_{L^2} - \langle a, \overline{M}^{(i)}_m\rangle_{L^2}\big)\psi\big(\overline{W}_m(\delta t_{\mathcal{D}_m}),\cdots,\overline{W}_m(i\,\delta t_{\mathcal{D}_m})\big)\big] = 0 \tag{62}$$

for all $0 \leqslant i \leqslant j \leqslant N_m$ and any bounded continuous function $\psi : (L^2)^i \to \mathbb{R}$. Furthermore, we obtain

$$\mathbb{E}\Bigg[\Bigg(\langle a, \overline{M}_m^{(j)}\rangle_{L^2}\langle b, \overline{M}_m^{(j)}\rangle_{L^2} - \langle a, \overline{M}_m^{(i)}\rangle_{L^2}\langle b, \overline{M}_m^{(i)}\rangle_{L^2} - \sum_{i+1 \leqslant k \leqslant j} \delta t_{\mathcal{D}_m}\Big\langle \big(f(\Pi_{\mathcal{D}_m}\widetilde{u}_m^{(k)})\mathcal{Q}^{1/2}\big)^*(a), \big(f(\Pi_{\mathcal{D}_m}\widetilde{u}_m^{(k)})\mathcal{Q}^{1/2}\big)^*(b)\Big\rangle_{\mathcal{K}}\Bigg) \psi\big(\overline{W}_m(\delta t_{\mathcal{D}_m}), \cdots, \overline{W}_m(i\, \delta t_{\mathcal{D}_m})\big)\Bigg] = 0. \tag{63}$$

<u>Proof that $\overline{M}$ is a martingale:</u> We have to show that for almost all $0 \leqslant s \leqslant t \leqslant T$, all $K \in \mathbb{N}$, any bounded continuous function ϕ defined on $(L^2)^K$, and for any choice of times $0 \leqslant s_1 < s_2 < \cdots < s_K \leqslant s$, the following relation holds

$$\mathbb{E}\big[\big(\langle a, \overline{M}(t)\rangle_{L^2} - \langle a, \overline{M}(s)\rangle_{L^2}\big)\phi\big(\overline{W}(s_1), \cdots, \overline{W}(s_K)\big)\big] = 0. \tag{64}$$

Let $\lfloor x \rfloor$ denote the floor of x for any $x \geqslant 0$. For all $0 \leqslant i \leqslant K$ we have

$$\lfloor \frac{s_i}{\delta t_{\mathcal{D}_m}} \rfloor \delta t_{\mathcal{D}_m} \to s_i \quad \text{as} \quad m \to \infty.$$

It follows from (48) and the continuity of ϕ that

$$\phi\left(\overline{W}(\lfloor \frac{s_1}{\delta t_{\mathcal{D}_m}} \rfloor \delta t_{\mathcal{D}_m}), \cdots, \overline{W}(\lfloor \frac{s_K}{\delta t_{\mathcal{D}_m}} \rfloor \delta t_{\mathcal{D}_m})\right) \to \phi\big(\overline{W}(s_1), \cdots, \overline{W}(s_K)\big) \tag{65}$$

as $m \to \infty$, $\overline{\mathbb{P}}$-a.s. in $(L^2)^K$. For any $m \in \mathbb{N}$ and $\delta t_{\mathcal{D}_m} > 0$ there exist $l_1, l_2 \in \{0, \ldots, N_m - 1\}$ such that $s \in (t^{(l_1)}, t^{(l_1+1)}]$ and $t \in (t^{(l_2)}, t^{(l_2+1)}]$. From (62) we obtain that

$$\mathbb{E}\big[\big(\langle a, \overline{M}_m(t)\rangle_{L^2} - \langle a, \overline{M}_m(s)\rangle_{L^2}\big)\psi\big(\overline{W}_m(\delta t_{\mathcal{D}_m}), \cdots, \overline{W}_m(l_1\, \delta t_{\mathcal{D}_m})\big)\big] = 0, \tag{66}$$

for any bounded continuous function ψ defined on $(L^2)^{l_1}$. Since $\lfloor \frac{s_K}{\delta t_{\mathcal{D}_m}} \rfloor \leqslant l_1$, we can choose ψ in (66) such that

$$\psi\big(\overline{W}_m(\delta t_{\mathcal{D}_m}), \cdots, \overline{W}_m(l_1\, \delta t_{\mathcal{D}_m})\big) = \phi\left(\overline{W}_m(\lfloor \frac{s_1}{\delta t_{\mathcal{D}_m}} \rfloor \delta t_{\mathcal{D}_m}), \cdots, \overline{W}_m(\lfloor \frac{s_K}{\delta t_{\mathcal{D}_m}} \rfloor \delta t_{\mathcal{D}_m})\right).$$

We obtain (64) by taking limit of (66) as m tends to infinity and using the convergences (50) and (65).

<u>Proof of (61):</u> From the definition of the quadratic variation [14, page 75], in order to prove (61), we have to show that

$$\mathbb{E}\Bigg[\Bigg(\langle a, \overline{M}(t)\rangle_{L^2}\langle b, \overline{M}(t)\rangle_{L^2} - \langle a, \overline{M}(s)\rangle_{L^2}\langle b, \overline{M}(s)\rangle_{L^2} - \int_s^t \Big\langle \big(f(\overline{u})\mathcal{Q}^{1/2}\big)^*(a), \big(f(\overline{u})\mathcal{Q}^{1/2}\big)^*(b)\Big\rangle_{\mathcal{K}}\Bigg)\phi\big(\overline{W}(s_1), \cdots, \overline{W}(s_K)\big)\Bigg] = 0. \tag{67}$$

The above identity can be obtained by using the same arguments as in the proof of (64) with the continuity of f, (56) and (63).

The continuity and square integrability of $\overline{M}$ follows from Lemma 4.3 and (55).

■

We now apply the continuous martingale representation [14, Theorem 8.2]. We have showed that the limit process $\overline{M}$ satisfies its hypotheses. Hence, there exists an enlarged probability space $(\widetilde{\Omega}, \widetilde{\mathbb{F}}, \widetilde{\mathbb{P}})$, with $\overline{\Omega} \subset \widetilde{\Omega}$ and a $\mathcal{Q}$-Wiener process $\widetilde{W}$

defined on $(\widetilde{\Omega},\widetilde{\mathbb{F}},\widetilde{\mathbb{P}})$ such that $\overline{M}$, $\overline{u}$ can be extended to random variables on this space and, for every $t\geqslant 0$,

$$\overline{M}(t,\cdot)=\int_0^t f(\overline{u})(s,\cdot)d\widetilde{W}(s). \tag{68}$$

We are ready to prove the main theorem.

Proof of Theorem 2.11.

For any $t\in[0,T]$, there exists $k\in\{0,\cdots,N_m-1\}$ such that $t\in(t^{(k)},t^{(k+1)}]$. For any $\psi\in W_0^{1,p}(\Theta)\cap L^{\hat{p}}(\Theta)$, we take the sum of (53) from $n=0$ to $n=k$ with test function $\phi:=P_{\mathcal{D}_m}\psi$ (recall the definition (58) of $P_{\mathcal{D}_m}$) to obtain, $\overline{\mathbb{P}}$ a.s.,

$$\begin{aligned}\big\langle\Pi_{\mathcal{D}_m}\widetilde{u}_m(t),\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}-\big\langle\Pi_{\mathcal{D}_m}u^{(0)},\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}&\\ +\sum_{n=0}^{k}\delta t_{\mathcal{D}}\big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m^{(n+1)},\nabla_{\mathcal{D}_m}\widetilde{u}_m^{(n+1)}),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}&\\ =\big\langle\overline{M}_m(t),\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}.&\end{aligned}\tag{69}$$

By consistency of $(\mathcal{D}_m)_{m\in\mathbb{N}}$ (Definition 2.3) we have $\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\to\psi$ in $L^{\hat{p}}$. Hence, Equations (49) and (50) show that, for almost every t,

$$\begin{aligned}\big\langle\Pi_{\mathcal{D}_m}\widetilde{u}_m(t),\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}&\to\big\langle\overline{u}(t),\psi\big\rangle_{L^2}\quad\text{in } L^p(\overline{\Omega})\\ \big\langle\overline{M}_m(t),\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}&\to\big\langle\overline{M}(t),\psi\big\rangle_{L^2}\quad\text{in } L^r(\overline{\Omega}).\end{aligned}\tag{70}$$

Moreover, we also have

$$\big\langle\Pi_{\mathcal{D}_m}u^{(0)},\Pi_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}\to\big\langle u_0,\psi\big\rangle_{L^2}. \tag{71}$$

It remains to prove the convergence of the last term in the left hand side of (69). We first note that

$$\begin{aligned}&\sum_{n=0}^{k}\delta t_{\mathcal{D}}\big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m^{(n+1)},\nabla_{\mathcal{D}_m}\widetilde{u}_m^{(n+1)}),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^2}\\ &\quad=\int_0^t\big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s),\nabla_{\mathcal{D}_m}\widetilde{u}_m(s)),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^{p'},L^p}\,ds\\ &\quad\quad+\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil\delta t_{\mathcal{D}_m}}\big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s),\nabla_{\mathcal{D}_m}\widetilde{u}_m(s)),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^{p'},L^p}\,ds.\end{aligned}\tag{72}$$

Since $\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\to\nabla\psi$ in L^p, the a.s. convergences (45) and (46) enable us to apply the standard Minty argument (as in, e.g., [19, Proof of Theorem 5.19 (Step 3)]) to get the a.s. convergence of the first term in the right hand side of (72): for any $t\in[0,T]$, $\overline{\mathbb{P}}$-a.s.,

$$\begin{aligned}&\int_0^t\big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s),\nabla_{\mathcal{D}_m}\widetilde{u}_m(s)),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^{p'},L^p}\,ds\\ &\qquad\to\int_0^t\big\langle a(\overline{u}(s),\nabla\overline{u}(s)),\nabla\psi\big\rangle_{L^{p'},L^p}\,ds.\end{aligned}\tag{73}$$

The expectation of the last term in the right hand side of (72) tends to zero as $m\to\infty$. Indeed, by using (3), Hölder inequality and (55) we obtain

$$\mathbb{E}\Big[\Big|\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil\delta t_{\mathcal{D}_m}}\big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s),\nabla_{\mathcal{D}_m}\widetilde{u}_m(s)),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\big\rangle_{L^{p'},L^p}\,ds\Big|\Big]$$

$$
\begin{aligned}
&\leqslant C\mathbb{E}\Big[\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil \delta t_{\mathcal{D}_m}} \int_\Theta (1+|\nabla_{\mathcal{D}_m}\widetilde{u}_m^{(k+1)}|^{p-1})|\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi|\,dx\,ds\Big]\\
&\leqslant C\delta t_{\mathcal{D}_m}\|\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\|_{L^1}\\
&\quad + C\mathbb{E}\Big[\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil \delta t_{\mathcal{D}_m}} \|\nabla_{\mathcal{D}_m}\widetilde{u}_m^{(k+1)}\|_{L^p}^{p-1}\|\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\|_{L^p}\,ds\Big]\\
&\leqslant C\delta t_{\mathcal{D}_m}\\
&\quad + C\mathbb{E}\Big[\Big(\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil \delta t_{\mathcal{D}_m}} \|\nabla_{\mathcal{D}_m}\widetilde{u}_m^{(k+1)}\|_{L^p}^{p}\,ds\Big)^{(p-1)/p}\Big(\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil \delta t_{\mathcal{D}_m}} ds\Big)^{1/p}\Big]\\
&\leqslant C\delta t_{\mathcal{D}_m} + C(\delta t_{\mathcal{D}_m})^{1/p}\mathbb{E}\Big[\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil \delta t_{\mathcal{D}_m}} \|\nabla_{\mathcal{D}_m}\widetilde{u}_m^{(k+1)}\|_{L^p}^{p}\,ds\Big]^{(p-1)/p}\\
&\leqslant C\delta t_{\mathcal{D}_m} + C(\delta t_{\mathcal{D}_m})^{1/p}\mathbb{E}\Big[\int_0^T \|\nabla_{\mathcal{D}_m}\widetilde{u}_m(s)\|_{L^p}^{p}\,ds\Big]^{(p-1)/p}\\
&\leqslant C(\delta t_{\mathcal{D}_m} + (\delta t_{\mathcal{D}_m})^{1/p}),
\end{aligned}
$$

which implies

$$
\int_t^{\lceil t/\delta t_{\mathcal{D}_m}\rceil \delta t_{\mathcal{D}_m}} \Big\langle a(\Pi_{\mathcal{D}_m}\widetilde{u}_m(s),\nabla_{\mathcal{D}_m}\widetilde{u}_m(s)),\nabla_{\mathcal{D}_m}P_{\mathcal{D}_m}\psi\Big\rangle_{L^{p'},L^p}\,ds \to 0 \quad \text{in } L^1(\overline{\Omega}) \tag{74}
$$

Using (70)–(74) and (68), we pass to the limit in (69) to see that $\overline{u}$ satisfies (4) in Definition 2.9, with $\widetilde{W}$ instead of W. ∎

6. Appendix

Lemma 6.1. *Let $\alpha > 0$, $q > 0$ and (E, d_E) be a metric space. Assume that $g : [0,T] \to E$ is piecewise constant with respect to the partition $(t^{(n)})_{n=0,\dots,N}$ and that, for all $\ell = 1,\dots,N-1$, denoting by $g^{(n)}$ the constant value of g on $(t^{(n)},t^{(n+1)}]$,*

$$
\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} d_E(g^{(n+\ell)},g^{(n)})^q \leqslant C(t^{(\ell)})^\alpha. \tag{75}
$$

Then, there exists a constant C' not depending on g or $\delta t_{\mathcal{D}}$ such that

$$
\int_0^{T-\rho} d_E(g(t+\rho),g(t))^q\,dt \leqslant C'\sigma(\rho,\delta t_{\mathcal{D}}),
$$

for any $\rho \in [0,T]$, where

$$
\sigma(\rho,\delta t_{\mathcal{D}}) = \begin{cases} \rho^\alpha & \text{if } \alpha \in (0,1] \\ \rho^\alpha + (\delta t_{\mathcal{D}})^{\alpha-1}\rho & \text{if } \alpha > 1. \end{cases}
$$

Proof. (i) $\rho \in (0,\delta t_{\mathcal{D}}]$.
For any $t \in [0,T-\rho]$, there exists $n \in \{1,\cdots,N\}$ such that $t \in (t^{(n-1)},t^{(n)}]$. If $t \in (t^{(n-1)},t^{(n)}-\rho]$, then $t+\rho \in (t^{(n-1)},t^{(n)}]$ and $g(t+\rho)=g(t)=g^{(n)}$, so that $d_E(g(t+\rho),g(t))=0$. If $t \in (t^{(n)}-\rho,t^{(n)}]$, then $t+\rho \in (t^{(n)},t^{(n+1)}]$ and $g(t)=g^{(n)}$, $g(t+\rho)=g^{(n+1)}$, so that $d_E(g(t+\rho),g(t)) = d_E(g^{(n+1)},g^{(n)})$. Therefore, from (75) with $\ell=1$ we have

$$
\int_0^{T-\rho} d_E(g(t+\rho),g(t))^q\,dt = \rho\sum_{n=1}^{N-1} d_E(g^{(n+1)},g^{(n)})^q \leqslant \rho C(t^{(1)})^\alpha \delta t_{\mathcal{D}}^{-1} = C\rho\,\delta t_{\mathcal{D}}^{\alpha-1}
$$

$$\leqslant \begin{cases} C\rho^\alpha & \text{if } \alpha \in (0,1] \\ C\delta t_{\mathcal{D}}^{\alpha-1}\rho & \text{if } \alpha > 1. \end{cases}$$

Above, in the case $\alpha \leqslant 1$, we have concluded by writing $\rho\delta t_{\mathcal{D}}^{\alpha-1} = (\rho/\delta t_{\mathcal{D}})^{1-\alpha}\rho^\alpha \leqslant \rho^\alpha$, since $\rho \leqslant \delta t_{\mathcal{D}}$.

(ii) $\rho > \delta t_{\mathcal{D}}$.

In this case, we can find $1 \leqslant \ell \leqslant N-1$ and $\epsilon \in (0,1)$ such that $\rho = \delta t_{\mathcal{D}}(\ell+\epsilon)$. For any $t \in [0, T-\rho]$, there exists $n \in \{1, \cdots, N-\ell\}$ such that $t \in (t^{(n-1)}, t^{(n)}]$. If $t \in (t^{(n-1)}, t^{(n)} - \delta t_{\mathcal{D}}\epsilon]$, then $t + \delta t_{\mathcal{D}}\epsilon \in (t^{(n-1)}, t^{(n)}]$ and $t + \rho \in (t^{(n-1+\ell)}, t^{(n+\ell)}]$. If $t \in (t^{(n)} - \delta t_{\mathcal{D}}\epsilon, t^{(n)}]$, then $t + \delta t_{\mathcal{D}}\epsilon \in (t^{(n)}, t^{(n+1)}]$ and $t + \rho \in (t^{(n+\ell)}, t^{(n+\ell+1)}]$. Therefore, from (75) we have

$$\begin{aligned}\int_0^{T-\rho} d_E(g(t+\rho), g(t))^q\,dt &= \sum_{n=1}^{N-\ell-1}\Big[\int_{t^{(n-1)}}^{t^{(n)}-\delta t_{\mathcal{D}}\epsilon} d_E(g^{(n+\ell)}, g^{(n)})^q\,dt \\ &\quad + \int_{t^{(n)}-\delta t_{\mathcal{D}}\epsilon}^{t^{(n)}} d_E(g^{(n+\ell+1)}, g^{(n)})^q\,dt\Big] \\ &\quad + \int_{t^{(N-\ell-1)}}^{t^{(N-\ell)}-\delta t_{\mathcal{D}}\epsilon} d_E(g^{(N)}, g^{(N-\ell)})^q\,dt \\ &\leqslant \delta t_{\mathcal{D}}(1-\epsilon)C\delta t_{\mathcal{D}}^{-1}(t^{(\ell)})^\alpha + \delta t_{\mathcal{D}}\epsilon C\delta t_{\mathcal{D}}^{-1}(t^{(\ell+1)})^\alpha \\ &\leqslant C(1+2^\alpha)(t^{(\ell)})^\alpha \\ &\leqslant C(1+2^\alpha)\rho^\alpha,\end{aligned}$$

which concludes the proof of this lemma. ■

The following lemma is a consequence of Lemma 6.1.

Lemma 6.2. *Let $0 < \alpha \leqslant 1$, $q > 0$ and $0 < \beta < \alpha/q$. Let $g : [0,T] \to E$ be piecewise constant with respect to the partition $(t^{(n)})_{n=0,\ldots,N}$, and let $g^{(n)}$ be its constant value on $(t^{(n)}, t^{(n+1)}]$. Assume that, for all $\ell = 1, \ldots, N-1$,*

$$\mathbb{E}\Big[\delta t_{\mathcal{D}} \sum_{n=1}^{N-\ell} \|g^{(n+\ell)} - g^{(n)}\|_{L^2}^q\Big] \leqslant C(t^{(\ell)})^\alpha.$$

Then, there exists a constant C' not depending on g neither on $\delta t_{\mathcal{D}}$ such that

$$\mathbb{E}\big[\|g\|^q_{W^{\beta,q}([0,T];L^2)}\big] \leqslant C'.$$

Proof. Using the same arguments as in Lemma 6.1 and adding the expectation on estimates, we also obtain from the assumption on g that

$$\mathbb{E}\Big[\int_0^{T-\rho} \|g(t+\rho) - g(t)\|_{L^2}^q\,dt\Big] \leqslant C\rho^\alpha. \tag{76}$$

This implies that

$$\begin{aligned}\mathbb{E}\big[\|g\|^q_{W^{\beta,q}([0,T];L^2)}\big] &= \mathbb{E}\Big[\int_0^T \Big(\int_0^{T-\rho} \|g(s+\rho) - g(s)\|_{L^2}^q\,ds\Big)\frac{d\rho}{\rho^{1+\beta q}}\Big] \\ &\leqslant C\int_0^T \rho^{\alpha-\beta q-1}d\rho = CT^{\alpha-\beta q}.\end{aligned}$$

■

Lemma 6.3. *Let $\beta \in (0,1)$. For any $r \geqslant 1$, the following embedding is compact:*

$$H^\beta(0,T;L^2) \cap L^\infty(0,T;L^2) \overset{c}{\hookrightarrow} L^r(0,T;L^2_{\mathrm{w}})$$

where the space $L^r(0,T;L^2_{\mathrm{w}})$ and its topology are defined in Section 4.

Proof. For any bounded sequence $\{w_m\}_{m\in\mathbb{N}}$ in $H^\beta(0,T;L^2) \cap L^\infty(0,T;L^2)$, there exists $w \in H^\beta(0,T;L^2) \cap L^\infty(0,T;L^2)$ such that

$$w_m \to w \quad \text{weakly in } H^\beta(0,T;L^2) \cap L^2(0,T;L^2)$$

up to a subsequence. Let $v_m = w_m - w$. It is sufficient to prove that $\{v_m\}_{m\in\mathbb{N}}$ converges to zero in $L^r(0,T;L^2_{\mathrm{w}})$.
For any $L \in \mathbb{N}$, let $\eta := T/L$. We define the picewise constant function v_m^η by

$$v_m^\eta\big|_{[\ell\eta,(\ell+1)\eta)} := \frac{1}{\eta}\int_{\ell\eta}^{(\ell+1)\eta} v_m(s)\,ds$$

We note that $\{v_m\}_{m\in\mathbb{N}}$ is bounded in $H^\beta(0,T;L^2)$. By using the Minkowski's integral inequality, we deduce

$$\begin{aligned}
\|v_m^\eta - v_m\|^2_{L^2(0,T;L^2)} &= \sum_{\ell=0}^{L-1}\int_{\ell\eta}^{(\ell+1)\eta}\int_\Theta\left(\frac{1}{\eta}\int_{\ell\eta}^{(\ell+1)\eta} v_m(s,x)-v_m(t,x)\,ds\right)^2 dx\,dt\\
&\leqslant \sum_{\ell=0}^{L-1}\int_{\ell\eta}^{(\ell+1)\eta}\int_{\ell\eta}^{(\ell+1)\eta}\|v_m(s)-v_m(t)\|^2_{L^2}\,ds\,dt\\
&\leqslant T\eta^{2\beta}\sum_{\ell=0}^{L-1}\int_{\ell\eta}^{(\ell+1)\eta}\int_{\ell\eta}^{(\ell+1)\eta}\frac{\|v_m(s)-v_m(t)\|^2_{L^2}}{|t-s|^{2\beta+1}}\,ds\,dt\\
&\leqslant T\eta^{2\beta}\|v_m\|^2_{H^\beta(0,T;L^2)} \leqslant C\eta^{2\beta}.
\end{aligned}$$

Using the boundedness of $v_m^\eta - v_m$ in $L^\infty(0,T;L^2)$ and an interpolation inequality of $L^r(0,T)$ between $L^\infty(0,T)$ and $L^2(0,T)$, we infer

$$\|v_m^\eta - v_m\|_{L^r(0,T;L^2)} \leqslant C\eta^{\frac{2\beta}{r}}. \tag{77}$$

On the other side,

$$d_{L^r(L^2_{\mathrm{w}})}(v_m^\eta,0)^r = \int_0^T d_{L^2_{\mathrm{w}}}(v_m^\eta(s),0)^r\,ds = \sum_{\ell=0}^{L-1}\eta\, d_{L^2_{\mathrm{w}}}(v_m^\eta\big|_{[\ell\eta,(\ell+1)\eta)},0)^r, \tag{78}$$

and, for any $0 \leqslant \ell \leqslant L-1$ and $\phi \in L^2$, by weak convergence of v_m in $L^2(0,T;L^2)$,

$$\int_\Theta v_m^\eta\big|_{[\ell\eta,(\ell+1)\eta)}(x)\,\phi(x)\,dx = \frac{1}{\eta}\int_0^T\int_\Theta v_m(t,x)\mathbb{1}\big|_{[\ell\eta,(\ell+1)\eta)}(t)\phi(x)\,dt\,dx \to 0$$

as m tends to infinity. Plugged into (78), this implies, for all η,

$$d_{L^r(L^2_{\mathrm{w}})}(v_m^\eta,0) \to 0 \quad \text{as } m \to \infty. \tag{79}$$

Using (77), we obtain

$$d_{L^r(L^2_{\mathrm{w}})}(v_m,0) \leqslant d_{L^r(L^2_{\mathrm{w}})}(v_m,v_m^\eta) + d_{L^r(L^2_{\mathrm{w}})}(v_m^\eta,0) \leqslant C\eta^{\frac{2\beta}{r}} + d_{L^r(L^2_{\mathrm{w}})}(v_m^\eta,0).$$

We first take the superior limit as m tends to infinity of the above inequality, use (79) and then let η tend to zero to obtain $d_{L^r(L^2_{\mathrm{w}})}(v_m,0) \to 0$ as $m \to \infty$, which completes the proof. ∎

Lemma 6.4. *Let A be a complete metric space and $\{K_m\}_{m\in\mathbb{N}}$ be a sequence of compact sets in A. Then $\bigcup_{m\in\mathbb{N}} K_m$ is relatively compact in A if and only if, for any sequence $\{x_m\}_{m\in\mathbb{N}}$ such that $x_m \in K_m$ for all m, the set $\{x_m : m \in \mathbb{N}\}$ is relatively compact in A.*

Proof. Let $Z := \bigcup_{m\in\mathbb{N}} K_m$. If Z is relatively compact in A, then $\{x_m : m \in \mathbb{N}\}$ is also relatively compact in A since it is included in Z. We now prove the converse statement, by way of contradiction.
Let $\varepsilon > 0$ and assume that Z is not covered by a finite number of balls of radius ε. Since each K_m is compact it has a finite covering $K_m \subset \bigcup_{i\in I_m} B_i$ by balls of radius ε. Let $m_1 = 1$ and take $x_{m_1} \in K_{m_1}$. By assumption, Z is not covered by $\bigcup_{i\in I_1} B_i \cup B(x_1, \varepsilon)$ so there is $m_2 \in \mathbb{N}$ and $x_{m_2} \in K_{m_2}$ such that $x_{m_2} \notin \bigcup_{i\in I_1} B_i \cup B(x_{m_1}, \varepsilon)$; in particular, $x_{m_2} \notin K_{m_1}$ so $m_2 > m_1 = 1$ and $d(x_{m_1}, x_{m_2}) \geqslant \varepsilon$. Still using the assumption $Z \not\subset \bigcup_{\ell=1}^{m_2} \bigcup_{i\in I_\ell} B_i \cup B(x_{m_1}, \varepsilon) \cup B(x_{m_2}, \varepsilon)$ so we can find $m_3 \in \mathbb{N}$ and $x_{m_3} \in K_{m_3}$ such that $x_{m_3} \notin \bigcup_{\ell=1}^{m_2} \bigcup_{i\in I_\ell} B_i \cup B(x_{m_1}, \varepsilon) \cup B(x_{m_2}, \varepsilon)$; since each K_ℓ, for $\ell = 1, \ldots, m_2$, is contained in $\bigcup_{i\in I_\ell} B_i$, we infer that $x_{m_3} \notin \bigcup_{\ell=1}^{m_2} K_\ell$, and thus that $m_3 > m_2$; additionally, $d(x_{m_1}, x_{m_3}) \geqslant \varepsilon$ and $d(x_{m_2}, x_{m_3}) \geqslant \varepsilon$.
Continuing the construction, we design a strictly increasing sequence $(m_k)_{k\in\mathbb{N}}$ of natural numbers and a sequence $(x_{m_k})_{k\in\mathbb{N}}$ such that $x_{m_k} \in K_{m_k}$ for all $k \in \mathbb{N}$, and

$$d(x_{m_k}, x_{m_j}) \geqslant \varepsilon \qquad \forall k \neq j. \tag{80}$$

The sequence $(x_{m_k})_{k\in\mathbb{N}}$ is incomplete, but can easily be completed into a sequence $(x_m)_{m\in\mathbb{N}}$ with $x_m \in K_m$ for all $m \in \mathbb{N}$. The assumption then tell us that $\{x_{m_k} : k \in \mathbb{N}\} \subset \{x_m : m \in \mathbb{N}\}$ is relatively compact. We should then be able to extract from $(x_{m_k})_{k\in\mathbb{N}}$ a converging subsequence, which contradicts the property (80) and completes the proof. ∎

Acknowledgement: this research was supported by the Australian Government through the Australian Research Council's Discovery Projects funding schemes (project number DP170100605 and DP160101755).

References

[1] C. Atkinson and C. W. Jones. Similarity solutions in some non-linear diffusion problems and in boundary-layer flow of a pseudo-plastic fluid. *The Quarterly Journal of Mechanics and Applied Mathematics*, 27(2):193–211, 05 1974.

[2] B. Ayuso de Dios, K. Lipnikov, and G. Manzini. The nonconforming virtual element method. *ESAIM: Math. Model Numer. Anal.*, 50(3):879–904, 2016.

[3] L. Banas, Z. Brzeźniak, M. Neklyudov, and A. Prohl. A convergent finite-element-based discretization of the stochastic Landau–Lifshitz–Gilbert equation. *IMA Journal of Numerical Analysis*, 34(2):502–549, April 2014.

[4] L. Banas, Z. Brzeźniak, and A. Prohl. Computational studies for the stochastic Landau–Lifshitz–Gilbert equation. *SIAM J. Sci. Comput.*, 35(1):B62–B81, 2013.

[5] L. Banas, Z. Brzeźniak, A. Prohl, and M. Neklyudov. A convergent finite-element-based discretization of the stochastic Landau–Lifshitz–Gilbert equation. *IMA Journal of Numerical Analysis*, 2013.

[6] J. W. Barrett and W. B. Liu. Finite element approximation of the parabolic p-Laplacian. *SIAM Journal on Numerical Analysis*, 31(2):413–428, 1994.

[7] L. Beirão da Veiga, F. Brezzi, A. Cangiani, G. Manzini, L. D. Marini, and A. Russo. Basic principles of virtual element methods. *Math. Models Methods Appl. Sci. (M3AS)*, 199(23):199–214, 2013.

[8] D. Breit. Regularity theory for nonlinear systems of spdes. *Manuscripta Math.*, 146:329–349, 2015.

[9] Z. Brzeźniak. On stochastic convolution in banach spaces and applications. *Stochastics and Stochastic Reports*, 61(3-4):245–295, 1997.

[10] Z. Brzeźniak, E. Carelli, and A. Prohl. Finite-element-based discretizations of the incompressible navierstokes equations with multiplicative random forcing. *IMA Journal of Numerical Analysis*, 33(3):771–824, 01 2013.

[11] E. Carelli and A. Prohl. Rates of convergence for discretizations of the stochastic incompressible navier–stokes equations. *SIAM Journal on Numerical Analysis*, 50(5):2467–2496, 2012.

[12] C. Carstensen, W. Liu, and N. Yan. A posteriori error estimates for finite element approximation of parabolic p-Laplacian. *SIAM Journal on Numerical Analysis*, 43(6):2294–2319, 2006.

[13] B. Cockburn, B. Dong, J. Guzmán, M. Restelli, and R. Sacco. A hybridizable discontinuous Galerkin method for steady-state convection-diffusion-reaction problems. *SIAM J. Sci. Comput.*, 31(5):3827–3846, 2009.

[14] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume Second edition. Cambridge: Cambridge University Press, 2014.

[15] A. Debussche, M. Hofmanova, and J. Vovelle. Degenerate parabolic stochastic partial differential equations: quasilinear case. *Ann. Probab.*, 44(3):1916–1955, 2016.

[16] D. A. Di Pietro and J. Droniou. *The Hybrid High-Order Method for Polytopal Meshes: Design, Analysis, and Applications*, volume 19 of *Modeling, Simulation and Applications*. Springer International Publishing, 2020.

[17] D. A. Di Pietro, J. Droniou, and G. Manzini. Discontinuous skeletal gradient discretisation methods on polytopal meshes. *J. Comput. Phys.*, 355:397–425, 2018.

[18] J. Droniou. Finite volume schemes for diffusion equations: introduction to and review of modern methods. *Math. Models Methods Appl. Sci.*, 24(8):1575–1619, 2014.

[19] J. Droniou, R. Eymard, T. Gallouët, C. Guichard, and R. Herbin. *The gradient discretisation method*, volume 82 of *Mathematics & Applications*. Springer, 2018.

[20] J. Droniou, R. Eymard, T. Gallouët, and R. Herbin. Gradient schemes: a generic framework for the discretisation of linear, nonlinear and nonlocal elliptic and parabolic equations. *Mathematical Models and Methods in Applied Sciences*, 23(13):2395–2432, 2013.

[21] J. Droniou, R. Eymard, and R. Herbin. Gradient schemes: generic tools for the numerical analysis of diffusion equations. *M2AN Math. Model. Numer. Anal.*, 50(3):749–781, 2016. Special issue – Polyhedral discretization for PDE.

[22] R. Eymard, C. Guichard, and R. Herbin. Small-stencil 3D schemes for diffusive flows in porous media. *ESAIM Math. Model. Numer. Anal.*, 46(2):265–290, 2012.

[23] B. Goldys, J. F. Grotowski, and K.-N. Le. Weak martingale solutions to the stochastic Landau–Lifshitz–Gilbert equation with multi-dimensional noise via a convergent finite-element scheme. *Stochastic Processes and their Applications*, 130(1):232 – 261, 2020.

[24] B. Goldys, K.-N. Le, and T. Tran. A finite element approximation for the stochastic Landau–Lifshitz–Gilbert equation. *Journal of Differential Equations*, 260(2):937 – 970, 2016.

[25] M. Hofmanova and T. Zhang. Quasilinear parabolic stochastic partial differential equations: existence, uniqueness. *Stochastic Process. Appl*, 127(10):3354–3371, 2017.

[26] L. Hornung. Quasilinear parabolic stochastic evolution equations via maximal L^p-regularity. *Potential Anal.*, 50(2):279–326, 2019.

[27] A. Ichikawa. Stability of semilinear stochastic evolution equations. *Journal of Mathematical Analysis and Applications*, 90(1):12 – 44, 1982.

[28] A. Jakubowski. The almost sure Skorokhod representation for subsequences in nonmetric spaces. *Theory of Probability & Its Applications*, 42(1):167–174, 1998.

[29] N. Ju. Numerical analysis of parabolic p-laplacian: Approximation of trajectories. *SIAM Journal on Numerical Analysis*, 37(6):1861–1884, 2000.

[30] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag Berlin Heidelberg, 1992.

[31] R. Kruse. *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*. Springer International Publishing, 2014.

[32] J. R. Philip. N-diffusion. *Australian Journal of Physics*, 14:1–13, 1961.

[33] C. Prévot and M. Röckner. *A concise course on stochastic partial differential equations*. Lecture Notes in Mathematics, 1905. Springer, 2007.

[34] D. C. Wilcox. *Turbulence modeling for CFD*. DWC Industries, La Canada, 1998.

[35] Z. Zhang and G. E. Karniadakis. *Numerical Methods for Stochastic Partial Differential Equations with White Noise*. Springer International Publishing, 2017.

School of Mathematics, Monash University, Clayton, Victoria 3800, Australia. jerome.droniou@monash.edu

School of Mathematics and Statistics, The University of Sydney, Sydney 2006, Australia beniamin.goldys@sydney.edu.au

School of Mathematics, Monash University, Clayton, Victoria 3800, Australia. ngan.le@monash.edu