



HAL
open science

Strong rates of convergence of space-time discretization schemes for the 2D Navier-Stokes equations with additive noise

Hakima Bessaih, Annie Millet

► **To cite this version:**

Hakima Bessaih, Annie Millet. Strong rates of convergence of space-time discretization schemes for the 2D Navier-Stokes equations with additive noise. *Stochastics and Dynamics*, 2022, Special Issue on Modern Topics on Stochastic Dynamics, to celebrate Bjorn Schmalfuss' birthday, 22 (2), pp.2240005. hal-03133228

HAL Id: hal-03133228

<https://hal.science/hal-03133228>

Submitted on 18 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SPACE-TIME EULER DISCRETIZATION SCHEMES FOR THE STOCHASTIC 2D NAVIER-STOKES EQUATIONS

HAKIMA BESSAIH AND ANNIE MILLET

ABSTRACT. We prove that the implicit time Euler scheme coupled with finite elements space discretization for the 2D Navier-Stokes equations on the torus subject to a random perturbation converges in $L^2(\Omega)$, and describe the rate of convergence for an H^1 -valued initial condition. This refines previous results which only established the convergence in probability of these numerical approximations. Using exponential moment estimates of the solution of the stochastic Navier-Stokes equations and convergence of a localized scheme, we can prove strong convergence of this space-time approximation. The speed of the $L^2(\Omega)$ -convergence depends on the diffusion coefficient and on the viscosity parameter. In case of Scott-Vogelius mixed elements and for an additive noise, the convergence is polynomial.

1. INTRODUCTION

Numerical schemes and algorithms have been introduced to best approximate and construct solutions for PDEs. A similar approach has started to emerge for stochastic models and in particular SPDEs. Many algorithms based on either finite difference, finite element or spectral Galerkin methods (for the space discretization), and on either Euler schemes, Crank-Nicolson or Runge-Kutta schemes (for the time discretization) have been introduced for both the linear and nonlinear cases. Their rates of convergence have been widely investigated. The literature on numerical analysis for SPDEs is now very extensive. In [1] the models are either linear, have global Lipschitz properties, or more generally some monotonicity property. In this case the convergence is proven to be in mean square. When nonlinearities are involved that are not of Lipschitz or monotone type, then a rate of convergence in mean square is more difficult to obtain. Indeed, because of the stochastic perturbation, one may not use the Gronwall lemma after taking the expectation of the error bound since it involves a nonlinear term which is often quadratic. One way to get around this problem is to localize the nonlinear term in order to get a linear inequality, and then use the Gronwall lemma. This gives rise to a rate of convergence in probability, that was first introduced by J. Printems [19].

In this paper, we focus on the stochastic 2D Navier-Stokes equations. Our goal is to implement a space-time discretization scheme, prove its $L^2(\Omega)$ convergence, and get appropriate rates of convergence.

Our model is given by

$$\partial_t u - \nu \Delta u + (u \cdot \nabla)u + \nabla \pi = G(u)dW \quad \text{in } (0, T) \times D, \quad (1.1)$$

$$\operatorname{div} u = 0 \quad \text{in } (0, T) \times D, \quad (1.2)$$

2000 *Mathematics Subject Classification.* Primary 60H15, 60H35; Secondary 76D06, 76M35.

Key words and phrases. Stochastic Navier-Stokes equations, Euler schemes, finite elements, strong convergence, implicit time discretization, exponential moments.

Hakima Bessaih was partially supported by Simons Foundation Grant 582264 .

where $D = [0, L]^2$, $T > 0$. The process $u : \Omega \times (0, T) \times D \rightarrow \mathbb{R}^2$ is the velocity field with initial condition $u_0 \in \mathbb{W}^{1,2}(D)$, periodic boundary conditions $u(t, x + Lv_i) = u(t, x)$ on $(0, T) \times \partial D$, where v_i , $i = 1, 2$ denotes the canonical basis of \mathbb{R}^2 , and $\pi : \Omega \times (0, T) \times D \rightarrow \mathbb{R}$ is the pressure. Here G is a diffusion coefficient with global Lipschitz conditions. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ denote a filtered probability space and W be a Wiener process to be precisely defined later. Unique global weak and strong solutions (in the PDE sense) for (1.1) are constructed for both additive and multiplicative noise, and without being exhaustive, we refer to [6, 10].

Various space-time numerical schemes have been studied for the stochastic Navier-Stokes equations with a multiplicative noise (1.1). We refer to [8, 13, 6, 9, 7], where convergence in probability is stated with various rates of convergence. As stated previously, the main tool to get the convergence in probability is the localization of the nonlinear term over a space of large probability. Our previous paper [3] describes most of the results contained in these papers. The first result on an strong, that is $L^2(\Omega)$ -convergence, rate is proved in [3] for an H^1 -valued initial condition. The method is based on the fact that the solution (and the scheme) have finite moments (bounded uniformly on the mesh). When the noise is additive, the solution has exponential moments; we used this property in [3] to get an explicit polynomial strong rate of convergence. Let us mention the result by Duan-Yang [12], where strong convergence results are stated, and proved using semi-group techniques; in that paper the initial condition $u_0 \in H^2(D)$ is more regular than in our setting.

In our previous paper [3], only a time discretization scheme for the model (1.1) was studied. In the current paper, we implement a space-time discretization. More specifically, we have an implicit time discretization coupled with a finite elements spatial discretization. A similar algorithm has been studied in [9], where convergence in probability was proven. Using the tools introduced in our previous paper [3], which are formalized in some general framework, we improve these results by proving rates of $L^2(\Omega)$ -convergence for a space-time scheme.

This paper deals with the fully implicit time Euler scheme with constant time mesh $\frac{T}{N}$ and the finite elements discretization with space parameter h used in [9]. For general finite elements, in order to ensure stability, we have to deal with the velocity and the pressure which satisfy the discrete LBB condition. As in the results on convergence in probability proved in [9], we compare the space-time discretization \mathbf{U}^l and the fully implicit time discretization \mathbf{u}^l in $\mathbb{L}^2 := \mathbb{L}^2(D)$ uniformly on the time grid $\{l\frac{T}{N}\}_{l=0, \dots, N}$. The strong speed of convergence we obtain for the velocity includes a term containing the pressure.

The proof relies first on a $L^2(\Omega)$ convergence result of $\max_{1 \leq l \leq N} 1_{\Omega_{l-1}(M)} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2$ localized by a set $\tilde{\Omega}_l(M)$. On that set, some power of the $\mathbb{W}^{1,2}$ -norm of the time discretization \mathbf{u}^l is controlled. However, due to the fact that the classical non-linear term $[u, \nabla]u$ has to be coupled with $\frac{1}{2}[\operatorname{div} u]u$, for general finite elements the localization set is $\tilde{\Omega}_l(M) := \{ \max_{j=0, \dots, l} \|\mathbf{u}^j\|_{\mathbb{W}^{1,2}}^4 \leq M \}$. Since we can control moments of the \mathbb{L}^2 -norm of both discretizations \mathbf{u}^l and \mathbf{U}^l , choosing the bound $M(N, h)$ in an appropriate way, we can prove a strong speed of convergence. In general, the error contains a term including the pressure and its discretization. To get more precise results, we have to make some assumptions either on the pressure or on the relation between h and T/N , or to strengthen the assumptions on the diffusion coefficient. In case of a multiplicative stochastic perturbation, for a constant divergence-free $\mathbb{W}^{1,2}$ -valued initial condition (or more generally for a Gaussian $\mathbb{W}^{1,2}$ -valued initial random variable), the strong speed of convergence of

$\max_l |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2$ is any negative power of $\ln(h^2 + T/N)$. If the noise is additive the speed of convergence is $C \exp(c\sqrt{\ln(h^2 + T/N)})$. The fact that the convergence is much faster is related to the existence of exponential moments of $\max_{0 \leq l \leq N} \|\mathbf{u}^l\|_{\mathbb{W}^{1,2}}^2$. However, the speed is not polynomial due to the fact that the localization involves the fourth power of the $\mathbb{W}^{1,2}$ norm of the time scheme $\{\mathbf{u}^l\}_l$.

In the particular case of divergence-free finite elements, such as the Scott-Vogelius mixed elements [20, 23] the localization set is $\Omega_l(M) := \{\max_{j=0, \dots, l} \|\mathbf{u}^j\|_{\mathbb{W}^{1,2}}^2 \leq M\}$. If the speed of convergence is similar to that in the case of general finite elements (with a higher power of $\ln(h^2 + T/N)$), the situation is different for an additive stochastic perturbation. As it was already observed in [3], where the existence of exponential moments for $\sup_{s \leq T} \|u(s)\|_{\mathbb{W}^{1,2}}^2$ was used, the localization by the square of $\|\mathbf{u}^l\|_{\mathbb{W}^{1,2}}$, and the existence of exponential moments for the time scheme provide a “polynomial” speed of convergence of $\max_{1 \leq l \leq N} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2$ in terms of $h^2 + T/N$. The exponent depends on the viscosity ν and on the “strength” of the noise. If the strength of the noise converges to 0, this exponent converges to 1.

In all cases, coupling the above results with the strong convergence of $\max_{1 \leq l \leq N} |u(lT/N) - \mathbf{u}^l|_{\mathbb{L}^2}^2$ from [3], we deduce the strong speed of convergence of our space-time Euler scheme. Note that due to the parabolic feature of the Navier-Stokes equations, the coupling of h^2 and T/N is to be expected.

The paper is organized as follows. In section 2 we define precisely our model, the stochastic perturbation, and recall classical results for the solution u to the stochastic 2D Navier-Stokes equations. Section 3 describes the fully implicit time Euler scheme; it recalls some bounds for moments proved in [8] and [9], as well as the strong speed of convergence of this time scheme proved in [3] for multiplicative and additive stochastic perturbations. It also recalls some results about averages of the time discretization of the pressure proved in [9]. In section 4, we introduce the finite elements, recall the definition of \mathbf{U}^l from [8, 9], as well as various bounds of moments for \mathbf{U}^l . Some technical estimates about the non-linear terms are proved, among which terms using some localization on abstract subsets of Ω . Choosing these subsets in an appropriate way and using a discrete version of Gronwall’s lemma, we prove the localized $L^2(\Omega)$ -convergence of the error between space-time Euler scheme \mathbf{U}^l and the fully implicit time scheme \mathbf{u}^l in section 5. In section 6 we introduce a general framework to deduce the strong speed of convergence from localized $L^2(\Omega)$ estimates and moment bounds. We then apply the general results to deduce the strong speed of convergence of $\max_{1 \leq l \leq N} |\mathbf{U}^l - u(lT/N)|_{\mathbb{L}^2}^2$ in the case of general finite elements, and then in that of divergence-free ones in section 7. Finally the existence of exponential moments for $\max_{1 \leq l \leq N} \|\mathbf{u}^l\|_{\mathbb{W}^{1,2}}^2$ is proven in the Appendix, both for a deterministic and random initial condition.

As usual, throughout the paper, C will denote a constant which can change from one line to the next, and we let $C(a)$ denote a constant depending on some parameter a .

2. NOTATIONS AND PRELIMINARY RESULTS

In what follows, we will consider velocity fields that have mean zero over $[0, L]^2$. Let $\mathbb{L}^p := [L_{per}^p(D)]^2$ (resp. $\mathbb{W}^{k,p} := [W_{per}^{k,p}(D)]^2$) denote the usual Lebesgue and Sobolev spaces of periodic vector-valued functions with mean zero over $[0, L]^2$ endowed with the norms $|\cdot|_{\mathbb{L}^p}$ (resp. $\|\cdot\|_{\mathbb{W}^{k,p}}$). To ease notations, we will denote by $\|\cdot\|_k$ the $\mathbb{W}^{k,p}$ -norm.

Let

$$H := \{u \in \mathbb{L}^2 : \operatorname{div} u = 0 \text{ weakly in } D\}, \quad V := H \cap \mathbb{W}^{1,2};$$

these are separable Hilbert spaces. The space H inherits its inner product denoted by (\cdot, \cdot) and its norm from \mathbb{L}^2 . The norm in V , inherited from $\mathbb{W}^{1,2}$, is denoted by $\|\cdot\|_V$. Moreover, let V' be the dual space of V with respect to the Gelfand triple, $\langle \cdot, \cdot \rangle$ denotes the duality between V' and V . Let $A = -\Delta$ with its domain $\operatorname{Dom}(A) = \mathbb{W}^{2,2} \cap H$.

Let $b : [\mathbb{W}^{1,2}]^3 \rightarrow \mathbb{R}$ denote the trilinear map defined by

$$b(u_1, u_2, u_3) := \int_D (u_1(x) \cdot \nabla u_2(x)) \cdot u_3(x) dx,$$

which by the incompressibility condition satisfies

$$b(u_1, u_2, u_3) = -b(u_1, u_3, u_2), \quad b(u_1, u_2, u_2) = 0, \quad \forall u_1 \in V, \forall u_2, u_3 \in \mathbb{W}^{1,2}. \quad (2.1)$$

There exists a continuous bilinear map $B : V \times V \mapsto V'$ such that

$$\langle B(u_1, u_2), u_3 \rangle = b(u_1, u_2, u_3), \quad \text{for all } u_i \in V, i = 1, 2, 3.$$

Furthermore, the Gagliardo-Nirenberg inequality implies

$$\|u\|_{\mathbb{L}^4}^2 \leq \bar{C} \|u\|_{\mathbb{L}^2} \|\nabla u\|_{\mathbb{L}^2} \leq \frac{\bar{C}}{2} \|u\|_{\mathbb{W}^{1,2}}^2 \quad (2.2)$$

for some positive constant \bar{C} . Recall some well-known properties of b , which easily follow from the Hölder and Young inequalities: given any $\beta > 0$ we have for $X := H \cap \mathbb{L}^4(D)$

$$|\langle B(u_1, u_2), u_3 \rangle| \leq \beta \|u_3\|_V^2 + \frac{1}{4\beta} \|u_1\|_X \|u_2\|_X, \quad (2.3)$$

$$|\langle B(u_1, u_1) - B(u_2, u_2), u_1 - u_2 \rangle| \leq \beta \|u_1 - u_2\|_V^2 + C_\beta \|u_1 - u_2\|_{\mathbb{L}^2}^2 \|u_1\|_X^4, \quad (2.4)$$

for $u_i \in V$, $i = 1, 2, 3$, and \bar{C} is defined by (2.2). Finally, recall that the Poincaré inequality implies the existence of a constant $\tilde{C} > 0$ such that if we set $\|u\|^2 = \|\nabla u\|_{\mathbb{L}^2}^2 + \|Au\|_{\mathbb{L}^2}^2$ for $u \in \operatorname{Dom}(A)$, then

$$\|u\|_V^2 = \|u\|_{\mathbb{L}^2}^2 + \|\nabla u\|_{\mathbb{L}^2}^2 \leq \tilde{C} \|u\|^2. \quad (2.5)$$

Finally, recall that since the domain D is the torus with periodic boundary conditions, the following identity involving the Stokes operator A and the bilinear term holds (see e.g. [21] Lemma 3.1):

$$\langle B(u, u), Au \rangle = 0, \quad \forall u \in \operatorname{Dom}(A). \quad (2.6)$$

Let K be a separable Hilbert space and Q be a symmetric, positive trace-class operator on K . Let $(W(t), t \in [0, T])$ be a K -valued Wiener process with covariance operator Q , defined on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. Let $\{e_j\}_{j \geq 1}$ denote an orthonormal basis of K made of eigenfunctions of Q , with eigenvalues $\{q_j\}_{j \geq 1}$. Then

$$W(t) = \sum_{j=1}^{\infty} \sqrt{q_j} \beta^j(t) e_j, \quad \forall t \in [0, T],$$

where $\{\beta_j\}_{j \geq 1}$ are independent one-dimensional Brownian motions defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. For technical reasons, we assume that the initial condition u_0 belongs to $L^p(\Omega; V)$ for some $p \in [2, \infty]$, and only consider *strong solutions* in the PDE sense. Given two Hilbert spaces H_1 and H_2 , let $\mathcal{L}(H_1, H_2)$ denote the set of linear operators from H_1 to H_2 . The diffusion coefficient G satisfies the following assumption:

Condition (G1) Assume that $G : \mathbb{L}^2 \rightarrow \mathcal{L}(K, \mathbb{W}^{1,2})$ is Lipschitz-continuous with linear growth. More precisely, there exist positive constants K_i , $i = 0, 1$ and L_1 such that

$$\|G(u)\|_{\mathcal{L}(K, \mathbb{L}^2)}^2 \leq K_0 + K_1 |u|_{\mathbb{L}^2}^2, \quad \forall u \in \mathbb{L}^2, \quad (2.7)$$

$$\|G(u)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \leq K_0 + K_1 |u|_{\mathbb{W}^{1,2}}^2, \quad \forall u \in \mathbb{W}^{1,2}, \quad (2.8)$$

$$\|G(u) - G(v)\|_{\mathcal{L}(K, \mathbb{L}^2)}^2 \leq L_1 |u - v|_{\mathbb{L}^2}^2, \quad \forall u, v \in \mathbb{L}^2. \quad (2.9)$$

We also consider the following variant of the assumptions on the diffusion coefficient G in terms of divergence-free fields.

Condition (G2) Assume that $G : H \rightarrow \mathcal{L}(K, V)$ is Lipschitz-continuous with linear growth. More precisely, there exist positive constants K_i , $i = 0, 1$ and L_1 such that

$$\|G(u)\|_{\mathcal{L}(K, H)}^2 \leq K_0 + K_1 |u|_H^2, \quad \forall u \in H, \quad (2.10)$$

$$\|G(u)\|_{\mathcal{L}(K, V)}^2 \leq K_0 + K_1 |u|_V^2, \quad \forall u \in V, \quad (2.11)$$

$$\|G(u) - G(v)\|_{\mathcal{L}(K, H)}^2 \leq L_1 |u - v|_H^2, \quad \forall u, v \in H. \quad (2.12)$$

Let \mathbf{P}_H denote the Leray projection. Note that the conditions (2.7) and (2.8) imply a linear growth of the trace of $\mathbf{P}_H G(u) Q G^*(u) \mathbf{P}_H$ and $A^{\frac{1}{2}} \mathbf{P}_H G(u) Q G^*(u) \mathbf{P}_H A^{\frac{1}{2}}$. A similar result holds under condition **(G2)**.

We define a strong solution of (1.1) as follows (see Definition 2.1 in [9]):

Definition 2.1. *We say that equation (1.1) has a strong solution if:*

- u is an adapted V -valued process,
- \mathbb{P} a.s. we have $u \in C([0, T]; V) \cap L^2(0, T; \text{Dom}(A))$,
- \mathbb{P} a.s.

$$\begin{aligned} (u(t), \phi) + \nu \int_0^t (\nabla u(s), \nabla \phi) ds + \int_0^t \langle [u(s) \cdot \nabla] u(s), \phi \rangle ds \\ = (u_0, \phi) + \int_0^t (\phi, G(u(s)) dW(s)) \end{aligned}$$

for every $t \in [0, T]$ and every $\phi \in V$.

As usual, by projecting (1.1) on divergence free fields when dealing with the velocity, the pressure term can be disregarded and the velocity is implicitly in the space V . Lemma 2.1 in [9] (see also [2], Theorem 4.1) shows the following:

Theorem 2.2. *Assume that u_0 is a V -valued, \mathcal{F}_0 -measurable random variable such that $\mathbb{E}(\|u_0\|_V^{2p}) < \infty$ for some real number $p \in [2, \infty)$. Assume that the condition **(G1)** or **(G2)** is satisfied. Then there exists a unique solution u to equation (1.1). Furthermore, for some positive constant C we have*

$$\mathbb{E} \left(\sup_{t \in [0, T]} \|u(t)\|_V^{2p} + \int_0^T |Au(s)|_{\mathbb{L}^2}^2 (1 + \|u(s)\|_V^{2(p-1)}) ds \right) \leq C [1 + \mathbb{E}(\|u_0\|_V^{2p})]. \quad (2.13)$$

3. TIME EULER DISCRETIZATION SCHEME

In this section, we describe the fully implicit time Euler scheme of the stochastic 2D Navier-Stokes equations introduced by E. Carelli and A. Prohl in [9], and recall the strong convergence proved in [3]. Fix $N \geq 1$, let $k = \frac{T}{N}$ denote the constant time mesh, and let $t_l = l \frac{T}{N}$, $l = 0, \dots, N$, denote the time grid.

Fully implicit Euler scheme Let u_0 be a V -valued, \mathcal{F}_0 -measurable random variable; set $\mathbf{u}^0 = u_0$. Fix $N \geq 1$ and for $l = 1, \dots, N$, find pairs $(\mathbf{u}^l, \pi^l) \in V \times L^2_{per}$ such that \mathbb{P} a.s. for all $\phi \in \mathbb{W}^{1,2}$ and $\psi \in L^2_{per}$,

$$(\mathbf{u}^l - \mathbf{u}^{l-1}, \phi) + \frac{T}{N} \left[\nu(\nabla \mathbf{u}^l, \nabla \phi) + \langle (\mathbf{u}^l \cdot \nabla \mathbf{u}^l), \phi \rangle \right] - \frac{T}{N} (\pi^l, \operatorname{div} \phi) = (G(\mathbf{u}^{l-1})) \Delta_l W, \phi), \quad (3.1)$$

$$(\operatorname{div} \mathbf{u}^l, \psi) = 0, \quad (3.2)$$

where $\Delta_l W = W(t_l) - W(t_{l-1})$.

In this section, our aim is to recall bounds for the strong error of this Euler time scheme. Since we are looking for a V -valued process, we define the scheme for the velocity projected on divergence free fields and reformulate the algorithm as follows (see [9, Section 3]).

$$(\mathbf{u}^l - \mathbf{u}^{l-1}, \phi) + \frac{T}{N} \left[\nu(\nabla \mathbf{u}^l, \nabla \phi) + \langle (\mathbf{u}^l \cdot \nabla \mathbf{u}^l), \phi \rangle \right] = (G(\mathbf{u}^{l-1}) \Delta_l W, \phi), \quad \forall \phi \in V. \quad (3.3)$$

The following result proves the existence and uniqueness of the solution $\{\mathbf{u}^l, l = 0, \dots, N\}$ of (3.3); it provides moment estimates for this solution. Note that here only dyadic moments are computed because an the induction argument which relates two consecutive dyadic numbers (see step 4 of the proof of [8, Lemma 3.1]).

Lemma 3.1. [9, Lemma 3.1] Let $u_0 \in L^{2q}(\Omega; V)$ for some integer $q \in [1, \infty)$ be \mathcal{F}_0 -measurable, such that $\mathbb{E}(\|u_0\|_V^{2q}) \leq C$. Assume that G satisfies condition **(G1)**. Then there exists a unique solution $\{\mathbf{u}^l\}_{l=0}^N$ to (3.3) with $\mathbf{u}^0 = u_0$. The random variables \mathbf{u}^l are \mathcal{F}_{t_l} -measurable and belong to $L^2(\Omega; V)$ a.s. Furthermore,

$$\mathbb{E} \left(\max_{0 \leq l \leq N} \|\mathbf{u}^l\|_V^{2q} + 2\nu \frac{T}{N} \sum_{l=1}^N \|\mathbf{u}^l\|_V^{2q-2} |A\mathbf{u}^l|_{\mathbb{L}^2}^2 \right) \leq C_1(T, q), \quad (3.4)$$

$$\mathbb{E} \left(\sum_{l=1}^N \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2 \|\mathbf{u}^l\|_V^2 \right) \leq C_2(T, 2), \quad (3.5)$$

$$\mathbb{E} \left[\left(\sum_{l=1}^N \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2 \right)^{2q} + \left(\nu \frac{T}{N} \sum_{l=1}^N |A\mathbf{u}^l|_{\mathbb{L}^2}^2 \right)^{2q} \right] \leq C_3(T, q), \quad q = 1, 2, \quad (3.6)$$

where for $i = 1, 2, 3$, $C_i(T, q) := C_i(T, q, \operatorname{Tr}(Q), K_0, K_1, L_1, \mathbb{E}(\|u_0\|_V^{2q}))$ is a constant which does not depend on N .

The following result about the pressure term will be used in the study of space-time discretization.

Lemma 3.2. [9, Lemma 3.2] Let $q \in [2, \infty)$ and $\{\mathbf{u}^l, \pi^l\}_{l=0, \dots, N}$ be the solution to (3.1) and (3.2).

(i) Suppose that G satisfies the growth and Lipschitz conditions **(G1)**. Then

$$\mathbb{E} \left(\frac{T}{N} \sum_{l=1}^N |\nabla \pi^l|_{\mathbb{L}^2}^2 \right) \leq C(T)N. \quad (3.7)$$

(ii) Suppose that G satisfies the growth and Lipschitz conditions **(G2)**. Then

$$\mathbb{E}\left(\frac{T}{N} \sum_{l=1}^N |\nabla \pi^l|_{\mathbb{L}^2}^2\right) \leq C(T). \quad (3.8)$$

The following theorem gives the strong rate of convergence for a multiplicative noise with linear growth. If u_0 is deterministic or Gaussian, then the speed of convergence is any negative power of $\ln(N)$.

Theorem 3.3. [3, Theorem 4.4] *Let u_0 be such that $\mathbb{E}(\|u_0\|_V^{2q}) < \infty$ for some $q \geq 3$, G satisfy assumption **(G1)** or **(G2)**. Then the fully implicit scheme $\{\mathbf{u}^l\}_l$ solution of (3.3) converges in $L^2(\Omega)$ to the solution u of (1.1). More precisely, for N large enough we have*

$$\mathbb{E}\left(\max_{1 \leq l \leq N} |u(t_l) - \mathbf{u}^l|_{\mathbb{L}^2}^2 + \frac{T}{N} \sum_{l=1}^N |\nabla(u(t_l) - \mathbf{u}^l)|_{\mathbb{L}^2}^2\right) \leq C[\ln(N)]^{-(2q-1-1)}. \quad (3.9)$$

The next result proves a better rate of convergence when the noise is additive. This is due to the existence of finite exponential moments for $\sup_{t \in [0, T]} \|u(t)\|_V$.

Theorem 3.4. [3] [Theorem 4.6] *Let $u_0 \in V$, G satisfy assumption **(G1)** with $K_1 = 0$. Let u denote the solution of (1.1) and $\{\mathbf{u}^l\}_l$ be the fully implicit scheme solution of (3.3). Then for N large enough, \bar{C} (resp. \tilde{C}) defined by (2.2) (resp. (2.5)),*

$$\mathbb{E}\left(\max_{1 \leq l \leq N} |u(t_l) - \mathbf{u}^l|_{\mathbb{L}^2}^2 + \frac{T}{N} \sum_{l=1}^N |\nabla[u(t_l) - \mathbf{u}^l]|_{\mathbb{L}^2}^2\right) \leq C\left(\frac{T}{N}\right)^\beta, \quad (3.10)$$

where for $\tilde{\alpha}_0 := \frac{\nu}{2K_0 \text{Tr} Q \bar{C}}$ we have

$$\beta < \beta_0 := \frac{1}{2} \left(\frac{\tilde{\alpha}_0}{\tilde{\alpha}_0 + \frac{\bar{C}^2 T}{2\nu}} \right). \quad (3.11)$$

In reference [3], we had different assumptions for the splitting scheme and the Euler one. The exponential moments were proven for the splitting scheme setting and when applying the result for the Euler scheme, we forgot to insert the trace of Q . Furthermore, using Proposition [3, Proposition 4.3], Theorem 8.2, Theorem 6.1 and Corollary 6.2, we can change $\tilde{\alpha}_0$ given in [3] to the above constant.

4. SPACE-TIME DISCRETIZATION

4.1. Description of the finite elements method. When studying a space time discretization using finite elements, one needs to have a stable pairing of the velocity and the pressure which satisfy the discrete LBB-condition (see e.g. [9], page 2469 and pages 2487-2489). Stability issues are crucial, and the pressure has to be discretized together with the velocity.

Let \mathcal{T}_h be a quasi-uniform triangulation of the domain $D \subset \mathbb{R}^2$, using triangles of maximal diameter $h > 0$, and set $\bar{D} = \cup_{K \in \mathcal{T}_h} \bar{K}$. Let $\mathbb{P}_i(K) := [P_i(K)]^2$ denote the space of polynomial vector fields on K of degree less than or equal to i . We define finite elements function spaces

$$\begin{aligned} \mathbb{H}_h &:= \{\mathbf{U} \in C^0(\bar{D}) \cap \mathbb{W}_{per}^{1,2}(D) : \mathbf{U} \in \mathbb{P}_1(K), \quad \forall K \in \mathcal{T}_h\}, \\ L_h &:= \{\Pi \in L_{per}^2(D) : \Pi \in P_1(K), \quad \forall K \in \mathcal{T}_h\}, \end{aligned}$$

which satisfy the discrete LBB-condition

$$\sup_{\Phi \in \mathbb{H}_h} \frac{(\operatorname{div} \Phi, \Pi)}{|\nabla \Phi|_{\mathbb{L}^2}} \geq C |\Pi|_{\mathbb{L}^2}, \quad \forall \Pi \in L_h, \quad (4.1)$$

with a constant $C > 0$ independent of the mesh size $h > 0$. Here $C^0(\bar{D})$ denotes the set of continuous vector fields on \bar{D} .

Define the subset $\mathbb{V}_h \subset \mathbb{H}_h$ of discrete divergence-free vector fields

$$\mathbb{V}_h := \{\Phi \in \mathbb{H}_h : (\operatorname{div} \Phi, \Lambda) = 0, \quad \forall \Lambda \in L_h\}. \quad (4.2)$$

Note that in general $\mathbb{V}_h \not\subset V$.

A way around this problem is to choose a space approximation such that $\mathbb{V}_h \subset V$, such as the Scott-Vogelius mixed elements (see [20] and [23]). This particular case yields a better approximation. Indeed, on one hand the pressure will not appear in the upper estimate, and on the other hand a different localization will provide a polynomial error in the case of an additive noise.

Let $\mathbf{Q}_h^0 : \mathbb{L}^2 \mapsto \mathbb{V}_h$ (resp. $P_h^0 : L_{per}^2(D) \mapsto L_h$) denote the orthogonal projection defined by

$$(\mathbf{z} - \mathbf{Q}_h^0 \mathbf{z}, \Phi) = 0, \quad \forall \Phi \in \mathbb{V}_h, \quad (\text{resp. } (z - P_h^0 z, \Lambda) = 0, \quad \forall \Lambda \in L_h). \quad (4.3)$$

The following estimates are standard (see e.g. [15])

$$|\mathbf{z} - \mathbf{Q}_h^0 \mathbf{z}|_{\mathbb{L}^2} + h |\nabla(\mathbf{z} - \mathbf{Q}_h^0 \mathbf{z})|_{\mathbb{L}^2} \leq C h^2 |A\mathbf{z}|_{\mathbb{L}^2}, \quad \forall \mathbf{z} \in V \cap \mathbb{W}^{2,2}(D), \quad (4.4)$$

$$|\mathbf{z} - \mathbf{Q}_h^0 \mathbf{z}|_{\mathbb{L}^2} \leq C h |\nabla \mathbf{z}|_{\mathbb{L}^2}, \quad \forall \mathbf{z} \in V, \quad (4.5)$$

$$|z - P_h^0 z|_{L^2} \leq C h |\nabla z|_{L^2}, \quad \forall z \in W_{per}^{1,2}(D). \quad (4.6)$$

Using the Gagliardo-Nirenberg inequality (2.2), we deduce from (4.4) and (4.5) the following upper estimates for $\mathbf{z} \in V \cap \mathbb{W}^{2,2}(D)$

$$\|\mathbf{z} - \mathbf{Q}_h^0 \mathbf{z}\|_{\mathbb{L}^4} \leq C h^{\frac{3}{2}} |A\mathbf{z}|_{\mathbb{L}^2}, \quad \text{and } \|\mathbf{z} - \mathbf{Q}_h^0 \mathbf{z}\|_{\mathbb{L}^4} \leq C h |A\mathbf{z}|_{\mathbb{L}^2}^{\frac{1}{2}} |\nabla \mathbf{z}|_{\mathbb{L}^2}^{\frac{1}{2}}. \quad (4.7)$$

4.2. Description of the space-time schemes. Unlike the time discretization, we will need to approximate the initial condition u_0 and replace it by an \mathbb{H}_h -valued random variable \mathbf{U}^0 . As in [9], we suppose that

$$\mathbb{E}(|u_0 - \mathbf{U}^0|_{\mathbb{L}^2}^2) \leq C h^2, \quad \mathbb{E}(|\nabla \mathbf{U}^0|_{\mathbb{L}^2}^2) \leq C \quad (4.8)$$

for some positive constant C . As it is usual in this framework, to ease notations we let $k := \frac{T}{N}$ denote the constant time mesh.

For general finite elements satisfying the discrete LBB condition (4.1), one has to change the tri-linear term $b(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) = ([\mathbf{U}_1 \cdot \nabla] \mathbf{U}_2, \mathbf{U}_3)$ to control the nonlinear effect in the presence of discretely divergence-free velocity iterates, and thus to allow stability of the scheme. Thus, we set

$$\tilde{b}(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) := ([\mathbf{U}_1 \cdot \nabla] \mathbf{U}_2, \mathbf{U}_3) + \frac{1}{2} ([\operatorname{div} \mathbf{U}_1] \mathbf{U}_2, \mathbf{U}_3), \quad \forall \mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3 \in \mathbb{W}^{1,2}. \quad (4.9)$$

Note that this trilinear term is anti-symmetric with respect to the last two variables, i.e.,

$$\tilde{b}(\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3) = -\tilde{b}(\mathbf{U}_1, \mathbf{U}_3, \mathbf{U}_2), \quad \forall \mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3 \in \mathbb{W}^{1,2}. \quad (4.10)$$

Therefore,

$$\tilde{b}(\mathbf{U}, \Phi, \Phi) = 0, \quad \forall \mathbf{U}, \Phi \in \mathbb{W}^{1,2}.$$

Algorithm 1. Let \mathbf{U}^0 be an \mathcal{F}_0 -measurable, \mathbb{H}_h -valued random variable. For every $l = 1, \dots, N$, we consider a pair of $\mathbb{H}_h \times L_h$ random variables (\mathbf{U}^l, Π^l) such that for every pair $(\Phi, \Lambda) \in \mathbb{H}_h \times L_h$, we have a.s.

$$\begin{aligned} (\mathbf{U}^l - \mathbf{U}^{l-1}, \Phi) + k\nu(\nabla\mathbf{U}^l, \nabla\Phi) + k([\mathbf{U}^{l-1} \cdot \nabla]\mathbf{U}^l, \Phi) + \frac{k}{2}([\operatorname{div}\mathbf{U}^{l-1}]\mathbf{U}^l, \Phi) \\ - k(\Pi^l, \operatorname{div}\Phi) = (\mathbf{G}(\mathbf{U}^{l-1})\Delta_l W, \Phi), \end{aligned} \quad (4.11)$$

$$(\operatorname{div}\mathbf{U}^l, \Lambda) = 0. \quad (4.12)$$

The following result, which states the existence and uniqueness of the pairs $(\mathbf{U}^l, \Pi^l)_{l=1, \dots, N}$ and provides moments of the solution, has been proven in [8, Lemma 3.1] when \mathbf{U}^0 is deterministic (see also [9, Lemma 4.1] for a random initial condition). Once more the exponents are dyadic numbers because of an induction argument which enables to deduce results when doubling the exponent.

Lemma 4.1. Let $q \in [1, \infty)$ and \mathbf{U}^0 be an \mathcal{F}_0 -measurable, \mathbb{H}_h -valued random variable which satisfies (4.8), and such that $\mathbb{E}(|\mathbf{U}^0|_{\mathbb{L}^2}^{2q}) \leq C$ for some positive constant C independent of $h > 0$. Suppose that the coefficient G satisfies condition **(G1)**. Then for every $l = 1, \dots, N$, there exists a unique pair (\mathbf{U}^l, Π^l) of \mathcal{F}_{lk} -measurable, $\mathbb{V}_h \times L_h$ -valued random variables which satisfy (4.11)-(4.12). Furthermore,

$$\mathbb{E}\left(\max_{0 \leq l \leq N} |\mathbf{U}^l|_{\mathbb{L}^2}^{2q} + \nu k \sum_{l=1}^N |\mathbf{U}^l|_{\mathbb{L}^2}^{2q-1} |\nabla\mathbf{U}^l|_{\mathbb{L}^2}^2\right) \leq \mathbf{C}_1(T, q), \quad (4.13)$$

$$\mathbb{E}\left[\left(k \sum_{l=1}^N |\nabla\mathbf{U}^l|_{\mathbb{L}^2}^2\right)^{2q-1}\right] \leq \mathbf{C}_2(T, q), \quad (4.14)$$

$$\mathbb{E}\left(\sum_{l=1}^N |\mathbf{U}^l - \mathbf{U}^{l-1}|_{\mathbb{L}^2}^2\right) \leq \mathbf{C}_3(T, 1), \quad (4.15)$$

where the constants $\mathbf{C}_i(T, q) := \mathbf{C}_i(T, q, \operatorname{Tr} Q, K_0, K_1, L_1, \mathbb{E}(|\mathbf{U}^0|_{\mathbb{L}^2}^{2q}))$, $i = 1, 2, 3$ do not depend on N and $h > 0$.

We can reformulate the algorithm (4.11)-(4.12) as follows, using divergence-free test functions (see [9, (4.4)]).

Algorithm 2. We have a.s. for $l = 1, \dots, N$

$$\begin{aligned} (\mathbf{U}^l - \mathbf{U}^{l-1}, \Phi) + \nu k(\nabla\mathbf{U}^l, \nabla\Phi) + k([\mathbf{U}^{l-1} \cdot \nabla]\mathbf{U}^l, \Phi) + \frac{k}{2}([\operatorname{div}\mathbf{U}^{l-1}]\mathbf{U}^l, \Phi) \\ = (\mathbf{G}(\mathbf{U}^{l-1})\Delta_l W, \Phi), \quad \forall \Phi \in \mathbb{V}_h. \end{aligned} \quad (4.16)$$

As in [9], we will compare the space-time scheme \mathbf{U}^l and the fully implicit time scheme \mathbf{u}^l . For $l = 0, \dots, N$, let $\mathbf{E}^l := \mathbf{u}^l - \mathbf{U}^l$. For every $l = 1, \dots, N$ and $\Phi \in \mathbb{H}_h$,

$$\begin{aligned} (\mathbf{E}^l - \mathbf{E}^{l-1}, \Phi) + \nu k(\nabla\mathbf{E}^l, \nabla\Phi) + k\tilde{b}(\mathbf{u}^l, \mathbf{u}^l, \Phi) - k\tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \Phi) \\ - k(\pi^l - \Pi^l, \operatorname{div}\Phi) = ([G(\mathbf{u}^{l-1}) - G(\mathbf{U}^{l-1})]\Delta_l W, \Phi). \end{aligned}$$

Since $\mathbf{U}^l \in \mathbb{V}_h$ for $l = 1, \dots, N$, we have $\mathbf{Q}_h^0 \mathbf{U}^l = \mathbf{U}^l$; in the above identity, choose

$$\Phi := \mathbf{Q}_h^0 \mathbf{E}^l = \mathbf{E}^l - (\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l).$$

Then, since $\mathbf{Q}_h^0 \mathbf{E}^l \in \mathbb{V}_h$ and $\Pi^l \in L_h$, using (4.2) we deduce that $(\Pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l) = 0$ for $l = 1, \dots, N$. Since $(a, a - b) = \frac{1}{2}(|a|_{\mathbb{L}^2}^2 - |b|_{\mathbb{L}^2}^2 + |a - b|_{\mathbb{L}^2}^2)$, we deduce

$$\begin{aligned} & \frac{1}{2} \left[|\mathbf{Q}_h^0 \mathbf{E}^l|_{\mathbb{L}^2}^2 - |\mathbf{Q}_h^0 \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + |\mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1})|_{\mathbb{L}^2}^2 \right] + \nu k |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \\ & \quad + k \tilde{b}(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) + k \tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) - k \tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l) \\ & = \nu k (\nabla \mathbf{E}^l, \nabla (\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l)) + k (\pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l) + \left([G(\mathbf{u}^{l-1})] - G(\mathbf{U}^{l-1}) \right) \Delta_l W, \mathbf{Q}_h^0 \mathbf{E}^l. \end{aligned} \quad (4.17)$$

4.3. Intermediate results. In this section, we will prove a series of estimates that will be crucial for the convergence results later on.

For $l = 0, \dots, N$, let $\Omega_l \subset \Omega$ be a decreasing family of \mathcal{F}_{t_l} -adapted sets, that is

$$\Omega_{l+1} \subset \Omega_l, \quad l = 0, \dots, N-1, \quad \text{and} \quad \Omega_l \in \mathcal{F}_{t_l}, \quad l = 0, \dots, N. \quad (4.18)$$

Part 1: Estimate of the error term $|(\nabla \mathbf{E}^l, \nabla [\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l])|$

Using the Cauchy-Schwarz and Young inequalities, and then (4.4), we obtain for $\epsilon > 0$

$$\begin{aligned} |(\nabla \mathbf{E}^l, \nabla (\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l))| & \leq |\nabla \mathbf{E}^l|_{\mathbb{L}^2} |\nabla (\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l)|_{\mathbb{L}^2} \\ & \leq \epsilon |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\epsilon) |\nabla (\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l)|_{\mathbb{L}^2}^2 \\ & \leq \epsilon |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\epsilon) h^2 |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2. \end{aligned}$$

This implies for $\epsilon > 0$ and $m = 1, \dots, N$

$$\begin{aligned} & \mathbb{E} \left(\max_{1 \leq n \leq m} \nu k \sum_{l=1}^n |1_{\Omega_{l-1}} (\nabla \mathbf{E}^l, \nabla (\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l))| \right) \\ & \leq \epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\epsilon) \nu k h^2 \mathbb{E} \left(\sum_{l=1}^N \nu |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2 \right) \\ & \leq \epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) h^2 C_1(T, 1), \end{aligned} \quad (4.19)$$

where in the last upper estimate we have used (3.4) in Lemma 3.1 with $q = 1$.

Part 2: Estimate of the term $\tilde{b}(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l)$

Using the antisymmetry properties (2.1) and (4.10), the Hölder inequality and the inequality $\|u\|_{\mathbb{L}^4} \leq C \|\nabla u\|_{\mathbb{L}^2}$ coming from the Sobolev embedding theorem, we deduce

$$\begin{aligned} & |b(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| + |\tilde{b}(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| \\ & \leq 2 |b(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l, \mathbf{u}^l)| + \frac{1}{2} |([\operatorname{div}(\mathbf{u}^l - \mathbf{u}^{l-1})] \mathbf{Q}_h^0 \mathbf{E}^l, \mathbf{u}^l)| \\ & \leq 2 \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_{\mathbb{L}^4} \|\nabla \mathbf{Q}_h^0 \mathbf{E}^l\|_{\mathbb{L}^2} \|\mathbf{u}^l\|_{\mathbb{L}^4} + \frac{1}{2} \|\nabla(\mathbf{u}^l - \mathbf{u}^{l-1})\|_{\mathbb{L}^2} \|\mathbf{Q}_h^0 \mathbf{E}^l\|_{\mathbb{L}^4} \|\mathbf{u}^l\|_{\mathbb{L}^4} \\ & \leq \epsilon \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon) \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_{\mathbb{V}}^2 \|\mathbf{u}^l\|_{\mathbb{V}}^2. \end{aligned}$$

Therefore, the upper estimate (3.5) in Lemma 3.1 imply

$$\mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} \left[|b(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| + |\tilde{b}(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| \right] \right)$$

$$\leq \epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) C_2(T, 2) k. \quad (4.20)$$

Part 3: Estimate of the term $-\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) + \tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l)$

For every $l = 1, \dots, N$, we have

$$\begin{aligned} & -\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) + \tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l) \\ &= -\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l - \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l) - \tilde{b}(\mathbf{u}^{l-1} - \mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l). \end{aligned}$$

The antisymmetry properties of b (resp. \tilde{b}) in (2.1) (resp. (4.10)) implies for $l = 1, \dots, N$

$$-\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) + \tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l) = \sum_{i=1}^4 \tilde{T}_i(l), \quad (4.21)$$

$$-b(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) + b(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l) = \sum_{i=1}^4 T_i(l), \quad (4.22)$$

where

$$\begin{aligned} \tilde{T}_1(l) &= \tilde{b}(\mathbf{u}^{l-1}, \mathbf{E}^l, \mathbf{E}^l - \mathbf{Q}_h^0 \mathbf{E}^l), & \tilde{T}_2(l) &= \tilde{b}(\mathbf{E}^{l-1}, \mathbf{E}^l, \mathbf{Q}_h^0 \mathbf{E}^l - \mathbf{E}^l), \\ \tilde{T}_3(l) &= -\tilde{b}(\mathbf{E}^{l-1}, \mathbf{u}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l), & \tilde{T}_4(l) &= \tilde{b}(\mathbf{E}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l, \mathbf{u}^l - \mathbf{u}^{l-1}), \end{aligned}$$

and the terms $T_i(l)$, $i = 1, \dots, 4$ are similar to the corresponding terms $\tilde{T}_i(l)$ replacing \tilde{b} by b .

Terms $\tilde{T}_1(l)$ and $T_1(l)$. The antisymmetry of \tilde{b} and the identity

$$\mathbf{E}^l - \mathbf{Q}_h^0 \mathbf{E}^l = \mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l \quad (4.23)$$

imply that $\tilde{T}_1(l) = -\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l, \mathbf{E}^l)$. Using the Hölder and Gagliardo-Nirenberg inequalities, (2.2) and (4.4), we deduce

$$\begin{aligned} |b(\mathbf{u}^{l-1}, \mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l, \mathbf{E}^l)| &\leq \|\mathbf{u}^{l-1}\|_{\mathbb{L}^4} |\nabla(\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l)|_{\mathbb{L}^2} \|\mathbf{E}^l\|_{\mathbb{L}^4} \\ &\leq \sqrt{C} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} |\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \|\mathbf{u}^{l-1}\|_V h |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2} \\ &\leq \epsilon \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + |\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon) h^2 \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2, \end{aligned}$$

where the last upper estimate is deduced using Young's inequality. Furthermore, using once more the Gagliardo-Nirenberg inequality, (4.7) and then Young's inequality, we obtain

$$\begin{aligned} |([\operatorname{div} \mathbf{u}^{l-1}](\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l), \mathbf{E}^l)| &\leq |\nabla \mathbf{u}^{l-1}|_{\mathbb{L}^2} \|\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l\|_{\mathbb{L}^4} \|\mathbf{E}^l\|_{\mathbb{L}^4} \\ &\leq \sqrt{C} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} |\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \|\mathbf{u}^{l-1}\|_V h^{\frac{3}{2}} |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2} \\ &\leq \epsilon \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + |\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon) h^3 \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2. \end{aligned}$$

Therefore, the Cauchy-Schwarz inequality, the upper estimates (3.4) with $q = 2$ and (3.6) with $q = 1$ imply for $m = 1, \dots, N$

$$\begin{aligned} \mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} |\tilde{T}_1(l)| \right) &\leq 2\epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + 2\mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + C(\nu, \epsilon) h^2 \left\{ \mathbb{E} \left(\max_{0 \leq l \leq N} \|\mathbf{u}^l\|_V^4 \right) \right\}^{\frac{1}{2}} \left\{ \mathbb{E} \left(\nu k \sum_{l=1}^N |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2 \right) \right\}^{\frac{1}{2}} \end{aligned}$$

$$\leq 2\epsilon\nu\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) + 2\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\mathbf{E}^l|_{\mathbb{L}^2}^2\right) + C(\nu, \epsilon)h^2. \quad (4.24)$$

A similar computation implies

$$\begin{aligned} \mathbb{E}\left(\max_{1\leq n\leq m} k\sum_{l=1}^n 1_{\Omega_{l-1}}|T_1(l)|\right) &\leq \epsilon\nu\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) + \mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \\ &\quad + C(\nu, \epsilon)h^2. \end{aligned} \quad (4.25)$$

Terms $\tilde{T}_2(l)$ and $T_2(l)$. Using once more (4.23), we can replace the difference $\mathbf{E}^l - \mathbf{Q}_h^0\mathbf{E}^l$ by $\mathbf{u}^l - \mathbf{Q}_h^0\mathbf{u}^l$. The Gagliardo-Nirenberg inequality, (4.7) and Young's inequality imply

$$\begin{aligned} |b(\mathbf{E}^{l-1}, \mathbf{E}^l, \mathbf{u}^l - \mathbf{Q}_h^0\mathbf{u}^l)| &\leq \sqrt{C}|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}}|\mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}\|\mathbf{u}^l - \mathbf{Q}_h^0\mathbf{u}^l\|_{\mathbb{L}^4} \\ &\leq \epsilon\nu|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon\nu|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon)h^4|\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2|\mathbf{A}\mathbf{u}^l|_{\mathbb{L}^2}^2|\nabla\mathbf{u}^l|_{\mathbb{L}^2}^2 \\ &\leq \epsilon\nu|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon\nu|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon)h^4(|\mathbf{u}^{l-1}|_{\mathbb{L}^2}^2 + |\mathbf{U}^{l-1}|_{\mathbb{L}^2}^2)|\mathbf{A}\mathbf{u}^l|_{\mathbb{L}^2}^2|\nabla\mathbf{u}^l|_{\mathbb{L}^2}^2. \end{aligned}$$

Furthermore, the Hölder and Gagliardo-Nirenberg inequalities together with (4.7) and Young's inequality yield

$$\begin{aligned} |([\operatorname{div}\mathbf{E}^{l-1}]\mathbf{E}^l, \mathbf{u}^l - \mathbf{Q}_h^0\mathbf{u}^l)| &\leq \sqrt{C}|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}}|\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}}\|\mathbf{u}^l - \mathbf{Q}_h^0\mathbf{u}^l\|_{\mathbb{L}^4} \\ &\leq \epsilon\nu|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon\nu|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon)h^4|\mathbf{E}^l|_{\mathbb{L}^2}^2|\mathbf{A}\mathbf{u}^l|_{\mathbb{L}^2}^2|\nabla\mathbf{u}^l|_{\mathbb{L}^2}^2 \\ &\leq \epsilon\nu|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon\nu|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon)h^4(|\mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{U}^l|_{\mathbb{L}^2}^2)|\mathbf{A}\mathbf{u}^l|_{\mathbb{L}^2}^2|\nabla\mathbf{u}^l|_{\mathbb{L}^2}^2. \end{aligned}$$

Since the family of sets Ω_l is decreasing, we have $1_{\Omega_{l-1}}|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 \leq 1_{\Omega_{l-2}}|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2$ for $l = 2, \dots, N$. Therefore, using (4.8), the Cauchy-Schwarz and Young inequalities, (3.4) and (4.13) with $q = 3$, and (3.6) with $q = 1$, we deduce

$$\begin{aligned} \mathbb{E}\left(\max_{1\leq n\leq m} k\sum_{l=1}^n 1_{\Omega_{l-1}}|\tilde{T}_2(l)|\right) &\leq 2\nu\epsilon k\mathbb{E}(|\nabla\mathbf{E}^0|_{\mathbb{L}^2}^2) + 4\nu\epsilon\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \\ &\quad + C(\nu, \epsilon)h^4\mathbb{E}\left[\left(\max_{0\leq l\leq N}\{|\mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{U}^l|_{\mathbb{L}^2}^2\}\right)\left(\max_{1\leq l\leq N}\|\mathbf{u}^l\|_V^2\right)\left(\nu k\sum_{l=1}^N|\mathbf{A}\mathbf{u}^l|_{\mathbb{L}^2}^2\right)\right] \\ &\leq 2\nu\epsilon Ck + 4\nu\epsilon\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \\ &\quad + C(\nu, \epsilon)h^4\left[\left\{\mathbb{E}\left(\max_{0\leq l\leq N}\|\mathbf{u}^l\|_V^8\right)\right\}^{\frac{1}{2}} + \left\{\mathbb{E}\left(\max_{0\leq l\leq N}|\mathbf{U}^l|_{\mathbb{L}^2}^8\right)\right\}^{\frac{1}{2}}\right]\left\{\mathbb{E}\left[\left(\nu k\sum_{l=1}^N|\mathbf{A}\mathbf{u}^l|_{\mathbb{L}^2}\right)^2\right]\right\}^{\frac{1}{2}} \\ &\leq 4\nu\epsilon\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) + C(\nu, \epsilon)[k + h^4]. \end{aligned} \quad (4.26)$$

A similar computation yields

$$\mathbb{E}\left(\max_{1\leq n\leq m} k\sum_{l=1}^n 1_{\Omega_{l-1}}|T_2(l)|\right) \leq 2\nu\epsilon\mathbb{E}\left(k\sum_{l=1}^m 1_{\Omega_{l-1}}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) + C(\nu, \epsilon)(k + h^4). \quad (4.27)$$

Terms $\tilde{T}_3(l)$ and $T_3(l)$. The Gagliardo-Nirenberg inequality (2.2) implies

$$|b(\mathbf{E}^{l-1}, \mathbf{u}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l)| \leq \bar{C} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} |\nabla \mathbf{u}^{l-1}|_{\mathbb{L}^2} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} |\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}}.$$

If $l = 1$, Young's inequality implies

$$\begin{aligned} |b(\mathbf{E}^0, \mathbf{u}^0, \mathbf{Q}_h^0 \mathbf{E}^1)| &\leq \bar{C} (|\nabla \mathbf{E}^0|_{\mathbb{L}^2}^{\frac{1}{2}} \|\mathbf{u}^0\|_V^{\frac{1}{2}} |\mathbf{E}^0|_{\mathbb{L}^2}^{\frac{1}{2}}) (\|\mathbf{u}^0\|_V^{\frac{1}{2}} |\mathbf{E}^1|_{\mathbb{L}^2}^{\frac{1}{2}}) |\nabla \mathbf{E}^1|_{\mathbb{L}^2}^{\frac{1}{2}} \\ &\leq \epsilon \nu |\nabla \mathbf{E}^1|_{\mathbb{L}^2}^2 + \|\mathbf{u}^0\|_V^2 |\mathbf{E}^1|_{\mathbb{L}^2}^2 + C(\nu, \epsilon) \|\mathbf{u}^0\|_V [|\nabla \mathbf{u}^0|_{\mathbb{L}^2} + |\nabla \mathbf{U}^0|_{\mathbb{L}^2}] |\mathbf{E}^0|_{\mathbb{L}^2}. \end{aligned}$$

Hence Hölder's inequality, the upper estimates (4.8), (3.4) for $q = 2$ and (4.13) for $q = 2$ and imply

$$\begin{aligned} k\mathbb{E}(1_{\Omega_0} |T_3(1)|) &\leq k\epsilon \nu \mathbb{E}(1_{\Omega_0} |\nabla \mathbf{E}^1|_{\mathbb{L}^2}^2) + Ck \left\{ \mathbb{E}(\|\mathbf{u}^0\|_V^4) \right\}^{\frac{1}{2}} \left\{ \mathbb{E}(|\mathbf{u}^1|_{\mathbb{L}^2}^2 + |\mathbf{U}^1|_{\mathbb{L}^2}^2) \right\}^{\frac{1}{2}} \\ &\quad + kC(\nu, \epsilon) \left\{ \mathbb{E}(|\mathbf{E}^0|_{\mathbb{L}^2}^2) \right\}^{\frac{1}{2}} \left\{ \mathbb{E}(\|\mathbf{u}^0\|_V^4) \right\}^{\frac{1}{4}} \left\{ 2^3 \mathbb{E}(|\mathbf{u}^0|_{\mathbb{L}^2}^4) + 2^3 \mathbb{E}(|\mathbf{U}^0|_{\mathbb{L}^2}^4) \right\}^{\frac{1}{4}} \\ &\leq k\epsilon \nu \mathbb{E}(1_{\Omega_0} |\nabla \mathbf{E}^1|_{\mathbb{L}^2}^2) + C(\nu, \epsilon) k(1+h), \end{aligned} \quad (4.28)$$

for some constant $C(\nu, \epsilon)$ depending on \bar{C} , ϵ , ν , $C_1(T, 2)$ and $\mathbf{C}_1(T, 2)$.

For $l = 2, \dots, N$, using (2.2), the Hölder and Young inequalities we obtain for $\epsilon, \epsilon_1 > 0$

$$\begin{aligned} |b(\mathbf{E}^{l-1}, \mathbf{u}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l)| &\leq \bar{C} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \left(\|\mathbf{u}^{l-1}\|_V^{\frac{1}{2}} |\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \right) \left(\|\mathbf{u}^{l-2}\|_V^{\frac{1}{2}} |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} \right) \\ &\quad + \bar{C} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \left(\|\mathbf{u}^{l-1}\|_V^{\frac{1}{2}} |\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \right) \left(\|\mathbf{u}^{l-1} - \mathbf{u}^{l-2}\|_V^{\frac{1}{2}} |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} \right) \\ &\leq \epsilon_1 \nu |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon_1 \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + \frac{\bar{C}^2}{2^4 \epsilon_1 \nu} \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{E}^l|_{\mathbb{L}^2}^2 + \frac{\bar{C}^2}{2^4 \epsilon_1 \nu} \|\mathbf{u}^{l-2}\|_V^2 |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 \\ &\quad + \epsilon \nu |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{E}^l|_{\mathbb{L}^2}^2 + \frac{\bar{C}^4}{2^8 \epsilon^2 \nu^2} \|\mathbf{u}^{l-1} - \mathbf{u}^{l-2}\|_V^2 |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2. \end{aligned}$$

We next use the upper estimates (3.4) and (4.13) with $q = 2$, and (3.6) with $q = 1$ and the inequality $1_{\Omega_{l-1}} \leq 1_{\Omega_{l-2}}$. This yields for $m = 2, \dots, N$

$$\begin{aligned} \mathbb{E} \left(\max_{2 \leq n \leq m} k \sum_{l=2}^n 1_{\Omega_{l-1}} |T_3(l)| \right) &\leq 2(\epsilon_1 + \epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + \left[1 + \frac{2\bar{C}^2}{2^4 \epsilon_1 \nu} \right] \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + C(\epsilon, \nu) k \left\{ \mathbb{E} \left(\max_{1 \leq l \leq N} |\mathbf{u}^l|_{\mathbb{L}^2}^4 + \max_{1 \leq l \leq N} |\mathbf{U}^l|_{\mathbb{L}^2}^4 \right) \right\}^{\frac{1}{2}} \left\{ \mathbb{E} \left[\left(\sum_{l=1}^N \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2 \right)^2 \right] \right\}^{\frac{1}{2}} \\ &\leq 2(\epsilon_1 + \epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + \left[1 + \frac{\bar{C}^2}{2^3 \epsilon_1 \nu} \right] \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + C(\nu, \epsilon) k. \end{aligned} \quad (4.29)$$

The upper estimates (4.28) and (4.29) imply for $\epsilon, \epsilon_1 > 0$ and $h \in (0, 1]$

$$\begin{aligned} \mathbb{E} \left(k \max_{1 \leq n \leq m} \sum_{l=1}^n 1_{\Omega_{l-1}} |T_3(l)| \right) &\leq (2\epsilon_1 + 3\epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + \left[1 + \frac{\bar{C}^2}{2^3 \epsilon_1 \nu} \right] \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} \|\mathbf{u}^{l-1}\|_V^2 |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) k. \end{aligned} \quad (4.30)$$

Furthermore, the Gagliardo-Nirenberg (2.2) and Young inequalities imply for $\epsilon_2 > 0$

$$\begin{aligned} |([\operatorname{div} \mathbf{E}^{l-1}] \mathbf{u}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l)| &\leq \frac{\bar{C}}{\sqrt{2}} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2} \|\mathbf{u}^{l-1}\|_V |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} |\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \\ &\leq \epsilon_2 \nu |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon_2 \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + \frac{\bar{C}^4}{2^8 (\epsilon_2 \nu)^3} \|\mathbf{u}^{l-1}\|_V^4 |\mathbf{E}^l|_{\mathbb{L}^2}^2. \end{aligned}$$

Using once more the inequality $1_{\Omega_{l-1}} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 \leq 1_{\Omega_{l-2}} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2$, we deduce for any $m = 1, \dots, N$

$$\begin{aligned} \mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} |([\operatorname{div} \mathbf{E}^{l-1}] \mathbf{u}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l)| \right) &\leq 2\epsilon_2 \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + \frac{\bar{C}^4}{2^8 (\epsilon_2 \nu)^3} k \mathbb{E} \left(\sum_{l=1}^m 1_{\Omega_{l-1}} \|\mathbf{u}^{l-1}\|_V^4 |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right). \end{aligned} \quad (4.31)$$

The upper estimates (4.30) and (4.31) imply for $h \in (0, 1]$, $\epsilon, \epsilon_1, \epsilon_2 > 0$,

$$\begin{aligned} \mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} |\tilde{T}_3(l)| \right) &\leq (2\epsilon_1 + 2\epsilon_2 + 3\epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + Ck \\ &\quad + \mathbb{E} \left\{ k \sum_{l=1}^m 1_{\Omega_{l-1}} \left[\left(1 + \frac{\bar{C}^2}{2^3 \epsilon_1 \nu} \right) \|\mathbf{u}^{l-1}\|_V^2 + \frac{\bar{C}^4}{2^8 (\epsilon_2 \nu)^3} \|\mathbf{u}^{l-1}\|_V^4 \right] |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right\}. \end{aligned} \quad (4.32)$$

Terms $\tilde{T}_4(l)$ and $T_4(l)$. The Gagliardo-Nirenberg and Young inequalities imply

$$\begin{aligned} |b(\mathbf{E}^{l-1}, \mathbf{Q}_h^0 \mathbf{E}^l, \mathbf{u}^l - \mathbf{u}^{l-1})| &\leq \frac{\bar{C}}{\sqrt{2}} \|\nabla \mathbf{E}^{l-1}\|_{\mathbb{L}^2}^{\frac{1}{2}} |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^{\frac{1}{2}} |\nabla \mathbf{Q}_h^0 \mathbf{E}^l|_{\mathbb{L}^2} \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V \\ &\leq \epsilon \nu |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon \nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\epsilon, \nu) |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^4 \\ &\leq \epsilon \nu (|\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2) + C(\epsilon, \nu) (|\mathbf{u}^{l-1}|_{\mathbb{L}^2}^2 + |\mathbf{U}^{l-1}|_{\mathbb{L}^2}^2) (\|\mathbf{u}^l\|_V^2 + \|\mathbf{u}^{l-1}\|_V^2) \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2. \end{aligned}$$

For $l = 1$ the inequality (4.8) implies

$$k \epsilon \nu \mathbb{E} (|\nabla \mathbf{E}^0|_{\mathbb{L}^2}^2) \leq C k \nu \epsilon.$$

Furthermore, the Cauchy-Schwarz inequality together with the upper estimates (3.4), (4.13) for $q = 3$, (3.6) for $q = 1$, and the inequality $1_{\Omega_{l-1}} \leq 1_{\Omega_{l-2}}$ imply

$$\begin{aligned} \mathbb{E} \left(k \max_{1 \leq n \leq m} \sum_{l=1}^n 1_{\Omega_{l-1}} |T_4(l)| \right) &\leq 2\epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) k \\ &\quad + C(\epsilon, \nu) k \left\{ \mathbb{E} \left(\left| \sum_{l=1}^N \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2 \right|^2 \right) \right\}^{\frac{1}{2}} \left\{ \mathbb{E} \left(\max_{1 \leq l \leq N} \|\mathbf{u}^l\|_V^8 + \max_{1 \leq l \leq m} |\mathbf{U}^l|_{\mathbb{L}^2}^8 \right) \right\}^{\frac{1}{2}} \\ &\leq 2\epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) k. \end{aligned} \quad (4.33)$$

On the other hand, the Hölder, Gagliardo-Nirenberg and Young inequalities imply

$$\begin{aligned} |([\operatorname{div} \mathbf{E}^{l-1}] \mathbf{Q}_h^0 \mathbf{E}^l, \mathbf{u}^l - \mathbf{u}^{l-1})| &\leq |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2} \|\mathbf{E}^l\|_{\mathbb{L}^4} \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_{\mathbb{L}^4} \\ &\leq \bar{C} |\nabla \mathbf{E}^{l-1}|_{\mathbb{L}^2} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} \left(|\mathbf{E}^l|_{\mathbb{L}^2}^{\frac{1}{2}} |\nabla(\mathbf{u}^l - \mathbf{u}^{l-1})|_{\mathbb{L}^2}^{\frac{1}{2}} \max_{0 \leq l \leq N} \|\mathbf{u}^l\|_V^{\frac{1}{2}} \right) \end{aligned}$$

$$\leq \epsilon\nu|\nabla\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 + \epsilon\nu|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2 + C(\nu, \epsilon) \left[\max_{0 \leq l \leq N} \|\mathbf{u}^l\|_V^4 + \max_{0 \leq l \leq N} |\mathbf{U}^l|_{\mathbb{L}^2}^4 \right] \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2.$$

A similar argument as for the upper estimate of $T_4(l)$, based on (4.8), (3.4) and (4.13) with $q = 3$, (3.6) with $q = 1$, and $1_{\Omega_{l-1}} \leq 1_{\Omega_l}$ yield

$$\begin{aligned} & \mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} |([\operatorname{div} \mathbf{E}^{l-1}] \mathbf{Q}_h^0 \mathbf{E}^l, \mathbf{u}^l - \mathbf{u}^{l-1})| \right) \\ & \leq 2\epsilon\nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + \epsilon\nu C k \mathbb{E} (|\nabla \mathbf{E}^0|_{\mathbb{L}^2}^2) \\ & \quad + C(\nu, \epsilon) k \left[\left\{ \mathbb{E} \left(\max_{1 \leq l \leq N} \|\mathbf{u}^l\|_V^8 \right) \right\}^{\frac{1}{2}} + \left\{ \mathbb{E} \left(\max_{1 \leq l \leq N} |\mathbf{U}^l|_{\mathbb{L}^2}^8 \right) \right\}^{\frac{1}{2}} \right] \left\{ \mathbb{E} \left(\sum_{l=1}^N \|\mathbf{u}^l - \mathbf{u}^{l-1}\|_V^2 \right) \right\}^{\frac{1}{2}} \\ & \leq 2\epsilon\nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) k. \end{aligned} \quad (4.34)$$

The upper estimates (4.33) and (4.34) imply

$$\mathbb{E} \left(\max_{1 \leq n \leq m} \sum_{l=1}^m 1_{\Omega_{l-1}} |\tilde{T}_4(l)| \right) \leq 4\epsilon\nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) k. \quad (4.35)$$

The upper estimates (4.24), (4.26), (4.32) and (4.35) imply that for any $m = 1, \dots, N$,

$$\begin{aligned} & \mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} |\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) - \tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| \right) \\ & \leq C(\nu, \epsilon) (k + h^2 + h^4) + (2\epsilon_1 + 2\epsilon_2 + 13\epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ & \quad + \mathbb{E} \left[k \sum_{l=1}^m 1_{\Omega_{l-1}} \left\{ 2 + \left(1 + \frac{\bar{C}^2}{2^3 \epsilon_1 \nu} \right) \|\mathbf{u}^{l-1}\|_V^2 + \frac{\bar{C}^4}{2^8 \epsilon_2^3 \nu^3} \|\mathbf{u}^{l-1}\|_V^4 \right\} |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right]. \end{aligned} \quad (4.36)$$

Similarly, the upper estimates (4.25), (4.27) and (4.33) imply that for $m = 1, \dots, N$,

$$\begin{aligned} & \mathbb{E} \left(\max_{1 \leq n \leq m} k \sum_{l=1}^n 1_{\Omega_{l-1}} |b(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) - b(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| \right) \\ & \leq C(\nu, \epsilon) (k + h^2 + h^4) + (2\epsilon_1 + 8\epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ & \quad + \mathbb{E} \left[k \sum_{l=1}^m 1_{\Omega_{l-1}} \left\{ 1 + \left(1 + \frac{\tilde{C}^2}{2^3 \epsilon_1 \nu} \right) \|\mathbf{u}^{l-1}\|_V^2 \right\} |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right]. \end{aligned} \quad (4.37)$$

Part 4: Estimate of $(\pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l)$

Since $\mathbf{Q}_h^0 \mathbf{E}^l \in \mathbb{V}_h$ and $P_h^0 \pi^l \in L_h$, using (4.2) we deduce $(\pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l) = (\pi^l - P_h^0 \pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l)$. Therefore, the Cauchy-Schwarz and Young inequalities coupled with (4.6) imply

$$\mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |(\pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l)| \right) \leq \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} \left[\epsilon\nu |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 + \frac{1}{4\epsilon\nu} |\pi^l - P_h^0 \pi^l|_{L^2}^2 \right] \right)$$

$$\leq \epsilon \nu \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + C(\nu, \epsilon) h^2 \mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}} |\nabla \pi^l|_{L^2}^2 \right). \quad (4.38)$$

5. CONVERGENCE IN PROBABILITY

This section contains results on strong convergence localized on a subset of Ω due to the non linear term for both algorithms 1 and 2. This is classical when dealing with a non linear term. However, unlike [9], our bound in the localization only involves the time discretization $\{\mathbf{u}^l\}_l$. In the general case, it will require an upper bound of the fourth power of its V norm. Arguments similar to that used in [9] or [2], we could deduce a rate of convergence in probability. We do not provide this rate since our paper focuses on $L^2(\Omega)$ rates of convergence.

The following proposition is one of the main results of this section.

Proposition 5.1. *Let G satisfy the growth and Lipschitz conditions **(G1)**. Let $u_0 \in L^8(\Omega; V)$ and $\mathbf{U}^0 \in L^8(\Omega; \mathbb{H}_h)$ be \mathcal{F}_0 -measurable random variables such that (4.8) holds. For every $M > 0$ and $l = 0, \dots, N$, set*

$$\tilde{\Omega}_l(M) := \left\{ \omega \in \Omega : \max_{0 \leq j \leq l} \|\mathbf{u}^j\|_V^4 \leq M \right\}. \quad (5.1)$$

There exists $C_0 > 0$ such that for $k = \frac{T}{N} > 0$ small enough to have $kM \leq C_0$, we have for any $\delta > 0$ and $m = 1, \dots, N$

$$\mathbb{E} \left(1_{\tilde{\Omega}_{m-1}(M)} \max_{1 \leq n \leq m} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) \leq C(\nu) e^{\tilde{C}_1(M)T} \left[k + h^2 + h^2 \mathbb{E} \left(k \sum_{l=1}^m |\nabla \pi^l|_{L^2}^2 \right) \right], \quad (5.2)$$

for some constant

$$\begin{aligned} \tilde{C}_1(M) &= (1 + \delta) \frac{\bar{C}^4}{2^4 \nu^3} M + C(\nu, L_1 \text{Tr } Q, \delta) \\ &\equiv (1 + \delta) \frac{\bar{C}^4}{2^4 \nu^3} M \quad \text{for "large" } M. \end{aligned} \quad (5.3)$$

Furthermore,

$$\mathbb{E} \left(k \sum_{l=1}^m 1_{\tilde{\Omega}_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \leq C(\nu) \tilde{C}_1(M) T e^{\tilde{C}_1(M)T} \left[k + h^2 + h^2 \mathbb{E} \left(k \sum_{l=1}^m |\nabla \pi^l|_{L^2}^2 \right) \right]. \quad (5.4)$$

Proof. Note that the definition of $\tilde{\Omega}_l(M)$ clearly implies that for fixed $M > 0$, $l = 0, \dots, N-1$, $\tilde{\Omega}_{l+1}(M) \subset \tilde{\Omega}_l(M)$. Hence, as proved in [9] (see also [3]), we have

$$\begin{aligned} \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} \left(|\mathbf{Q}_h^0 \mathbf{E}^l|_{\mathbb{L}^2}^2 - |\mathbf{Q}_h^0 \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 \right) &= \left(1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 \right. \\ &\quad \left. + \sum_{l=2}^n (1_{\tilde{\Omega}_{l-2}(M)} - 1_{\tilde{\Omega}_{l-1}(M)}) |\mathbf{Q}_h^0 \mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 \right) - 1_{\tilde{\Omega}_0(M)} |\mathbf{Q}_h^0 \mathbf{E}^0|_{\mathbb{L}^2}^2 \\ &\geq 1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 - 1_{\tilde{\Omega}_0(M)} |\mathbf{Q}_h^0 \mathbf{E}^0|_{\mathbb{L}^2}^2, \quad \forall n = 1, \dots, N. \end{aligned} \quad (5.5)$$

This inequality and (4.17) imply for $n = 1, \dots, N$

$$\frac{1}{2} 1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 + \frac{1}{2} \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1})|_{\mathbb{L}^2}^2 + \nu k \sum_{l=1}^m 1_{\tilde{\Omega}_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2$$

$$\leq \sum_{i=1}^3 \tilde{R}_i(n, k, h, M), \quad (5.6)$$

where

$$\begin{aligned} \tilde{R}_1(n, k, h, M) &= 1_{\tilde{\Omega}_0(M)} |\mathbf{Q}_h^0 \mathbf{E}^0|_{\mathbb{L}^2}^2 + k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |\tilde{b}(\mathbf{u}^l - \mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| \\ &\quad + k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |\tilde{b}(\mathbf{u}^{l-1}, \mathbf{u}^l, \mathbf{Q}_h^0 \mathbf{E}^l) - \tilde{b}(\mathbf{U}^{l-1}, \mathbf{U}^l, \mathbf{Q}_h^0 \mathbf{E}^l)| \\ &\quad + k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |(\nabla \mathbf{E}^l, \nabla[\mathbf{u}^l - \mathbf{Q}_h^0 \mathbf{u}^l])| + k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |(\pi^l, \operatorname{div} \mathbf{Q}_h^0 \mathbf{E}^l)|, \\ \tilde{R}_2(n, k, h, M) &= \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} \left([G(\mathbf{u}^{l-1}) - G(\mathbf{U}^{l-1})] \Delta_l W, \mathbf{Q}_h^0 \mathbf{E}^{l-1} \right), \\ \tilde{R}_3(n, k, h, M) &= \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} \left([G(\mathbf{u}^{l-1}) - G(\mathbf{U}^{l-1})] \Delta_l W, \mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1}) \right). \end{aligned} \quad (5.7)$$

The upper estimates (4.8), (4.20), (4.36), (4.19) and (4.38) imply for any $\epsilon_1, \epsilon_2, \epsilon > 0$ and small $h > 0$

$$\begin{aligned} E(\tilde{R}_1(n, k, h, M)) &\leq C(\nu, \epsilon)(k + h^2) + C(\nu, \epsilon) h^2 \mathbb{E} \left(k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |\nabla \pi^l|_{L^2}^2 \right) \\ &\quad + (2\epsilon_1 + 2\epsilon_2 + 16\epsilon) \nu \mathbb{E} \left(k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ &\quad + \mathbb{E} \left(k \sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} \left[2 + \left(1 + \frac{\bar{C}^2}{2^3 \epsilon_1 \nu} \right) \sqrt{M} + \frac{\bar{C}^4}{2^8 \epsilon_2^3 \nu^3} M \right] |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right). \end{aligned} \quad (5.8)$$

Since $1_{\tilde{\Omega}_{l-1}(M)}, \mathbf{u}^{l-1}, \mathbf{U}^{l-1}$ and \mathbf{E}^{l-1} are independent of $\Delta_l W$, we have $\mathbb{E}(\tilde{R}_2(n, k, h, M)) = 0$.

Using the Lipschitz property of G (2.9), the Cauchy-Schwarz and Young inequalities, we obtain

$$\begin{aligned} \mathbb{E}(\tilde{R}_3(n, k, h, M)) &\leq \mathbb{E} \left[\sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |(G(\mathbf{u}^{l-1}) - G(\mathbf{U}^{l-1})) \Delta_l W|_{\mathbb{L}^2} |\mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1})|_{\mathbb{L}^2} \right] \\ &\leq \sum_{l=1}^n \left\{ \mathbb{E} \left(1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1})|_{\mathbb{L}^2}^2 \right) \right\}^{\frac{1}{2}} \\ &\quad \times \left\{ \mathbb{E} \left(1_{\tilde{\Omega}_{l-1}(M)} \|G(\mathbf{u}^{l-1}) - G(\mathbf{U}^{l-1})\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 k \operatorname{Tr} Q \right) \right\}^{\frac{1}{2}} \\ &\leq \frac{1}{2} \mathbb{E} \left(\sum_{l=1}^n 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1})|_{\mathbb{L}^2}^2 \right) + \frac{L_1 \operatorname{Tr} Q}{2} \mathbb{E} \left(k \sum_{l=1}^{n-1} 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + Ch^2 k, \end{aligned} \quad (5.9)$$

where in the last upper estimate we have used the inequalities $1_{\tilde{\Omega}_{l-1}(M)} \leq 1_{\tilde{\Omega}_{l-2}(M)}$ for $l = 2, \dots, N$, and (4.8) for $l = 0$.

Fix $\lambda \in (0, \frac{1}{2})$, let $\epsilon_1 = \frac{\lambda}{2}$, $\epsilon_2 = \frac{1}{2}(1 - 2\lambda)$ and $\epsilon = \frac{\lambda}{2^5}$. Then $2\epsilon_1 + 2\epsilon_2 + 16\epsilon = 1 - \frac{\lambda}{2}$. The upper estimates (5.6) – (5.9) yield

$$\begin{aligned} & \mathbb{E}\left(1_{\tilde{\Omega}_{n-1}(M)}|\mathbf{Q}_h^0\mathbf{E}^n|_{\mathbb{L}^2}^2\right) + \lambda\nu\mathbb{E}\left(k\sum_{l=1}^n1_{\tilde{\Omega}_{l-1}(M)}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \\ & \leq C(\nu)(k+h^2) + 2h^2\mathbb{E}\left(k\sum_{l=1}^n1_{\tilde{\Omega}_{l-1}(M)}|\nabla\pi^l|_{\mathbb{L}^2}^2\right) + L_1\text{Tr}Q\mathbb{E}\left(k\sum_{l=1}^{n-1}1_{\tilde{\Omega}_{l-1}(M)}|\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \\ & \quad + 2\mathbb{E}\left(k\sum_{l=1}^n1_{\tilde{\Omega}_{l-1}(M)}\left[2 + C(\nu, \bar{C}, \lambda)\sqrt{M} + \tilde{\alpha}(\nu, \bar{C}, \lambda)M\right]|\mathbf{E}^l|_{\mathbb{L}^2}^2\right), \end{aligned} \quad (5.10)$$

where $\tilde{\alpha}(\nu, \bar{C}, \lambda) := \frac{\bar{C}^4}{2^5(1-2\lambda)^3\nu^3}$. Using (4.5), we deduce that

$$|\mathbf{Q}_h^0\mathbf{E}^n - \mathbf{E}^n|_{\mathbb{L}^2}^2 = |\mathbf{Q}_h^0\mathbf{u}^n - \mathbf{u}^n|_{\mathbb{L}^2}^2 \leq Ch^2|\nabla\mathbf{u}^n|_{\mathbb{L}^2}^2.$$

Hence, (3.4) with $q = 1$ yields for $\lambda > 0$

$$\mathbb{E}\left(1_{\tilde{\Omega}_{n-1}(M)}|\mathbf{E}^n|_{\mathbb{L}^2}^2\right) \leq (1+\lambda)\mathbb{E}\left(1_{\tilde{\Omega}_{n-1}(M)}|\mathbf{Q}_h^0\mathbf{E}^n|_{\mathbb{L}^2}^2\right) + C(\lambda)h^2. \quad (5.11)$$

Plugging this upper estimate in (5.10), we deduce that for N large enough (that is k small enough) to ensure

$$2(1+\lambda)k\left[2 + C(\nu, \bar{C}, \lambda)\sqrt{M} + \tilde{\alpha}(\nu, \bar{C}, \lambda)M\right] \leq \lambda, \quad (5.12)$$

we have

$$\begin{aligned} & \mathbb{E}\left(1_{\tilde{\Omega}_{n-1}(M)}|\mathbf{E}^n|_{\mathbb{L}^2}^2\right) + \lambda\mathbb{E}\left(k\sum_{l=1}^n1_{\tilde{\Omega}_{l-1}(M)}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \leq C(\nu, \lambda)(k+h^2) \\ & \quad + C(\lambda)h^2\mathbb{E}\left(k\sum_{l=1}^n|\nabla\pi^l|_{L^2}^2\right) + \mathbb{E}\left(k\sum_{l=1}^{n-1}1_{\tilde{\Omega}_{l-1}(M)}\tilde{C}_2(M, \lambda)|\mathbf{E}^l|_{\mathbb{L}^2}^2\right), \end{aligned} \quad (5.13)$$

where using the Young inequality, we set

$$\begin{aligned} \tilde{C}_2(M, \lambda) &= \frac{(1+\lambda)\bar{C}^4}{2^4(1-\lambda)(1-2\lambda)^3\nu^3}M + C(\nu, \bar{C}, \lambda)\sqrt{M} + C(L_1\text{Tr}Q, \lambda) \\ &\leq \frac{(1+2\lambda)\bar{C}^4}{2^4(1-\lambda)(1-2\lambda)^3\nu^3}M + C(L_1\text{Tr}Q, \nu, \bar{C}, \lambda). \end{aligned}$$

Neglecting the gradient term and using the discrete version of Gronwall's lemma, we deduce for $\lambda \in (0, \frac{1}{2})$

$$\max_{0 \leq n \leq N} \mathbb{E}\left(1_{\tilde{\Omega}_{n-1}(M)}|\mathbf{E}^n|_{\mathbb{L}^2}^2\right) \leq \exp(T\tilde{C}_2(M, \lambda))\left[C(\nu)(k+h^2) + Ch^2\mathbb{E}\left(k\sum_{l=1}^N|\nabla\pi^l|_{L^2}^2\right)\right].$$

Plugging this upper estimate in (5.13), we obtain for $n \leq N$

$$\begin{aligned} & \mathbb{E}\left(k\sum_{l=1}^n1_{\tilde{\Omega}_{l-1}(M)}|\nabla\mathbf{E}^l|_{\mathbb{L}^2}^2\right) \leq C(\nu)\left[1 + \tilde{C}_2(M, \lambda)Te^{\tilde{C}_2(M, \lambda)T}\right] \\ & \quad \times \left[k+h^2 + h^2\mathbb{E}\left(k\sum_{l=1}^N|\nabla\pi^l|_{L^2}^2\right)\right]. \end{aligned} \quad (5.14)$$

We next study $\mathbb{E}(\max_{1 \leq n \leq m} 1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2)$. Using (5.6), we deduce

$$\begin{aligned} & \frac{1}{2} \mathbb{E} \left(\max_{1 \leq n \leq m} 1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 \right) + \mathbb{E} \left(\sum_{l=1}^m 1_{\tilde{\Omega}_{l-1}(M)} \left[\frac{1}{2} |\mathbf{Q}_h^0 (\mathbf{E}^l - \mathbf{E}^{l-1})|_{\mathbb{L}^2}^2 + \nu k |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right] \right) \\ & \leq 2 \sum_{i \in \{1,3\}} \mathbb{E}(\tilde{R}_i(m, k, h, M)) + 2 \mathbb{E} \left(\max_{1 \leq n \leq m} \tilde{R}_2(n, k, h, M) \right), \quad \forall m = 1, \dots, N. \end{aligned} \quad (5.15)$$

We rewrite $\tilde{R}_2(n, k, h, M)$ as a stochastic integral. For every $s \in [t_{l-1}, t_l]$ with $l = 1, \dots, N$, set $\underline{s} = t_{l-1}$, and write accordingly $\mathbf{u}(\underline{s}) = \mathbf{u}^{l-1}$, $\mathbf{U}(\underline{s}) = \mathbf{U}^{l-1}$, $\mathbf{E}(\underline{s}) = \mathbf{E}^{l-1}$ and $\tilde{\Omega}_{\underline{s}}(M) = \tilde{\Omega}_{l-1}(M)$. Then

$$\tilde{R}_2(n, k, h, M) = \int_0^{t_n} 1_{\tilde{\Omega}_{\underline{s}}(M)} ([G(\mathbf{u}(\underline{s})) - G(\mathbf{U}(\underline{s}))] dW(s), \mathbf{Q}_h^0 \mathbf{E}(\underline{s}))$$

The Davis inequality implies for $m = 2, \dots, N$

$$\begin{aligned} & \mathbb{E} \left(\max_{1 \leq n \leq m} \tilde{R}_2(n, k, h, M) \right) \\ & \leq 3 \mathbb{E} \left(\left\{ \int_0^{t_m} 1_{\tilde{\Omega}_{\underline{s}}(M)} \|G(\mathbf{u}(\underline{s})) - G(\mathbf{U}(\underline{s}))\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \text{Tr } Q |\mathbf{Q}_h^0 \mathbf{E}(\underline{s})|_{\mathbb{L}^2}^2 ds \right\}^{\frac{1}{2}} \right) \end{aligned}$$

(see e.g. [11, Theorem 3.14]). Using the Young inequality together with the upper estimates $1_{\tilde{\Omega}_{l-1}(M)} \leq 1_{\tilde{\Omega}_{l-1}(M)}$ for $l = 2, \dots, N$, and (4.8), we deduce for any $\lambda > 0$,

$$\begin{aligned} & \mathbb{E} \left(\max_{1 \leq n \leq m} \tilde{R}_2(n, k, h, M) \right) \\ & \leq 3 \mathbb{E} \left(\max_{2 \leq n < m} 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^{l-1}|_{\mathbb{L}^2} \left\{ \sum_{l=2}^m L_1 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{E}^{l-1}|_{\mathbb{L}^2}^2 k \text{Tr } Q \right\}^{\frac{1}{2}} \right) \\ & \quad + 3 \sqrt{L_1 \text{Tr } Q} \sqrt{k} \mathbb{E}(|\mathbf{E}^0|_{\mathbb{L}^2}^2) \\ & \leq \lambda \mathbb{E} \left(\max_{1 \leq n \leq m-1} 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + \frac{9L_1 \text{Tr } Q}{4\lambda} \mathbb{E} \left(k \sum_{l=1}^{m-1} 1_{\tilde{\Omega}_{l-1}(M)} |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) + Ch^2 \sqrt{k}. \end{aligned} \quad (5.16)$$

Using the above upper estimates of $\mathbb{E}(\max_{1 \leq n \leq m} \tilde{R}_i(n, k, h, M))$, $i = 1, 2, 3$, an argument similar to the previous ones yields for $\lambda \in (0, \frac{1}{2})$ and N large enough

$$\begin{aligned} \mathbb{E} \left(\max_{1 \leq n \leq m} 1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) & \leq \mathbb{E} \left(k \sum_{l=1}^{m-1} 1_{\tilde{\Omega}_{l-1}(M)} \tilde{C}_1(M, \lambda) |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \\ & \quad + C(\nu, \lambda) \left[k + h^2 + h^2 \mathbb{E} \left(k \sum_{l=1}^m 1_{\tilde{\Omega}_{l-1}(M)} |\nabla \pi^l|_{\mathbb{L}^2}^2 \right) \right], \end{aligned}$$

where

$$\tilde{C}_1(M, \lambda) = \frac{(1 + 2\lambda)\bar{C}^4}{2^4 (1 - 2\lambda)^4 \nu^3} M + C(L_1 \text{Tr } Q, \lambda, \nu, \lambda).$$

Using once more the discrete Gronwall lemma, we obtain

$$\mathbb{E} \left(\max_{1 \leq n \leq m} 1_{\tilde{\Omega}_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \lambda) \left[k + h^2 + h^2 \mathbb{E} \left(k \sum_{l=1}^m |\nabla \pi^l|_{\mathbb{L}^2}^2 \right) \right] e^{T \tilde{C}_1(M, \lambda)}.$$

Let $\delta \in (0, 1)$; one may choose $\lambda \in (0, \frac{1}{2})$ small enough to have $\frac{1+2\lambda}{(1-2\lambda)^4} \leq (1 + \delta)$. Since the sets $\{\tilde{\Omega}_l(M)\}_l$ decrease, (5.2) holds with $\tilde{C}_1(M)$ defined by (5.3).

The upper estimate (5.4) is a straightforward consequence of (5.14) for $\lambda \in (0, \frac{1}{2})$ small enough.

Finally, observe that for “large M ”, we may replace $\tilde{C}_1(M)$ by $(1 + \delta)\frac{\bar{C}^4}{2^4\nu^3}M$, provided that the constraint (5.12), which enables us to use the discrete Gronwall lemma, is satisfied for some constant λ defined in terms of δ . This completes the proof. \square

A similar argument proves the following localized convergence result for Algorithm 2.

Proposition 5.2. *Let $(\mathbf{U}^l)_l$ be solution of Algorithm 2, that is $\mathbf{V}_h \subset V$, and G satisfy the growth and Lipschitz conditions **(G2)**. Let $u_0 \in L^8(\Omega; V)$ and $\mathbf{U}^0 \in L^8(\Omega; \mathbb{H}_h)$ be \mathcal{F}_0 -measurable random variables such that (4.8) holds. For every $M > 0$ and $l = 0, \dots, N$, set*

$$\Omega_l(M) := \left\{ \omega \in \Omega : \max_{0 \leq j \leq l} \|\mathbf{u}^j\|_V^2 \leq M \right\}. \quad (5.17)$$

There exists $C_0 > 0$ such that for $k = \frac{T}{N} > 0$ small enough to have $kM \leq C_0$, we have for every $\delta > 0$ and for every $m = 1, \dots, N$

$$\mathbb{E} \left(1_{\Omega_{m-1}(M)} \max_{1 \leq n \leq m} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \delta) (k + h^2) e^{\tilde{C}_3(M)T}, \quad (5.18)$$

where

$$\begin{aligned} \tilde{C}_3(M) &:= (1 + \delta) \left[\frac{\bar{C}^2}{4\nu} + 1 \right] M + C(L_1 \text{Tr } Q, \delta) \\ &\equiv (1 + \delta) \left[\frac{\bar{C}^2}{4\nu} + 1 \right] M \quad \text{for “large” } M, \end{aligned} \quad (5.19)$$

and \bar{C} is the constant appearing in the Gagliardo-Nirenberg inequality (2.2). Furthermore,

$$\mathbb{E} \left(k \sum_{l=1}^m 1_{\Omega_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \delta) (k + h^2) T \tilde{C}_3(M) e^{\tilde{C}_3(M)T}. \quad (5.20)$$

Proof. We briefly sketch the argument. The identities (4.17) with b instead of \tilde{b} and $\text{div } \mathbf{Q}_h^0 \mathbf{E}^l = 0$, the estimates (5.5) and (5.11) imply for $n, m = 1, \dots, N$

$$\begin{aligned} &\mathbb{E} \left(1_{\Omega_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 \right) + 2\nu \mathbb{E} \left(\sum_{l=1}^n k 1_{\Omega_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \leq 2\mathbb{E} [R_1(n, k, h, M) + R_3(n, k, h, M)], \\ &\mathbb{E} \left(\max_{1 \leq n \leq m} 1_{\Omega_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 \right) \leq 2\mathbb{E} \left[R_1(n, k, h, M) + R_3(n, k, h, M) + \max_{1 \leq n \leq m} R_2(n, k, h, M) \right], \end{aligned}$$

where $R_j(n, k, h, M)$ is deduced from $\tilde{R}_j(n, k, h, M)$ replacing $\tilde{\Omega}_{l-1}(M)$ by $\Omega_{l-1}(M)$ for $j = 1, 2, 3$ (and where in $R_2(n, k, h, M)$ the term containing the pressure is omitted).

The upper estimates (4.8), (4.19), (4.20), (4.37) and (5.9) with $\Omega_{l-1}(M)$ instead of $\tilde{\Omega}_{l-1}(M)$ imply for $\epsilon_1 \in (0, \frac{1}{2})$ and $\epsilon > 0$

$$\begin{aligned} &\mathbb{E} \left(1_{\Omega_{n-1}(M)} |\mathbf{Q}_h^0 \mathbf{E}^n|_{\mathbb{L}^2}^2 \right) + 2\nu \mathbb{E} \left(k \sum_{l=1}^n 1_{\Omega_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \epsilon, \epsilon_1) (k + h^2) \\ &\quad + C(\nu, \epsilon_1, M) \mathbb{E} \left(k 1_{\Omega_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) + 2(10\epsilon + 2\epsilon_1)\nu \mathbb{E} \left(k \sum_{l=1}^n 1_{\Omega_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \end{aligned}$$

$$+ \mathbb{E} \left(k \sum_{l=1}^{n-1} 1_{\Omega_{l-1}(M)} \left[C(L_1 \operatorname{Tr} Q) + \left(1 + \frac{\tilde{C}^2}{2^3 \epsilon_1 \nu} \right) M \right] \mathbf{E}^l \Big|_{\mathbb{L}^2}^2 \right).$$

Let $\lambda \in (0, 1)$; set $\epsilon_1 = \frac{1}{2}(1 - \lambda)$ and $\epsilon = \frac{\lambda}{20}$; then $2\epsilon_1 + 10\epsilon = 1 - \frac{\lambda}{2}$. Using (4.23), we deduce for $\lambda \in (0, 1)$ and N large enough (that is k small enough)

$$\begin{aligned} \mathbb{E} \left(1_{\Omega_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) + \lambda \mathbb{E} \left(k \sum_{l=1}^n 1_{\Omega_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) &\leq C(\nu, \lambda)(k + h^2) \\ &+ \mathbb{E} \left(k \sum_{l=1}^{n-1} 1_{\tilde{\Omega}_{l-1}(M)} \tilde{C}_4(M, \nu, \lambda) |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right), \end{aligned} \quad (5.21)$$

where

$$\tilde{C}_4(M, \nu, \lambda) := (1 + \lambda) \left(\frac{\tilde{C}^2}{2^2(1 - \lambda)^2 \nu} + 1 \right) M + C(L_1 \operatorname{Tr} Q, \lambda).$$

Using the discrete version of Gronwall's lemma, we obtain an upper estimate of $\mathbb{E}(1_{\Omega_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2)$. Plugging this result in the above upper estimate (5.21), we deduce

$$\mathbb{E} \left(k \sum_{l=1}^N 1_{\Omega_{l-1}(M)} |\nabla \mathbf{E}^l|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \lambda)(k + h^2) T \tilde{C}_4(M, \nu, \lambda) \exp \left(T \tilde{C}_4(M, \nu, \lambda) \right). \quad (5.22)$$

On easily sees that the upper estimate (5.16) holds for $R_2(n, k, h, M)$ and $\Omega_{l-1}(M)$ instead of $\tilde{R}_2(n, k, h, M)$ and $\tilde{\Omega}_{l-1}(M)$ respectively. Hence, we deduce that for $\lambda \in (0, \frac{1}{2})$ and k small enough, we have

$$\mathbb{E} \left(\max_{1 \leq n \leq m} 1_{\Omega_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \lambda)(k + h^2) + \mathbb{E} \left(\tilde{C}_3(M, \lambda) k \sum_{l=1}^{m-1} 1_{\Omega_{l-1}(M)} |\mathbf{E}^l|_{\mathbb{L}^2}^2 \right),$$

where

$$\tilde{C}_3(M, \lambda) := (1 + \lambda) \left[\frac{\tilde{C}^2}{2^2(1 - 2\lambda)(1 - \lambda)\nu} + 1 \right] M + C(L_1 \operatorname{Tr} Q, \lambda).$$

The discrete Gronwall lemma implies for $m = 1, \dots, N$

$$\mathbb{E} \left(\max_{1 \leq n \leq m} 1_{\Omega_{n-1}(M)} |\mathbf{E}^n|_{\mathbb{L}^2}^2 \right) \leq C(\nu, \lambda)(k + h^2) \exp \left(T \tilde{C}_3(M, \lambda) \right).$$

Fix $\delta > 0$; one can choose $\lambda \in (0, \frac{1}{2})$ small enough to ensure $\frac{1+\lambda}{(1-\lambda)(1-2\lambda)} \leq 1 + \delta$. Since the sets $\{\Omega_l(M)\}_l$ are decreasing, this concludes the proof of (5.18). The proof of (5.20), which is similar to that in the proof of (5.4), is omitted. \square

6. A GENERAL FRAMEWORK FOR STRONG CONVERGENCE

This section is devoted to a very general setting to prove $L^2(\Omega)$ -convergence for a real-valued family $\{X(\eta)\}_\eta$ of random variables with a specific rate of convergence. These results will be applied in the next section to obtain rates of convergence of $|\mathbf{E}_n|_{\mathbb{L}^2}$.

We first prove that the speed of convergence of $\{X(\eta)\}_\eta$ can be deduced from localized $L^2(\Omega)$ estimates of $X(\eta)$ in terms of another family $\{Y(\eta)\}_\eta$, and from a control of moments of $X(\eta)$ and $Y(\eta)$ uniformly in η .

Theorem 6.1. *Let $\{X(\eta)\}_\eta$ and $\{Y(\eta)\}_\eta$ be families of non-negative random variables indexed by some parameter $\eta \in (0, \eta_0]$, and let $\varphi : (0, \eta_0] \rightarrow \mathbb{R}$ be a function such that $\varphi(\eta) \rightarrow 0$ as $\eta \rightarrow 0$. Suppose that for some exponents $a \geq 1$ and $p \in (1, \infty)$, we have*

- Given positive constants M_0, C_0 there exist positive constant C_1 and $\eta_1 \in (0, \eta_0]$ such that if $M(\eta) \geq M_0$ and $\varphi(\eta)M(\eta) \leq C_0$ for $\eta \in (0, \eta_1]$, we have

$$\mathbb{E}\left(X(\eta)1_{\{Y(\eta)^a \leq M(\eta)\}}\right) \leq \varphi(\eta) \exp [C_1 M(\eta)], \quad \eta \in (0, \eta_1]; \quad (6.1)$$

-

$$\sup_{\eta \in (0, \eta_0]} \mathbb{E}(X(\eta)^p) = C(p) < \infty. \quad (6.2)$$

(i) Suppose that

$$\sup_{\eta \in (0, \eta_0]} \mathbb{E}(Y(\eta)^q) = \tilde{C}(q) < \infty \quad (6.3)$$

for some exponent $q > 0$. Then for η small enough

$$\mathbb{E}(X(\eta)) \leq C |\ln(\varphi(\eta))|^{-\frac{q(p-1)}{ap}}. \quad (6.4)$$

If condition (6.2) holds for any $p \in (1, \infty)$, the upper bound of (6.4) can be replaced by $C |\ln(\varphi(\eta))|^{-\gamma_1}$ for any $\gamma_1 < \frac{q}{a}$.

(ii) Suppose that $a > 1$ and that for some positive constant $\tilde{\alpha}_0$

$$\mathbb{E}[\exp(\alpha Y(\eta))] = \bar{C}(\alpha) < \infty, \quad \text{for } \alpha \in (0, \tilde{\alpha}_0) \quad (6.5)$$

Then for η small enough

$$\mathbb{E}(X(\eta)) \leq C \exp\left(-\gamma_2 |\ln(\varphi(\eta))|^{\frac{1}{a}}\right) \quad \text{for } \gamma_2 < \frac{p-1}{p} \tilde{\alpha}_0 C_1^{-\frac{1}{a}}. \quad (6.6)$$

Furthermore, if (6.2) holds for every $p \in (1, \infty)$, then the upper estimate (6.6) holds for $\gamma_2 < \tilde{\alpha}_0 C_1^{-\frac{1}{a}}$.

(iii) Suppose that $a = 1$ and that (6.5) holds. Then for η small enough we have

$$\mathbb{E}(X(\eta)) \leq \varphi(\eta)^\gamma \quad \text{for } \gamma_3 < \frac{\tilde{\alpha}_0(p-1)}{\tilde{\alpha}_0(p-1) + C_1 p}. \quad (6.7)$$

Furthermore, if (6.2) holds for every $p \in (1, \infty)$, then the upper estimate (6.7) holds for $\gamma_3 < \frac{\tilde{\alpha}_0}{\tilde{\alpha}_0 + C_1}$.

Proof. For $\eta > 0$ and $M(\eta) > 0$, let $\Omega(\eta) := \{Y(\eta)^a \leq M(\eta)\}$.

(i) Using (6.3) we deduce

$$\mathbb{P}(\Omega(\eta)^c) = \mathbb{P}(Y(\eta)^a > M(\eta)^{\frac{q}{a}}) \leq \tilde{C}(q) M(\eta)^{-\frac{q}{a}}.$$

Furthermore, (6.2) and Hölder's inequality with conjugate exponents p and $\frac{p}{p-1}$ imply

$$\mathbb{E}(1_{\Omega(\eta)^c} X(\eta)) \leq \{\mathbb{E}(X(\eta)^p)\}^{\frac{1}{p}} \{\mathbb{P}(\Omega(\eta)^c)\}^{\frac{p-1}{p}} \leq C(p, q) M(\eta)^{-\frac{q(p-1)}{ap}}. \quad (6.8)$$

Let η_1 be small enough to have $\eta_1 M_0 < C_0$, $M(\eta) \geq M_0$ and $\varphi(\eta)M(\eta) \leq C_0$ for $\eta \in (0, \eta_1]$. Using (6.1) we deduce

$$\mathbb{E}(1_{\Omega(\eta)} X(\eta)) \leq \varphi(\eta) \exp [C_1 M(\eta)]. \quad (6.9)$$

Choose $M(\eta)$ such that, up to some bounded multiplicative term, the right hand sides of (6.8) and (6.9) agree. Taking logarithms, this comes down to

$$C_1 M(\eta) - |\ln(\varphi(\eta))| = -c \ln(M(\eta)) + \bar{c}(\eta)$$

where $c = \frac{q(p-1)}{ap}$, and $\bar{c}_1 \leq \bar{c}(\eta) \leq \bar{c}_2$ holds for some constants \bar{c}_1, \bar{c}_2 and any $\eta \in (0, \eta_1]$. Set

$$M(\eta) := \frac{1}{C_1} [|\ln(\varphi(\eta))| - c \ln(|\ln(\varphi(\eta))|)] \sim C |\ln(\varphi(\eta))|$$

for some positive constant C as $\eta \rightarrow 0$. Then $\varphi(\eta)M(\eta) \rightarrow 0$ as $\eta \rightarrow 0$; furthermore, for some constant \bar{C} we have

$$C_1 M(\eta) - |\ln(\varphi(\eta))| + c \ln(M(\eta)) = c \ln\left(\frac{M(\eta)}{|\ln(\varphi(\eta))|}\right) \rightarrow \bar{C} \quad \text{as } \eta \rightarrow 0.$$

Hence, for η small enough, both upper estimates (6.8) and (6.9) agree - up to some (bounded) multiplicative term - with $|\ln(\varphi(\eta))|^{-\frac{q(p-1)}{ap}}$; this concludes the proof of (6.4). If (6.2) holds for any $p \in (1, \infty)$, the bound on γ_1 is a straightforward consequence of (6.4).

(ii) Using (6.5) and the Markov inequality, we deduce for $\alpha \in (0, \tilde{\alpha}_0)$

$$\mathbb{P}(\Omega(\eta)^c) = \mathbb{P}[\exp(\alpha Y(\eta)) > \exp(\alpha M(\eta)^{\frac{1}{a}})] \leq \bar{C}(\alpha) \exp(-\alpha M(\eta)^{\frac{1}{a}}).$$

Using again (6.2) and Hölder's inequality, we deduce

$$\mathbb{E}(1_{\Omega(\eta)^c} X(\eta)) \leq C(p, \alpha) \exp\left(-\alpha \frac{p-1}{p} M(\eta)^{\frac{1}{a}}\right). \quad (6.10)$$

We next choose $M(\eta) = \frac{1}{C} |\ln(\varphi(\eta))|$ for some constant $C > C_1$. Then $\varphi(\eta)M(\eta) \rightarrow 0$ as $\eta \rightarrow 0$; thus for $\eta > 0$ small enough, we have $M(\eta) \geq M_0$ and $\varphi(\eta)M(\eta) \leq C_0$. For this choice of η the inequality (6.1) implies

$$\mathbb{E}(1_{\Omega(\eta)} X(\eta)) \leq \exp\left[-\left(1 - \frac{C_1}{C}\right) |\ln(\varphi(\eta))|\right].$$

Since $a > 1$ and $M(\eta) \rightarrow \infty$ as $\eta \rightarrow 0$, the right hand side of (6.10) converges to 0 slower than the above one. Splitting $\mathbb{E}(X(\eta))$ on $\Omega(\eta)$ and $\Omega(\eta)^c$, we deduce that the largest term in this sum is the expected value on $\Omega(\eta)^c$. Hence, using (6.10) for this choice of $M(\eta)$, we deduce that for η small enough,

$$\mathbb{E}(X(\eta)) \leq C \exp\left(-\alpha \frac{p-1}{p} C^{-\frac{1}{a}} |\ln(\varphi(\eta))|^{\frac{1}{a}}\right) \quad \text{for } C > C_1.$$

Fix $\gamma_2 < \tilde{\alpha}_0 \frac{p-1}{p} C_1^{-\frac{1}{a}}$; then choose $\alpha < \tilde{\alpha}_0$ close to $\tilde{\alpha}_0$ and $C > C_1$ very close to C_1 to have

$$\gamma_2 < \alpha \frac{p-1}{p} C_1^{-\frac{1}{a}} < \tilde{\alpha}_0 \frac{p-1}{p} C^{-\frac{1}{a}}.$$

This inequality clearly yields (6.6).

Note that if (6.2) holds for all $p \in (0, 1)$, given $\gamma_2 < \tilde{\alpha}_0 C_1^{-\frac{1}{a}}$ we can choose p large enough to have $\gamma_2 < \alpha \frac{p-1}{p} C_1^{-\frac{1}{a}}$.

(iii) Using the upper estimates (6.1) with $a = 1$, and $M := M(\eta)$ and (6.10), we choose $M(\eta)$ such that the right hand sides of both inequalities agree up to some multiplicative constant. Taking logarithms, we deduce that $M(\eta)$ should be such that

$$C_1 M(\eta) - |\ln(\varphi(\eta))| = -\alpha \frac{p-1}{p} M(\eta) \quad \text{for some } \alpha \in (0, \tilde{\alpha}_0).$$

Thus, if we set $\beta := \alpha \frac{p-1}{p}$ and $M(\eta) := \frac{|\ln(\varphi(\eta))|}{C_1 + \beta}$, we have $M(\eta) \rightarrow \infty$, $\varphi(\eta)M(\eta) \rightarrow 0$ as $\eta \rightarrow 0$, and for η small enough, $\varphi(\eta) < 1$; thus for η small enough,

$$\mathbb{E}(X(\eta)) = C \exp\left(-\frac{\beta |\ln(\varphi(\eta))|}{\beta + C_1}\right) = C \varphi(\eta)^{\frac{\beta}{C_1 + \beta}} \quad \text{for } \alpha \in (0, \tilde{\alpha}_0).$$

Set $\beta_0 := \tilde{\alpha}_0 \frac{p-1}{p}$ and let $\gamma_3 < \frac{\beta_0}{\beta_0 + C_1}$; choose $\alpha < \tilde{\alpha}_0$ close enough to $\tilde{\alpha}_0$ to have $\gamma_3 < \frac{\beta}{\beta + C_1} < \frac{\beta_0}{\beta_0 + C_1}$. We deduce (6.7).

If (6.2) is satisfied for any $p \in (0, 1)$, given $\gamma_3 < \frac{\tilde{\alpha}_0}{\tilde{\alpha}_0 + C_1}$ we can choose p large enough to have $\gamma_3 < \frac{\beta_0}{\beta_0 + C_1} < \frac{\tilde{\alpha}_0}{\tilde{\alpha}_0 + C_1}$; therefore, we conclude that (6.7) holds with the exponent γ_3 in the right hand side. \square

We next deduce similar results if the assumption (6.1) is slightly weakened as follows.

Corollary 6.2. *Let $\{Z(\eta)\}_\eta$ and $\{Y(\eta)\}_\eta$ be families of non-negative random variables indexed by some parameter $\eta \in (0, \eta_0]$, and let $\varphi : (0, \eta_0] \rightarrow \mathbb{R}$ be a function such that $\varphi(\eta) \rightarrow 0$ as $\eta \rightarrow 0$. Suppose that for some exponents $a \geq 1$ and $p \in (1, \infty)$*

- *Given positive constants M_0, C_0 there exists positive constants C_1 and $\eta_1 \in (0, \eta_0]$ such that if $M(\eta) \geq M_0$ satisfies $\varphi(\eta)M(\eta) \leq C_0$ for any $\eta \in (0, \eta_1]$, we have*

$$\mathbb{E}\left(Z(\eta)1_{\{Y(\eta)^a \leq M(\eta)\}}\right) \leq \varphi(\eta) C_1 M(\eta) \exp[C_1 M(\eta)], \quad \eta \in (0, \eta_1]; \quad (6.11)$$

-

$$\sup_{\eta \in (0, \eta_0]} \mathbb{E}(Z(\eta)^p) = C(p) < \infty. \quad (6.12)$$

(i) *Let $\{Y(\eta)\}_\eta$ satisfy the assumption (6.3). Then for η small enough, (6.4) holds for $Z(\eta)$ instead of $X(\eta)$.*

(ii) *Let $\{Y(\eta)\}_\eta$ satisfy the assumption (6.5) and $a > 1$. Then for η small enough, (6.5) holds with $Z(\eta)$ instead of $X(\eta)$.*

(iii) *$\{Y(\eta)\}_\eta$ satisfy the assumption (6.5) and $a = 1$. Then for η small enough, (6.6) holds with $Z(\eta)$ instead of $X(\eta)$.*

Proof. The proof is a straightforward consequence of the arguments used in the proof of Theorem 6.1, and, given any $\delta > 0$ arbitrary small, of the inequality $x \exp(x) \leq \exp[(1 + \delta)x]$ valid for large enough $x > 0$. \square

7. CONVERGENCE IN $L^2(\Omega)$

In this section, we prove a speed of convergence in $L^2(\Omega)$ -norm of the difference \mathbf{E}^l between the time and space-time discretizations. Coupled with the results of Theorems 3.3 and 3.4, this will provide a strong (that is $L^2(\Omega)$) speed of convergence of the space-time Euler scheme \mathbf{U}^l to $u(t_l)$ in $\mathbb{L}^2(D)$ uniformly on the time grid $t_l = l \frac{T}{N}$.

7.1. Strong convergence of Algorithm 1. In this subsection, we focus on the solution defined in Algorithm 1, that is (4.11) and (4.12). We will give results in the case of a multiplicative and of an additive stochastic perturbation.

7.1.1. *Multiplicative noise.* We suppose that the diffusion coefficient G satisfies the growth and Lipschitz conditions **(G1)** or **(G2)**. The following theorem is the main result of this section for general finite elements and diffusion coefficients. As in [9], for general coefficients we have either to keep the gradient of the pressure (with some growth condition), or to impose some relation between the space mesh h and the time mesh $k = \frac{T}{N}$.

Theorem 7.1. *Fix an integer $q_0 \geq 3$; suppose that $u_0 \in L^{2q_0}(\Omega; V)$ and $U_0 \in L^{2q_0}(\Omega; L^2_{per})$.*

(i) *Let G satisfy condition **(G1)** and suppose that $h^2 \mathbf{E}\left(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2\right) \rightarrow 0$ as $k, h \rightarrow 0$; then for $k + h^2 + h^2 \mathbf{E}\left(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2\right)$ small enough,*

$$\begin{aligned} & \mathbb{E}\left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2\right) \\ & \leq C \left| \ln \left[k + h^2 + h^2 \mathbf{E}\left(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2\right) \right] \right|^{-(2q_0-2-\frac{1}{2})}. \end{aligned} \quad (7.1)$$

(ii) *Let G satisfy condition **(G1)** and suppose that as $h, k \rightarrow 0$ we have $h^2 k^{-1} \rightarrow 0$; then for k and $h^2 k^{-1}$ small enough,*

$$\mathbb{E}\left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2\right) \leq C \left| \ln [k + h^2 k^{-1}] \right|^{-(2q_0-2-\frac{1}{2})}. \quad (7.2)$$

(iii) *Suppose that G satisfies condition **(G2)**. Then for k and h small enough (without any restriction), we have*

$$\mathbb{E}\left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2\right) \leq C \left| \ln (k + h^2) \right|^{-(2q_0-2-\frac{1}{2})}. \quad (7.3)$$

Remark 7.2. *If u_0 and \mathbf{U}^0 are deterministic or have moments of all orders (such as Gaussian random variables), in all cases the exponent of the logarithm is arbitrary large.*

Proof. (i) Let $\eta = k + h^2$ and $\varphi(\eta) = k + h^2 + h^2 \mathbf{E}(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2)$; by assumption, $\varphi(\eta) \rightarrow 0$ as $\eta \rightarrow 0$. Set

$$X(\eta) = \max_{0 \leq l \leq N} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2, \quad Y(\eta) = \max_{0 \leq l \leq N-1} \|\mathbf{u}^l\|_V^2 \quad \text{and} \quad Z(\eta) = k \sum_{l=1}^N |\nabla \mathbf{u}^l - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2. \quad (7.4)$$

Using (5.2) – (5.4) in Proposition 5.1, we deduce that (6.1) and (6.11) hold for $a = 2$ and $C_1 := (1 + \delta) \frac{\tilde{C}^4 T}{2^4 \nu^3}$ for some $\delta > 0$. Furthermore, the upper estimates (3.4), (4.13) and (4.14) imply that (6.2) and (6.12) are true for $p = 2^{q_0-1}$, while (3.4) implies that (6.3) holds for $q = 2^{q_0-1}$. Since $\frac{q(p-1)}{ap} = 2^{q_0-2} - \frac{1}{2}$, using part (i) in Theorem 6.1 and Corollary 6.2, we deduce

$$\mathbb{E}\left(\max_{0 \leq l \leq N} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2\right) + k \sum_{l=1}^N |\nabla \mathbf{u}^l - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \leq C \left| \ln \left[k + h^2 + h^2 \mathbf{E}\left(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2\right) \right] \right|^{-(2q_0-2-\frac{1}{2})}.$$

We next use the upper estimate of $\mathbb{E}(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2)$ proved in (3.9) and compare the right hand sides of the above inequality and of (3.9).

Let $\eta > 0$ be small enough to ensure $\varphi(\eta) < e^{-1}$. Since $k \leq \varphi(\eta)$ and $\ln(\varphi(\eta)) < -1$, $|\ln(k)|^2 > |\ln(k)| > |\ln(\varphi(\eta))| > 1$. Furthermore, $\ln(N) = |\ln(k)| - \ln(T) \geq |\ln(k)|$ if

$T \in (0, 1)$, while $\ln(N) \geq \frac{1}{2} |\ln(k)|$ if $T > 1$ for small enough k . Therefore, $|\ln(N)|^{-2A} \leq C |\ln(\varphi(\eta))|^{-A}$ for $A := 2^{q_0-2} - \frac{1}{2}$. This concludes the proof of (7.1).

(ii) Using the upper estimate (3.7) we deduce

$$k + h^2 + h^2 \mathbf{E} \left(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2 \right) \leq k + h^2 + C(T) \frac{h^2}{k} \leq C(T) \left(k + \frac{h^2}{k} \right),$$

where $k + h^2/k \equiv 0$. Hence, (7.1) implies (7.2).

(iii) Recall that if G satisfies condition **(G2)**, (3.8) implies

$$k + h^2 + h^2 \mathbf{E} \left(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2 \right) \leq k + h^2 (1 + C(T)) \leq C(T) (k + h^2).$$

Thus (7.1) implies (7.3). \square

7.1.2. Additive noise. As in [3], we replace the Hölder inequality for a power of $\max_l \|\mathbf{u}^l\|_V^2$ by an exponential Markov inequality for $\max_l \|\mathbf{u}^l\|_V^2$; recall that the maximal exponent of the exponential moments proved in Theorem 8.1 is $\tilde{\alpha}_0$. This yields the following

Theorem 7.3. *Let G satisfy condition **(G1)** with $K_1 = 0$, i.e., $\|G(u)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \leq K_0$. Let $u_0 \in V$ and $U_0 \in L^{2q_0}(\Omega; L_{per}^2)$ for some $q_0 \in [3, \infty)$. Set $\tilde{\alpha}_0 := \frac{\nu}{2\tilde{C}K_0 \text{Tr}Q}$, where \tilde{C} is defined by $|\nabla u|_{\mathbb{L}^2}^2 \leq \tilde{C} |Au|_{\mathbb{L}^2}^2$. Let $\kappa_0 := \frac{2^{q_0-1}-1}{2^{q_0-1}} \tilde{\alpha}_0 \frac{4}{\tilde{C}^2} \sqrt{\frac{\nu^3}{T}}$.*

(i) *Suppose that $k + h^2 + h^2 \mathbf{E}(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2)$ is “small”. Then for $\gamma < \kappa_0$*

$$\begin{aligned} \mathbf{E} \left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \\ \leq C \exp \left(-\gamma \left| \ln \left[k + h^2 + h^2 \mathbf{E} \left(k \sum_{j=1}^N |\nabla \pi^j|_{L^2}^2 \right) \right] \right|^{\frac{1}{2}} \right). \end{aligned} \quad (7.5)$$

(ii) *Suppose that G satisfies condition **(G1)** and that $h^2 k^{-1} \rightarrow 0$ as $h, k \rightarrow 0$; then for “small” $k + h^2 k^{-1}$ we have for $\gamma < \kappa_0$*

$$\mathbf{E} \left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \leq C \exp \left(-\gamma \left| \ln [k + h^2 k^{-1}] \right|^{\frac{1}{2}} \right). \quad (7.6)$$

(iii) *Suppose that G satisfies condition **(G2)**. Then, when $k \rightarrow 0$ and $h \rightarrow 0$ (without any restriction), we have for $\gamma < \kappa_0$*

$$\mathbf{E} \left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \leq C \exp \left(-\gamma \left| \ln [k + h^2] \right|^{\frac{1}{2}} \right). \quad (7.7)$$

Remark 7.4. (i) *The speed of convergence is an increasing function of the viscosity ν , and a decreasing function of the “strenght” of the noise $K_0 \text{Tr}Q$, and of the length of the time interval.*

(ii) *If \mathbf{U}^0 is deterministic such that $|u_0 - \mathbf{u}^0|_{\mathbb{L}^2} \leq Ch^2$ for “small” h , the upper estimates (7.5)–(7.7) hold for $\gamma < \tilde{\alpha}_0 \frac{4}{\tilde{C}^2} \sqrt{\frac{\nu^3}{T}}$.*

Proof. (i) Let $\eta = k + h^2$, $\varphi(\eta) = k + h^2 + h^2 \mathbb{E}(k \sum_{l=1}^N |\nabla \pi^l|_{L^2}^2)$, and suppose that $\varphi(\eta) \rightarrow 0$ as $\eta \rightarrow 0$. Let $X(\eta)$, $Y(\eta)$ and $Z(\eta)$ be defined by (7.4). Then, as in the proof of Theorem 7.1, (6.1) and (6.11) are satisfied with $C_1 := (1 + \delta) \frac{\tilde{C}^4 T}{2^4 \nu^3}$ for $\delta > 0$ arbitrary small and $a = 2$, while (6.2) and (6.12) hold with $p = 2^{q_0-1}$. Furthermore, Theorem 8.1 implies that (6.5) holds for $\tilde{\alpha}_0$ defined above.

Therefore, using part (ii) of Theorem 6.1 and Corollary 6.2, we deduce that

$$\begin{aligned} & \mathbb{E} \left(\max_{0 \leq l \leq N} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla \mathbf{u}^l - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \\ & \leq C \exp \left(-\gamma \left| \ln \left[k + h^2 + h^2 \mathbb{E} \left(k \sum_{j=1}^N |\nabla \pi^j|_{L^2}^2 \right) \right] \right|^{\frac{1}{2}} \right) \end{aligned} \quad (7.8)$$

holds for $\gamma < \frac{2^{q_0-1}-1}{2^{q_0-1}} \tilde{\alpha}_0 C_1^{-\frac{1}{2}}$. Furthermore, (3.10) implies $\mathbb{E}(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{u}^l|_{\mathbb{L}^2}^2) \leq Ck^{-\beta}$. For k small enough we have $\beta |\ln(k)| \geq \gamma |\ln(k)|^{\frac{1}{2}} \geq \gamma |\ln(\varphi(\eta))|^{\frac{1}{2}}$ for any $\beta, \gamma > 0$. Therefore, (3.10) and (7.8) imply (7.5).

Note that unlike the convergence for the time discretization \mathbf{u}^l to the true solution $u(t_l)$ described in Theorem 3.4, the final result does not provide a polynomial speed. This is due to the fact that in Proposition 5.1, the localization involves the fourth power of $\|\mathbf{u}\|_V$.

(ii) As in the proof of Theorem 7.1, the inequality (3.7) implies $h^2 \mathbb{E}(\sum_{j=1}^N |\nabla \pi^j|_{L^2}^2) \leq C(T)h^2/k$. Therefore, given k and h^2k^{-1} small enough, (7.5) implies (7.6).

(ii) If G satisfies **(G2)**, the inequality (3.8) implies $k + h^2 + h^2 \mathbb{E}(\sum_{j=1}^N |\nabla \pi^j|_{L^2}^2) \leq C(T)(k + h^2)$. Thus, (7.5) implies (7.7). \square

Corollary 7.5. *If u_0 is random, independent of W and such that $\mathbb{E}[\exp(\gamma_0 \|u_0\|_V^2)] < \infty$, then the statement of Theorem 7.6 is valid with $\tilde{\alpha}_0$ replaced by $\tilde{\beta}_0 := \tilde{\alpha}_0 \frac{\gamma_0}{\gamma_0 + \tilde{\alpha}_0}$. Indeed, it suffices to use (8.12) in the proof of Theorem 7.3.*

7.2. Strong convergence of Algorithm 2. In this subsection, we focus on Algorithm 2, that is divergence-free finite elements, and suppose that $\mathbb{V}_h \subset V$. We will state the convergence results for both multiplicative and additive stochastic perturbations.

7.2.1. Multiplicative noise. In this case, the localized convergence result from section 5 only involves the square of the V norm of the time discretization. The following result is similar to Theorem 7.1.

Theorem 7.6. *Suppose that the finite elements are divergence free, so that $\mathbb{V}_h \subset V$. Fix an integer $q_0 \geq 3$; suppose that $u_0 \in L^{2^{q_0}}(\Omega; V)$ and $U_0 \in L^{2^{q_0}}(\Omega; L_{per}^2)$. Let the coefficient G satisfy the assumptions **(G1)**. Let \tilde{C} be the constant defined in (2.2). Then if k, l are “small”*

$$\mathbb{E} \left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \leq C(\nu, T) |\ln(k + h^2)|^{-(2^{q_0-1}-1)}. \quad (7.9)$$

Remark 7.7. *Once more, if u_0 and \mathbf{U}^0 are deterministic or have moments of all order, for example are Gaussians, the exponent of the logarithm in the right hand side of (7.9) is arbitrary large.*

Proof. Let $\eta = k + h^2$, $\varphi(\eta) = \eta$, $X(\eta)$, $Y(\eta)$ and $Z(\eta)$ be defined by (7.4); then Proposition 5.2 implies that the upper estimate (6.1) is satisfied with $a = 1$ and $C_1 := (1 + \delta) \left[\frac{\bar{C}^2}{4\nu} + 1 \right] T$ for $\delta > 0$ arbitrary small. Using the upper estimates (3.4), (4.13) and (4.14), we deduce that (6.2) and (6.12) are true for $p = 2^{q_0-1}$, while (3.4) implies that (6.3) holds for $q = 2^{q_0-1}$. Therefore, parts (i) of Theorem 6.1 and Corollary 6.2 imply

$$\mathbb{E} \left(\max_{0 \leq l \leq N} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla \mathbf{u}^l - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \leq C(\nu, T) |\ln(k + h^2)|^{-(2^{q_0-1}-1)}.$$

Since $k \leq k + h^2 \ll e^{-1}$, we have $\{\ln N\}^{-(2^{q_0-1}-1)} \leq C |\ln(k + h^2)|^{-(2^{q_0-1}-1)}$; the upper estimate (3.9) completes the proof of (7.9). \square

7.2.2. Additive noise. Assume condition **(G1)** holds with $K_1 = 0$; then the strong speed of convergence is polynomial.

Theorem 7.8. *Let G satisfy condition **(G1)** with $K_1 = 0$, i.e., $\|G(u)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \leq K_0$. Let $u_0 \in V$ be deterministic, and $U_0 \in L^{2^{q_0}}(\Omega; L_{per}^2)$ for some $q_0 \in [3, \infty)$. Set*

$$\tilde{\alpha}_0 := \frac{\nu}{2\tilde{C}K_0\text{Tr}Q}, \quad C_0 := \frac{\bar{C}^2 T}{2\nu} \quad \text{and} \quad C_1 := \left[\frac{\bar{C}^2}{4\nu} + 1 \right] T, \quad (7.10)$$

where \bar{C} is defined by (2.2) and \tilde{C} is defined by $|\nabla u|_{\mathbb{L}^2}^2 \leq \tilde{C}|Au|_{\mathbb{L}^2}^2$.

Define the critical exponents as follows:

$$\beta_0 := \frac{1}{2} \left(\frac{\tilde{\alpha}_0}{\tilde{\alpha}_0 + C_0} \right) \quad \text{and} \quad \kappa_0 := \frac{\tilde{\alpha}_0(2^{q_0-1} - 1)}{\tilde{\alpha}_0(2^{q_0-1} - 1) + C_1 2^{q_0-1}}. \quad (7.11)$$

Then for $\gamma_1 < \kappa_0$ and $\gamma < \beta_0$, setting $k := \frac{T}{N}$, we have for k and h small enough

$$\mathbb{E} \left(\max_{0 \leq l \leq N} |u(t_l) - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla u(t_l) - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \leq C[(k + h^2)^{\gamma_1} + k^\gamma]. \quad (7.12)$$

Remark 7.9. *The exponent β_0 (resp. κ_0), which reflects the speed for the time (resp. the finite elements) approximation, is an increasing function of the viscosity $\nu > 0$ and decreasing functions of $K_0\text{Tr}Q$. For a given viscosity, as $K_0\text{Tr}Q \rightarrow 0$, κ_0 converges to 1. This limit is twice the corresponding one $\frac{1}{2}$ of β_0 ; this is consistent with the scaling between the space and time regularity for the heat kernel, which is behind this model. Note that these upper bounds are also approached for a given noise if the viscosity is “large”. The exponent β_0 cannot be better than the time regularity. The maximal exponents cannot be improved.*

Proof. Let $\eta = k + h^2$ and $\varphi(\eta) = \eta$. As in the proof of Theorem 7.6, (6.1) and (6.11) are satisfied for $a = 1$ and $(1 + \delta)C_1$ where C_1 is defined in (7.10) and $\delta > 0$ is arbitrary small. Furthermore, (6.2) and (6.12) hold for $p = 2^{q_0-1}$. Since the noise is additive, (6.5) is satisfied for the parameter $\tilde{\alpha}_0$ defined in (7.10). Using parts (iii) of Theorem 6.1 and Corollary 6.2, we deduce

$$\mathbb{E} \left(\max_{0 \leq l \leq N} |\mathbf{u}^l - \mathbf{U}^l|_{\mathbb{L}^2}^2 + k \sum_{l=1}^N |\nabla \mathbf{u}^l - \nabla \mathbf{U}^l|_{\mathbb{L}^2}^2 \right) \leq C(k + h^2)^{\gamma_1},$$

for $\gamma_1 < \kappa_0$, where κ_0 is defined in (7.11). Coupling this upper estimate with (3.10) and choosing $\delta > 0$ small enough we deduce (7.12); this completes the proof. \square

Corollary 7.10. *If u_0 is random, independent of W and such that $\mathbb{E}[\exp(\gamma_0 \|u_0\|_V^2)] < \infty$, then the statement of Theorem 7.8 is valid with $\tilde{\alpha}_0$ replaced by $\tilde{\beta}_0 := \tilde{\alpha}_0 \frac{\gamma_0}{\gamma_0 + \tilde{\alpha}_0}$. Indeed, it suffices to use (8.12) in the proof of Theorem 7.8.*

8. APPENDIX

In this section, for an additive noise, we prove the existence of exponential moments of the V norm for the time discretization $u_N(t_l)$ uniformly on the time grid $\{t_l = \frac{lT}{N}, l = 0, \dots, N\}$ and with a bound independent of N . This is similar to a similar result proved in [14] for the solution u to the stochastic 2D Navier-Stokes equation (see also [3]).

8.1. Deterministic initial condition. We first suppose that $u_0 \in V$ is deterministic.

Theorem 8.1. *Let G satisfy condition **(G1)** with $K_1 = 0$, that is $\|G(u)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \leq K_0$ and set $\tilde{\alpha}_0 := \frac{\nu}{2\tilde{C}K_0 \text{Tr}Q}$, where the constant \tilde{C} is defined by $|\nabla u|_{\mathbb{L}^2}^2 \leq \tilde{C}|Au|_{\mathbb{L}^2}^2$.*

Let $u_0 \in V$; then for $0 < \alpha < \tilde{\alpha}_0$, there exists a positive constant $C(\alpha)$ such that for N large enough,

$$\mathbb{E} \left[\exp \left(\alpha \max_{0 \leq l \leq N} \|u^l\|_V^2 \right) \right] = C(\alpha) < \infty. \quad (8.1)$$

Proof. Since Lemma 3.1 implies that $\mathbb{E}(\frac{T}{N} \sum_{l=1}^N |Au^l|_{\mathbb{L}^2}^2) < \infty$, using integration by parts we may write (3.1) with $\phi = -\Delta u^l$; this yields a.s.

$$\begin{aligned} & -(\mathbf{u}^l - \mathbf{u}^{l-1}, \Delta \mathbf{u}^l) + \frac{T}{N} \left[\nu(\Delta \mathbf{u}^l, \Delta \mathbf{u}^l) - \langle (\mathbf{u}^l \cdot \nabla \mathbf{u}^l), \Delta \mathbf{u}^l \rangle + (\pi^l, \text{div} \Delta \mathbf{u}^l) \right] \\ & = -(G(\mathbf{u}^{l-1})\Delta_l W, \Delta \mathbf{u}^l). \end{aligned}$$

Using (2.6), $\text{div} \mathbf{u}^l = 0$, and integration by parts in the stochastic term, we deduce

$$(\nabla \mathbf{u}^l - \nabla \mathbf{u}^{l-1}, \nabla \mathbf{u}^l) + \nu \frac{T}{N} |Au^l|_{\mathbb{L}^2}^2 = (\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla \mathbf{u}^l).$$

The identity $(a - b, a) = \frac{1}{2}(|a|^2 - |b|^2 + |a - b|^2)$ implies

$$\begin{aligned} & |\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 - |\nabla \mathbf{u}^{l-1}|_{\mathbb{L}^2}^2 + |\nabla(\mathbf{u}^l - \mathbf{u}^{l-1})|_{\mathbb{L}^2}^2 \\ & = 2(\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla \mathbf{u}^l) - 2\nu \frac{T}{N} |Au^l|_{\mathbb{L}^2}^2 \end{aligned} \quad (8.2)$$

$$= 2(\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla \mathbf{u}^{l-1}) - 2\frac{T}{N}\nu |Au^l|_{\mathbb{L}^2}^2 + 2(\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla[\mathbf{u}^l - \mathbf{u}^{l-1}]). \quad (8.3)$$

A similar computation, based on (3.1) with $\phi = \mathbf{u}^l$ implies a.s.

$$|\mathbf{u}^l|_{\mathbb{L}^2}^2 - |\mathbf{u}^{l-1}|_{\mathbb{L}^2}^2 + |(\mathbf{u}^l - \mathbf{u}^{l-1})|_{\mathbb{L}^2}^2 = 2(G(\mathbf{u}^{l-1})\Delta_l W, \mathbf{u}^l) - 2\nu \frac{T}{N} |\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 \quad (8.4)$$

$$= 2(G(\mathbf{u}^{l-1})\Delta_l W, \mathbf{u}^{l-1}) - 2\frac{T}{N}\nu |\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + 2(G(\mathbf{u}^{l-1})\Delta_l W, \mathbf{u}^l - \mathbf{u}^{l-1}). \quad (8.5)$$

For $l = 1$, condition **(G1)** with $K_1 = 0$, the Cauchy-Schwarz and Young inequalities imply for $\lambda \in (0, 1)$

$$2|(\nabla[G(\mathbf{u}^0)\Delta_l W], \nabla \mathbf{u}^1)| \leq \frac{1}{\lambda} \|G(\mathbf{u}^0)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \|\Delta_1 W\|_K^2 + \lambda |\nabla \mathbf{u}^1|_{\mathbb{L}^2}^2.$$

Hence (8.2) implies

$$|\nabla \mathbf{u}^1|_{\mathbb{L}^2}^2 \leq \frac{1}{1-\lambda} |\nabla u_0|_{\mathbb{L}^2}^2 + \frac{1}{\lambda(1-\lambda)} K_0 \|\Delta_1 W\|_K^2 - 2 \frac{T}{N} \nu |\mathbf{A} \mathbf{u}^1|_{\mathbb{L}^2}^2.$$

For $l \geq 2$, a similar argument using the Cauchy-Schwarz and Young inequalities implies

$$\begin{aligned} 2(\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla[\mathbf{u}^l - \mathbf{u}^{l-1}]) &\leq 2|\nabla(G(\mathbf{u}^{l-1})\Delta_l W)|_{\mathbb{L}^2} |\nabla(\mathbf{u}^l - \mathbf{u}^{l-1})|_{\mathbb{L}^2} \\ &\leq \|G(\mathbf{u}^{l-1})\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \|\Delta_l W\|_K^2 + |\nabla(\mathbf{u}^l - \mathbf{u}^{l-1})|_{\mathbb{L}^2}^2. \end{aligned}$$

Therefore, (8.3) and condition (2.8) in **(G1)** with $K_1 = 0$ imply

$$|\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 - |\nabla \mathbf{u}^{l-1}|_{\mathbb{L}^2}^2 \leq 2(\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla \mathbf{u}^{l-1}) - 2 \frac{T}{N} \nu |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2 + K_0 \|\Delta_l W\|_K^2.$$

A similar argument, based on (2.7) in Condition **(G1)** with $K_1 = 0$, implies

$$\begin{aligned} |\mathbf{u}^l|_{\mathbb{L}^2}^2 &\leq \frac{1}{1-\lambda} |u_0|_{\mathbb{L}^2}^2 + \frac{1}{\lambda(1-\lambda)} K_0 \|\Delta_1 W\|_K^2 - 2 \frac{T}{N} \nu |\nabla \mathbf{u}^1|_{\mathbb{L}^2}^2, \\ |\mathbf{u}^l|_{\mathbb{L}^2}^2 - |\mathbf{u}^{l-1}|_{\mathbb{L}^2}^2 &\leq 2(G(\mathbf{u}^{l-1})\Delta_l W, \mathbf{u}^{l-1}) - 2 \frac{T}{N} \nu |\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + K_0 \|\Delta_l W\|_K^2, \quad l = 2, \dots, N. \end{aligned}$$

Adding these inequalities for $l = 1, \dots, n$, we deduce for $n = 2, \dots, N$

$$\begin{aligned} \|\mathbf{u}^n\|_V^2 &\leq \frac{1}{1-\lambda} \|u_0\|_V^2 + \frac{2}{\lambda(1-\lambda)} K_0 \sum_{l=1}^n \|\Delta_l W\|_K^2 - 2\nu \frac{T}{N} \sum_{l=1}^n (|\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2) \\ &\quad + \sum_{l=2}^n 2 \left[(G(\mathbf{u}^{l-1})\Delta_l W, \mathbf{u}^{l-1}) + (\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla \mathbf{u}^{l-1}) \right]. \end{aligned} \quad (8.6)$$

Let Y be a K -valued centered Gaussian random variable with covariance operator Q . Using the independence of the time increments $\Delta_l W$ we deduce that for any $\beta > 0$,

$$\mathbb{E} \left[\exp \left(\beta \sum_{l=1}^N \|\Delta_l W\|_K^2 \right) \right] = \left\{ \mathbb{E} \left[e^{\beta \frac{T}{N} \|Y\|_K^2} \right] \right\}^N.$$

Proposition 2.16 in [11] implies that if $\tilde{\gamma} \in (0, \frac{1}{2\text{Tr}Q})$ and $\gamma \in (0, \tilde{\gamma})$, we have

$$\mathbb{E}(e^{\gamma \|Y\|_K^2}) \leq \exp \left(\frac{1}{2} \sum_{i=1}^{\infty} \frac{(2\gamma)^i}{i} \text{Tr}(Q^i) \right) \leq \exp \left(\frac{1}{2} \ln(1 + 2\tilde{\gamma} \text{Tr}Q) \right) < \infty.$$

Hence, if $\frac{2\beta K_0 T}{\lambda(1-\lambda)N} < \frac{1}{2\text{Tr}Q}$ (which is satisfied for any $\beta > 0$ provided that N is large enough), we obtain

$$\left\{ \mathbb{E} \left(\exp \left(\frac{2\beta K_0 T}{\lambda(1-\lambda)N} \|Y\|_K^2 \right) \right) \right\}^N \leq \sqrt{2}. \quad (8.7)$$

Given $\alpha > 0$ and $n = 2, \dots, N$, set

$$M_n = 2\alpha \sum_{l=2}^n [(G(\mathbf{u}^{l-1})\Delta_l W, \mathbf{u}^{l-1}) + (\nabla[G(\mathbf{u}^{l-1})\Delta_l W], \nabla \mathbf{u}^{l-1})].$$

Then $(M_n, \mathcal{F}_{t_n}, n = 1, \dots, N)$ is a discrete martingale. For $s \in [t_l, t_{l+1})$, $l = 1, \dots, N-1$, set $\underline{s} = t_l$ and $\mathbf{u}^{\underline{s}} = \mathbf{u}^l$. With these notations, $M_n = \tilde{M}_{t_n}$, where

$$\tilde{M}_t = 2\alpha \int_{t_1}^t [(G(\mathbf{u}^{\underline{s}}) dW(s), \mathbf{u}^{\underline{s}}) + (\nabla G(\mathbf{u}^{\underline{s}}) dW(s), \nabla \mathbf{u}^{\underline{s}})], \quad t \in [t_1, T].$$

The processes $G(\mathbf{u}^\varepsilon)$, $\nabla G(\mathbf{u}^\varepsilon)$, \mathbf{u}^ε and $\nabla \mathbf{u}^\varepsilon$ are \mathcal{F}_s -adapted, and $(\tilde{M}_t, \mathcal{F}_t, t \in [t_1, T])$ is a square integrable martingale, such that

$$\begin{aligned} \langle \tilde{M} \rangle_{t_n} &\leq 4\alpha^2 \int_{t_1}^{t_n} 2[\|G(\mathbf{u}^\varepsilon)\|_{\mathcal{L}(K, \mathbb{L}^2)}^2 \operatorname{Tr} Q |\mathbf{u}^\varepsilon|_{\mathbb{L}^2}^2 + \|G(\mathbf{u}^\varepsilon)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \operatorname{Tr} Q |\nabla \mathbf{u}^\varepsilon|_{\mathbb{L}^2}^2] ds \\ &\leq 8\alpha^2 K_0 \operatorname{Tr} Q \tilde{C} \frac{T}{N} \sum_{l=1}^{n-1} [|\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2]. \end{aligned} \quad (8.8)$$

Using (8.6) we deduce that for $\lambda \in (0, 1)$, $\alpha > 0$ and $\mu > 1$,

$$\begin{aligned} \exp\left(\alpha \max_{1 \leq n \leq N} \|\mathbf{u}^n\|_V^2\right) &\leq \exp\left(\frac{\alpha}{1-\lambda} \|u_0\|_V^2\right) \exp\left(\frac{2K_0\alpha}{\lambda(1-\lambda)} \sum_{l=1}^N \|\Delta_l W\|_K^2\right) \\ &\times \exp\left(\max_{2 \leq n \leq N} \left[M_n - \frac{\mu}{2} \langle \tilde{M} \rangle_{t_n}\right]\right) \exp\left(\max_{2 \leq n \leq N} \left[\frac{\mu}{2} \langle \tilde{M} \rangle_{t_n} - 2k\nu \sum_{l=1}^n (|\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2)\right]\right). \end{aligned} \quad (8.9)$$

Let $\tilde{\alpha}_0 = \frac{\nu}{2\tilde{C}K_0\operatorname{Tr}Q}$; for $\alpha \in (0, \tilde{\alpha}_0)$, we may choose $\mu > 1$ such that $\mu\alpha < \tilde{\alpha}_0$; using (8.8) we deduce that for such a choice of α and μ we have a.s.

$$\begin{aligned} \max_{2 \leq n \leq N} \left[\mu \langle \tilde{M} \rangle_{t_n} - 2k\nu \sum_{l=1}^n [|\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2]\right] \\ \leq \left(\frac{\mu}{2} 8\alpha K_0 \operatorname{Tr} Q \tilde{C} - 2\nu\right) \alpha \frac{T}{N} \sum_{l=2}^{N-1} [|\nabla \mathbf{u}^l|_{\mathbb{L}^2}^2 + |\mathbf{A} \mathbf{u}^l|_{\mathbb{L}^2}^2] \leq 0. \end{aligned}$$

Thus, Hölder's inequality with conjugate exponents μ and $\frac{\mu}{\mu-1}$ implies for $\lambda = \frac{1}{2}$

$$\begin{aligned} \mathbb{E}\left[\exp\left(\alpha \max_{0 \leq l \leq N} \|\mathbf{u}^l\|_V^2\right)\right] &\leq \exp\left(2\alpha \|u_0\|_V^2\right) \left\{\mathbb{E}\left[\exp\left(\frac{8\mu K_0\alpha}{\mu-1} \sum_{l=1}^N \|\Delta_l W\|_K^2\right)\right]\right\}^{\frac{\mu-1}{\mu}} \\ &\times \left\{\mathbb{E}\left[\max_{2 \leq n \leq N} \exp\left(\mu \tilde{M}_{t_n} - \frac{\mu^2}{2} \langle \tilde{M} \rangle_{t_n}\right)\right]\right\}^{\frac{1}{\mu}}. \end{aligned} \quad (8.10)$$

Since $\{\exp(\mu \tilde{M}_t - \frac{\mu^2}{2} \langle \tilde{M} \rangle_t)\}_{t \in [t_1, T]}$ is an exponential martingale, choosing N large enough to ensure $\frac{8\mu\tilde{\alpha}_0 K_0 T}{(\mu-1)N} < \frac{1}{2\operatorname{Tr}Q}$, (8.7) implies (8.1). This completes the proof. \square

The following theorem proves a similar result about the existence of exponential moments for the solution u to the stochastic Navier-Stokes equations. Its proof, which is similar to the above one, with a slight modification of that of [3, Lemma 3.8], is omitted.

Theorem 8.2. *Let G satisfy condition **(G1)** with $K_1 = 0$, that is $\|G(u)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \leq K_0$ and set $\tilde{\alpha}_0 := \frac{\nu}{2\tilde{C}K_0\operatorname{Tr}Q}$, where the constant \tilde{C} is defined by $|\nabla u|_{\mathbb{L}^2}^2 \leq \tilde{C}|\mathbf{A}u|_{\mathbb{L}^2}^2$.*

Let $u_0 \in V$; then for $0 < \alpha < \tilde{\alpha}_0$, there exists a positive constant $C(\alpha)$ such that

$$\mathbb{E}\left[\exp\left(\alpha \sup_{t \in [0, T]} \|u(t)\|_V^2\right)\right] = C(\alpha) < \infty. \quad (8.11)$$

8.2. Random initial condition. In this section, we extend Theorems 8.1 and 8.2 to a random initial condition u_0 such that its V norm has exponential moments.

Theorem 8.3. *Let G satisfy condition (G1) with $K_1 = 0$, that is $\|G(u)\|_{\mathcal{L}(K, \mathbb{W}^{1,2})}^2 \leq K_0$ and suppose that u_0 is a v -valued random variable independent of W and such that $\mathbb{E}[\exp(\gamma_0 \|u_0\|_V^2)] < \infty$ for some $\gamma_0 > 0$. Set $\tilde{\alpha}_0 := \frac{\nu}{2\tilde{C}K_0 \text{Tr}Q}$, where the constant \tilde{C} is defined by $|\nabla u|_{\mathbb{L}^2}^2 \leq \tilde{C}|Au|_{\mathbb{L}^2}^2$ and $\tilde{\beta}_0 := \tilde{\alpha}_0 \frac{\gamma_0}{\gamma_0 + \tilde{\alpha}_0}$. Then for $0 < \alpha < \tilde{\beta}_0$, there exists a positive constant $C(\beta)$ such that for N large enough,*

$$\mathbb{E}\left[\exp\left(\alpha \max_{0 \leq l \leq N} \|u_N(t_l)\|_V^2\right)\right] + \mathbb{E}\left[\exp\left(\alpha \sup_{t \in [0, T]} \|u(t)\|_V^2\right)\right] = C(\beta) < \infty. \quad (8.12)$$

Remark 8.4. *Note that when u_0 is deterministic, the exponential moment estimate of $\|u_0\|_V^2$ holds for every $\gamma_0 > 0$. As $\gamma_0 \rightarrow \infty$, we have $\tilde{\beta}_0 \rightarrow \tilde{\alpha}_0$.*

Proof. We only prove the exponential moment estimate for $\max_{0 \leq l \leq N} \|u_N(t_l)\|_V^2$; the corresponding one for $\sup_{0 \leq t \leq T} \|u(t)\|_V^2$ is proven in a similar (simpler) way.

We want to use Hölder's inequality in (8.9) with exponents $p_1 \in (1, \infty)$, $p_2 = \mu \in (1, \infty)$ and $p_3 \in (1, \infty)$ such that $\frac{1}{p_1} + \frac{1}{\mu} + \frac{1}{p_3} = 1$; this requires $\alpha\mu < \tilde{\alpha}_0$. Suppose that for some $\lambda \in (0, 1)$, $\frac{p_1\alpha}{1-\lambda} \leq \gamma_0$ and $\alpha\mu < \tilde{\alpha}_0$, that is $\frac{1}{p_1} > \frac{\alpha}{\gamma_0}$ and $\frac{1}{\mu} > \frac{\alpha}{\tilde{\alpha}_0}$. Since $\frac{1}{p_1} + \frac{1}{\mu} < 1$, we deduce that $\alpha < \tilde{\alpha}_0 \frac{\gamma_0}{\gamma_0 + \tilde{\alpha}_0} := \tilde{\beta}_0$. For $\alpha < \tilde{\beta}_0$, set $\bar{p}_1 = \frac{\gamma_0}{\alpha}$ and $\mu = \frac{\tilde{\alpha}_0}{\alpha}$. Then choose $\lambda \in (0, 1)$ small enough to have $p_1 := (1-\lambda)\bar{p}_1 = \frac{\gamma_0(1-\lambda)}{\alpha}$ such that $\frac{1}{p_1} + \frac{1}{\mu} < 1$, and define $p_3 \in (1, \infty)$ by $\frac{1}{p_3} = 1 - \frac{1}{p_1} - \frac{1}{\mu}$. This yields

$$\begin{aligned} \mathbb{E}\left[\exp\left(\alpha \max_{0 \leq l \leq N} \|u^l\|_V^2\right)\right] &\leq \left\{\mathbb{E}\left(\exp\left(\frac{p_1\alpha\|u_0\|_V^2}{1-\lambda}\right)\right)\right\}^{\frac{1}{p_1}} \left\{\mathbb{E}\left[\max_{2 \leq n \leq N} \exp\left(\mu\tilde{M}_{t_n} - \frac{\mu^2}{2}\langle\tilde{M}\rangle_{t_n}\right)\right]\right\}^{\frac{1}{\mu}} \\ &\times \left\{\mathbb{E}\left[\exp\left(\frac{p_3 2K_0\alpha}{\lambda(1-\lambda)} \sum_{l=1}^N \|\Delta_l W\|_K^2\right)\right]\right\}^{\frac{1}{p_3}}. \end{aligned} \quad (8.13)$$

Let N be large enough to ensure $\frac{p_3 2\tilde{\beta}_0 K_0 T}{N\lambda(1-\lambda)} < \frac{1}{2\text{Tr}Q}$. Then (8.7) implies (8.12); this concludes the proof. \square

Acknowledgements This work started while Hakima Bessaih and Annie Millet stayed at the Mathematisches Forschung Institute Oberwolfach, during a Research In Pairs program in March 2019. They gratefully thank the MFO for the financial support and excellent working conditions. It was completed while A. Millet visited the University of Wyoming in October 2019; she thanks this University for the hospitality and a very friendly atmosphere. A. Millet's research has been conducted within the FP2M federation (CNRS FR 2036). Hakima Bessaih is partially supported by Simons Foundation grant 582264. The authors wish to thank an anonymous referee for a careful reading which helped them to improve the presentation of the paper.

REFERENCES

- [1] Bensoussan A., Some existence results for stochastic partial differential equations, Pitman Res. Notes Math. Ser., 268, Longman Sci. Tech., Harlow, (Trento, 1990), p. 37–53.
- [2] Bessaih, H., Brzeźniak, Z. & Millet, A., Splitting up method for the 2D stochastic Navier-Stokes equations, *Stochastic PDE: Analysis and Computations* **2-4**, (2014), 433-470.
- [3] Bessaih, H. & Millet, A., Stong L^2 convergence of time numerical schemes for the stochastic two-dimensional Navier-Stokes equations, *IMA Journal of Numerical Analysis* **39-4**, (2019), 2135-2167.
- [4] Bessaih, H., Hausenblas, E., Randrianasolo, T. & Razafimandimby, P. A., Numerical approximation of stochastic evolution equations: Convergence in scale of Hilbert spaces, *J. Comput. Appl. Math.* **343**, (2018), 250-274.

- [5] Blömker, D. & Jentzen, A., Galerkin approximations for the stochastic Burgers equation, *SIAM J. Numer. Anal.* 51(1), (2013), 694–715.
- [6] Breckner, H., Galerkin approximation and the strong solution of the Navier-Stokes equation, *J. Appl. Math. Stochastic Anal.*, 13(3), (2000), 239–259.
- [7] Breit, D. & Dogson, A., Convergence rates for the numerical approximation of the 2D Navier-Stokes equations, arXiv:1906.11778v2, July 2019.
- [8] Brzeźniak, Z., Carelli, E. & Prohl, A., Finite element base discretizations of the incompressible Navier-Stokes equations with multiplicative random forcing, *IMA J. Numer. Anal.*, **33-3**, (2013), 771–824.
- [9] Carelli, E. & Prohl, A., Rates of convergence for discretizations of the stochastic incompressible Navier-Stokes equations, *SIAM J. Numer. Anal.* **50 (5)**, (2012), 2467-2496.
- [10] Chueshov, I. & Millet, A., Stochastic 2D hydrodynamical type systems: Well posedness and large deviations, *Appl. Math. Optim.*, **61-3**, (2010), 379–420.
- [11] Da Prato, G. & Zabczyk, J., Stochastic equations in infinite dimensions. Second edition. Encyclopedia of Mathematics and its Applications, 152. Cambridge University Press, Cambridge, 2014.
- [12] Duan, Y. & Yang, X., The finite element method of a Euler scheme for stochastic NavierStokes equations involving the turbulent component, *International Journal of Numerical Analysis and Modeling*, **10-3**, 727-744.
- [13] Dörsek, P., Semigroup splitting and cubature approximations for the stochastic Navier-Stokes Equations, *SIAM J. Numer. Anal.* **50-2**, (2012), 729-746.
- [14] Hairer, M. & Mattingly, J., Ergodicity of the 2D Navier-Stokes equations with degenerate forcing, *Annals of Mathematics* **164**, (2006), 993-1032.
- [15] Heywood, J.G. & Rannacher, R., G-Finite element approximation of the non-linear Navier-Stokes problem, I. Regularity of solutions and second order error estimates for spatial discretization, *SIAM J. Numer. Anal.* **19** (1982), 275-311.
- [16] Huttenhaller, M. & Jentzen, A., Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients, *Mem. Amer. Math. Soc.* 236 (2015), no. 1112.
- [17] Jentzen, A., Salimova, D. & Welti, T., Strong convergence for explicit space-time discrete numerical approximation methods for stochastic Burgers equations, *J. Math. Anal. Appl.* **469-2**, (2019), 493-548.
- [18] Langa, J.A., Real, J. & Simon, J., Existence and regularity of the pressure for the stochastic Navier-Stokes equations, *Appl. Math. Optim.* **48** (2003), 195-210.
- [19] Printems, J., On the discretization in time of parabolic stochastic partial differential equations, *M2AN Math. Model. Numer. Anal.* **35-6**, (2001) 1055-1078.
- [20] Scott, L.R. & Vogelius, M., Conforming finite element methods for incompressible and nearly compressible continua, *Lectures in Applied Mathematics* **22** (1985), 221-244.
- [21] Temam, R., Navier-Stokes equations. Theory and numerical analysis. Studies in Mathematics and its Applications 2, North-Holland Publishing Co., Amsterdam - New York (1979).
- [22] Temam, R., Navier-Stokes equations and Nonlinear Functional Analysis, CBMS-NSF Regional Conference Series in Applied Mathematics, 66. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, (1995).
- [23] Zhang, S., A new family of stable mixed finite elements for the 3D Stokes equations, *Math. Comp.* **74** (2005), 543-554.

UNIVERSITY OF WYOMING, DEPARTMENT OF MATHEMATICS AND STATISTICS, DEPT. 3036, 1000 EAST UNIVERSITY AVENUE, LARAMIE WY 82071, UNITED STATES

Email address: `bessaih@uwyo.edu`

SAMM, EA 4543, UNIVERSITÉ PARIS 1 PANTHÉON SORBONNE, 90 RUE DE TOLBIAC, 75634 PARIS CEDEX FRANCE *and* LABORATOIRE DE PROBABILITÉS, STATISTIQUE ET MODÉLISATION, UMR 8001, UNIVERSITÉS PARIS 6-PARIS 7

Email address: `amillet@univ-paris1.fr`