IMA Journal of Numerical Analysis (2018) **00**, 1–33
doi:10.1093/imanum/dry058

Strong L^2 convergence of time numerical schemes for the stochastic two-dimensional Navier–Stokes equations

Hakima Bessaih
Department of Mathematics and Statistics, University of Wyoming, Dept. 3036, 1000 East University Avenue, Laramie, WY 82071, USA
bessaih@uwyo.edu

and

Annie Millet*
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, France
*Corresponding author: amillet@univ-paris1.fr

[Received on 11 January 2018; revised on 7 July 2018]

We prove that some time discretization schemes for the two-dimensional Navier–Stokes equations on the torus subject to a random perturbation converge in $L^2(\Omega)$. This refines previous results that established the convergence only 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 fully implicit and semiimplicit temporal Euler discretizations and of a splitting scheme. The speed of the $L^2(\Omega)$ convergence depends on the diffusion coefficient and on the viscosity parameter.

Keywords: stochastic Navier–Sokes equations; numerical schemes; strong convergence; implicit time discretization; splitting scheme; exponential moments.

1. Introduction

Incompressible fluid flow dynamics is described by the so-called incompressible Navier–Stokes equations. The fluid flow is defined by a velocity field and a pressure term that evolve in a very particular way. These equations are parametrized by the viscosity coefficient $\nu > 0$. Their quantitative and qualitative properties depend on the dimensional setting. For example, while the well-posedness of global weak solutions of the 2D Navier–Stokes equations is well known and established, the uniqueness of global weak solutions for the three-dimensional case is completely open. In this paper we will focus on the 2D incompressible Navier–Stokes equations in a bounded domain $D = [0, L]^2$, subject to an external forcing defined as

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

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

where $T > 0$. The process $u : \Omega \times (0,T) \times D \to \mathbb{R}^2$ is the velocity field with initial condition u_0 in D and 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 $\pi : \Omega \times (0,T) \times D \to \mathbb{R}$ is the pressure.

The external force is described by a stochastic perturbation and will be defined in detail later. Here G is a diffusion coefficient satisfying a global Lipschitz condition. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ denote a filtered probability space and let W be a Wiener process to be precisely defined later.

There is an extensive literature concerning deterministic models and we refer to the books of Temam: see Temam (1979, 1995) for known results. The stochastic case has also been widely investigated; see Flandoli & Gatarek (1995) for some very general results and the references therein. For the 2D case, unique global weak and strong solutions (in the Partial Differential Equation (PDE) sense) are constructed for both additive and multiplicative noise, and without being exhaustive, we refer to Breckner (2000) and Chueshov & Millet (2010).

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 Stochastic Partial Differential Equations (SPDEs) and has been of significant interest to the probability community. Many algorithms based on finite difference, finite element or spectral Galerkin methods (for the space discretization), and on Euler schemes, Crank–Nicolson or Runge–Kutta schemes (for the temporal discretization) have been introduced for both the linear and nonlinear case. Their rates of convergence have been widely investigated. The literature on the numerical analysis of SPDEs is now extensive. When the models are linear, have global Lipschitz properties or more generally some monotonicity property then there is an extensive literature; see Bensoussan (1990) and Bensoussan *et al.* (1992). Moreover, in this case 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 difficult to obtain. Indeed, because of the stochastic perturbation, there is no way of using the Gronwall lemma after taking the expectation of the error bound because it involves a nonlinear term that is usually in a quadratic form. One way of getting around this 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 Printems (2001).

The stochastic Navier–Stokes equations with multiplicative noise (1.1) have been investigated by Brzeźniak *et al.* (2013). There, space discretization based on finite elements and an Euler scheme for the time discretization has been implemented. The numerical scheme was proven to converge in probability with a particular rate. A similar problem has been investigated by Carelli & Prohl (2012), with more focus on various Euler schemes including semiimplicit and fully implicit ones. This gave rise to a slightly different rate of convergence, although still in probability. Again, the main tool used is the localization of the nonlinear term over a probability space of ‘large’ probability. In Bessaih *et al.* (2014), the authors used a splitting method, based on the Lie–Trotter formula, proving again a certain rate of convergence in probability of the numerical scheme. Dörsek (2012) studied a semigroup splitting and used cubature approximations, obtaining interesting results for the case of additive noise. When the noise is additive a pathwise argument was used by Breckner (2000); almost sure convergence and convergence in the mean were obtained, although no rate of convergence was explicitly given. To the best of our knowledge there is no result about a strong speed of convergence for the stochastic Navier–Stokes equations in the current literature.

Numerical schemes for stochastic nonlinear models with locally Lipschitz nonlinearities related with the Navier–Stokes equations have been studied by several authors. For nonlinear parabolic SPDEs Blömker & Jentzen (2013) proved a speed of convergence in probability of Galerkin approximations of the stochastic Burgers equation, which is a simpler nonlinear PDE that has some similarity with the

Navier–Stokes equation. In Bessaih *et al.* (2016), an abstract stochastic nonlinear evolution equation in a separable Hilbert space was investigated, including the Gledzer-Ohkitani-Yamada (GOY) and Sabra shell models. These nondimensional models are phenomenological approximations of the Navier–Stokes equations. The authors proved the convergence in probability in a fractional Sobolev space H^s, $0 \le s < \frac{1}{4}$, of a space–time numerical scheme defined in terms of a Galerkin approximation in space and a semiimplicit Euler–Maruyama scheme in time. For the Burgers equation as well as more general nonlinear SPDEs subject to space–time white-noise-driven perturbation Jentzen *et al.* (2017) proved the strong convergence of the scheme but did not give a rate of convergence.

In this paper we focus on the stochastic 2D Navier–Stokes equations and would like to go one step further, that is, obtain a strong speed of convergence in mean square instead of the convergence in probability. In fact, the main goal is twofold. On the one hand, we will improve the convergence from convergence in probability to $L^2(\Omega)$ convergence, the so-called strong convergence in mean square. On the other hand, we will also improve the rate of convergence from logarithmic to almost polynomial.

To explain the method the paper will deal with two different algorithms: the splitting scheme used in Bessaih *et al.* (2014) and the implicit Euler schemes used in Carelli & Prohl (2012). In the case of a diffusion coefficient G with linear growth conditions, which may depend on the solution and its gradient for the Euler schemes, we prove that the speed of convergence of both schemes is any negative power of the logarithm of the time mesh $\frac{T}{N}$ when the initial condition belongs to $\mathbb{W}^{1,2}$ and is divergence-free. In the case of an additive noise—or under a slight generalization of such a noise—we prove that the strong $L^2(\Omega)$ speed of convergence of the fully or semiimplicit Euler schemes introduced by Carelli & Prohl (2012) is polynomial in the time mesh. This speed depends on the viscosity coefficient ν and on the length of the time interval T. When T is small, or when ν is large, this speed is close to the best one that can be achieved in time, that is, almost $\frac{1}{4}$. This is consistent with the time regularity of the strong solution to the stochastic Navier–Stokes equations, due to the scaling between the time and space variables in the heat kernel, and to the stochastic integral.

Let us try to explain the steps of our method here before going into more detail later on in the paper. As we explained earlier the main difficulty that prevents us from getting the strong convergence in mean square is due to the nonlinear term $(u \cdot \nabla)u$. Indeed, in order to bound the error $e_k := u(t_k) - u_N(t_k)$ over the grid points $t_k,\ k = 1, \dots, N$, in an implicit Euler method, one has to find an upper bound for

$$\mathbb{E}\|(u(t_k) \cdot \nabla)u(t_k) - (u_N(t_k) \cdot \nabla)u_N(t_k)\|_{V'}.$$

To close the estimates and use a Gronwall lemma, the tool used in Carelli & Prohl (2012) (as well as in Bessaih *et al.*, 2014) is to localize on a subspace of Ω. However, in both previous results, the localization set depends on the discretization. In this work we perform slightly different computations, based on the antisymmetry of the bilinear term, and localize on sets that depend only on the solution to the stochastic Navier–Stokes equations (1.1), such as Ω_N^M defined by (4.6) for the Euler schemes. Hence, one obtains, for example,

$$\mathbb{E}\Big(1_{\Omega_N^M} \max_{1\le k\le N} |e_k|^2_{\mathbb{L}^2}\Big) \le C\ \exp\big[C_1(M)T\big]\ \Big(\frac{T}{N}\Big)^{\eta},$$

where $\eta < \frac{1}{2}$ and $C_1(M)$ is a constant depending on the bound M of the $\mathbb{L}^2$-norm of ∇u imposed on the localization set Ω_N^M. The exponent η is natural and related to the time regularity of the solution u when the initial condition u_0 belongs to $\mathbb{W}^{1,2}$.

In order to prove the strong speed of convergence we will use the partition of Ω into Ω_N^M and its complement for some threshold M depending on N. More precisely, we have to balance the upper bound on the moments localized on the set $\Omega_N^{M(N)}$ for some well-chosen sequence $M(N)$, going to infinity as N does, and a similar upper bound on the $L^2(\Omega)$ moment of the error localized on the complement of the set $\Omega_N^{M(N)}$. This is performed by finding upper bounds on moments of $u(t_k)$ and of $u_N(t_k)$ uniformly in N and k, estimating $\mathbb{P}((\Omega_N^{M(N)})^c)$, and using the Hölder inequality

$$\mathbb{E}\Big(1_{(\Omega_N^{M(N)})^c} \max_{1\le k\le N} |e_k|_{\mathbb{L}^2}^2\Big) \le \Big(\mathbb{P}\big((\Omega_N^{M(N)})^c\big)\Big)^{\frac{1}{p}} \left[\mathbb{E}\left(\sup_{0\le s\le T} |u(s)|_{\mathbb{L}^2}^{2q} + \max_{0\le k\le N} |u_N(t_k)|_{\mathbb{L}^2}^{2q}\right)\right]^{\frac{1}{q}},$$

where p and q are conjugate exponents. A similar bound was previously used in Hutzenthaler & Jentzen (2015) in a different numerical framework. The upper bound on the probability of the 'bad' set depends on the assumptions about the diffusion coefficient. Note that since we are localizing on a set that does not depend on the discretization scheme, only moments of the solution to the stochastic Navier–Stokes equation (1.1) have to be dealt with.

In the case of a globally Lipschitz coefficient G we use bounds on various moments of u in $\mathbb{W}^{1,2}$. For both schemes the strong speed of convergence is again in the logarithmic scale; when the initial condition is deterministic it is any negative power of $\ln(N)$.

In the case of additive noise we use a slight extension of exponential moments of the solution of (1.1) in vorticity formulation proved previously by Hairer & Mattingly (2006), given by $\mathbb{E}\big(\sup_{t\in[0,T]} \exp(\alpha_0 |\nabla u(t)|_{\mathbb{L}^2}^2)\big) < \infty$ for some $\alpha_0 > 0$. This yields a better speed of convergence, due to the fact that the polynomial Markov inequality is replaced by an exponential one.

For the implicit Euler scheme the strong speed of convergence is polynomial with exponent $\gamma < \frac{1}{2}$ which depends on the viscosity ν. Note that for large ν, γ approaches $\frac{1}{2}$. For the splitting scheme the strong speed of convergence we obtain in this paper is better than that of the convergence in probability proven in Bessaih *et al.* (2014), although not polynomial; it is of the form $c\exp(-C\sqrt{N})$.

In this paper we deal with time discretization only, unlike in Carelli & Prohl (2012) where a space–time discretization is studied. Furthermore, in order to keep the paper at a reasonable length and present arguments that are simple for the reader to follow, we assume that G does not depend on time. We add relevant comments and remarks on the assumptions to be added in the case of time-dependent coefficients for implicit Euler schemes (see Section 4.5).

The paper is organized as follows. Section 2 recalls basic properties of the 2D Navier–Stokes equations, function spaces and strong solutions. We formulate the assumptions on the noise. The splitting scheme is described in Section 3 and various moment estimates previously used in Bessaih *et al.* (2014) are recalled. The strategy for proving the strong speed of convergence is described and explained in detail. The same strategy is used for the Euler schemes in Section 4, which is devoted to the fully implicit and semiimplicit Euler schemes studied previously in Carelli & Prohl (2012). Their strong speed of convergence is proved with a rate of convergence that is polynomial when the exponential moment is used. Finally, Section 5 provides on the one hand an improved moment estimate for an auxiliary process used in the splitting scheme, and on the other hand the proof for the exponential moment estimates of the gradient of the solution to (1.1).

As usual, except if specified otherwise, C denotes a positive constant that may change throughout the paper, and $C(a)$ denotes a positive constant depending on the parameter a.

2. Notation and preliminary results

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

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

be 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 $\mathrm{Dom}(A) = \mathbb{W}^{2,2} \cap H$.

Let $b : V^3 \to \mathbb{R}$ denote the trilinear map defined by

$$b(u_1,u_2,u_3) := \int_D \big(u_1(x)\cdot\nabla u_2(x)\big)\cdot u_3(x)\,\mathrm{d}x,$$

which by the incompressibility condition satisfies $b(u_1,u_2,u_3) = -b(u_1,u_3,u_2)$ for $u_i \in V$, $i=1,2,3$. There exists a continuous bilinear map $B : V\times V \mapsto V'$ such that

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

The map B satisfies the following antisymmetry relations:

$$\langle B(u_1,u_2),u_3\rangle = -\langle B(u_1,u_3),u_2\rangle, \quad \langle B(u_1,u_2),u_2\rangle = 0 \qquad \text{for all} \quad u_i \in V. \tag{2.1}$$

Furthermore, the Gagliardo–Nirenberg inequality implies that for $X := H \cap \mathbb{L}^4(D)$ we have

$$\|u\|_X^2 \le \bar{C}\,|u|_{\mathbb{L}^2}\,|\nabla u|_{\mathbb{L}^2} \le \frac{\bar{C}}{2}\|u\|_V^2 \tag{2.2}$$

for some positive constant $\bar{C}$. For $u \in V$ set $B(u) := B(u,u)$ and recall some well-known properties of B, which easily follow from the Hölder and Young inequalities: given any $\beta > 0$ we have

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

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

for $u_i \in V$, $i=1,2,3$, where

$$C_\beta = \frac{\bar{C}^2 3^3}{4^4\beta^3}, \tag{2.5}$$

and $\bar{C}$ is defined by (2.2).

Let K be a separable Hilbert space and $(W(t), t\in[0,T])$ be a K-cylindrical Wiener process defined on the probability space $(\Omega,\mathcal{F},(\mathcal{F}_t),\mathbb{P})$. For technical reasons we assume that the initial condition u_0

belongs to $L^p(\Omega; V)$ for some $p \in [2, \infty]$ and consider only *strong solutions* in the PDE sense. Given two Hilbert spaces H_1 and H_2 let $\mathcal{L}_2(H_1, H_2)$ denote the set of Hilbert–Schmidt operators from H_1 to H_2. The diffusion coefficient G satisfies the following assumption.

Condition (G1). Assume that $G : V \to \mathcal{L}_2(K, H)$ is continuous and there exist positive constants K_i, $i = 0, 1$ and L_1 such that for $u, v \in V$,

$$\|G(u)\|^2_{\mathcal{L}_2(K,H)} \le K_0 + K_1 |u|^2_{\mathbb{L}^2}, \tag{2.6}$$

$$\|G(u) - G(v)\|^2_{\mathcal{L}_2(K,H)} \le L_1 |u - v|^2_{\mathbb{L}^2}. \tag{2.7}$$

Finally, note that the following identity involving the Stokes operator A and the bilinear term holds (see, e.g., Temam, 1979, Lemma 3.1):

$$\langle B(u), Au \rangle = 0, \quad u \in \mathrm{Dom}(A). \tag{2.8}$$

We also suppose that G satisfies the following assumptions.

Condition (G2). The coefficient $G :\ \mathrm{Dom}(A) \to \mathcal{L}_2(K, V)$ and there exist positive constants K_i, $i = 0, 1$ and L_1 such that for every $u, v \in\ \mathrm{Dom}(A)$,

$$\|G(u)\|^2_{\mathcal{L}_2(K,V)} \le K_0 + K_1 \|u\|^2_V, \tag{2.9}$$

$$\|G(u) - G(v)|^2_{\mathcal{L}_2(K,V)} \le L_1 \|u - v\|^2_V. \tag{2.10}$$

We define a strong solution of (1.1) as follows (see Carelli & Prohl, 2012, Definition 2.1).

DEFINITION 2.1 We say that equation (1.1) has a strong solution if

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

$$\begin{aligned}\big(u(t), \phi\big) + \nu \int_0^t \big(\nabla u(s), \nabla \phi\big)\, \mathrm{d}s + \int_0^t \big\langle [u(s) \cdot \nabla] u(s), \phi \big\rangle\, \mathrm{d}s \\ = \Big(u_0, \phi) + \int_0^t (\phi, G(u(s))\, \mathrm{d}W(s) \Big)\end{aligned}$$

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

As usual, by projecting (1.1) on divergence-free fields, the pressure term is implicitly in the space V and can be recovered afterwards. Bessaih & Millet (2012), Proposition 2.2 (see also Bessaih *et al.*, 2014, 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 $\mathbb{E}\big(\|u_0\|^{2p}_V\big) < \infty$ for some real number $p \in [2, \infty)$. Assume that the conditions **(G1)** and **(G2)** are satisfied.

Then there exists a unique solution u to equation (1.1). Furthermore, for some positive constant C we have

$$\mathbb{E}\left(\sup_{t\in[0,T]} \|u(t)\|_V^{2p} + \int_0^T |Au(s)|_{\mathbb{L}^2}^2 \big(1 + \|u(s)\|_V^{2(p-1)}\big)\, \mathrm{d}s\right) \le C\left[1 + \mathbb{E}\left(\|u_0\|_V^{2p}\right)\right]. \tag{2.11}$$

3. Time-splitting scheme

In this section we prove the strong $L^2(\Omega)$ convergence of the splitting scheme introduced in Bessaih *et al.* (2014).

3.1 *Description of the splitting scheme*

Let $N > 1$, $h = \frac{T}{N}$ denote the time mesh and $t_i = \frac{iT}{N}$, $i = 0, \ldots, N$ denote a partition of the time interval $[0, T]$. Let $F : V \to V'$ be defined by

$$F(u) = \nu Au + B(u, u). \tag{3.1}$$

Note that the formulation of **(G2)** is slighty different from that used in Bessaih *et al.* (2014). They are equivalent due to the inequality $|\nabla u|_{\mathbb{L}^2} \le C|\mathrm{curl}\ u|_{\mathbb{L}^2}$ for $u \in V$, where $\mathrm{curl}\ u = \partial_{x_1} u_2 - \partial_{x_2} u_1$.

Set $t_{-1} = -\frac{T}{N}$. For $t \in [t_{-1}, 0)$ set $y^N(t) = u^N(t) = u_0$ and $\mathcal{F}_t = \mathcal{F}_0$. The approximation (y^N, u^N) is defined by induction as follows. Suppose that the processes $u^N(t)$ and $y^N(t)$ are defined for $t \in [t_{i-1}, t_i)$ and that $y^N(t_i^-)$ is H-valued and $\mathcal{F}_{t_i}$-measurable. Then for $t \in [t_i, t_{i+1})$, $u^N(t)$ with initial condition $y^N(t_i^-)$ at time t_i, is the unique solution of equation

$$\frac{\mathrm{d}}{\mathrm{d}t} u^N(t) + F(u^N(t)) = 0, \quad t \in [t_i, t_{i+1}), \quad u^N(t_i) = u^N\big(t_i^+\big) = y^N\big(t_i^-\big). \tag{3.2}$$

Then $u^N(t_{i+1}^-)$ is well defined, H-valued and $\mathcal{F}_{t_i}$-measurable. Then set $y^N(t_i) = u^N(t_{i+1}^-)$, and for $t \in [t_i, t_{i+1})$ define $y^N(t)$ as the unique solution of equation

$$\mathrm{d}y^N(t) = G(y^N(t))\, \mathrm{d}W(t), \quad t \in [t_i, t_{i+1}), \quad y^N(t_i) = y^N\big(t_i^+\big) = u^N\big(t_{i+1}^-\big). \tag{3.3}$$

Finally, set $u^N(T) = y^N(T) = y^N(T^-)$. The processes u^N and y^N are well defined and have finite moments as proved in Bessaih *et al.* (2014, Lemma 4.2).

THEOREM 3.1 Let u_0 be a V-valued, $\mathcal{F}_0$ random variable such that $\mathbb{E}(\|u_0\|_V^{2p}) < \infty$ for some real number $p \ge 2$. Suppose that G satisfies conditions **(G1)** and **(G2)**. Then there exists a positive constant C such that for every integer $N \ge 1$,

$$\sup_{t\in[0,T]} \mathbb{E}\big(\|u^N(t)\|_V^{2p} + \|y^N(t)\|_V^{2p}\big) + \mathbb{E}\int_0^T \big(1 + \|u^N(t)\|_V^{2(p-1)}\big)|Au^N(t)|_{\mathbb{L}^2}^2\, \mathrm{d}t \le C. \tag{3.4}$$

The following result gives a bound of the difference between u^N and y^N (see Bessaih *et al.*, 2014, Proposition 4.3).

PROPOSITION 3.2 Let u_0 be $\mathcal{F}_0$-measurable such that $\mathbb{E}(\|u_0\|_V^4) < \infty$ and let G satisfy conditions **(G1)** and **(G2)**. Then there exists a positive constant $C := C(T)$ such that for every integer $N \geq 1$,

$$\mathbb{E}\int_0^T \|y^N(t) - u^N(t)\|_V^2 \,\mathrm{d}t \leq \frac{C}{N}. \tag{3.5}$$

For technical reasons let us consider the process $(z^N(t), t \in [0,T])$ which mixes u^N and y^N and is defined by

$$z^N(t) = u_0 - \int_0^t F(u^N(s))\,\mathrm{d}s + \int_0^t G(y^N(s))\,\mathrm{d}W(s). \tag{3.6}$$

The process z^N is a.s. continuous on $[0,T]$. Note that $z^N(t_k) = y^N(t_k^-) = u^N(t_k)$ for $k = 0, 1, \ldots, N$. The following result gives bounds of the V-norm of $z^N - u^N$ and $z^N - y^N$. It is an extension of in Bessaih *et al.* (2014, Lemma 4.4).

PROPOSITION 3.3 Let u_0 be V-valued, $\mathcal{F}_0$-measurable such that $\mathbb{E}(\|u_0\|_V^{2p}) < \infty$ for some integer $p \geq 2$ and let G satisfy conditions **(G1)** and **(G2)**. Then there exists a constant $C := C(p,T)$ such that for every integer $N \geq 1$,

$$\mathbb{E}\left(\sup_{t\in[0,T]} \|z^N(t) - u^N(t)\|_V^{2p}\right) \leq \frac{C}{N^{p-1}} \quad \text{and} \quad \sup_{t\in[0,T]} \mathbb{E}\big(\|z^N(t) - u^N(t)\|_V^{2p}\big) \leq \frac{C}{N^p}. \tag{3.7}$$

Note that combining (3.5) and (3.7) we deduce that if $\mathbb{E}(\|u_0\|_V^4) < \infty$,

$$\mathbb{E}\int_0^T \|z^N(t) - y^N(t)\|_V^2 \,\mathrm{d}t \leq \frac{C}{N} \tag{3.8}$$

for some constant $C := C(T)$ independent of N.

Proof. For $t \in [t_k, t_{k+1})$, $k = 0, \ldots, N-1$ we have

$$z^N(t) - u^N(t) = \int_{t_k}^t G(y^N(s))\,\mathrm{d}W(s).$$

Since $z^N(T) = u^N(T)$, for any $p \geq 2$ the Burkholder–Davies–Gundy inequality implies

$$\begin{aligned}
\mathbb{E}\left(\sup_{t\in[0,T]} \|z^N(t) - u^N(t)\|_V^{2p}\right) &= \mathbb{E}\left(\sup_{0\leq k<N}\ \sup_{t\in[t_k,t_{k+1})} \|z^N(t) - u^N(t)\|_V^{2p}\right) \\
&\leq C_p \sum_{k=0}^{N-1} \mathbb{E}\left(\left|\int_{t_k}^{t_{k+1}} \|G\big(s, y^N(s)\big)\|_{\mathcal{L}_2(K,V)}^2 \,\mathrm{d}s\right|^p\right) \\
&\leq C_p \sum_{k=0}^{N-1} \left(\frac{T}{N}\right)^p \sup_{t\in[0,T]} \big(K_0 + K_1\mathbb{E}(\|y^N(t)\|_V^{2p})\big).
\end{aligned}$$

Inequality (3.4) concludes the proof of the first part of (3.7). Furthermore,

$$\sup_{t\in[0,T]} \mathbb{E}\big(\|z^N(t)-u^N(t)\|_V^{2p}\big) = \sup_{0\le k<N}\ \sup_{t\in[t_k,t_{k+1})} \mathbb{E}\big(\|z^N(t)-u^N(t)\|_V^{2p}\big).$$

A similar argument concludes the proof. □

3.2 *A localized $L^2(\Omega)$ convergence*

Recall that $X=\mathbb{L}^4(D)\cap H$ is an interpolation space between H and V such that (2.2) holds. For every $M>0$ set

$$\tilde{\Omega}_M(t) := \Big\{\omega\in\Omega\ :\ \sup_{s\in[0,t]}\|u(s)(\omega)\|_X^4 \le M\Big\}. \tag{3.9}$$

Note that once more, and unlike Bessaih *et al.* (2014), this set depends only on the solution u of (1.1) and does not depend on the scheme. Let $\tau_M =: \inf\{t\ge 0: \|u(t)\|_X^4\ge M\}\wedge T$; then $\tau_M=T$ on $\Omega_M(T)$. The following result improves Bessaih *et al.* (2014, Proposition 5.1). Note that in our case both processes u and z^N have a.s. continuous trajectories.

PROPOSITION 3.4 Let u_0 be V-valued, $\mathcal{F}_0$-measurable such that $\mathbb{E}(\|u_0\|_V^8)<\infty$ and suppose that G satisfies conditions **(G1)** and **(G2)**. Then there exist constants C and $\widetilde{C(M)}$ such that

$$\mathbb{E}\left(\sup_{t\in[0,\tau_M]}|z^N(t)-u(t)|_{\mathbb{L}^2}^2+\int_0^{\tau_M}\big[\|u^N(t)-u(t)\|_V^2+\|y^N(t)-u(t)\|_V^2\big]\,\mathrm{d}t\right)\le \frac{C}{N}e^{T\widetilde{C(M)}}, \tag{3.10}$$

where

$$\widetilde{C(M)} := \frac{3^3\bar{C}^2}{2^5\beta^3\nu^3}M + C(\nu,L_1,\beta,\epsilon), \tag{3.11}$$

and $\bar{C}$ is the constant defined in (2.2) and $\beta<1$.

Proof. Let us apply the Itô formula to $|z^N(t\wedge\tau_M)-u(t\wedge\tau_M)|_{\mathbb{L}^2}^2$. This is possible even if z^N and u are not regular enough. Indeed, we can use the Yosida approximation $e^{-\delta A}(z^N(t)-u(t))$ for some $\delta>0$, apply the Itô formula to this smooth process and then pass to the limit as $\delta\to 0$ (see, e.g., Chueshov & Millet, 2010, step 4 of the proof of Theorem 2.4). This implies

$$|z^N(t\wedge\tau_M)-u(t\wedge\tau_M)|_{\mathbb{L}^2}^2=\sum_{i=1}^2 T_i(t\wedge\tau_M)+I(t\wedge\tau_M), \tag{3.12}$$

where

$$
\begin{aligned}
T_1(t\wedge\tau_M) &= -\,2\int_0^{t\wedge\tau_M} \big\langle F(u^N(s)) - F(u(s))\,,\, z^N(s)-u(s)\big\rangle\,\mathrm{d}s,\\
T_2(t\wedge\tau_M) &= \int_0^{t\wedge\tau_M} \|G(y^N(s)) - G(u(s))\|^2_{\mathcal{L}_2(K,H)}\,\mathrm{d}s,\\
I(t\wedge\tau_M) &= 2\int_0^{t\wedge\tau_M} \big(z^N(s)-u(s)\,,\,\big[G(y^N(s)) - G(u(s))\big]\,\mathrm{d}W(s)\big).
\end{aligned}
$$

The Lipschitz property **(G1)** implies

$$
T_2(t\wedge\tau_M) \le 2L_1\left[\int_0^{t\wedge\tau_M} |z^N(s)-u(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s + \int_0^{t\wedge\tau_M} |z^N(s)-y^N(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s\right]. \tag{3.13}
$$

Furthermore, $T_1(t\wedge\tau_M) = \sum_{i=1}^4 T_{1,i}(t\wedge\tau_M)$, where

$$
\begin{aligned}
T_{1,1}(t\wedge\tau_M) &= -\,2\nu\int_0^{t\wedge\tau_M} \langle Au^N(s) - Au(s),\, u^N(s)-u(s)\rangle\,\mathrm{d}s,\\
T_{1,2}(t\wedge\tau_M) &= -\,2\nu\int_0^{t\wedge\tau_M} \langle Au^N(s) - Au(s),\, z^N(s)-u^N(s)\rangle\,\mathrm{d}s,\\
T_{1,3}(t\wedge\tau_M) &= -\,2\int_0^{t\wedge\tau_M} \langle B(u^N(s),u^N(s)) - B(u(s),u(s))\,,\, u^N(s)-u(s)\rangle\,\mathrm{d}s,\\
T_{1,4}(t\wedge\tau_M) &= -\,2\int_0^{t\wedge\tau_M} \langle B(u^N(s),u^N(s)) - B(u(s),u(s))\,,\, z^N(s)-u^N(s)\rangle\,\mathrm{d}s.
\end{aligned}
$$

The definition of A implies

$$
T_{1,1}(t\wedge\tau_M) = -2\nu\int_0^{t\wedge\tau_M} |\nabla u^N(s) - \nabla u(s)|^2_{\mathbb{L}^2}\mathrm{d}s. \tag{3.14}
$$

The Cauchy–Schwarz and Young inequalities imply that for any $\beta_2>0$,

$$
T_{1,2}(t\wedge\tau_M) \le \nu\beta_2\int_0^{t\wedge\tau_M} |\nabla u^N(s) - \nabla u(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s + \frac{\nu}{\beta_2}\int_0^t \|z^N(s)-u^N(s)\|^2_V\,\mathrm{d}s. \tag{3.15}
$$

Using inequality (2.4), we deduce that for any $\beta_3 > 0$ and $\epsilon > 0$,

$$\begin{aligned}T_{1,3}(t\wedge\tau_M) &\le 2\nu\beta_3 \int_0^{t\wedge\tau_M} \big(|\nabla u^N(s) - \nabla u(s)|_{\mathbb{L}^2}^2 + |u^N(s) - u(s)|_{\mathbb{L}^2}^2\big)\, \mathrm{d}s \\ &\quad + 2C_{\nu\beta_3} \int_0^{t\wedge\tau_M} \|u(s)\|_X^4\, |u^N(s) - u(s)|_{\mathbb{L}^2}^2\, \mathrm{d}s \\ &\le 2\nu\beta_3 \int_0^{t\wedge\tau_M} |\nabla[u^N(s) - u(s)]|_{\mathbb{L}^2}^2\, \mathrm{d}s + 2(\nu\beta_3 + C_{\nu\beta_3} M)(1+\epsilon) \int_0^{t\wedge\tau_M} |z^N(t) - u(t)|_{\mathbb{L}^2}^2\, \mathrm{d}s \\ &\quad + C(\nu,\beta_3,\epsilon) \int_0^{t\wedge\tau_M} \big(1 + \|u(s)\|_X^4\big) |z^N(t) - u^N(t)|_{\mathbb{L}^2}^2\, \mathrm{d}s. \end{aligned} \tag{3.16}$$

Finally, since B is bilinear, the Hölder inequality implies

$$\begin{aligned}T_{1,4}(t\wedge\tau_M) &= 2\int_0^{t\wedge\tau_N} \Big[\big\langle B\big(u^N(s) - u(s), z^N(s) - u^N(s)\big)\,,\, u^N(s)\big\rangle \\ &\quad + \big\langle B\big(u(s), z^N(s) - u^N(s)\big)\,,\, u^N(s) - u(s)\big\rangle\Big]\, \mathrm{d}s \\ &\le 2\int_0^{t\wedge\tau_M} \|u^N(s) - u(s)\|_X \big[\|u^N(s)\|_X + \|u(s)\|_X\big] \|z^N(s) - u^N(s)\|_V\, \mathrm{d}s.\end{aligned}$$

Using the interpolation inequality (2.2) and the Young inequality, we deduce that for any $\beta_4 > 0$, there exists a positive constant $C(\nu,\beta_4)$ such that

$$\begin{aligned}&\sqrt{C}\big|\nabla\big(u^N(s) - u(s)\big)\big|_{\mathbb{L}^2}^{\frac12} |u^N(s) - u(s)|_{\mathbb{L}^2}^{\frac12} \|z^N(s) - u^N(s)\|_V \big[\|u^N(s)\|_X + \|u(s)\|_X\big] \\ &\le \nu\beta_4 \big|\nabla\big(u^N(s) - u(s)\big)\big|_{\mathbb{L}^2}^2 + C|z^N(s) - u(s)|_{\mathbb{L}^2}^2 + C|z^N(s) - u^N(s)|_{\mathbb{L}^2}^2 \\ &\quad + C(\nu,\beta_4)\big[\|u^N(s)\|_X^2 + \|u(s)\|_X^2\big] \|z^N(s) - u^N(s)\|_V^2.\end{aligned}$$

This implies

$$\begin{aligned}T_{1,4}(t\wedge\tau_M) &\le \nu\beta_4 \int_0^{t\wedge\tau_M} \big|\nabla\big[u^N(s) - u(s)\big]\big|_{\mathbb{L}^2}^2\, \mathrm{d}s + C\int_0^{t\wedge\tau_M} |z^N(s) - u(s)|_{\mathbb{L}^2}^2\, \mathrm{d}s \\ &\quad + C(\nu,\beta_4)\int_0^{t\wedge\tau_M} \Big(|u^N(s) - z^N(s)|_{\mathbb{L}^2}^2 + \big[\|u^N(s)\|_X^2 + \|u(s)\|_X^2\big] \|z^N(s) - u^N(s)\|_V^2\Big)\, \mathrm{d}s.\end{aligned} \tag{3.17}$$

Collecting the upper estimates (3.13–3.17), we deduce that for $\beta_2 + 2\beta_3 + \beta_4 < 2$ and $t \in [0, T]$,

$$\sup_{0\le s\le t} |z^N(s\wedge\tau_N) - u(s\wedge\tau_N)|^2_{\mathbb{L}^2} + \nu(2-\beta_2-2\beta_3-\beta_4)\int_0^{t\wedge\tau_N} |\nabla[u^N(s)-u(s)]|^2_{\mathbb{L}^2}\,\mathrm{d}s$$
$$\le R(t) + \sup_{s\in[0,t]} I(s\wedge\tau_N) + \Big[2(1+\epsilon)C_{\nu\beta_3}M + 2L_1 + C(\nu,\beta_3,\epsilon)\Big]\int_0^{t\wedge\tau_M} |z^N(s)-u(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s, \tag{3.18}$$

where, gathering all error terms, we let

$$R(t) = \int_0^{t\wedge\tau_M} \Big(2L_1|z^N(s)-y^N(s)|^2_{\mathbb{L}^2} + C(\nu,\beta_2,\beta_3,\beta_4)\|z^N(s)-u^N(s)\|^2_V\Big)\,\mathrm{d}s$$
$$+ C(\nu,\beta_3,\epsilon)\left\{\int_0^{t\wedge\tau_M} \|u(s)\|^8_X\,\mathrm{d}s\right\}^{\frac12}\left\{\int_0^{t\wedge\tau_M} |z^N(s)-u^N(s)|^4_{\mathbb{L}^2}\,\mathrm{d}s\right\}^{\frac12}$$
$$+ C(\nu,\beta_4)\int_0^{t\wedge\tau_M} \big[\|u^N(s)\|^2_X + \|u(s)\|^2_X\big]\|z^N(s)-u^N(s)\|^2_V\mathrm{d}s.$$

The Cauchy–Schwarz inequality, the integrability property $\mathbb{E}(\|u_0\|^8_V) < \infty$ and the upper estimates (2.11), (3.4), (3.7) and (3.8) imply the existence of some positive constant C (depending on the parameter β_i and ϵ) such that for every N, M,

$$\mathbb{E}\left(\sup_{t\in[0,T]} R(t)\right) \le \frac{C}{N}. \tag{3.19}$$

Using the Burkholder–Davis–Gundy inequality, then the Young inequality and the Lipschitz condition **(G1)** on G, we deduce that for any $\delta > 0$,

$$\mathbb{E}\left(\sup_{t\in[s\le t]} I(t)\right) \le 6\mathbb{E}\left(\left\{\int_0^{t\wedge\tau_M} \|G(s,y^N(s)) - G(s,u(s))\|^2_{\mathcal{L}_2(K,H)}|u^N(s)-u(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s\right\}^{\frac12}\right)$$
$$\le 6\mathbb{E}\left(\sup_{s\in[0,t\wedge\tau_M]} |z^N(s)-u(s)|_{\mathbb{L}^2}\left\{\int_0^{t\wedge\tau_M} L_1|y^N(s)-u(s)|^2_{\mathbb{L}^2}\mathrm{d}s\right\}^{\frac12}\right)$$
$$\le \delta\mathbb{E}\left(\sup_{s\in[0,t]} |z^N(s\wedge\tau_M) - u(s\wedge\tau_M)|^2_{\mathbb{L}^2}\right) + C(\delta)\mathbb{E}\int_0^{t\wedge\tau_M} |z^N(s)-u(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s$$
$$+ C(\delta)\mathbb{E}\int_0^{t\wedge\tau_M} |y^N(s)-z^N(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s$$
$$\le \delta\mathbb{E}\left(\sup_{s\in[0,t]} |z^N(s\wedge\tau_M) - u(s\wedge\tau_M)|^2_{\mathbb{L}^2}\right) + C(\delta)\mathbb{E}\int_0^{t\wedge\tau_M} |z^N(s)-u(s)|^2_{\mathbb{L}^2}\,\mathrm{d}s + \frac{C(\delta)}{N}, \tag{3.20}$$

where the last inequality is a consequence of (3.8). The upper estimates (3.18–3.20) imply that if $\delta < 1$ and $\bar{C}(\beta) := 2 - \beta_2 + 2\beta_3 + \beta_4 > 0$,

$$
\begin{aligned}
&(1-\delta)\mathbb{E}\left(\sup_{s\in[0,t]} |z^N(s\wedge\tau_M) - u(s\wedge\tau_M)|^2_{\mathbb{L}^2}\right) + \nu\bar{C}(\beta)\mathbb{E}\int_0^{t\wedge\tau_M} |\nabla[u^N(s) - u(s)]|^2_{\mathbb{L}^2}\,\mathrm{d}s \\
&\quad\leq \frac{C}{N} + \Big[2(1+\epsilon)C_{\nu\beta_3}M + C(L_1,\nu,\beta_3,\epsilon,\delta)\Big]\mathbb{E}\int_0^t |z^N(s\wedge\tau_M) - u(s\wedge\tau_M)|^2_{\mathbb{L}^2}\,\mathrm{d}s. \qquad (3.21)
\end{aligned}
$$

The Gronwall inequality (disregarding the second term on the left-hand side of (3.21)) proves that

$$
\mathbb{E}\left(\sup_{s\in[0,t]} |z^N(s\wedge\tau_M) - u(s\wedge\tau_M)|^2_{\mathbb{L}^2}\right) \leq \frac{C}{N}\exp\big(T\,\widetilde{C(M)}\big),
$$

where $\widetilde{C(M)}$ is defined by (3.11). Indeed, the term $\beta_3 < 1$ has to be chosen first since the other positive constants β_2 and β_4, which have to satisfy $\beta_2 + \beta_4 < 2(1-\beta_3)$, appear only in the 'error' term $R(t)$. Using this inequality in (3.21), the definition of $C_{\nu\beta_3}$ given in (2.5), (3.7) and (3.8) and changing the ratio $\frac{1+\epsilon}{(1-\delta)\beta_3^3}$ into $\frac{1}{\beta^3}$ for some $\beta \in (0,1)$ we conclude the proof of (3.10). □

3.3 *Strong $L^2(\Omega)$ speed of convergence of the splitting scheme*

Since $\tau_M = T$ on $\tilde{\Omega}_M$ we can rewrite (3.10) as

$$
\mathbb{E}\Big(1_{\tilde{\Omega}_M}\Big\{\sup_{t\in[0,T]} |z^N(t) - u(t)|^2_{\mathbb{L}^2} + \int_0^T \big[\|u^N(t) - u(t)\|^2_V + \|y^N(t) - u(t)\|^2_V\big]\,\mathrm{d}t\Big\}\Big) \leq \frac{C}{N}e^{T\tilde{C}(M)}.
$$

The previous section provides an upper bound for the $L^2(\Omega)$-norm of the maximal error on each time step localized on the set Ω_M. In order to deduce the strong speed of convergence we next have to analyse this error on the complement of this localization set.

The first step is described in the following simple upper estimates, which follow from the Hölder inequality. Hence, we suppose that $\mathbb{E}(\|u_0\|_V^{2q}) < \infty$ and that p and q are conjugate exponents; then

$$
\mathbb{E}\left(1_{\tilde{\Omega}^c_M}\sup_{t\in[0,T]} |z^N(t) - u(t)|^2_{\mathbb{L}^2}\right) \leq C\Big[\mathbb{P}(\tilde{\Omega}^c_M)\Big]^{\frac{1}{p}}\left[\mathbb{E}\left(\sup_{t\in[0,T]}\big(|z^N(t)|^{2q}_{\mathbb{L}^2} + |u(t)|^{2q}_{\mathbb{L}^2}\big)\right)\right]^{\frac{1}{q}}. \qquad (3.22)
$$

Note that we have a similar upper estimate

$$
\begin{aligned}
&\mathbb{E}\left(1_{\tilde{\Omega}^c_M}\int_0^T \|z^N(t) - u(t)\|^2_V\,\mathrm{d}t\right) \\
&\quad\leq C\Big[\mathbb{P}\big(\tilde{\Omega}^c_M\big)\Big]^{\frac{1}{p}}\left[\mathbb{E}\int_0^T \big(\|z^N(t) - u^N(t)\|^{2q}_V + \|u^N(t)\|^{2q}_V + \|u(t)\|^{2q}_V\big)\,\mathrm{d}t\right]^{\frac{1}{q}} \leq C\Big[\mathbb{P}\big(\tilde{\Omega}^c_M\big)\Big]^{\frac{1}{p}},
\end{aligned}
$$

where the last upper estimate is deduced from (2.11), (3.4) and (3.7). Theorem 2.2 shows that $\mathbb{E}(\sup_{t\in[0,T]}\|u(t)\|_V^2) < \infty$. We next prove that $\mathbb{E}(\sup_{t\in[0,T]}|z^N(t)|_{\mathbb{L}^2}^{2q})$ is bounded by a constant independent of N.

Lemma 3.5 Let $p \geq 2$ and u_0 be V-valued, $\mathcal{F}_0$-measurable such that $\mathbb{E}(\|u_0\|_V^{2p+1}) < \infty$. Suppose that G satisfies conditions **(G1)** and **(G2)**. Then there exists a positive constant C_p such that for every integer $N \geq 1$,

$$\mathbb{E}\left(\sup_{t\in[0,T]}|z^N(t)|_{\mathbb{L}^2}^{2p} + \nu\int_0^T \|z^N(t)\|_V^{2p}\,\mathrm{d}t\right) \leq C_p.$$

The proof of this lemma is given in Appendix A.1.

We next have to make sure that the threshold $M(N)$ is chosen to balance the upper estimates (3.10) and (3.22).

Case 1: Linear growth diffusion coefficient. Suppose that $\mathbb{E}(\|u_0\|_V^{2q}) < \infty$. Then using the Gagliardo–Nirenberg inequality (2.2) we deduce

$$\begin{aligned}\mathbb{P}\big(\tilde{\Omega}^c_{M(N)}\big) =&\mathbb{P}\left(\sup_{t\in[0,T]}\|u(t)\|_X^4 \geq M(N)\right) = \mathbb{P}\left(\sup_{t\in[0,T]}\|u(t)\|_V^4 \geq \frac{4M(N)}{\bar{C}^2}\right)\\ \leq&\frac{\bar{C}^q}{2^q M(N)^{\frac{q}{2}}}\mathbb{E}\left(\sup_{t\in[0,T]}\|u(t)\|_V^{2q}\right) \leq \frac{C}{M(N)^{\frac{q}{2}}},\end{aligned}$$

where the last upper estimate is a consequence of (2.11). To balance the right-hand sides of (3.10) and (3.22) we choose $M(N) \to \infty$ as $N \to \infty$ such that

$$\frac{1}{N}\exp(T\widetilde{C(M(N))}) \asymp C(q)\frac{1}{M(N)^{\frac{q-1}{2}}}. \tag{3.23}$$

Taking logarithms and using (3.11) this comes down to

$$-\ln(N) + 2(1+\epsilon)C_{\nu\beta}M(N)T \asymp -\frac{q-1}{2}\ln(M(N)),$$

for some $\beta \in (0,1)$ and $\epsilon \in (0,1)$, where $C_{\nu\beta}$ is defined in (2.5). Let

$$M(N) := \frac{1}{2(1+\epsilon)C_{\nu\beta}T}\Big[\ln(N) - \frac{q-1}{2}\ln\big(\ln(N)\big)\Big] \asymp C(\nu,\beta,\epsilon,T)\ln(N). \tag{3.24}$$

Then $M(N) \to \infty$ as $N \to \infty$ and (3.23) is satisfied. This yields the following result, where the first upper estimate follows from (3.23) and (3.7), while the second one is deduced from the fact that $y^N(t_k^+) = u^N(t_{k+1}^-)$.

THEOREM 3.6 Suppose that u_0 is V-valued such that $\mathbb{E}(\|u_0\|_V^{2q+1}) < \infty$ for some $q \geq 4$ and that G satisfies conditions **(G1)** and **(G2)**. Then there exists a constant $C > 0$ such that

$$\mathbb{E}\Bigg(\sup_{t\in[0,T]} \big[|z^N(t) - u(t)|_{\mathbb{L}^2}^2 + |u^N(t) - u(t)|_{\mathbb{L}^2}^2\big] + \int_0^T \big[\|z^N(s) - u(s)\|_V^2 + \|u^N(s) - u(s)\|_V^2 + \|y^N(s) - u(s)\|_V^2\big]\,\mathrm{d}s\Bigg) \leq \frac{C}{\ln(N)^{\frac{q-1}{2}}}, \tag{3.25}$$

$$\mathbb{E}\Bigg(\sup_{k=1,\ldots,N} \big[|u^N(t_k^+) - u(t_k)|_{\mathbb{L}^2}^2 + |y^N(t_k^+) - u(t_k)|_{\mathbb{L}^2}^2\big]\Bigg) \leq \frac{C}{\ln(N)^{\frac{q-1}{2}}}. \tag{3.26}$$

REMARK 3.7 Note that if u_0 is a deterministic element of V, or more generally if $\|u_0\|_V$ has moments of all orders (for example, if u_0 is a V-valued Gaussian random variable independent of the noise W), then the speed of convergence of the current splitting scheme is any negative power of $\ln(N)$.

Case 2: Additive noise. Suppose that $G(u) := G \in \mathcal{L}_2(K, V)$, that is, the noise is additive, or more generally that G satisfies conditions **(G1)** and **(G2)** with $K_1 = 0$, that is, $\|G(u)\|_{\mathcal{L}_2(K,V)}^2 \leq K_0$. Then for any constant $\alpha > 0$ using an exponential Markov inequality and the Gagliardo–Nirenberg inequality (2.2) we deduce

$$\mathbb{P}(\tilde{\Omega}_M^c) \leq \mathbb{P}\Bigg(\sup_{t\in[0,T]} \|u(t)\|_X^2 \geq \sqrt{M}\Bigg) \leq \mathbb{P}\Bigg(\sup_{t\in[0,T]} \|u(t)\|_V^2 \geq \frac{2\sqrt{M}}{\bar{C}}\Bigg). \tag{3.27}$$

We next prove that for an additive noise, or in a slightly more general setting, for $\alpha > 0$ small enough, $\mathbb{E}[\exp(\alpha \sup_{t\in[0,T]} \|u(t)\|_V^2)] < \infty$. In the case of an additive noise this result is a particular case of Hairer & Mattingly (2006, Lemma A.1). In this reference the periodic Navier–Stokes equation is written in vorticity formulation $\xi(t) = \partial_1 u_2(t) - \partial_2 u_1(t)$. The velocity u projected on divergence-free fields can be deduced from ξ by the Biot–Savart kernel, and $|\nabla u(t)|_{\mathbb{L}^2} \leq C|\xi(t)|_{\mathbb{L}^2}$. We extend this result to the case of a more general diffusion coefficient whose Hilbert–Schmidt norm is bounded. 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 $u \in \mathrm{Dom}(A)$ then we have

$$\|u\|_V^2 = |u|_{\mathbb{L}^2}^2 + |\nabla u|_{\mathbb{L}^2}^2 \leq \tilde{C}|||u|||^2. \tag{3.28}$$

LEMMA 3.8 Let $u_0 \in V$ and let G satisfy conditions **(G1)** and **(G2)** with $K_1 = 0$, that is, $\|G(t,u)\|_{\mathcal{L}_2(K,V)}^2 \leq K_0$. Then the solution u to (1.1) satisfies

$$\mathbb{E}\Big\{\exp\Big[\alpha\Big(\sup_{0\leq s\leq T} \|u(s)\|_V^2\Big) + \nu \int_0^T |Au(s)|_{\mathbb{L}^2}^2\,\mathrm{d}s\Big)\Big]\Big\} \leq 3\exp\big(\alpha[\|u_0\|_V^2 + TK_0]\big) \tag{3.29}$$

for $\alpha \in (0, \alpha_0]$ and $\alpha_0 = \frac{\nu}{4K_0\tilde{C}}$, where $\tilde{C}$ is defined in (3.28).

In order to make this paper as self-contained as possible we prove this lemma in Appendix A.2.

Let $u_0 \in V$; then for $\alpha_0 = \frac{\nu}{4K_0\bar{C}}$ the inequalities (3.27) and (3.29) imply

$$\mathbb{P}(\tilde{\Omega}_M^c) \le 3e^{\alpha_0 K_0 T} \exp\Big(-2\alpha_0 \frac{\sqrt{M}}{\bar{C}}\Big),$$

where $\bar{C}$ is defined by (2.2). We then have to choose $M(N) \to \infty$ as $N \to \infty$ to balance the right-hand sides of (3.10) and (3.22), that is, such that for some $p > 1$,

$$\exp\Big(-\frac{2\alpha_0\sqrt{M(N)}}{p\bar{C}}T\Big) \asymp c_2 \frac{T}{N} \exp\big(\widetilde{C(M(N))}T\big)$$

for some positive constant c_2. Taking logarithms we look for $M(N)$ such that

$$-\frac{2\alpha_0}{p\bar{C}}\sqrt{M(N)} \asymp -\ln(N) + 2(1+\epsilon)C_{\nu\beta}M(N)T,$$

where $\beta \in (0,1)$, $\epsilon > 0$ and $C_{\nu\beta}$ is defined by (2.5). Set $X = \sqrt{M(N)}$, $a_2 = 2(1+\epsilon)C_{\nu\beta}T$ and $a_1 = \frac{2\alpha_0}{p\bar{C}}$. We have to solve the equation $a_2X^2 + a_1X - \ln(N) = 0$. The positive root of this polynomial is equal to $\sqrt{\frac{\ln(N)}{a_2}} + \mathcal{O}(1)$ as $N \to \infty$ and $\exp\big(-\frac{2\alpha_0\sqrt{M(N)}}{p\bar{C}}\big) \asymp C\exp\big(-2\alpha_0\frac{\sqrt{\ln(N)}}{\sqrt{a_2}p\bar{C}}\big)$. Thus, we deduce the following rate of convergence of the splitting scheme.

THEOREM 3.9 Let $u_0 \in V$ and let G satisfy conditions **(G1)** and **(G2)** with $K_1 = 0$, that is, $\|G(s,u)\|^2_{\mathcal{L}(K,V)} \le K_0$. Then

$$\mathbb{E}\Big(\sup_{t\in[0,T]}\big[|z^N(t)-u(t)|^2_{\mathbb{L}^2} + |u^N(t)-u(t)|^2_{\mathbb{L}^2}\big] \tag{3.30}$$

$$+\int_0^T \big[\|z^N(s)-u(s)\|^2_V + \|u^N(s)-u(s)\|^2_V + \|y^N(s)-u(s)\|^2_V\big]\,\mathrm{d}s\Big) \le Ce^{-\gamma\sqrt{\ln(N)}},$$

$$\mathbb{E}\Big(\sup_{k=1,\dots,N}\big[|u^N(t_k^+)-u(t_k)|^2_{\mathbb{L}^2} + |y^N(t_k^+)-u(t_k)|^2_{\mathbb{L}^2}\big]\Big) \le Ce^{-\gamma\sqrt{\ln(N)}}, \tag{3.31}$$

where

$$\gamma < \frac{\alpha_0}{\bar{C}^2}\sqrt{\frac{2^9\nu^3}{3^3T}}.$$

Note that when ν increases, the upper bound of the exponent γ increases.

REMARK 3.10 Note that the statements of Theorems 3.6 and 3.9 are valid if the diffusion coefficient G depends on the time parameter $t \in [0,T]$ and satisfies the global growth and Lipschitz versions of **(G1)** and **(G2)**. We have removed the time dependence of G to focus on the main arguments used to obtain strong convergence results.

4. Euler time schemes

4.1 *Description of the fully implicit scheme and first results*

In this section, we have to be more specific in the definition of the noise. Let $\mathcal{K}$ be a Hilbert space, Q be a trace-class operator in $\mathcal{K}$ and $W := (W(t), t \in [0,T])$ be a $\mathcal{K}$-valued Wiener process with covariance Q. Let $K = Q^{\frac{1}{2}}\mathcal{K}$ denote the Reproducing Kernel Hilbert Space (RKHS) of the Gaussian process W. Let $\mathcal{G} : \mathcal{K} \to H$ be a linear operator and suppose that analogs of conditions **(G1)** and **(G2)** are satisfied with $\mathcal{G}$ instead of G and the operator norms $\mathcal{L}(\mathcal{K}, H)$ (resp. $\mathcal{L}(\mathcal{K}, V)$) instead of the Hilbert–Schmidt norms $\mathcal{L}_2(K, H)$ (resp. $\mathcal{L}_2(\mathcal{K}, V)$), with constants $\bar{K}_i$, $i = 0, 1$ and $\bar{L}_1$. Then the diffusion coefficient $G = \mathcal{G} \circ Q^{-\frac{1}{2}}$ satisfies conditions **(G1)** and **(G2)** with constants $K_i = \mathrm{Trace}(Q)\bar{K}_i$ and $L_1 = \mathrm{Trace}(Q)\bar{L}_1$.

Let us first recall the fully implicit time discretization scheme of the stochastic 2D Navier–Stokes equations introduced by Carelli & Prohl (2012). As in the previous section let $t_k = \frac{kT}{N}$, $k = 0, \ldots, N$ denote the time grid. When studying a space–time discretization using finite elements one needs to have a stable pairing of the velocity and the pressure that satisfies the discrete Ladyzenskaja-Babuška-Brezzi (LBB)-condition (see, e.g., Carelli & Prohl, 2012, p. 2469 and pp. 2487–9). Stability issues are crucial and the pressure has to be discretized together with the velocity. In this section our aim is to obtain bounds for the strong error of an Euler time scheme. Thus, as in the previous section, we may define the scheme for the velocity projected on divergence-free fields (see Carelli & Prohl, 2012, Section 3).

Fully implicit Euler scheme. Let u_0 be a V-valued, $\mathcal{F}_0$-measurable random variable and set $u_N(t_0) = u_0$. For $k = 1, \ldots, N$, find $u_N(t_k) \in V$ such that $\mathbb{P}$ a.s. for all $\phi \in V$,

$$\begin{aligned}&\big(u_N(t_k) - u_N(t_{k-1}), \phi\big) + \frac{T}{N}\Big[\nu\big(\nabla u_N(t_k), \nabla\phi\big) + \big\langle B(u_N(t_k), u_N(t_k)), \phi\big\rangle\Big]\\ &\quad = \big(G(u_N(t_{k-1}))\, \Delta_k W, \phi\big), \qquad (4.1)\end{aligned}$$

where $\Delta_k W = W(t_k) - W(t_{k-1})$.

In the study of the Euler discretization schemes we will need some Hölder regularity of the solution. This is proved by means of semigroup theory; see Printems (2001, Proposition 3.4) and Carelli & Prohl (2012, Lemma 2.3).

Proposition 4.1 Let u_0 be $\mathcal{F}_0$-measurable such that $\mathbb{E}(\|u_0\|_V^{2p}) < \infty$ for some $p \in [2, 4]$. Let G satisfy conditions **(G1)** and **(G2)**. Then for $\eta \in (0, \frac{1}{2})$ we have

$$\mathbb{E}\big(\|u(t) - u(s)\|_{\mathbb{L}^4}^p\big) \le C\, |t - s|^{\eta p}, \qquad (4.2)$$

$$\mathbb{E}\big(\|u(t) - u(s)\|_V^p\big) \le C\, |t - s|^{\frac{\eta p}{2}}. \qquad (4.3)$$

Let us recall Carelli & Prohl (2012, Lemma 3.1), which proves moment estimates of the solution to (4.1). Note that here only dyadic moments are computed because of the induction argument that relates two consecutive dyadic numbers (see Brzeźniak *et al.*, 2013, step 4 of the proof of Lemma 3.1).

Lemma 4.2 Let u_0 be $\mathcal{F}_0$-measurable such that $\mathbb{E}\big(\|u_0\|_V^{2^q}\big) < \infty$ for some integer $q \in [2, \infty)$. Assume that G satisfies conditions **(G1)** and **(G2)**. Then there exists a $\mathbb{P}$ a.s. unique sequence of solutions $\{u_N(t_k)\}_{k=1}^N$ of (4.1), such that each random variable $u_N(t_k)$ is $\mathcal{F}_{t_k}$-measurable and satisfies

$$\sup_{N\geq 1} \mathbb{E}\Big(\max_{1\leq k\leq N} \|u_N(t_k)\|_V^{2q} + \nu \frac{T}{N} \sum_{k=1}^N \|u_N(t_k)\|_V^{2q-2} |Au_N(t_k)|_{\mathbb{L}^2}^2 \Big) \leq C(T,q), \tag{4.4}$$

where $C(T,q)$ is a constant, which depends on T, the constants K_i, $i=0,1$ in conditions **(G1)** and **(G2)** and also depends on $\mathbb{E}(\|u_0\|_V^{2q})$.

For $k=0,\ldots,N$, let $e_k := u(t_k) - u_N(t_k)$ denote the error of this scheme (note that $e_0=0$). Then for any $\phi \in V$ and $j=1,\ldots,N$ we have

$$\begin{aligned}(e_j - e_{j-1}, \phi) &+ \int_{t_{j-1}}^{t_j} \Big[\nu\big(\nabla u(s) - \nabla u_N(t_j)\,,\, \nabla\phi\big) + \big\langle B(u(s)) - B(u_N(t_j))\,,\, \phi\big\rangle \Big] \mathrm{d}s \\ &= \Big(\phi\,,\, \int_{t_{j-1}}^{t_j} \big[G(u(s)) - G(u_N(t_{j-1}))\big] \mathrm{d}W(s)\Big). \end{aligned} \tag{4.5}$$

4.2 *A localized convergence result*

The first result states localized upper bounds of the error terms. This is due to the nonlinear term, but unlike Carelli & Prohl (2012), it depends on u and not on u_N. Given $M>0$ and $k=1,\ldots,N$ set

$$\Omega_k^M := \Big\{ \omega\in\Omega \;:\; \max_{1\leq j\leq k} |\nabla u(t_j)|_{\mathbb{L}^2}^2 \leq M \Big\} \in \mathcal{F}_{t_k}. \tag{4.6}$$

The following proposition is one of the main results of this section. The modification with respect to Carelli & Prohl (2012, Theorem 3.1) is the localization set that does not depend on the approximation. This will be crucial to obtain a speed of $L^2(\Omega)$ strong convergence and not only that the scheme converges in probability.

Proposition 4.3 Let G satisfy the growth and Lipschitz conditions **(G1)** and **(G2)**. Let u_0 be such that $\mathbb{E}(\|u_0\|_V^8) < \infty$. Then for Ω_k^M defined by (4.6) and N large enough we have the following for every $k=1,\ldots,N$:

$$\mathbb{E}\Big(1_{\Omega_{k-1}^M} \max_{1\leq j\leq k} \Big[|e_j|_{\mathbb{L}^2}^2 + \nu\frac{T}{N}\sum_{j=1}^k |\nabla e_j|_{\mathbb{L}^2}^2 \Big]\Big) \leq C\, \exp\big[C_1(M)T\big] \Big(\frac{T}{N}\Big)^\eta \tag{4.7}$$

for some constant $C>0$, $\eta\in(0,\frac12)$ and

$$C_1(M) = \frac{(1+\bar\epsilon)\bar C^2}{2\nu} M + C(\bar\epsilon)\, L_1, \tag{4.8}$$

where $\bar C$ is defined in (2.2), and $\bar\epsilon$ is arbitrarily close to 0.

Proof. We follow the scheme of the arguments in Carelli & Prohl (2012, pp. 2480–4), but the upper estimate of the duality involving the difference of the bilinear terms is dealt with differently, which leads to a different localization set. Furthermore, in order to describe the strong speed of convergence of the scheme, we need a more precise control of various constants appearing in some upper estimates. Hence, we give a detailed proof below.

Step 1: Upper estimates for the bilinear term. Let us consider the duality between the difference of bilinear terms and e_j, that is, the upper estimate of $\int_{t_{j-1}}^{t_j} \langle B(u(s)) - B(u_N(t_j)), e_j\rangle \, ds$. For every $s \in (t_{j-1}, t_j]$, using the bilinearity of B and the antisymmetry property (2.1), we deduce

$$\big\langle B\big(u(s), u(s)\big) - B\big(u_N(t_j), u_N(t_j)\big), e_j\big\rangle = \sum_{i=1}^{3} T_i(s), \tag{4.9}$$

where, since $\big\langle B\big(v, u_N(t_j)\big), e_j\big\rangle = \big\langle B\big(v, u(t_j)\big), e_j\big\rangle$ for every $v \in V$,

$$\begin{aligned}
T_1(s) &:= \big\langle B\big(e_j, u_N(t_j)\big), e_j\big\rangle = \big\langle B\big(e_j, u(t_j)\big), e_j\big\rangle, \\
T_2(s) &:= \big\langle B\big(u(s) - u(t_j), u(t_j)\big), e_j\big\rangle, \\
T_3(s) &:= \big\langle B\big(u(s), u(s) - u(t_j)\big), e_j\big\rangle = -\big\langle B\big(u(s), e_j\big), u(s) - u(t_j)\big\rangle.
\end{aligned}$$

Note that, unlike the first formulation of $T_1(s)$, the second one depends only on the error and on the solution to (1.1) and not on the approximation scheme. The Hölder inequality and (2.2) yield for every $\delta_1 > 0$ and $\bar{C}$ defined in the interpolation inequality (2.2),

$$\begin{aligned}
\int_{t_{j-1}}^{t_j} |T_1(s)| \, ds &\le \bar{C} \frac{T}{N} |e_j|_{\mathbb{L}^2} \, |\nabla e_j|_{\mathbb{L}^2} \, |\nabla u(t_j)|_{\mathbb{L}^2} \\
&\le \delta_1 \nu \frac{T}{N} |\nabla e_j|^2_{\mathbb{L}^2} + \frac{\bar{C}^2}{4\delta_1 \nu} \frac{T}{N} |e_j|^2_{\mathbb{L}^2} |\nabla u(t_j)|^2_{\mathbb{L}^2},
\end{aligned}$$

where the last upper estimate follows from the Young inequality (with conjugate exponents 2 and 2). A similar argument using the Hölder and Young inequalities with exponents 4, 4 and 2 implies that for any $\delta_2 > 0$ and $\gamma_2 > 0$,

$$|T_2(s)| \le \delta_2 \nu |\nabla e_j|^2_{\mathbb{L}^2} + \gamma_2 |e_j|^2_{\mathbb{L}^2} + C(\nu, \delta_2, \gamma_2) \|u(t_j) - u(s)\|^2_{\mathbb{L}^4} |\nabla u(t_j)|^2_{\mathbb{L}^2}.$$

Using the Cauchy–Schwarz inequality we deduce

$$\begin{aligned}
\int_{t_{j-1}}^{t_j} |T_2(s)| \, ds \le\; & \delta_2 \nu \frac{T}{N} |\nabla e_j|^2_{\mathbb{L}^2} + \gamma_2 \frac{T}{N} |e_j|^2_{\mathbb{L}^2} \\
&+ C(\nu, \delta_2, \gamma_2) |\nabla u(t_j)|^2_{\mathbb{L}^2} \int_{t_{j-1}}^{t_j} \|u(t_j) - u(s)\|^2_{\mathbb{L}^4} \, ds.
\end{aligned}$$

Similar computations using the Hölder and Young inequalities imply

$$\int_{t_{j-1}}^{t_j} |T_3(s)| \, ds \le \delta_3 \nu \frac{T}{N} |\nabla e_j|^2_{\mathbb{L}^2} + \frac{1}{4\nu\delta_3} \int_{t_{j-1}}^{t_j} \|u(s)\|^2_{\mathbb{L}^4} \, \|u(t_j) - u(s)\|^2_{\mathbb{L}^4} \, ds \tag{4.10}$$

for any $\delta_3 > 0$. Note that

$$\nu \int_{t_{j-1}}^{t_j} \big(\nabla(u(s) - u_N(t_j))\,,\ \nabla e_j\big)\,\mathrm{d}s = \nu \frac{T}{N} |\nabla e_j|_{\mathbb{L}^2}^2 + \nu \int_{t_{j-1}}^{t_j} \big(\nabla(u(s) - u(t_j)), \nabla e_j\big)\,\mathrm{d}s.$$

Using the Cauchy–Schwarz and Young inequalities we deduce

$$\nu \int_{t_{j-1}}^{t_j} \big|\big(\nabla(u(s) - u(t_j)), \nabla e_j\big)\big|\,\mathrm{d}s \le \delta_0 \nu \frac{T}{N} |\nabla e_j|_{\mathbb{L}^2}^2 + \frac{\nu}{4\delta_0} \int_{t_{j-1}}^{t_j} \big|\nabla\big(u(s) - u(t_j)\big)\big|_{\mathbb{L}^2}^2\,\mathrm{d}s$$

for any $\delta_0 > 0$. Hence, using the above upper estimates in (4.5) with $\phi = e_j$, we deduce

$$\begin{aligned}\big(e_j - e_{j-1}\,, e_j\big) + \nu \frac{T}{N} |\nabla e_j|_{\mathbb{L}^2}^2 \le{}& \nu \sum_{r=0}^{3} \delta_r \frac{T}{N} |\nabla e_j|_{\mathbb{L}^2}^2 + \Big(\gamma_2 + \frac{\bar{C}^2}{4\delta_1 \nu} |\nabla u(t_j)|_{\mathbb{L}^2}^2\Big) \frac{T}{N} |e_j|_{\mathbb{L}^2}^2 \\ &+ \sum_{l=1}^{3} \tilde{T}_j(l) + \Big(\int_{t_{j-1}}^{t_j} \big[G(u(s)) - G(u_N(t_{j-1}))\big]\,\mathrm{d}W(s)\,,\ e_j\Big),\end{aligned} \tag{4.11}$$

where

$$\begin{aligned}\tilde{T}_j(1) ={}& \frac{\nu}{4\delta_0} \int_{t_{j-1}}^{t_j} \big|\nabla\big(u(s) - u(t_j)\big)\big|_{\mathbb{L}^2}^2\,\mathrm{d}s,\\ \tilde{T}_j(2) ={}& C(\nu, \delta_2, \gamma_2)\,|\nabla u(t_j)|_{\mathbb{L}^2}^2 \int_{t_{j-1}}^{t_j} \|u(s) - u(t_j)\|_{\mathbb{L}^4}^2\,\mathrm{d}s,\\ \tilde{T}_j(3) ={}& \frac{1}{4\nu\delta_3} \int_{t_{j-1}}^{t_j} \|u(s)\|_{\mathbb{L}^4}^2 \|u(t_j) - u(s)\|_{\mathbb{L}^4}^2\,\mathrm{d}s.\end{aligned}$$

Using the time regularity (4.3) with $p = 2$ we deduce

$$\mathbb{E}\big(\tilde{T}_j(1)\big) \le C \frac{\nu}{4\delta_0} \Big(\frac{T}{N}\Big)^{1+\eta}. \tag{4.12}$$

The Cauchy–Schwarz inequality, (2.11) with $p = 2$ and (4.2) imply

$$\mathbb{E}\big(\tilde{T}_j(2)\big) \le\ C(\nu, \delta_2, \gamma_2) \Big(\frac{T}{N}\Big)^{1+2\eta}, \tag{4.13}$$

$$\mathbb{E}\big(\tilde{T}_j(3)\big) \le C\, \frac{1}{4\nu\delta_3} \Big(\frac{T}{N}\Big)^{1+2\eta}. \tag{4.14}$$

Step 2: Localization. In order to use a discrete version of the Gronwall lemma to find an upper estimate for $|e_j|^2_{\mathbb{L}^2}$, due to the factor $|\nabla u(t_j)|^2_{\mathbb{L}^2}$ on the right-hand side of (4.11), we have to localize on the random set Ω^M_{j-1} defined in (4.6). The shift of index is due to the fact that, in order to deal with the stochastic integral, we have to make sure that the localization set is $\mathcal{F}_{t_{j-1}}$-measurable. This set depends on j, but we will need to add the localized inequalities (4.11) and take expected values.

Note that for $1 \le j \le k$, $\Omega^M_k \subset \Omega^M_j$. Hence, since $e_0 = 0$, as proved in Carelli & Prohl (2012, estimate (3.25)), we have

$$\max_{1\le j\le k} \sum_{l=1}^{j} 1_{\Omega^M_{l-1}} \Big(|e_l|^2_{\mathbb{L}^2} - |e_{l-1}|^2_{\mathbb{L}^2}\Big) = \max_{1\le j\le k} \Big(1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2} + \sum_{l=2}^{j} \big(1_{\Omega^M_{l-2}} - 1_{\Omega^M_{l-1}}\big) |e_{l-1}|^2_{\mathbb{L}^2}\Big)$$
$$\ge \max_{1\le j\le k} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}. \tag{4.15}$$

Thus, we will localize $|e_j|^2_{\mathbb{L}^2}$ on the set Ω^M_{j-1} and—shifting the index by 1—control some 'error term' $|e_j - e_{j-1}|^2_{\mathbb{L}^2}$ localized on the same set. Note that this localization se depends only on the projection of the solution u of equation (1.1) on divergence-free fields and not on its approximation.

Adding the inequalities (4.11) with $\phi = e_j$ localized on the set Ω^M_{j-1}, using $e_0 = 0$ and the identity $(a, a-b) = \frac{1}{2}\big[|a|^2_{\mathbb{L}^2} - |b|^2_{\mathbb{L}^2} + |a-b|^2_{\mathbb{L}^2}\big]$, we deduce for $k = 1, \dots, N$,

$$\max_{1\le j\le k} \Big(\frac{1}{2} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2} + \frac{1}{2} \sum_{l=1}^{j} 1_{\Omega^M_{l-1}} |e_l - e_{l-1}|^2_{\mathbb{L}^2}\Big)$$
$$\le \frac{1}{2}\Big(\max_{1\le j\le k} \sum_{l=1}^{j} 1_{\Omega^M_{l-1}} \Big(|e_l|^2_{\mathbb{L}^2} - |e_{l-1}|^2_{\mathbb{L}^2}\Big) + \sum_{l=1}^{j} 1_{\Omega^M_{l-1}} |e_l - e_{l-1}|^2_{\mathbb{L}^2}\Big)$$
$$\le \max_{1\le j\le k} \sum_{1\le l\le j} 1_{\Omega^M_{l-1}} \big(e_l - e_{l-1}, e_l\big).$$

The upper estimates (4.11) for $j = 1, \dots, k$ imply for any $\epsilon > 0$,

$$\max_{1\le j\le k} \Big[\frac{1}{2} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2} + \sum_{l=1}^{j} 1_{\Omega^M_{l-1}} |e_l - e_{l-1}|^2_{\mathbb{L}^2} + \nu\Big(1 - \sum_{r=0}^{3} \delta_r\Big) \frac{T}{N} \sum_{l=1}^{j} 1_{\Omega^M_{l-1}} |\nabla e_l|^2_{\mathbb{L}^2}\Big]$$
$$\le \Big[\gamma_2 + (1+\epsilon) \frac{\bar{C}^2 M}{4\delta_1 \nu}\Big] \frac{T}{N} \sum_{j=1}^{k} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2} + \sum_{i=1}^{3} \sum_{j=1}^{k} \tilde{T}_j(i)$$
$$+ C(\nu, \delta_1, \epsilon) \frac{T}{N} \sum_{j=1}^{k} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2} \Big|\nabla\big[u(t_j) - u(t_{j-1})\big]\Big|^2_{\mathbb{L}^2} + M_k(1) + M_k(2), \tag{4.16}$$

where

$$M_k(1) = \sum_{j=1}^{k} 1_{\Omega_{j-1}^M} \Big(e_{j-1}\,,\, \int_{t_{j-1}}^{t_j} \big[G(u(s)) - G(u_N(t_{j-1})) \big]\, \mathrm{d}W(s) \Big),$$

$$M_k(2) = \sum_{j=1}^{k} 1_{\Omega_{j-1}^M} \Big(e_j - e_{j-1}\,,\, \int_{t_{j-1}}^{t_j} \big[G(u(s)) - G(u_N(t_{j-1})) \big]\, \mathrm{d}W(s) \Big).$$

The inequalities (4.12–4.14) imply the existence of a constant C depending on T, ν, δ_i, $i = 0, \ldots, 3$ and γ_2 such that

$$\sum_{i=1}^{3} \sum_{j=1}^{N} \mathbb{E}\big(\tilde{T}_j(i)\big) \le C \left(\frac{T}{N}\right)^{\eta}. \tag{4.17}$$

The Cauchy–Schwarz inequality, (2.11) and (4.4) for $p = q = 2$ and the time regularity (4.3) for $p = 4$ imply the existence of a constant C such that

$$\frac{T}{N} \sum_{j=1}^{N} \mathbb{E}\Big(|e_j|_{\mathbb{L}^2}^2 |\nabla\big[u(t_j) - u(t_{j-1})\big]|_{\mathbb{L}^2}^2 \Big) \le C\Big(\frac{T}{N}\Big)^{\eta}. \tag{4.18}$$

We next find an upper estimate for $\mathbb{E}\big(\max_{1\le k\le N} M_k(2)\big)$. The Cauchy–Schwarz inequality, the Itô isometry and then the Young inequality imply that for any $\tilde{\delta}_2 > 0$,

$$\begin{aligned}
\mathbb{E}\Big(\max_{1\le j\le k} M_j(2)\Big) \le{}& \sum_{j=1}^{k} \Big\{ \mathbb{E}\big(1_{\Omega_{j-1}^M} |e_j - e_{j-1}|_{\mathbb{L}^2}^2 \big) \Big\}^{\frac{1}{2}} \\
&\times \Big\{ \mathbb{E}\Big(1_{\Omega_{j-1}^M} \int_{t_{j-1}}^{t_j} \| G(u(s)) - G(u_N(t_{j-1}) \|_{\mathcal{L}_2(K,H)}^2 \, \mathrm{d}s \Big) \Big\}^{\frac{1}{2}} \\
\le{}& \tilde{\delta}_2 \sum_{j=1}^{k} \mathbb{E}\big(1_{\Omega_{j-1}^M} |e_j - e_{j-1}|_{\mathbb{L}^2}^2 \big) \\
&+ \frac{1}{4\tilde{\delta}_2} \sum_{j=1}^{k} \int_{t_{j-1}}^{t_j} \mathbb{E}\Big(1_{\Omega_{j-1}^M} \| G(u(s)) - G(u_N(t_{j-1})) \|_{\mathcal{L}_2(K,H)}^2 \Big)\, \mathrm{d}s.
\end{aligned}$$

The Lipschitz condition **(G1)** and (4.3) imply, for any $\epsilon > 0$ and $s \in [t_{l-1}, t_l]$,

$$\begin{aligned}
&\mathbb{E}\big(1_{\Omega^M_{l-1}} \|G(u(s)) - G(u_N(t_{l-1}))\|^2_{\mathcal{L}_2(K,H)}\big) \\
&\quad \le (1+\epsilon)\mathbb{E}\big(1_{\Omega^M_{l-1}} \|G(u(t_{l-1})) - G(u_N(t_{l-1}))\|^2_{\mathcal{L}_2(K,H)}\big) \\
&\qquad + \Big(1+\frac{1}{\epsilon}\Big)\mathbb{E}\big(1_{\Omega^M_{l-1}} \|G(u(s)) - G(u(t_{l-1}))\|^2_{\mathcal{L}_2(K,H)}\big) \\
&\quad \le L_1\,(1+\epsilon)\,\mathbb{E}\Big(1_{\Omega^M_{l-1}} |e_{l-1}|^2_{\mathbb{L}^2}\Big) + C(\epsilon)\Big(\frac{T}{N}\Big)^\eta.
\end{aligned} \tag{4.19}$$

Since $\Omega^M_j \subset \Omega^M_{j-1}$ and $e_0 = 0$, we deduce for any $k = 2, \ldots, N$,

$$\begin{aligned}
\mathbb{E}\Big(\max_{1\le j\le k} M_j(2)\Big) \le{}& \tilde{\delta}_2 \frac{T}{N}\sum_{j=1}^k \mathbb{E}\big(1_{\Omega^M_{j-1}} |e_j - e_{j-1}|^2_{\mathbb{L}^2}\big) + \frac{1+\epsilon}{4\tilde{\delta}_2} L_1 \frac{T}{N}\sum_{j=1}^{k-1}\mathbb{E}\Big(1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\Big) \\
&+ C(\epsilon, \tilde{\delta}_2) T \Big(\frac{T}{N}\Big)^\eta.
\end{aligned} \tag{4.20}$$

Since $1_{\Omega^M_{l-1}}$ and e_{l-1} are $\mathcal{F}_{t_{l-1}}$-measurable, using the Burkholder–Davies–Gundy inequality, the Young inequality, (4.19) and using once more the inclusion $\Omega^M_j \subset \Omega^M_{j-1}$, we deduce that for any $\tilde{\delta}_1 > 0$,

$$\begin{aligned}
\mathbb{E}\Big(\max_{1\le j\le k} M_j(1)\Big) &\le 3\sum_{l=1}^k \mathbb{E}\Big[\Big\{1_{\Omega^M_{l-1}} \int_{t_{l-1}}^{t_l} \|G(u(s)) - G(u_N(t_{l-1}))\|^2_{\mathcal{L}_2(K,H)}\, |e_{l-1}|^2_{\mathbb{L}^2}\, \mathrm{d}s\Big\}^{\frac12}\Big] \\
&\le 3\mathbb{E}\Big[\Big(\max_{1\le l\le k} 1_{\Omega^M_{l-1}} |e_{l-1}|_{\mathbb{L}^2}\Big)\Big\{\sum_{l=1}^k 1_{\Omega^M_{l-1}} \int_{t_{l-1}}^{t_l} \|G(u(s)) - G(u_N(t_{l-1}))\|^2_{\mathcal{L}_2(K,H)}\, \mathrm{d}s\Big\}^{\frac12}\Big] \\
&\le \tilde{\delta}_1 \mathbb{E}\Big(\max_{1\le l\le k} 1_{\Omega^M_{l-1}} |e_{l-1}|^2_{\mathbb{L}^2}\Big) + \frac{9(1+\epsilon)}{4\tilde{\delta}_1} L_1 \frac{T}{N}\sum_{j=1}^{k-1} \mathbb{E}\Big(1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\Big) + C(\epsilon, \tilde{\delta}_1)\Big(\frac{T}{N}\Big)^\eta.
\end{aligned} \tag{4.21}$$

Collecting the upper estimates (4.16–4.21) and taking $\tilde{\delta}_2 = 1$ we deduce for $k = 2, \ldots, N$,

$$\begin{aligned}
&\mathbb{E}\bigg(\max_{1\le j\le k}\Big\{\frac12 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2} + \nu\Big(1 - \sum_{i=0}^3 \delta_i\Big)\frac{T}{N}\sum_{l=1}^j 1_{\Omega^M_{l-1}} |\nabla e_l|^2_{\mathbb{L}^2}\Big\}\bigg) \\
&\quad \le \Big[\tilde{\delta}_1 + \gamma_2 + (1+\epsilon)\frac{\bar{C}^2 M}{4\delta_1 \nu}\frac{T}{N}\Big]\mathbb{E}\Big(\max_{1\le j\le k} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\Big) \\
&\qquad + \Big[(1+\epsilon)\frac{\bar{C}^2 M}{4\delta_1 \nu} + \gamma_2 + \frac{1+\epsilon}{4}\Big(\frac{1}{\tilde{\delta}_2} + \frac{9}{\tilde{\delta}_1}\Big)L_1\Big]\frac{T}{N}\sum_{j=1}^{k-1}\mathbb{E}\big(1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\big) + C\Big(\frac{T}{N}\Big)^\eta.
\end{aligned} \tag{4.22}$$

Step 3: Discrete Gronwall lemma. Fix $\alpha \in (0,1)$ and choose $\delta_1 \in (0, 1-\alpha)$. Then

$$\nu(1-\delta_1) \geq \alpha\nu.$$

Fix $\tilde{\epsilon} \in (0,1)$ and $\gamma_2 \in \big(0, \frac{\tilde{\epsilon}}{2}(\frac{1}{2} - \tilde{\delta}_1)\big)$; suppose that N is large enough to imply

$$\frac{1}{2} - \tilde{\delta}_1 - \Big(\gamma_2 + (1+\epsilon)\frac{\bar{C}^2 M}{4\delta_1 \nu}\Big)\frac{T}{N} \geq (1-\tilde{\epsilon})\left(\frac{1}{2} - \tilde{\delta}_1\right),$$

and choose δ_i, $i = 0, 2, 3$ such that $\delta_0 + \delta_2 + \delta_3 < \frac{\alpha}{2}\nu$. Then for N large enough,

$$(1-\tilde{\epsilon})\left(\frac{1}{2} - \tilde{\delta}_1\right)\mathbb{E}\Big(\max_{1\leq j\leq k} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\Big) + \frac{\alpha\nu}{2}\frac{T}{N}\sum_{j=1}^{k} \mathbb{E}\big(1_{\Omega^M_{j-1}} |\nabla e_j|^2_{\mathbb{L}^2}\big)$$
$$\leq \Big[(1+\epsilon)\frac{\bar{C}^2 M}{4\delta_1 \nu} + C(\gamma_2, \tilde{\delta}_1, \epsilon)L_1\Big]\frac{T}{N}\sum_{j=1}^{k-1} \mathbb{E}\big(1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\big) + C\Big(\frac{T}{N}\Big)^{\eta}.$$

Set

$$C_1(M) := \frac{\frac{(1+\epsilon)\bar{C}^2}{4\delta_1 \nu} M + (1+\epsilon)C(\gamma_2, \tilde{\delta}_1)}{(1-\tilde{\epsilon})\big(\frac{1}{2} - \tilde{\delta}_1\big)}. \tag{4.23}$$

Neglecting the second term on the left-hand side and using a discrete version of the Gronwall lemma we deduce that

$$\mathbb{E}\Big(\max_{1\leq j\leq k} 1_{\Omega^M_{j-1}} |e_j|^2_{\mathbb{L}^2}\Big) \leq C\, e^{C_1(M)\,T}\left(\frac{T}{N}\right)^{\eta}.$$

Once these inequalities hold, choosing $\epsilon \sim 0$, $\tilde{\delta}_1 \sim 0$, $\gamma_2 \sim 0$, $\delta_1 \sim 1$, $\delta_i \sim 0$ for $i = 0, 2, 3$, we may take $C_1(M)$ such that

$$C_1(M) = \frac{(1+\bar{\epsilon})\,\bar{C}^2 M}{2\nu} + C(\bar{\epsilon})L_1,$$

where $\bar{C}$ is the constant defined in (2.2) and $\bar{\epsilon} > 0$ is arbitrarily close to 0. Indeed, given $\bar{\epsilon} > 0$, we choose the other constants so that $\frac{1+\epsilon}{\delta_1(2-4\tilde{\delta}_1)} \leq \frac{1+\bar{\epsilon}}{2}$. Plugging this into the previous upper estimate and using the inclusions $\Omega^M_k \subset \Omega^M_j$ for $j = 1, \ldots, k$ we deduce (4.7) and (4.8). □

4.3 *Strong speed of convergence of the implicit Euler scheme*

As in Section 3.3 let us use the Hölder inequality with conjugate exponents 2^{q-1} and $p = \frac{2^{q-1}}{2^{q-1}-1}$. We obtain

$$\mathbb{E}\Big(1_{(\Omega^M_N)^c} \max_{1\leq k\leq N} |e_k|^2_{\mathbb{L}^2}\Big) \leq C\Big[\mathbb{P}\big((\Omega^M_N)^c\big)\Big]^{\frac{1}{p}}$$
$$\times \Big[\mathbb{E}\Big(\sup_{0\leq s\leq T} |u(s)|^{2^q}_{\mathbb{L}^2} + \max_{0\leq k\leq N} |u_N(t_k)|^{2^q}_{\mathbb{L}^2}\Big)\Big]^{\frac{1}{2^{q-1}}}, \tag{4.24}$$

$$\mathbb{E}\Big(1_{(\Omega_N^M)^c}\frac{T}{N}\sum_{k=1}^{N}|\nabla e_k|^2_{\mathbb{L}^2}\Big) \le C\Big[\mathbb{P}\big((\Omega_N^M)^c\big)\Big]^{\frac{1}{p}}$$
$$\times\Big[\mathbb{E}\Big(\sup_{0\le s\le T}|\nabla u(s)|^{2q}_{\mathbb{L}^2}+\max_{0\le k\le N}|\nabla u_N(t_k)|^{2q}_{\mathbb{L}^2}\Big)\Big]^{\frac{1}{2^{q-1}}}. \tag{4.25}$$

The inequalities (2.11) and (4.4) prove that if $\mathbb{E}(\|u_0\|_V^{2^q})<\infty$ the second factors on the right-hand sides of (4.24) and (4.25) are bounded by a constant independent of N.We now find an upper estimate for the probability of the complement of the localization set and balance the upper estimates of the L^2 moments localized on the set Ω_N^M and its complement. Obtaining a strong speed of convergence will require the threshold M to depend on N. Two cases are studied.

Case 1: Linear growth diffusion coefficient. Suppose that G satisfies conditions **(G1)** and **(G2)** and that $\mathbb{E}(\|u_0\|^{2^q})<\infty$. Then (2.11) implies

$$\mathbb{P}\Big(\big(\Omega_N^{M(N)}\big)^c\Big) \le \mathbb{P}\Big(\sup_{0\le s\le T}|\nabla u(s)|^2_{\mathbb{L}^2}>M(n)\Big)$$
$$\le\Big(\frac{1}{M(N)}\Big)^{2^{q-1}}\mathbb{E}\Big(\sup_{0\le s\le T}\|u(s)\|_V^{2^q}\Big)\le C_q\,M(N)^{-2^{q-1}}. \tag{4.26}$$

If we suppose that $\mathbb{E}(\|u_0\|_V^{2^q})<\infty$, in order to balance the upper estimates (4.24), (4.25) with (4.26) and (4.7), we have to choose $M(N)\to\infty$ as $N\to\infty$, such that as $N\to\infty$,

$$\Big(\frac{T}{N}\Big)^{\eta}\exp[C_1\big(M(N)\big)\,T]\ \asymp\ C(q)M(N)^{-2^{q-1}+1},$$

where $C_1(M(N))$ is defined in (4.8). Fix $\bar\epsilon>0$; taking logarithms and neglecting constants leads to

$$-\eta\ln(N)+\frac{(1+\bar\epsilon)\bar C^2M(N)T}{2\nu}\asymp-\big(2^{q-1}-1\big)\ln(M(N))+\mathcal{O}(1)\quad\text{as } N\to\infty.$$

Let

$$M(N):=\frac{2\nu}{(1+\bar\epsilon)\bar C^2T}\Big\{\eta\ln(N)-\big(2^{q-1}-1\big)\ln\big(\ln(N)\big)\Big\}\asymp\frac{2\nu\eta\ln(N)}{(1+\bar\epsilon)\bar C^2T}. \tag{4.27}$$

Then for this choice of $M(N)$, we have

$$-\eta\ln(N)+C_1\big(M(N)\big)T=-\ln\Big[\big(\ln(N)\big)^{2^{q-1}-1}\Big]+\mathcal{O}(1),$$

which implies $\big(\frac{T}{N}\big)^{\eta}\exp[C_1(M(N))T]\asymp C\big(\ln(N)\big)^{-2^{q-1}+1}$ for some positive constant C. Furthermore, $M(N)^{-2^{q-1}+1}\asymp C\big(\ln(N)\big)^{-2^{q-1}+1}$ for some positive constant C. Similar computations with the sum of the V-norms of the error on the time grid yield

$$\mathbb{E}\Big(\max_{1\le k\le N}|e_k|^2_{\mathbb{L}^2}+\nu\frac{T}{N}\sum_{k=1}^N|\nabla e_k|^2_{\mathbb{L}^2}\Big)\le C\big(\ln(N)\big)^{-(2^{q-1}-1)}$$

for some constant C depending on T, q and the coefficients K_i, $i=0,1$. This completes the proof of the following.

THEOREM 4.4 Let u_0 be such that $\mathbb{E}(\|u_0\|_V^{2q})<\infty$ for some $q\ge 3$ and let G satisfy assumptions **(G1)** and **(G2)**. Then the fully implicit scheme solution u_N of (4.1) converges in $L^2(\Omega)$ to the solution u of (1.1). More precisely, for N large enough we have

$$\mathbb{E}\Big(\max_{1\le k\le N}|u(t_k)-u_N(t_k)|^2_{\mathbb{L}^2}+\frac{T}{N}\sum_{k=1}^N\big|\nabla\big[u(t_k)-u_N(t_k)\big]\big|^2_{\mathbb{L}^2}\Big)\le C\big[\ln(N)\big]^{-(2^{q-1}-1)}. \tag{4.28}$$

REMARK 4.5 Note that, as for the splitting scheme, if u_0 is a deterministic element of V and G satisfies conditions **(G1)** and **(G2)**, we have

$$\mathbb{E}\Big(\max_{1\le k\le N}|u(t_k)-u_N(t_k)|^2_{\mathbb{L}^2}+\frac{T}{N}\sum_{k=1}^N\big|\nabla\big[u(t_k)-u_N(t_k)\big]\big|^2_{\mathbb{L}^2}\Big)\le C\big[\ln(N)\big]^{-\gamma}$$

for any $\gamma>0$. This upper estimate is also true if $\|u_0\|_V$ has moments of all orders, for example, if u_0 is a V-valued Gaussian random variable independent of the noise W.

Case 2: Additive noise. Suppose that $G(u):=G\in\mathcal{L}_2(K,V)$, that is, the noise is additive, or more generally that conditions **(G1)** and **(G2)** are satisfied with $K_1=0$. Using an exponential Markov inequality we deduce that for any constant $\alpha>0$,

$$\mathbb{P}\Big(\big(\Omega_N^{M(N)}\big)^c\Big)\le\exp\big(-\alpha M(N)\big)\,\mathbb{E}\Big[\exp\Big(\alpha\sup_{0\le t\le T}|\nabla u(t)|^2_{\mathbb{L}^2}\Big)\Big]. \tag{4.29}$$

Recall that Lemma 3.8 implies that for $\alpha\in(0,\alpha_0]$, where $\alpha_0=\frac{\nu}{4K_0\tilde{C}}$ and $\tilde{C}$ is defined in (3.28), we have $\mathbb{E}\big[\sup_{t\in[0,T]}\exp(\alpha\|u(t)\|_V^2)\big]<\infty$. Using (4.7) with (4.8), (4.29) and (3.29) we choose $M(N)$ such that

$$\Big(\frac{T}{N}\Big)^{\eta}\exp\Big(\frac{(1+\bar{\epsilon})\,\bar{C}^2M(N)T}{2}\nu\Big)=c_2\exp\Big(-\frac{\nu M(N)}{p4K_0\tilde{C}}\Big) \tag{4.30}$$

for some $p>1$, $\bar{\epsilon}>0$ and some positive constant c_2, where $\bar{C}$ (resp. $\tilde{C}$) is defined by (2.2) (resp. (3.28)). For any $p\in(1,\infty)$ since u_0 is deterministic, $\mathbb{E}(\|u_0\|_V^q)<\infty$ for conjugate exponents p and q. Set

$$M(N):=\frac{\eta\ln(N)}{\frac{\nu}{p4K_0\tilde{C}}+\frac{(1+\bar{\epsilon})\bar{C}^2T}{2\nu}}$$

for some $\bar{\epsilon} > 0$. Then $M(N) \to \infty$ as $n \to \infty$, and both sides of (4.30) are equal to some constant multiple of $N^{-\beta\eta}$, where choosing p close enough to 1 and $\bar{\epsilon} \sim 0$, we have $\beta < \frac{\frac{\nu}{4K_0\bar{C}}}{\left(\frac{\nu}{4K_0\bar{C}} + \frac{\tilde{C}^2 T}{2\nu}\right)}$. Since $\eta < \frac{1}{2}$ can be chosen as close to $\frac{1}{2}$ as wanted this yields the following rate of convergence.

THEOREM 4.6 Let $u_0 \in V$ and let G satisfy assumptions **(G1)** and **(G2)** with $K_1 = L_1 = 0$. Let u denote the solution of (1.1) and u_N be the fully implicit scheme solution of (4.1). Then for N large enough, and $\bar{C}$ (resp. $\tilde{C}$) defined by (2.2) (resp. (3.28)),

$$\mathbb{E}\Big(\max_{1\le k\le N} |u(t_k) - u_N(t_k)|^2_{\mathbb{L}^2} + \frac{T}{N} \sum_{k=1}^{N} \big|\nabla\big[u(t_k) - u_N(t_k)\big]\big|^2_{\mathbb{L}^2}\Big) \le C\Big(\frac{T}{N}\Big)^{\gamma}, \tag{4.31}$$

where $\gamma < \frac{1}{2}\left(\frac{\frac{\nu}{4K_0\bar{C}}}{\frac{\nu}{4K_0\bar{C}} + \frac{\tilde{C}^2 T}{2\nu}}\right)$.

Note that if ν is large the speed of convergence of the H- and V-norms in Theorem 4.6 is 'close' to $C(T)N^{-\frac{1}{2}}$. Intuitively, it cannot be better because of the stochastic integral and the scaling between the time and space parameters in the heat kernel, which is behind the time regularity of the solution stated in (4.3).

4.4 *Semi-implicit Euler scheme*

In this section we prove the strong $L^2(\Omega)$ convergence of a discretization scheme with a linearized drift. Let v_N be defined on the time grid $(t_k, k = 0, \dots, N)$ as follows.

Semi-implicit Euler scheme. Let u_0 be a V-valued, $\mathcal{F}_0$-measurable random variable and set $v_N(0) = u_0$. For $k = 1, \dots, N$, let $v_N(t) \in V$ be such that $\mathbb{P}$ a.s. for all $\phi \in V$,

$$\begin{aligned}\big(v_N(t_k) - v_N(t_{k-1})\,,\,\phi\big) + \frac{T}{N}\Big[\nu\big(\nabla v_N(t_k)\,,\,\nabla\phi\big) + \big\langle B(v_N(t_{k-1}), v_N(t_k))\,,\,\phi\big\rangle \\ = \big(G(v_N(t_{k-1}))\,\Delta_k W\,,\,\phi\big),\end{aligned} \tag{4.32}$$

where $\Delta_k W = W(t_k) - W(t_{k-1})$.

Note that since in general $\langle B(u,v)\,,\,Av\rangle \neq 0$ for $u, v \in \mathrm{Dom}(A)$, the moments of v_N are bounded in a weaker norm than those of the fully implicit scheme u_N.

LEMMA 4.7 Let $u_0 \in L^{2^q}(\Omega, V)$ for some integer $q \ge 2$ be $\mathcal{F}_0$-measurable and let G satisfy condition **(G1)**. Then each random variable $v_N(t_k)$, $k = 0, \dots, N$ is $\mathcal{F}_{t_k}$-measurable such that

$$\sup_N \mathbb{E}\Big(\max_{1\le k\le N} |v_N(t_k)|^{2^q}_{\mathbb{L}^2} + \nu\frac{T}{N}\sum_{k=1}^{N} |v_N(t_k)|^{2^{q-1}}_{\mathbb{L}^2}\, \|v_N(t_k)\|^2_V\Big) \le C(T,q). \tag{4.33}$$

For $k = 0, \dots, N$ set $\bar{e}_k = u(t_k) - v_N(t_k)$. Unlike Carelli & Prohl (2012), we will not compare the schemes u_N and v_N since the norm of the difference would require a localization in terms of the gradient of u_N. Instead of that, we prove the following analog of Proposition 4.3.

PROPOSITION 4.8 Let G satisfy the growth and Lipschitz conditions **(G1)** and **(G2)** and u_0 be $\mathcal{F}_0$-measurable such that $\mathbb{E}(\|u_0\|_V^8) < \infty$. Then for Ω_k^M defined by (4.6) and N large enough we have for $k = 1, \dots, N$ and $\eta < \frac{1}{2}$

$$\mathbb{E}\Big(1_{\Omega_{k-1}^M}\Big[\max_{1\le j\le k}|\bar{e}_j|_{\mathbb{L}^2}^2 + \nu\frac{T}{N}\sum_{j=1}^{k}|\nabla\bar{e}_j|_{\mathbb{L}^2}^2\Big]\Big) \le C\Big(\frac{T}{N}\Big)^{\eta}\exp\big[C_1(M)T\big], \tag{4.34}$$

where $C > 0$ is some constant and $C_1(M)$ is defined by (4.8) for any $\bar{\epsilon} > 0$.

Proof. Many parts of the argument are similar to the corresponding ones in the proof of Proposition 4.3; we focus on the differences only.

We first consider the duality between the difference of bilinear terms and $\bar{e}_j$, that is, we find an upper estimate for $\int_{t_{j-1}}^{t_j}\langle B(u(s),u(s)) - B(v_N(t_{j-1}),v_N(t_j))\,,\,\bar{e}_j\rangle\,\mathrm{d}s$. For every $s\in[t_{j-1},t_j]$ using the bilinearity and antisymmetry of B we deduce

$$\langle B(u(s),u(s)) - B(v_N(t_{j-1}),v_N(t_j))\,,\,\bar{e}_j\rangle = \sum_{i=1}^{3}\bar{T}_i(s),$$

where

$$\begin{aligned}
\bar{T}_1(s) &:= \big\langle B\big(\bar{e}_{j-1}, v_N(t_j)\big)\,,\,\bar{e}_j\big\rangle = \big\langle B\big(\bar{e}_{j-1}, u(t_j)\big)\,,\,\bar{e}_j\big\rangle,\\
\bar{T}_2(s) &:= \big\langle B\big(u(s) - u(t_{j-1}), u(t_j)\big)\,,\,\bar{e}_j\big\rangle,\\
\bar{T}_3(s) &:= \big\langle B\big(u(s), u(s) - u(t_j)\big)\,,\,\bar{e}_j\big\rangle = -\big\langle B(u(s),\bar{e}_j)\,,\,u(s) - u(t_j)\big\rangle.
\end{aligned}$$

Using the Hölder inequality, (2.2) and the Young inequality, we deduce that for every $\delta_1 > 0$,

$$\begin{aligned}
\int_{t_{j-1}}^{t_j}|\bar{T}_1(s)|\,\mathrm{d}s \le\,& \bar{C}\frac{T}{N}|\bar{e}_{j-1}|_{\mathbb{L}^2}^{\frac12}|\bar{e}_j|_{\mathbb{L}^2}^{\frac12}|\nabla\bar{e}_{j-1}|_{\mathbb{L}^2}^{\frac12}|\nabla\bar{e}_j|_{\mathbb{L}^2}^{\frac12}|\nabla u(t_j)|_{\mathbb{L}^2}\\
\le\,& \frac{\delta_1}{2}\nu\frac{T}{N}|\nabla\bar{e}_{j-1}|_{\mathbb{L}^2}^2 + \frac{\delta_1}{2}\nu\frac{T}{N}|\nabla\bar{e}_j|_{\mathbb{L}^2}^2\\
&+ \frac{\delta_1\bar{C}^2}{8\delta_1\nu}\frac{T}{N}|\bar{e}_{j-1}|_{\mathbb{L}^2}^2|\nabla u(t_j)|_{\mathbb{L}^2}^2 + \frac{\delta_1\bar{C}^2}{8\delta_1\nu}\frac{T}{N}|\bar{e}_j|_{\mathbb{L}^2}^2|\nabla u(t_j)|_{\mathbb{L}^2}^2.
\end{aligned}\tag{4.35}$$

The upper estimates of $\int_{t_{j-1}}^{t_j}\bar{T}_i(s)\,\mathrm{d}s$, $i = 2,3$ are similar to the corresponding ones in the first step of the proof of Theorem 4.3. This yields the following analog of (4.11) with the same upper estimates (4.12–4.14) of the terms $\tilde{T}_j(i)$, $i = 1,2,3$:

$$\begin{aligned}
(\bar{e}_j - \bar{e}_{j-1}\,,\,\bar{e}_j) + \nu|\nabla\bar{e}_j|_{\mathbb{L}^2}^2 \le\,& \nu\left(\delta_0 + \frac12\delta_1 + \delta_2 + \delta_3\right)\frac{T}{N}|\nabla\bar{e}_j|_{\mathbb{L}^2}^2 + \frac12\delta_1\nu\frac{T}{N}|\nabla\bar{e}_{j-1}|_{\mathbb{L}^2}^2\\
&+ \Big(\gamma_2 + \frac{\bar{C}^2}{8\delta_1\nu}|\nabla u(t_j)|_{\mathbb{L}^2}^2\Big)\frac{T}{N}|\bar{e}_j|_{\mathbb{L}^2}^2 + \frac{\bar{C}^2}{8\delta_1\nu}|\nabla u(t_j)|_{\mathbb{L}^2}^2|\bar{e}_{j-1}|_{\mathbb{L}^2}^2\\
&+ \sum_{i=1}^{3}\tilde{T}_j(i) + \Big(\bar{e}_j\,,\,\int_{t_{j-1}}^{t_j}\big[G(u(s)) - G(v_N(t_{j-1}))\big]\,\mathrm{d}W(s)\Big).
\end{aligned}\tag{4.36}$$

Adding these estimates localized on the set $\Omega^M_{t_{j-1}}$ we deduce an upper estimate similar to (4.16), where e_j is replaced by $\bar{e}_j$. Following the same steps as in the proof of Proposition 4.3 we conclude the proof. $\square$

The arguments in Section 4.3 prove that the statements of Theorems 4.4 and 4.6 remain valid if we replace the solution $u_N(t_k)$ of the fully implicit Euler scheme by the solution $v_N(t_k)$ of the semi-implicit one.

4.5 *Time-dependent coefficients*

For the sake of simplicity we have supposed that the diffusion coefficient G does not depend on time. An easy modification of the proofs of this section shows that the statements of Theorems 4.4 and 4.6 for the fully or semi-implicit Euler schemes remain true if we suppose that $G : [0,T] \times V \to \mathcal{L}_2(K,H)$ (resp. $G : [0,T] \times \mathrm{Dom(A)} \to \mathcal{L}_2(K,V)$) satisfies the following global linear growth and Lipschitz conditions similar to those imposed in conditions **(G1)** and **(G2)**:

$$\|G(t,u)\|^2_{\mathcal{L}_2(K,H)} \le K_0 + K_1 |u|^2_{\mathbb{L}^2}, \quad \|G(t,u)\|^2_{\mathcal{L}_2(K,V)} \le K_0 + K_1 \|u\|^2_V,$$

$$\|G(t,u) - G(t,v)\|^2_{\mathcal{L}_2(K,H)^2} \le L_1 |u-v|^2_{\mathbb{L}^2},$$

$$\|G(t,u) - G(t,v)\|^2_{\mathcal{L}_2(K,V)^2} \le L_1 \|u-v\|^2_V,$$

for $u,v \in V$ (resp. $u,v \in \mathrm{Dom}(A)$). Furthermore, the diffusion coefficient G should also satisfy the following time regularity condition:
(G3) There exists a constant $C > 0$ such that for any $u,v \in V$ and $s,t \in [0,T]$,

$$\|G(t,u) - G(s,u)\|^2_{\mathcal{L}_2(K,H)} \le C|t-s|^{\frac{1}{2}}\,(1 + \|u\|^2_{\mathbb{L}^2}).$$

In that case the fully implicit scheme u_N (resp. semi-implicit scheme v_N) is defined replacing $G(u_N(t_{k-1}))$ by $G(t_{k-1}, u_N(t_{k-1}))$ (resp. by $G(t_{k-1}, v_N(t_{k-1}))$) on the right-hand side of (4.1).

Acknowledgements

This research was started in the fall 2017 while AM visited the University of Wyoming. She would like to thank this university for the hospitality and the very pleasant working conditions. Finally, we would like to thank the anonymous referee for the careful reading and valuable remarks, which helped improving this paper.

Funding

National Science Foundation (NSF) grant (DMS-1418838 to H.B.).

References

Bensoussan, A. (1990) Some existence results for stochastic partial differential equations. Pitman Research Notes in Mathematics Series, vol. 268. Harlow, Trento: Longman Sci. Tech., pp. 37–53.

Bensoussan, A., Glowinski, R. & Rascanu, A. (1992) Approximation of some stochastic differential equations by splitting up method. *Appl. Math. Optim.*, **25**, 81–106.

Bessaih, H., Brzeźniak, Z. & Millet, A. (2014) Splitting up method for the 2D stochastic Navier-Stokes equations. *Stochastic PDE Anal. Comput.,* **2–4**, 433–470.
Bessaih, H., Hausenblas, E., Randrianasolo, T. & Razafimandimby, P. A. (2018) Numerical approximation of stochastic evolution equations: convergence in scale of Hilbert spaces. *J. Comput. Appl. Math.,* **343**, 250–274.
Bessaih, H. & Millet, A. (2012) Large deviations and the zero viscosity limit for the 2D stochastic Navier-Stokes equations with free boundary. *SIAM J. Math. Anal.,* **44–43**, 1861–1893.
Blömker, D. & Jentzen, A. (2013) Galerkin approximations for the stochastic Burgers equation. *SIAM J. Numer. Anal.,* **51**, 694–715.
Breckner, H. (2000) Galerkin approximation and the strong solution of the Navier-Stokes equation. *J. Appl. Math. Stochastic Anal.*, **13**, 239–259.
Brzeźniak, Z., Carelli, E. & Prohl, A. (2013) Finite element base discretizations of the incompressible Navier-Stokes equations with multiplicative random forcing. *IMA J. Numer. Anal.*, **33**, 771–824.
Carelli, E. & Prohl, A. (2012) Rates of convergence for discretizations of the stochastic incompressible Navier-Stokes equations. *SIAM J. Numer. Anal.,* **50**, 2467–2496.
Chueshov, I. & Millet, A. (2010) Stochastic 2D hydrodynamical type systems: well posedness and large deviations. *Appl. Math. Optim.*, **61–63**, 379–420.
Dörsek, P. (2012) Semigroup splitting and cubature approximations for the stochastic Navier-Stokes equations. *SIAM J. Numer. Anal.,* **50–52**, 729–746.
Flandoli, F. & Gatarek, D. (1995) Martingale and stationary solutions for stochastic Navier-Stokes equations. *Probab. Theory Related Fields*, **102**, 367–391.
Hairer, M. & Mattingly, J. (2006) Ergodicity of the 2D Navier-Stokes equations with degenerate forcing. *Ann. Math.,* **164**, 993–1032.
Hutzenthaler, M. & Jentzen, A. (2015) *Numerical Approximations of Stochastic Differential Equations with Non-Globally Lipschitz Continuous Coefficients*. Memoirs of the American Mathematical Society, vol. 236, no. 1112.
Jentzen, A., Salimova, D. & Welti, T. (2017) Strong convergence for explicit space-time discrete numerical approximation methods for stochastic Burgers equations. arXiv:1710.07123.
Printems, J. (2001) On the discretization in time of parabolic stochastic partial differential equations. *Math. Model. Numer. Anal.,* **35–36**, 1055–1078.
Temam, R. (1979) *Navier-Stokes Equations: Theory and Numerical Analysis.* Studies in Mathematics and its Applications, vol. 2. Amsterdam-New York: North-Holland.
Temam, R. (1995) *Navier-Stokes Equations and Nonlinear Functional Analysis*. CBMS-NSF Regional Conference Series in Applied Mathematics,vol. 66. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Appendix

In this section we prove two technical lemmas used to obtain the strong convergence results.

A.1 *Proof of Lemma 3.5*

First note that using (3.4) and (3.7) we deduce that if $\mathbb{E}(\|u_0\|_V^{2p}) < \infty$ then

$$\sup_{N\geq 1} \mathbb{E} \int_0^T \|z^N(s)\|_V^{2p}\, \mathrm{d}t \leq C(p).$$

Thus, only moments of $\sup_{t\in[0,T]} |z^N(t)|_{\mathbb{L}^2}^{2p}$ have to be dealt with.

The process z^N defined by (3.6) is not regular enough to apply Itô's formula directly to $|z^N(t)|^2_{\mathbb{L}^2}$. Hence, as in the proof of Proposition 3.4, we need to apply Itô's formula on the smooth Galerkin approximations of the processes u^N, y^N and z^N and then pass to the limit. This yields, for every $t \in [0,T]$,

$$|z^N(t)|^2_{\mathbb{L}^2} = |u_0|^2_{\mathbb{L}^2} - 2\int_0^t \langle F(u^N(s)) \, , \, z^N(s)\rangle \, \mathrm{d}s + \int_0^t \|G(y^N(s)\|^2_{\mathcal{L}_2(K,H)} \, \mathrm{d}s$$
$$+ 2\int_0^t \big(z^N(s) \, , \, G(y^N(s)) \, \mathrm{d}W(s)\big).$$

Using the Itô formula once more we deduce

$$|z^N(t)|^{2p}_{\mathbb{L}^2} = |u_0|^{2p}_{\mathbb{L}^2} + I(t) + \sum_{i=1}^3 J_i(t), \tag{A.1}$$

where

$$\begin{aligned}
I(t) =& 2p \int_0^t \big(z^N(s) \, , \, G(y^N(s)) \, \mathrm{d}W(s)\big) |z^N(s)|^{2(p-1)}_{\mathbb{L}^2},\\
J_1(t) =& - 2p\nu \int_0^t |z^N(s)|^{2(p-1)}_{\mathbb{L}^2} \big(\nabla u^N(s) \, , \, \nabla z^N(s)\big) \, \mathrm{d}s,\\
J_2(t) =& - 2p \int_0^t |z^N(s)|^{2(p-1)}_{\mathbb{L}^2} \big\langle B(u^N(s), u^N(s)) \, , \, z^N(s)\big\rangle \, \mathrm{d}s,\\
J_3(t) =& + p \int_0^t |z^N(s)|^{2(p-1)}_{\mathbb{L}^2} \|G(y^N(s)\|^2_{\mathcal{L}_2(K,H)} \, \mathrm{d}s\\
& + 2p(p-1) \int_0^t \|G^*(y^N(s)) z^N(s)\|^2_K |z^N(s)|^{2(p-2)}_{\mathbb{L}^2} \, \mathrm{d}s.
\end{aligned}$$

The Hölder and Young inequalities imply

$$|J_1(t)| \le 2(p-1)\nu \int_0^t |z^N(s)|^{2p}_{\mathbb{L}^2} \, \mathrm{d}s + \nu \int_0^t \big[|\nabla z^N(s)|^{2p}_{\mathbb{L}^2} + |\nabla u^N(s)|^{2p}_{\mathbb{L}^2}\big] \, \mathrm{d}s.$$

Using again the Hölder and Young inequalities with exponents $\frac{2p+1}{2p-1}$ and $\frac{2p+1}{2}$ we deduce

$$\begin{aligned}
|J_2(t)| \le& 2p \int_0^t |z^N(s)|^{2(p-1)}_{\mathbb{L}^2} |\nabla z^N(s)|_{\mathbb{L}^2} \, \|u^N(s)\|^2_X \, \mathrm{d}s\\
\le& \frac{(2p-1)2p}{2p+1} \int_0^t \|z^N(s)\|^{2p+1}_V \, \mathrm{d}s + \frac{4p}{2p+1} \Big(\frac{\bar{C}}{2}\Big)^{\frac{2p+1}{2}} \int_0^t \|u^N(s)\|^{2p+1}_V \, \mathrm{d}s,
\end{aligned}$$

where $\bar{C}$ is the constant defined in (2.2). Finally, using the growth condition **(G1)**, we deduce

$$\begin{aligned}
|J_3(t)| \le& (2p^2 - p) \int_0^t |z^N(s)|^{2(p-1)}_{\mathbb{L}^2} \big[K_0 + K_1 |y^N(s)|^2_{\mathbb{L}^2}\big] \, \mathrm{d}s\\
\le& (2p-1)(p-1) \int_0^t |z^N(s)|^{2p}_{\mathbb{L}^2} \, \mathrm{d}s + C(p) K_1^p \int_0^t \|y^N(s)\|^{2p}_{\mathbb{L}^2} \, \mathrm{d}s + C(p) K_0 T,
\end{aligned}$$

where $C(p)$ is a constant depending on p. The inequalities (3.4) and (3.7), and the above estimates of $J_i(t)$ for $i = 1, 2, 3$, imply the existence of a positive constant $C(p)$ depending on p such that for every integer $N \geq 1$,

$$\sum_{i=1}^{3} \mathbb{E}\Big(\sup_{t\in[0,T]} |J_i(t)|\Big) \leq C(p). \tag{A.2}$$

Furthermore, the Burkholder–Davies–Gundy inequality, the growth condition in **(G1)** and the Young inequality imply

$$\begin{aligned}
\mathbb{E}\Big(\sup_{t\in[0,T]} |I(s)|\Big) &\leq 6p\mathbb{E}\Big(\Big\{ \int_0^T |z^N(s)|_{\mathbb{L}^2}^{4p-2} \|G(y^N(s))\|^2_{\mathcal{L}_2(K,H)}\, \mathrm{d}s\Big\}^{\frac{1}{2}}\Big) \\
&\leq \frac{1}{2}\mathbb{E}\Big(\sup_{t\in[0,T]} |z^N(t)|_{\mathbb{L}^2}^{2p}\Big) + C(p)\mathbb{E}\int_0^T \big[K_0^p + K_1^p |y^N(s)|_{\mathbb{L}^2}^{2p}\big]\, \mathrm{d}s \\
&\leq \frac{1}{2}\mathbb{E}\Big(\sup_{t\in[0,T]} |z^N(t)|_{\mathbb{L}^2}^{2p}\Big) + C(p),
\end{aligned} \tag{A.3}$$

where the last inequality is deduced from (3.4).

The upper estimates (A.1–A.3), (2.11), (3.4) and (3.7) conclude the proof. □

A.2 *Proof of Lemma 3.8*

We prove the existence of exponential moments for the square of the V-norm of the solution u of (1.1).

Proof of Lemma 3.8. Let $\alpha > 0$; we apply the Itô formula to the square of the V-norm of the process u solution of (1.1). As explained in the proofs of Proposition 3.4 and of the previous lemma the Itô formula has to be performed on smooth processes, for example, a Galerkin or a Yosida approximation of u, and then pass to the limit or use computations similar to those in Chueshov & Millet (2010, step 4 of the Appendix on p. 416). This yields

$$\begin{aligned}
\alpha\|u(t)\|_V^2 + \alpha\nu \int_0^t |||u(s)|||^2\, \mathrm{d}s &= \alpha\|u_0\|_V^2 + \alpha\int_0^t \|G(s,u(s))\|^2_{\mathcal{L}_2(K,V)}\, \mathrm{d}s \\
&\quad + 2\alpha\int_0^t \big(u(s)\,,\, G(s,u(s))\,\mathrm{d}W(s)\big)_V - \alpha\nu\int_0^t |||u(s)|||^2\, \mathrm{d}s,
\end{aligned} \tag{A.4}$$

where for $u, v \in V$ we set $(u,v)_V = (u,v) + (\nabla u, \nabla v)$ and recall that $|||u|||^2 := |\nabla u|^2_{\mathbb{L}^2} + |Au|^2_{\mathbb{L}^2}$.

Let $M(t) := 2\alpha \int_0^t \big(u(s)\,,\, G(s,u(s))\,\mathrm{d}W(s)\big)_V$; then M is a martingale with quadratic variation

$$\begin{aligned}
\langle M\rangle_t &\leq 4\alpha^2 \int_0^t \|u(s)\|_V^2 \|G(s,u(s))\|^2_{\mathcal{L}_2(K,V)}\, \mathrm{d}s \leq 4\alpha^2 K_0 \int_0^t \|u(s)\|_V^2\, \mathrm{d}s \\
&\leq 4\alpha^2 K_0 \tilde{C}\int_0^t |||u(s)|||^2\, \mathrm{d}s,
\end{aligned}$$

where the last inequality follows from (2.9) and the definition of $\tilde{C}$ in (3.28).

Let $\alpha_0 := \frac{\nu}{4K_0\tilde{C}}$; then we deduce

$$M_t - \alpha\nu\int_0^t |||u(s)|||^2\, \mathrm{d}s \leq M_t - \frac{\alpha_0}{\alpha}\langle M\rangle_t.$$

Therefore, using the previous inequality and classical exponential martingale arguments, we deduce

$$\begin{aligned}\mathbb{P}\Big[\sup_{0\le t\le T}\Big(M_t-\alpha\nu\int_0^t |||u(s)|||^2\,\mathrm{d}s\Big)\ge K\Big]&\le \mathbb{P}\Big[\sup_{0\le t\le T}\Big(M_t-\frac{\alpha_0}{\alpha}\langle M\rangle_t\Big)\ge K\Big]\\ \le&\mathbb{P}\Big[\sup_{0\le t\le T}\exp\Big(\frac{2\alpha_0}{\alpha}M_t-\frac{1}{2}\Big\langle\frac{2\alpha_0}{\alpha}M\Big\rangle_t\Big)\ge\exp\Big(\frac{2\alpha_0}{\alpha}K\Big)\Big]\\ \le&\exp\Big(-\frac{2\alpha_0}{\alpha}K\Big)\,\mathbb{E}\Big[\exp\Big(\frac{2\alpha_0}{\alpha}M_T-\frac{1}{2}\Big\langle\frac{\beta}{\alpha}M\Big\rangle_T\Big)\Big]\le\exp\Big(-\frac{2\alpha_0}{\alpha}K\Big)\end{aligned}$$

for any $K>0$. Set

$$X=\exp\Big(\alpha\Big\{\sup_{t\in[0,T]}2\int_0^t\big(u(s)\,,\,G(s,u(s))\big)_V-\nu\int_0^t|||u(s)|||^2\,\mathrm{d}s\Big\}\Big);$$

then for $\alpha\in(0,\alpha_0]$ we deduce that $\mathbb{P}(X\ge e^K)\le\exp(-K\frac{2\alpha_0}{\alpha})\le e^{-2K}=(e^{-K})^2$ for any $K>0$.

Using this inequality for any $C=e^K>1$ with $K>0$ we deduce $\mathbb{E}(X)\le 2+\int_0^\infty\mathbb{P}(X\ge C)\,\mathrm{d}C\le 3$. Since (A.4) implies

$$\begin{aligned}\alpha\Big(\sup_{t\in[0,T]}\Big[\|u(t)\|_V^2+\nu\int_0^t|||u(s)|||^2\,\mathrm{d}s\Big]\Big)\le{}&\alpha\|u_0\|_V^2+\alpha K_0T\\&+\sup_{t\in[0,T]}\Big(\alpha\Big\{2\int_0^t\big(u(s)\,,G(s,u(s))\big)_V-\nu\int_0^t|||u(s)|||^2\,\mathrm{d}s\Big\}\Big)\end{aligned}$$

we conclude the proof of (3.29). □