Chin. Ann. Math.
38B(2), 2017, 425–472
DOI: 10.1007/s11401-017-1077-6

Chinese Annals of Mathematics, Series B

Time Discrete Approximation of Weak Solutions to Stochastic Equations of Geophysical Fluid Dynamics and Applications*

Nathan GLATT-HOLTZ[1] Roger TEMAM[2] Chuntian WANG[3]

(*Dedicated to Haim Brézis on the occasion of his* 70*th birthday*)

Abstract As a first step towards the numerical analysis of the stochastic primitive equations of the atmosphere and the oceans, the time discretization of these equations by an implicit Euler scheme is studied. From the deterministic point of view, the 3D primitive equations are studied in their full form on a general domain and with physically realistic boundary conditions. From the probabilistic viewpoint, this paper deals with a wide class of nonlinear, state dependent, white noise forcings which may be interpreted in either the Itô or the Stratonovich sense. The proof of convergence of the Euler scheme, which is carried out within an abstract framework, covers the equations for the oceans, the atmosphere, the coupled oceanic-atmospheric system as well as other related geophysical equations. The authors obtain the existence of solutions which are weak in both the PDE and probabilistic sense, a result which is new by itself to the best of our knowledge.

Keywords Nonlinear stochastic partial differential equations, Geophysical fluid dynamics, Primitive equations, Discrete time approximation, Martingale solutions, Numerical analysis of stochastic PDEs
2010 MR Subject Classification 35Q86, 60H15, 35Q35

1 Introduction

The primitive equations (PEs for short) of the oceans and the atmosphere are a fundamental model for the large scale fluid flows forming the analytical core of the most advanced general circulation models (GCMs for short) in use today. In recent years, these systems have been a subject of considerable interest in the mathematical community not only because of their wide significance in geophysical applications but also for their delicate nonlinear, nonlocal, anisotropic structure and as a cousin to the other basic equations of mathematical fluid dynamics, namely the incompressible Navier-Stokes and Euler equations.

In this paper, we study a stochastic version of the PEs and develop techniques which may be viewed as a first step toward their numerical analysis. From the point of view of applications,

Manuscript received May 12, 2015. Revised February 3, 2016.
[1]Department of Mathematics, Tulane University, LA 70118, USA. E-mail: neigh@tulane.edu
[2]Department of Mathematics and the Institute for Scientific Computing and Applied Mathematics, Indiana University, Bloomington, IN 47405, USA. E-mail: temam@indiana.edu
[3]Department of Mathematics, University of Los Angeles, CA 90059, USA. E-mail: cwang@math.ucla.edu
*This work was supported by the National Science Foundation under the grants NSF-DMS-1206438 and NSF-DHS-1510249, the Research Fund of Indiana University and the National Science Foundation under the grants NSF-DMS-1004638 and NSF-DMS-1313272.

this work is motivated by a plea from the geophysical community to further develop the theory of nonlinear stochastic partial differential equations (SPDEs for short) in a large scale fluid dynamics context and in general (see [63]). Indeed, in view of the many sources of uncertainty both physical and numerical which are typically encountered by the modeler, stochastic techniques are playing an increasingly central role in the study of geophysical fluid dynamics (see, e.g., [9, 23, 45, 52, 55, 59, 62, 75] and also [32] for a small sampling of this vast literature).

The primitive equations trace their origins to the beginning of the 20th century with the seminal works of Bjerknes and Richardson [6, 61] and have played a central role in the development of climate modeling and weather prediction since that time (see [56]). To the best of our knowledge, the development of the mathematical theory for the deterministic PEs began in the early 1990's with a series of articles by Lions, Temam and Wang [46–48]. This direction in mathematical geophysics is now a fairly well developed subject with results guaranteeing the global existence of weak solutions which are bounded in $L^2_{\mathbf{x}}$ (see [47]), and the global existence and uniqueness of strong solutions, i.e., solutions evolving continuously in $H^1_{\mathbf{x}}$ (see [13, 38–39, 43]). Of course, these latter developments stand in striking contrast to the current state of the art for the Navier-Stokes equations as proving the global existence and uniqueness of strong solutions is tantamount to solving the famous Clay problem. For further background on the deterministic mathematical theory, see the recent surveys [60, 64].

Recently, significant efforts have been made to establish suitable analogues of the above (deterministic) mathematical results in a stochastic setting. In a series of works [15–16, 24, 27, 30–31, 35], the mathematical theory of strong, pathwise[1] solutions has been developed. These recent works more or less bring this aspect of the subject to the state of the art, that is they establish, in increasingly physically realistic settings, the global existence and uniqueness of solutions evolving continuously in $H^1_{\mathbf{x}}$.

Notwithstanding the above cited body of works, many aspects of the stochastic theory still need further consideration. In this paper, we develop existence results for weak solutions, which remain bounded in time only in $L^2_{\mathbf{x}}$. This is a direction which, to the best of our knowledge, remained unaddressed previously. Since such "weak solutions" are not expected to be unique, even in the deterministic setting, it is natural to work within the framework of martingale solutions. In other words, we consider below solutions which are weak in both the sense of PDE theory and stochastic analysis.

One particular advantage of this weak-martingale setting is that it allows us to consider physical situations unattainable so far in the above cited works on strong (or strong-pathwise) solutions. From the deterministic point of view, we obtain results for the case of inhomogenous, physically realistic boundary conditions. On the other hand, from the stochastic viewpoint, our results cover a very general class of state-dependent (multiplicative) noise structures. In particular, these noise terms may be interpreted in either the Itô or Stratonovich sense. The latter Stratonovich interpretation of noise is important as it may be more realistic in geophysical settings (see, e.g., [37, 57] for further details). Note that we develop our analysis in a slightly

[1] Here pathwise refers to the fact that solutions are found relative to a prescribed driving noise. In this paper, we use the terms "pathwise" and "martingale" as opposed to the alternate terminology of "weak" and "strong" solutions to avoid confusion with the typical PDE terminology for which weak solutions are, roughly speaking, those in $L^\infty_t(L^2_x)$ and strong solutions are those in $L^\infty_t(H^1_x)$.

abstract setting which at once allows us to treat the PEs of the oceans, the atmosphere and the coupled oceanic/atmospheric system.[2]

While the results established here take an important further step in the development of the analytical theory for the PEs, we believe that the main contribution of this article relates to numerical considerations. The approach below centers on an implicit Euler (i.e., time discrete) scheme, and we choose this set-up mainly because it may be seen as a mathematical setting suitable for the development of tools needed for the numerical analysis of the stochastic PEs and other nonlinear SPDEs arising in fluid dynamics. Note that while discrete time approximation was previously employed in [14, 17], these works treat hyperbolic type systems and only address the case of an additive noise. As such, a number of the techniques developed here, play a crucial role in a work related to the stability and consistency of a class of numerical schemes (both explicit and semi-implicit) for the 2D and 3D stochastic Navier-Stokes equations (see [33]).

Let us now finally turn to sketch some of the main technical challenges and contributions of the article. In fact, the first main difficulty is to justify the validity of the implicit scheme on which our analysis centers. While classical arguments involving the Brouwer fixed point theorem can be used to establish the existence of sequences satisfying the implicit scheme, we crucially need that these sequences are adapted to the driving noise. To address this concern, we rely on a specifically chosen filtration and a suitable measurable selection theorem from [10] (see also [11, 41]).

With suitable solutions to the semi-implicit scheme in hand, basic uniform estimates proceed analogously to the continuous time case with the use of martingale inequalities, etc. In contrast to previous works on Martingale solutions (see, e.g., [4, 15, 25, 34, 51]), we circumvent the need for higher moments with suitable stopping time arguments. Another difficulty related to the concern that solutions are adapted appears when we associate continuous time processes with the discrete time schemes in pursuit of compactness and the passage to the limit. In contrast to the deterministic case (see [53, 71]), we must introduce processes which are lagged by a time step. While these processes are indeed adapted, we obtain a time evolution equation with troublesome error terms. In turn, these error terms prevent us from addressing compactness directly from the equations and force us to carry out the compactness arguments for a series of interrelated processes.

1.1 Organization of this paper

The exposition is organized as follows. In Section 2, we outline an abstract, functional-analytic framework for the stochastic primitive equations (and related evolution systems) which may be seen as an "axiomatic" basis for the rest of the work. The section concludes by recalling the basic notion of Martingale solutions within the context of this framework. In Section 3, we introduce an implicit Euler scheme which discretizes the equations in time. The details of the existence of suitable solutions (adapted to the specific filtration) of this implicit scheme along with associated uniform estimates are given in Propositions 3.1 and 3.2, respectively. In Section 4, we study some continuous time processes associated with the implicit Euler scheme

[2] We have previously taken such an abstract approach in other work on the stochastic primitive equations (see [15]). There however our focus was on the local existence of strong, pathwise solutions and that framework was, by necessity, more restrictive with respect to domains, noise structures, etc.

introduced in Section 3. Section 5 then outlines the compactness (tightness) arguments that allow us to pass to the limit and derive the existence of solutions from these approximating continuous time processes. Finally, Section 7 provides extended details connecting the abstract results that we just derived with the concrete example of the primitive equations of the oceans. In this section, we also provide a number of examples of possible types of nonlinear state dependent noises covered under the main abstract results. In the interest of making the manuscript as self-contained as possible, an Appendix (Section A) collects various technical tools used in the course of our analysis.

After this work was completed, we heard of [3] which we regrettably overlooked. In this paper, the authors study the space and time discretization of the incompressible Navier-Stokes equations with multiplicative random forcing in space dimension 2 or 3. The space discretization of the equations is made by finite elements and the time discretization by an implicit Euler scheme. In this paper, we only perform discretization in time, also by an implicit Euler scheme. However, the issue of time and space discretization will be addressed in a forthcoming paper [33]. Note that [33] is still distinct from [3] because we also discuss in this paper the discretization of the Navier-Stokes equations by an explicit or semi-implicit scheme which raises issues of a.s. stability, a question not addressed in [3].

We continue with some additional remarks and comparisons between [3] and the present paper, and leave to [33] some further comparisons of [3] with our work.

(i) Regarding the equations considered, we study here a class of "abstract" fluid mechanics equations as in [16], and this class of equations covers the Navier-Stokes equations as well as the primitive equations of the atmosphere and the oceans (see, e.g., [46–48]). [3] dealt only with the Navier-Stokes equations. Because of the difficulty of constructing divergence free finite elements, the authors of [3] chose to deal with weakly incompressible finite elements, using the antisymmetrized form of the nonlinear term introduced in [67–68] to overcome the difficulties arising from handling approximate functions which are not exactly divergence free (see [3, 33] for further aspects of the spatial discretization).

(ii) In [3], the authors construct martingale solutions to the 3D Navier-Stokes equations and pathwise solutions to the 2D Navier-Stokes equations, also called weak and strong solutions in the probabilistic sense. All solutions are weak solutions in the PDE sense that is correspond to $L^\infty(L^2)$ and $L^2(H^1)$ solutions. In our case, the framework is general enough to include the 3D Navier-Stokes equations and therefore we only obtain martingale solutions; we do not specialize our results to the 2D case.

(iii) The tools are generally the same in both articles: Existence of approximate solutions $U^n \approx U(n\Delta t)$ by a fixed point method, energy a priori estimates, and compactness argument to pass to the limit. However, the construction of the approximate solutions $U^n \approx U(n\Delta t)$, raises a delicate question of measurability which we fully address in this paper. We did not see how this issue of measurability is addressed or bypassed in [3]. This issue of measurability was also overlooked in [14] to which [3] refers. The authors of this paper thank Debussche for helping them resolve this measurability difficulty.

(iv) In [3], the authors derived estimates on higher moments after assuming that U^0 is deterministic, which implies that U^0 is uniformly bounded in the probability space and in turn makes the derivation of the higher moments estimate possible. However, we assumed that the initial data belongs to only L^2 in the probability space and thereby were forced to develop some techniques to overcome the lack of higher moment estimates when e.g., establishing the compactness argument.

(v) In both papers, the passage to the limit is based on the construction of auxiliary approximate processes. We use very different arguments than that in [3]. However, it is not clear whether the methods are interchangeable in both circumstances, as again the lack of higher moment estimates in our case may matter, so much so that the more probabilistic approach of [3] may fail. Another difference is that, in 2D space, [3] provided the convergence to the unique solution using a monotonicity argument. This argument is inspired from the theory of montone operators of Minty and Browder [8, 49] (see also [7, 44] for pseudo monotone operators). The argument was extended to the stochastic context in [50] (see also [28, 51]). However, this argument implies uniqueness and therefore it cannot be applied to the general framework that we study which includes the 3D Navier Stokes equations.

This paper is dedicated to Haim Brézis on the occasion of his 70th birthday with admiration and friendship and (for RT) warm recollection of many years of interaction.

2 The Abstract Problem Set-up

We begin by describing the setting for the abstract evolution equation that we will study below (see (2.13) at the end of this section). As we note in the introduction, we take this point of view in order to systematically treat the existence of weak solutions to a class of geophysical fluids equations including but not limited to the example (7.1)–(7.4) developed below in Section 7. For further details about how to cast other related equations of geophysical fluid dynamics in the following abstract formulation, we refer the reader to [60] and the references therein.

Throughout what follows, we fix a Gelfand-Lions inclusion of Hilbert spaces

$$V_{(3)} \subset V_{(2)} \subset V \subset H \subset V' \subset V'_{(2)} \subset V'_{(3)}. \tag{2.1}$$

Each space is densely, continuously and compactly embedded in the next one. We denote the norms for H and V by $|\cdot|$ and $\|\cdot\|$, respectively, and the remaining spaces simply by e.g. $\|\cdot\|_{V'_{(2)}}$. When the context is clear, we denote the dual pairing between V', V, $V'_{(2)}, V_{(2)}$ or $V'_{(3)}, V_{(3)}$ by $\langle\cdot,\cdot\rangle$.

2.1 Basic operators

We now outline the main elements, a collection of abstract operators, which we use to build the stochastic evolution (2.13) below. We suppose that the following are given.

(1) A linear continuous operator $A : V \mapsto V'$ which defines a bilinear continuous form $a(U, U^\sharp) := \langle AU, U^\sharp\rangle_{V',V}$ on V. We assume that a is coercive, i.e.,

$$a(U, U) \geq c_1\|U\|^2 \quad \text{for all } U \in V. \tag{2.2}$$

This term will typically capture the diffusive terms in the concrete equations: Molecular and eddy viscosity, diffusion of heat, salt, humidity, etc.[3]

(2) A second linear operator E continuous on both H and V; E defines a bilinear continuous form $e(U, U^\sharp) := (EU, U^\sharp)$ on H (which is also continuous on V). We suppose furthermore that e is antisymmetric, that is,

$$e(U, U) = 0 \quad \text{for all } U \in H. \tag{2.3}$$

This term E appears in applications to account for the Coriolis (rotational) forces coming from the rotation of the earth.

(3) A bilinear form B which continuously maps $V \times V$ into $V'_{(2)}$; B gives rise to an associated trilinear form $b(U, U^\flat, U^\sharp) := \langle B(U, U^\flat), U^\sharp \rangle$ which satisfies the estimates

$$|b(U, U^\flat, U^\sharp)| \leq c_2 \|U\| |U^\flat|^{\frac{1}{2}} \|U^\flat\|^{\frac{1}{2}} \|U^\sharp\|_{V_{(2)}} \quad \text{for all } U, U^\flat \in V,\ U^\sharp \in V_{(2)}. \tag{2.4}$$

Moreover, we assume the antisymmetry property

$$b(U, \widetilde{U}, \widetilde{U}) = 0 \quad \text{for all } U \in V,\ \widetilde{U} \in V_{(2)}. \tag{2.5}$$

Note that, in particular, we may infer from (2.4) that

$$\|B(U)\|_{V'_{(2)}} \leq c_2 |U|^{\frac{1}{2}} \|U\|^{\frac{3}{2}} \quad \text{for any } U \in V. \tag{2.6}$$

Furthermore, from (2.4)–(2.5), we may assume that B is continuous from $V \times V_{(2)}$ into V' and satisfies

$$\|B(U)\|_{V'} \leq c_2 \|U\| \|U\|_{V_{(2)}} \quad \text{for all } U \in V_{(2)}. \tag{2.7}$$

Finally, we impose some additional technical convergence conditions on b. Firstly, we suppose that when U_k converges weakly to U in V then, up to a subsequence k',

$$b(U_{k'}, U_{k'}, U^\sharp) \to b(U, U, U^\sharp) \quad \text{for each } U^\sharp \in V_{(2)}. \tag{2.8}$$

Similarly, we assume that if, for some $T > 0$,

$$U_k \to U \quad \text{weakly in } L^2(0, T; V) \text{ and strongly in } L^2(0, T; H),$$

then, again up to a subsequence k',

$$\int_0^T b(U_{k'}, U_{k'}, U^\sharp) \mathrm{d}t \to \int_0^T b(U, U, U^\sharp) \mathrm{d}t \quad \text{for each } U^\sharp \in L^\infty(0, T; V_{(3)}). \tag{2.9}$$

B accounts for the main nonlinear (convective) terms in the equations.

(4) An externally given element ℓ. We consider ℓ to be random in general; it is specified only as a probability distribution on $L^2_{\text{loc}}(0, \infty; V')$ subject to the second moment condition (2.17)

[3] In previous works on the Stochastic PEs (see [15, 30–31]), we required that this a is symmetric. In particular, such a symmetry was strongly used in these previous works so that we could apply the spectral theorem to the inverse of an associated operator A^{-1}. This is not needed for the arguments presented here, and we therefore revert to the more general weak formulation of the PEs given in [60].

given below. This term ℓ captures various inhomogeneous elements, i.e., externally determined body forcings, boundary forcings, etc.

In order to define the operators involving the "stochastic terms" in the equations, we consider an auxiliary space $\mathfrak{U}$, on which the underlying driving noise, a cylindrical Brownian motion W evolves (see Subsection 2.2). We suppose that $\mathfrak{U}$ is a separable Hilbert space and use $L_2(\mathfrak{U}, X)$ to denote the space of Hilbert-Schmidt operators from $\mathfrak{U}$ into X, where, for example $X = H, V$ or $\mathbb{R}$. Sometimes, we abbreviate and write $L_2 := L_2(\mathfrak{U}, \mathbb{R})$.

Returning to the list of operators, we suppose that we have defined the following.

(1) A (possibly nonlinear) continuous map $\sigma : [0, \infty) \times H \mapsto L_2(\mathfrak{U}, H)$. We suppose that σ is uniformly sublinear, i.e.,

$$|\sigma(t, U)|_{L_2(\mathfrak{U},H)} \leq c_3(1 + |U|) \quad \text{for every } U \in H \text{ and } t \in \mathbb{R}^+, \tag{2.10}$$

where the constant $c_3 > 0$ is independent of $t \in [0, \infty)$. For economy of notation, we will frequently drop the dependence on t in the exposition below. We define $g : [0, \infty) \times H \times H \mapsto L_2$ according to $g(t, U, U^\sharp) = (\sigma(t, U), U^\sharp)$ for $U, U^\sharp \in H$. The element σ determines the structure of the (volumic) stochastic forcing applied to the equations. These stochastic terms typically appear to account for various sources of physical, empirical and numerical uncertainty as we described in the introduction.

(2) A continuous map $\xi : [0, \infty) \times H \mapsto H$ which is subject to the uniform sublinear condition

$$|\xi(t, U)| \leq c_4(1 + |U|) \quad \text{for every } U \in H \text{ and } t \in \mathbb{R}^+, \tag{2.11}$$

where $c_4 > 0$ does not depend on $t \geq 0$. We define $s : [0, \infty) \times H \times H \mapsto \mathbb{R}$ by

$$s(t, U, U^\sharp) = (\xi(t, U), U^\sharp) \tag{2.12}$$

for $U, U^\sharp \in H$. We include ξ in the abstract formulation to allow, in particular, for the treatment of a class of Stratonovich noises; ξ arises when we convert from a Stratonovich into an Itô type noise. This term S therefore allows us to carry out the forthcoming analysis entirely within the Itô framework (see Remarks 2.1, 7.3 below).

With the above abstract framework now in place, we may reduce the problem (7.1)–(7.4) below (and related equations) to studying the following abstract stochastic evolution equation in $V'_{(2)}$, namely,

$$\mathrm{d}U + (AU + B(U) + EU)\,\mathrm{d}t = (\ell + \xi(U))\,\mathrm{d}t + \sigma(U)\mathrm{d}W, \quad U(0) = U^0. \tag{2.13}$$

This system is to be interpreted in the Itô sense which we recall immediately below in Subsection 2.2.

Note that U^0 and ℓ in (7.1) are considered to be random in general. Indeed, since we are studying Martingale solutions to (2.13) where the underlying stochastic elements in the problem are considered as unknowns, we will specify U^0 and ℓ only as probability distributions on H and $L^2(0, T; V')$ (see Definition 2.1 and Remark 2.1). Note also that, for brevity of notation, we sometimes write

$$\mathcal{N}(t, U) := -(AU + B(U) + EU - \xi(t, U)) \tag{2.14}$$

in the course of the exposition below. When the context is clear, we sometimes drop the dependence on t and simply write $\mathcal{N}(U)$.

2.2 Some elements of stochastic analysis and abstract probability theory

Of course, (2.13) is understood relative to a stochastic basis $\mathcal{S} := (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, \{W^k\}_{k\geq 1})$, that is a filtered probability space with $\{W^k\}_{k\geq 1}$, a sequence of independent standard 1D Brownian motions relative to $\mathcal{F}_t$. Here we may define W on $\mathfrak{U}$ by considering an associated orthonormal basis $\{e_k\}_{k\geq 1}$ of $\mathfrak{U}$ and taking $W = \sum_k W_k e_k$; W is thus a "cylindrical Brownian" motion evolving over $\mathfrak{U}$.

Actually, this sum $W = \sum_k W_k e_k$ is only formal; it does not generally converge in $\mathfrak{U}$. For this reason, we will occasionally make use of a larger space $\mathfrak{U}_0 \supset \mathfrak{U}$ which we define according to

$$\mathfrak{U}_0 := \Big\{ v = \sum_{k\geq 0} \alpha_k e_k : |v|^2_{\mathfrak{U}_0} < \infty \Big\}, \quad \text{where } |v|^2_{\mathfrak{U}} := \sum_k \alpha_k^2 \text{ and } |v|^2_{\mathfrak{U}_0} := \sum_k \frac{\alpha_k^2}{k^2}. \tag{2.15}$$

Note that the embedding of $\mathfrak{U} \subset \mathfrak{U}_0$ is Hilbert-Schmidt. Moreover, using standard martingale arguments with the fact that each W_k is almost surely continuous, we have that, for almost every $\omega \in \Omega$, $W(\omega) \in \mathcal{C}([0,T], \mathfrak{U}_0)$.

Since (2.13) is actually short hand for a stochastic integral equation, we next briefly recall some elements of the theory of Itô stochastic integration in infinite dimensional spaces. We choose an arbitrary Hilbert space X and, as above, we use $L_2(\mathfrak{U}, X)$ to denote the collection of Hilbert-Schmidt operators from $\mathfrak{U}$ into X. Given an X-valued predictable[4] process $G \in L^2(\Omega; L^2_{\mathrm{loc}}(0,\infty, L_2(\mathfrak{U}, X)))$, the (Itô) stochastic integral

$$M_t := \int_0^t G \mathrm{d}W = \sum_k \int_0^t G_k \mathrm{d}W_k, \qquad \text{where } G_k = G e_k,$$

is defined as an element in $\mathcal{M}^2_X$, the space of all X-valued square integrable martingales (see [58, Subsections 2.2–2.3]). For further details on the general theory of infinite-dimensional stochastic integration and stochastic evolution equations, we refer the reader to [19, 58].

Since we will be working in the setting of Martingale solutions, where the data in the problem (2.13) are specified only as a probability distribution (over an appropriate function space), it is convenient to introduce some further notations around Borel probability measures. Let $(\mathcal{H}, \rho)$ be a complete metric space and denote the family of Borel probability measures on $\mathcal{H}$ by $\mathrm{Pr}(\mathcal{H})$. Given a Borel measurable function $f : \mathcal{H} \mapsto \mathbb{R}$ and an element $\mu \in \mathrm{Pr}(\mathcal{H})$, we sometimes write $\mu(f)$ for $\int_{\mathcal{H}} f(x) \mathrm{d}\mu(x)$ when the associated integral makes sense. In particular, we write

$$\mu(|f|) < \infty \iff \int_{\mathcal{H}} |f(x)| \mathrm{d}\mu(x) < \infty. \tag{2.16}$$

We review some basic properties related to convergence and compactness of subsets of $\mathrm{Pr}(\mathcal{H})$ in the Appendix below (see Section A.1). We refer the reader to [5] for an extended treatment of the general theory of probability measures on Polish spaces which include Hilbert spaces such as H and V.

[4]For a given stochastic basis $\mathcal{S}$, let $\Phi = \Omega \times [0,\infty)$ and take $\mathcal{G}$ to be the sigma algebra generated by the sets of the form

$$(s,t] \times F, \quad \text{with } 0 \leq s < t < \infty \text{ and } F \in \mathcal{F}_s; \quad \{0\} \times F; \quad F \in \mathcal{F}_0.$$

Recall that an X valued process U is called predictable (with respect to the stochastic basis $\mathcal{S}$) if it is measurable from $(\Phi, \mathcal{G})$ into $(X, \mathcal{B}(X))$ where $\mathcal{B}(X)$ denotes the family of Borelian subsets of X.

2.3 Definition of martingale solutions and statement of the main result

We turn now to give a rigorous meaning for the so-called weak-martingale solutions to (2.13) which are defined as follows.

Definition 2.1 (Weak-Martingale Solutions) *Fix μ_{U^0}, μ_ℓ Borel measures respectively on H and $L^2_{\rm loc}(0,\infty;V')$ with*

$$\mu_{U^0}(|\cdot|^2_H)<\infty \quad \text{and} \quad \mu_\ell(\|\cdot\|^2_{L^2(0,T;V')})<\infty \quad \text{for any } T>0. \tag{2.17}$$

A weak-martingale solution $(\widetilde{\mathcal{S}},\widetilde{U},\widetilde{\ell})$ to (2.13) *consists of a stochastic basis $\widetilde{\mathcal{S}}=(\widetilde{\Omega},\widetilde{\mathcal{F}},\{\widetilde{\mathcal{F}}_t\}_{t\geq 0},\widetilde{\mathbb{P}},\widetilde{W})$ and processes $\widetilde{U}$ and $\widetilde{\ell}$ (defined relative to $\widetilde{\mathcal{S}}$) adapted to $\{\widetilde{\mathcal{F}}_t\}_{t\geq 0}$. This triple $(\widetilde{\mathcal{S}},\widetilde{U},\widetilde{\ell})$ will enjoy the following properties:*

(i) *For every $T>0$,*

$$\begin{aligned}&\widetilde{U}\in L^2(\widetilde{\Omega};L^\infty(0,T;H)\cap L^2(0,T;V)), \quad \widetilde{U} \text{ is a.s. weakly continuous in } H,\\ &\widetilde{\ell}\in L^2(\widetilde{\Omega};L^2(0,T;V')).\end{aligned} \tag{2.18}$$

(ii) *For every $t>0$ and each test function $U^\sharp\in V_{(2)}$,*

$$\begin{aligned}&(\widetilde{U}(t),U^\sharp)+\int_0^t(a(\widetilde{U},U^\sharp)+b(\widetilde{U},\widetilde{U},U^\sharp)+e(\widetilde{U},U^\sharp))\mathrm{d}s\\ &=(\widetilde{U}(0),U^\sharp)+\int_0^t(\ell(U^\sharp)+s(\widetilde{U},U^\sharp))\mathrm{d}t+\int_0^t g(\widetilde{U},U^\sharp)\mathrm{d}\widetilde{W},\end{aligned} \tag{2.19}$$

almost surely.

(iii) *Finally, $\widetilde{U}(0)$ and $\widetilde{\ell}$ have the same laws as μ_{U^0}, μ_ℓ, i.e.,*

$$\widetilde{\mathbb{P}}(\widetilde{U}(0)\in\cdot)=\mu_{U^0}(\cdot),\quad \widetilde{\mathbb{P}}(\widetilde{\ell}\in\cdot)=\mu_\ell(\cdot). \tag{2.20}$$

With this definition in hand, we now state one of the main results of the work as follows.

Theorem 2.1 *Let μ_{U^0}, μ_ℓ be a given pair of Borel measures on respectively H and $L^2_{\rm loc}(0,\infty;V')$ which satisfy the moment conditions* (2.17). *Then, relative to this data, there exists a martingale solution $(\widetilde{\mathcal{S}},\widetilde{U},\widetilde{\ell})$ to* (2.13) *in the sense of Definition* 2.1.

Remark 2.1 Depending on the structure of σ the application of noise leads to a variety of different effects on the behavior of the solutions. In particular, σ can be chosen so that the noise either provides a damping or an exciting effect. It is therefore unsurprising that the structure of the stochastic terms in e.g. (7.1) remains a subject of ongoing debate among physicists and applied modelers. In any case, viewed as a proxy for physical and numerical uncertainty, the structure of the noise would be expected to vary by application. With this debate in mind we have therefore sought to treat a very general class of state-dependent noise structures in σ requiring only the sublinear condition (2.10). We have illustrated some interesting examples covered under this condition in Subsection 7.3 below.

Actually, the Stratonovich interpretation of white noise driven forcing may often be more appropriate for applications in geophysics (see, e.g., [37, 57] for extended discussions on this connection). Note that although (2.13) is considered in an Itô sense, an additional, state dependent drift term ξ is added to the equations which allows us to treat a class of Stratonovich noises with (2.13) via the standard "conversion formula" between Itô and Stratonovich evolutions (see, e.g., [1] and also Subsection 7.3 where we present one such example of Stratonovich forcing in detail).

3 A Discrete Time Approximation Scheme

We now describe in detail the semi-implicit Euler scheme, (3.3), which we use to approximate (2.13). This system is given rigorous meaning in Definition 3.1. We then recall a specific stochastic basis in Subsection 3.2 and establish the existence of solutions to (3.3) in Proposition 3.1 relative to this basis. We conclude this section by providing certain uniform bounds (energy estimates) independent of the time step of the discretization in Proposition 3.2.

3.1 The implicit scheme

Fix a stochastic basis $\mathcal{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, \{W^k\}_{k\geq 1})$ and elements $\ell \in L^2(\Omega; L^2_{\text{loc}}(0, \infty; V'))$, $U^0 \in L^2(\Omega; H)$ whose distributions correspond to the externally given μ_ℓ, μ_{U^0}. For a given $T > 0$ and any integer N, let

$$\Delta t = \frac{T}{N}, \quad t^n = t^n_N = n\Delta t \quad \text{for } n = 0, 1, \cdots, N, \tag{3.1}$$

along with the associated stochastic increments

$$\eta^n = \eta^n_N = W(t_n) - W(t_{n-1}) \quad \text{for } n = 1, \cdots, N. \tag{3.2}$$

Using an implicit Euler time discretization scheme, we would then like to approximate (2.13) by considering sequences $\{U^n_N\}_{n=1}^N$ satisfying

$$\frac{U^n_N - U^{n-1}_N}{\Delta t} + AU^n_N + B(U^n_N) + EU^n_N = \ell^n_N + \xi(t^n, U^n_N) + \sigma_N(t^{n-1}, U^{n-1}_N)\frac{\eta^n_N}{\Delta t}, \tag{3.3}$$

in $V'_{(2)}$ for $n = 1, \cdots, N$. For how to choose U^0_N, see Remark 3.1. The terms ℓ^n_N are given by

$$\ell^n_N(U^\sharp) = \frac{1}{\Delta t}\int_{(n-1)\Delta t}^{n\Delta t} \ell(t, U^\sharp)\mathrm{d}t \quad \text{for } n = 1, 2, \cdots, N, \tag{3.4}$$

and the operator $\sigma_N : [0, \infty) \times H \to L_2(\mathfrak{U}, V)$ is any approximation of σ which satisfies

$$\|\sigma_N(t, U)\|^2_{L_2(\mathfrak{U}, V)} \leq N|\sigma(t, U)|^2_{L_2(\mathfrak{U}, H)}, \tag{3.5}$$

$$|\sigma_N(t, U)|^2_{L_2(\mathfrak{U}, H)} \leq |\sigma(t, U)|^2_{L_2(\mathfrak{U}, H)} \tag{3.6}$$

for every $t \geq 0$ and every $U \in H$. Additionally, we suppose that, for any $t \geq 0$,

$$\lim_{N\to\infty} \sigma_N(t, U_N) = \sigma(t, U), \quad \text{whenever } U_N \to U \text{ in } H. \tag{3.7}$$

For the existence of such σ_N, see Remark 3.1. We write $g_N(t, U, U^\sharp) = (\sigma_N(t, U), U^\sharp)$.[5]

We make the notion of suitable solutions to (3.3) precise in the following definition.

[5]The choice of a "time explicit" term in $\sigma_N(t^{n-1}, U^{n-1}_N)$ is needed to obtain the correct (Itô) stochastic integral in the limit as $\Delta t \to 0$. Actually, this adaptivity (measurability) concern also leads us to introduce the approximations of σ in (3.3) (see Remark 3.1, (4.6), (4.21)). Note that, as explained in this remark approximations of σ satisfying (3.5)–(3.7) can always be found via an elementary functional-analytic construction.

Definition 3.1 *We consider a stochastic basis* $\mathcal{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, \{W^k\}_{k\geq 1})$. *Given* $N \geq 1$ *and an element* $U_N^0 \in L^2(\Omega, H)$ *which is* $(\mathcal{F}_0, \mathcal{B}(H))$ *measurable and a process* $\ell = \ell(t) \in L^2(\Omega; L^2(0, T; V'))$ *adapted to* $\{\mathcal{F}_t\}_{t\geq 0}$, *we say that a sequence* $\{U_N^n\}_{n=0}^N$ *is an admissible solution of the Euler Scheme* (3.3), *if*

(i) *for each* $n = 1, \cdots, N$, $U_N^n \in L^2(\Omega; V)$ *and* U_N^n *is* $\mathcal{F}_n$ *adapted, where* $\mathcal{F}_n := \mathcal{F}_{t^n}$, $n = 0, \cdots, N$;

(ii) *every pair* U_N^n, U_N^{n-1}, $n = 1, \cdots, N$, *satisfies*

$$\begin{aligned}&(U_N^n - U_N^{n-1}, U^\sharp) + (a(U_N^n, U^\sharp) + b(U_N^n, U_N^n, U^\sharp) + e(U_N^n, U^\sharp))\Delta t\\ &= (\ell_N^n(U^\sharp) + s(t^n, U_N^n, U^\sharp))\Delta t + g_N(t^{n-1}, U_N^{n-1}, U^\sharp)\eta_N^n,\end{aligned} \tag{3.8}$$

almost surely for all $U^\sharp \in V_{(2)}$;

(iii) *for each* $n = 1, \cdots, N$, U_n^N *and* U_{n-1}^N *satisfy the "energy inequality", almost surely on* Ω :

$$(U_N^n - U_N^{n-1}, U_N^n) + \Delta t c_1 \|U_N^n\|^2 \leq (\ell_N^n(U_N^n) + s(t^n, U_N^n, U_N^n))\Delta t + g_N(U_N^{n-1}, U_N^n)\eta_N^n \tag{3.9}$$

for $n = 1, \cdots, N$ *and where* c_1 *is the constant from* (2.2).

Remark 3.1 At first glance the dependence on N in both the initial condition and the noise term involving σ may seem strange. Indeed, in the deterministic setting, when we approximate (2.13) with (3.3), we would simply take U_N^0 to be equal to the initially given U^0 for all N. Similarly, if we were to add deterministic sublinear terms analogous to σ to the governing equations, no approximation as in (3.5)–(3.7) would be necessary. However, the situation is, in general, more complicated in the stochastic setting as we shall see in detail in Section 4, Proposition 4.1. This is essentially because we must construct continuous time processes from the U_N^n's which are adapted to a given filtration (see (4.6), (4.15)–(4.17) and (4.21) for specific details).

For now let us describe how we can achieve suitable approximations in the U_N^0 and σ_N's.

(1) For a given initial probability distributions μ_{U^0}, on H (with $\mu_{U^0}(|\cdot|_H^2) < \infty$) and having fixed a suitable stochastic basis and an element $U^0 \in L^2(\Omega; H)$, $\mathcal{F}_0$-measurable, with distribution μ_{U^0}, we then pick a sequence $U_N^0 \in L^2(\Omega; V_{(2)})$ such that $U_N^0 \to U^0$ as $N \to \infty$ in $L^2(\Omega; H)$ but subject to the restriction given in (4.3) below. Such a sequence can be found with a simple density argument. Indeed, since $V_{(2)}$ is dense in H, we may initially approximate U^0 in $L^2(\Omega, H)$ with a sequence $\overline{U}_M^0 \in L^\infty(\Omega; V_{(2)})$. We then define $M(N) = \max\{M \geq 1 : \|\overline{U}_M^0\|_{L^\infty(\Omega; V_{(2)})} \leq N^{\frac{1}{2}}\} \wedge N$ and define $U_N^0 = \overline{U}_{M(N)}^0$. Since $M(N) \to \infty$ as $N \to \infty$, U_N^0 approximates U^0 in $L^2(\Omega; H)$ while maintaining the constraint (4.3).

(2) We may construct elements σ_N from σ satisfying (3.5)–(3.7) according to the following general functional analytic construction. For any $U \in H$, via Lax-Milgram we define $\Psi(U)$ to be the unique solution in V of $(\Psi(U), U^\sharp) = (U, U^\sharp)$ for all $U^\sharp \in V$. Classically, Ψ is a compact, self-adjoint and injective linear operator on H. Thus, by the spectral theorem, we may find a complete orthonormal basis for H, $\{\Phi_j\}_{j\geq 1}$, which is made up of eigenfunctions of Ψ with a corresponding sequence of eigenvalues $\{\gamma_j\}_{j\geq 1}$ decreasing to zero. For any integer m, we let P_m to be the projection onto $H_m := \text{span}\{\Phi_1, \cdots, \Phi_m\}$. Now choose a sequence m_N increasing to

infinity but so that $\gamma_{m_N}^{-1} \le N$. It is not hard to see that defined in this way $\sigma_N(\cdot) = P_{m_N}\sigma(\cdot)$ satisfies the requirements given in (3.5)–(3.7).

3.2 Existence of the U_N^n's

While the existence for a.e. $\omega \in \Omega$ of solutions to (3.3) satisfying (3.9) follows along arguments similar to those found in [60, Lemma 2.3], some care is required to demonstrate the existence of sequences $\{U_N^n\}_{n=0}^N$ which are adapted to the underlying stochastic basis. For this complication, we will make use of a "measurable selection theorem" (see Theorem A.2) from [10] (see also the related earlier works [11, 41]). In order to apply this result, we use a specific stochastic basis defined around the canonical Wiener space whose definition we recall next.

3.2.1 The Wiener measure and its filtration

We recall the canonical Wiener space as follows (see [42] for further details). Let

$$\Omega = \mathcal{C}([0,T]; \mathfrak{U}_0),$$

equipped with the Borel σ-algebra denoted as $\mathcal{G}$. We equip $(\Omega, \mathcal{G})$ with the Wiener measure $\mathbb{P}$.[6] Then the evaluation map $W(\omega, t) := \omega(t)$, $\omega \in \Omega$, $t \in [0,T]$, is a cylindrical Wiener process on $\mathfrak{U}_0$. The filtration is given by $\mathcal{G}_t$ defined as follows:

the completion of the sigma algebra generated by the $W(s)$ for $s \in [0,t]$ with respect to $\mathbb{P}$.

Combining these elements $\mathcal{S}_\mathcal{G} = (\Omega, \mathcal{G}, \{\mathcal{G}_t\}_{t\ge 0}, \mathbb{P}, W)$ gives a stochastic basis suitable for applying Theorem A.2.

3.2.2 Existence of the U_N^n's adapted to $\mathcal{G}_{t_n}$

Proposition 3.1 *Suppose that*

$$N \ge N_0 := 4Tc_4 \quad (\textit{or equivalently that } 4\Delta t c_4 < 1), \tag{3.10}$$

where c_4 is the constant arising in (2.11). *Consider the stochastic basis $\mathcal{S}_\mathcal{G}$ defined as in Subsection* 3.2*, an $N \ge N_0$, and an element $U_N^0 \in L^2(\Omega; H)$ which is $\mathcal{G}_0$-measurable and a process $\ell = \ell(t) \in L^2(\Omega; L^2(0,T; V'))$ measurable with respect to the sigma algebra generated by the $W(s)$ for $s \in [0,t]$. Then there exists a sequence $\{U_N^n\}_{n=0}^N$ which is an admissible solution to the Euler scheme* (3.3) *in the sense of Definition* 3.1*.*

The rest of this subsection is devoted to the proof of Proposition 3.1. Below we will construct the sequence $\{U_N^n\}_{n=0}^N$ iteratively starting from U_N^0, but we first need to take the preliminary step of establishing the existence of a certain Borel measurable map $\Gamma : [0,T] \times V' \to V$ which is used at the heart of this construction.

We define the continuous map $\mathfrak{G} : [0,T] \times V \to V'_{(2)}$ according to

$$\mathfrak{G}(t, U) = U + \Delta t\big(AU + B(U) + EU - \xi(U,t)\big), \tag{3.11}$$

[6] Using the orthonormal basis $\{e_k\}_{k\ge 1}$ of $\mathfrak{U}$, $\mathbb{P}$ is obtained as the product of the independent Wiener measures each one defined on $\mathcal{C}([0,T]; \mathbb{R})$.

and, for each $t \in [0,T]$ and $F \in V'$ we set

$$\begin{aligned}\Lambda(t,F) =&\{U \in V : \langle \mathfrak{G}(t,U) - F, U^\sharp \rangle = 0,\ \forall U^\sharp \in V_{(2)} \ \text{ and} \\ &|U|^2 + \Delta t c_1 \|U\|^2 \le \langle F + \xi(t,U)\Delta t, U\rangle\}. \end{aligned} \tag{3.12}$$

Using this family of sets defined by (3.12), we now establish the following lemma.

Lemma 3.1 *There exists a map $\Gamma : (0,T) \times V' \to V$ which is universally Radon measurable (Radon measurable for every Radon measure on $(0,T) \times V'$), such that for every $t \in (0,T)$ and every $F \in V'$, $U := \Gamma(t,F) \in \Lambda(t,F)$.*

Proof We establish the existence of the desired Γ by showing that Λ satisfies the conditions of Theorem A.2. More precisely, we need to verify that[7]

(i) for each $t \in [0,T]$, $F \in V'$, the set $\Lambda(t,F)$ is non-empty,

(ii) $\Lambda(t,F)$ is closed. In other words, we need to show that, given any sequences

$$t_n \to t, \quad F_n \to F \quad \text{in } V', \quad U_n \to U \in V$$

such that, for every n,

$$\begin{aligned}&\langle \mathfrak{G}(t_n,U_n) - F_n, U^\sharp \rangle = 0 \quad \text{ for every } U^\sharp \in V_{(2)}, \\ &|U_n|^2 + \Delta t c_1 \|U_n\|^2 \le \langle F_n + \xi(t_n,U_n)\Delta t, U_n\rangle,\end{aligned}$$

we have

$$\langle \mathfrak{G}(t,U) - F, U^\sharp \rangle = 0 \quad \text{for every } U^\sharp \in V_{(2)}, \quad |U|^2 + \Delta t c_1 \|U\|^2 \le \langle F + \xi(t,U)\Delta t, U\rangle.$$

The first item, (i) may be established with a Galerkin scheme and the Brouwer fixed point theorem along standard arguments typically used to prove the existence of solutions to nonlinear elliptic equations of the type of Navier-Stokes and primitive equations (see Lemma 2.3 in [60, p. 26]). Since some specifics are different here, we briefly sketch some details of this argument. Fix any $t \in [0,T]$ and any $F \in V'$, and consider a family $\{\Psi_k\}_{k\ge1} \subset V_{(2)}$ which is free and total in V. For each $m \ge 1$, we seek an element $U_m = \sum_{j=1}^{m} \beta_{jm}\Psi_j$ such that

$$\langle \mathfrak{G}(t,U_m) - F \Psi_k \rangle = 0 \quad \text{ for every } k = 1, \cdots, m. \tag{3.13}$$

Observe that, for any U_m of this form, using (2.2)–(2.3), (2.5) and (2.11), we estimate

$$\begin{aligned}\langle \mathfrak{G}(t,U_m) - F, U_m\rangle &= |U_m|^2 + \Delta t(a(U_m,U_m) - (\xi(t,U_m),U_m)) - \langle F, U_m\rangle \\ &\ge |U_m|^2 + \Delta t(c_1\|U_m\|^2 - 2c_4(1+|U_m|^2)) - |F|_{V'}\|U_m\| \\ &\ge \frac{\Delta t c_1}{2}\|U_m\|^2 - \frac{1}{2}\Big(1 + \frac{1}{\Delta t c_1}|F|^2_{V'}\Big).\end{aligned}$$

The last inequality follows from the assumption (3.10) which implies that $2c_4\Delta t \le 1$. The existence of solutions to (3.13) for any given t, F of the form $U_m = \sum_{j=1}^{m} \beta_{jm}\Psi_j$ thus follows for each m from the Brouwer fixed point theorem.

[7] To apply Theorem A.2, we actually would like to define Λ on the Banach space $\mathbb{R} \times V'$. For this purpose, we may simply take $\Lambda(t,F) = \Lambda(T,F)$ when $t > T$, and when $t < 0$ we let $\Lambda(t,F) = \Lambda(0,F)$.

We next seek bounds on the resulting sequence of U_m's in V independent of m. Starting from (3.13), we find that

$$\begin{aligned}|U_m|^2 + c_1\Delta t\|U_m\|^2 &\le \Delta t(\xi(t,U_m),U_m) + \langle F, U_m\rangle \\ &\le 2c_4\Delta t(1+|U_m|^2) + \frac{1}{2\Delta t c_1}|F|_{V'}^2 + \frac{\Delta t c_1}{2}\|U_m\|^2. \end{aligned} \tag{3.14}$$

Using once again the standing assumption (3.10), we have that U_m is bounded in V independently of m. Passing to a subsequence as needed and using that V is compactly embedded in H, we infer the existence of an element U such that $U_m \to U$ weakly in V and strongly in H.

Returning to (3.14) and using the lower semicontinuity of weakly convergent sequences, we obtain that $|U|^2 + c_1\Delta t\|U\|^2 \le \langle \xi(t,U)+F, U\rangle$. To show that U satisfies $\langle \mathfrak{G}(t,U) - F, U^\sharp\rangle = 0$ for every $U^\sharp \in V_{(2)}$, we simply invoke (2.8) for B and the other continuity assumptions on A, E and ξ, and obtain this identity for $U^\sharp = \Psi_k$ for each $k \ge 1$. By linearity and density, we therefore infer the identify for arbitrary $U^\sharp \in V_{(2)}$. With this we now have established (i). The second item, (ii), to show that Λ is closed, follows immediately from the continuity of $\mathfrak{G}$ from $[0,T]\times V$ into $V'_{(2)}$ and the continuity of ξ from $[0,T]\times H$ into H. The proof of Lemma 3.1 is therefore complete.

Construction of an adapted solution

Step 1 We will build the desired sequence $\{U_N^n\}_{n=0}^N$ inductively as follows:

$$U_N^n = f_N^n(W|_{[0,t_n]}) \tag{3.15}$$

with $f_N^n : \mathcal{C}([0,t_n];\mathfrak{U}_0) \to V$ measurable for V equipped with $\mathcal{B}(V)$ and $\mathcal{C}([0,t_n];\mathfrak{U}_0)$ equipped with $\mathcal{G}_n := \mathcal{G}_{t_n}$ (defined as in Subsection 3.2).

Suppose that we have obtained U_N^{n-1} for some $n \ge 2$. Since $\mathcal{G}_{n-1}$ is the completion of $\mathcal{B}(\mathcal{C}([0,t_{n-1}];\mathfrak{U}))$ with respect to the Wiener measure $\mathbb{P}$,[8] f_N^{n-1} is $\mathbb{P}$-measurable. Now we define $\mathfrak{D}_N^n : V\times V'\times \mathcal{C}([0,t_n];\mathfrak{U}_0) \to V'$ by setting

$$\mathfrak{D}_N^n(x,y,z) = x + y\,\Delta t + \sigma_N(t^{n-1},U)z. \tag{3.16}$$

Then we can define

$$\begin{aligned} U_N^n &= \Gamma(t^n,\ \mathfrak{D}_N^n(U_N^{n-1},\, \ell_N^n,\, \eta_N^n)) \\ &:= \chi(t^n,\, U_N^{n-1},\, \ell_N^n,\, \eta_N^n). \end{aligned} \tag{3.17}$$

Since σ_N is a continuous map, clearly $\mathfrak{D}_N^n$ is a continuous map. Moreover, Γ is universally Radon measurable thanks to Lemma 3.1, hence Corollary A.1 applies and we infer that χ is universally Radon measurable from the Borel sigma algebra on $V\times V'\times\mathcal{C}([0,t_n];\mathfrak{U}_0)$ to the Borel sigma algebra on V.

Since $\ell = \ell(t)$ is a process assumed to be measurable with respect to the sigma algebra generated by the $W(s)$ for $s\in[0,t]$, ℓ_N^n is measurable with respect to the sigma algebra generated by the $W(s)$ for $s\in[0,t_n]$ thanks to (3.4). Hence by Theorem A.1 in the appendix

[8] We observe that the sigma algebra generated by the $W(s)$ for $s\in(0,t)$ is just $\phi_t^{-1}(\mathcal{B}(\mathcal{C}([0,T];\mathfrak{U}_0)))$, where $\phi_t : \mathcal{C}([0,T];\mathfrak{U}_0) \to \mathcal{C}([0,T];\mathfrak{U}_0)$ is the mapping $(\phi_t^{-1}\omega)(s) = \omega(t\wedge s)$, $0\le s\le T$ (see [42]).

with $\mathcal{X}$ as Ω, $(\mathcal{Y}, \mathcal{M})$ as $(\mathcal{C}([0,t_n];\mathfrak{U}_0), \mathcal{B}(\mathcal{C}([0,t_n];\mathfrak{U}_0)))$, ψ as $W|_{[0,t_n]}$, $\mathcal{H}$ as V, we see that there exists a function $L_N^n : \mathcal{C}([0,t_n];\mathfrak{U}_0) \to V$ which is Borel measurable, such that

$$\ell_N^n = L_N^n(W|_{[0,t_n]}). \tag{3.18}$$

From (3.17)–(3.18), we infer

$$\begin{aligned} U_N^n &= \kappa(t^n,\, f_N^{n-1}(W|_{[0,t_{n-1}]}),\, L_N^n(W|_{[0,t_n]}),\, \eta_N^n) \\ &:= f_N^n(W|_{[0,t_n]}). \end{aligned} \tag{3.19}$$

Since L_N^n and f_N^{n-1} are $\mathbb{P}$-measurable, and κ is universally Radon measurable, Theorem A.3 applies and we infer that f_N^n is $\mathbb{P}$-measurable, that is, f_N^n is measurable with respect to $\mathcal{G}_n$.

Step 2 We infer that $U_N^n : \Omega \to V$ is measurable with respect to $\mathcal{G}_n$ as desired.

Observe moreover that, according to Lemma 3.1 (see (3.12)), $\langle \mathfrak{G}(t_n, U_N^n), U^\sharp\rangle = \langle \mathfrak{D}_N^n(U_N^{n-1}, \ell_N^n, \eta_N^n), \widetilde{U}\rangle$, for every $U^\sharp \in V_{(2)}$ and $|U_N^n|^2 + \delta t c_1 \|U_N^n\|^2 \leq \langle \mathfrak{D}_N^n(U_N^{n-1}, \ell_N^n, \eta_N^n), U_N^n\rangle$ which is to say that U_N^{n-1} and U_N^n satisfy (3.3) and (3.9).

It remains to show that $U_N^n \in L^2(\Omega; V)$. We start from (3.9), now established for U_N^n and U_N^{n-1}, and use the elementary identity $2(U - U^\sharp, U) = |U|^2 - |U^\sharp|^2 + |U - U^\sharp|^2$ and obtain

$$\begin{aligned} &|U_N^n|^2 - |U_N^{n-1}|^2 + |U_N^n - U_N^{n-1}|^2 + 2\Delta t c_1 \|U_N^n\|^2 \\ &\leq 2\Delta t(\ell_N^n(U_N^n) + s(t^n, U_N^n, U_N^n)) + 2g_N(U_N^{n-1}, U_N^n)\eta_N^n, \end{aligned} \tag{3.20}$$

almost surely. To address the terms involving ℓ, we have that (see (3.4))

$$|2\Delta t \ell_N^n(U_N^n)| \leq 2\int_{(n-1)\Delta t}^{n\Delta t} \|\ell(t)\|_{V'} \|U_N^n\|_V \mathrm{d}t \leq c_1 \Delta t \|U_N^n\|^2 + c_1^{-1}\zeta_N^n,$$

where we define ζ_N^n according to

$$\zeta_N^n = \int_{(n-1)\Delta t}^{n\Delta t} \|\ell\|_{V'}^2 \mathrm{d}t. \tag{3.21}$$

For the terms involving s defined as in (2.12), we simply infer from (2.11)

$$2\Delta t|s(t^n, U_N^n, U_N^n)| \leq 4\Delta t c_4(1 + |U_N^n|^2). \tag{3.22}$$

With Hölder's inequality, we find

$$|2g_N(U_N^{n-1}, U_N^n - U_N^{n-1})\eta_N^n| \leq \frac{1}{2}|U_N^n - U_N^{n-1}|^2 + 2|\sigma_N(U_N^{n-1})\eta_N^n|^2. \tag{3.23}$$

Then using that g_N is linear in its second argument, we have

$$\begin{aligned} g_N(U_N^{n-1}, U_N^n)\eta_N^n &= g_N(U_N^{n-1}, U_N^{n-1})\eta_N^n + g_N(U_N^{n-1}, U_N^n - U_N^{n-1})\eta_N^n \\ &\leq g_N(U_N^{n-1}, U_N^{n-1})\eta_N^n + |g_N(U_N^{n-1}, U_N^n - U_N^{n-1})\eta_N^n| \quad \text{(thanks to (3.23))} \\ &\leq g_N(U_N^{n-1}, U_N^{n-1})\eta_N^n + \frac{1}{2}|U_N^n - U_N^{n-1}|^2 + 2|\sigma_N(U_N^{n-1})\eta_N^n|^2. \end{aligned} \tag{3.24}$$

Using these observations for g_N, ℓ_N^n and s, we rearrange and infer that, up to a set of measure zero,

$$\begin{aligned}&|U_N^n|^2 - |U_N^{n-1}|^2 + \frac{1}{2}|U_N^n - U_N^{n-1}|^2 + \Delta t c_1 \|U_N^n\|^2 \\ &\le c_1^{-1}\zeta_N^n + 4\Delta t c_4(1+|U_N^n|^2) + 2g_N(U_N^{n-1}, U_N^{n-1})\eta_N^n + 2|\sigma_N(U_N^{n-1})\eta_N^n|^2. \end{aligned} \tag{3.25}$$

Using (2.10), (3.6) and that U_N^{n-1} is $\mathcal{G}_{n-1}$-measurable, in $L^2(\Omega; H)$, we have

$$\begin{aligned}&\mathbb{E}g_N(U_N^{n-1}, U_N^{n-1})\eta_N^n = 0,\\ &\mathbb{E}|\sigma_N(U_N^{n-1})\eta_N^n|^2 = \Delta t\mathbb{E}|\sigma_N(U_N^{n-1})|^2_{L_2(\mathfrak{U},H)} \le 2\Delta t c_3^2\mathbb{E}(1+|U_N^{n-1}|^2).\end{aligned}$$

From this observation, (3.25) and (3.10), we infer

$$\mathbb{E}\Delta t c_1\|U_N^n\|^2 \le \mathbb{E}((4\Delta t c_4 - 1)|U_N^n|^2 + c(|U_N^{n-1}|^2 + \zeta_N^n + 1)) \le c\mathbb{E}(|U_N^{n-1}|^2 + \zeta_N^n + 1),$$

which implies that $U_N^n \in L^2(\Omega; V)$, as needed.

We have thus established the iterative step in the construction of $\{U_N^n\}_{n=0}^N$. The base case, $n = 1$, is established in an identical fashion to the iterative steps. The proof of Proposition 3.1 is now complete.

Remark 3.2 Although necessary for the establishment of the existence of the U_N^n's in Proposition 3.1, it is not necessary to assume the underlying stochastic basis to be $\mathcal{S}_\mathcal{G}$ (defined in Subsection 3.2) in the results throughout Subsection 3.3 to Subsection 5.1. The reason is that these results are true whenever such U_N^n's defined as in Definition 3.1 exist. In other words they are independent of the choice of the underlying stochastic basis. Similarly, it is not necessary at this point to assume that U^0 and ℓ have laws which coincide with those of the externally given μ_{U^0} and μ_ℓ for these results.

However, it is necessary that we resume these assumptions of $\mathcal{S}_\mathcal{G}$, μ_{U^0} and μ_ℓ starting in Subsection 5.2.

3.3 Uniform "energy" estimates for the U_N^n

Starting from (3.9) we next determine certain uniform bounds, independent of N, for (suitable) sequences $\{U_N^n\}_{n=1}^N$ satisfying (3.3) as follows.

Proposition 3.2 *Let*

$$N_1 := 12Tc_5 \quad \text{with } c_5 := 8c_4 + 80c_3^2, \tag{3.26}$$

where c_3 and c_4 are from (2.10) and (2.11), respectively. Let $\mathcal{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, \mathbb{P}, \{W^k\}_{k\ge1})$ be the given stochastic basis and assume that $\ell = \ell(t) \in L^2(\Omega; L^2(0,T;V'))$ is measurable with respect to $\mathcal{F}_t$. For each $N \ge N_1$, we assume that $U_N^0 \in L^2(\Omega, H)$, is $\mathcal{F}_0$ measurable and such that

$$\sup_{N\ge N_1} \mathbb{E}|U_N^0|^2 < \infty. \tag{3.27}$$

Then for each $N \geq N_1$, consider the sequences $\{U_N^n\}_{n=1}^N \subset L^2(\Omega; V)$ which satisfy (3.3) *starting from U_N^0 and relative to ℓ in the sense of Definition* 3.1. *Then*

$$\sup_{N \geq N_1} \mathbb{E}\Big(\max_{0 \leq l \leq N} |U_N^l|^2 + \sum_{k=1}^N (|U_N^k - U_N^{k-1}|^2 + \Delta t \|U_N^k\|^2) \Big) < \infty. \tag{3.28}$$

Proof The starting point for the estimates leading to (3.28) is of course (3.9) and from this inequality, we can use the same proof as in Proposition 3.1 to obtain (3.25). In order to make suitable estimates for the final two terms in (3.25), we need to take advantage of some martingale structure in the terms involving σ_N. For any $1 \leq m \leq n \leq N$, we define the stochastic processes

$$M_N^{m,n} := \sum_{k=m}^n g_N(U_N^{k-1}, U_N^{k-1}) \eta_N^k, \quad Q_N^{m,n} := \sum_{k=m}^n |\sigma_N(U_N^{k-1}) \eta_N^k|^2. \tag{3.29}$$

Summing (3.25) for $1 \leq m \leq n = k \leq l \leq N$, we find

$$\begin{aligned} &|U_N^l|^2 + \sum_{k=m}^l \Big(\frac{1}{2} |U_N^k - U_N^{k-1}|^2 + \Delta t c_1 \|U_N^k\|^2 \Big) \\ &\leq |U_N^{m-1}|^2 + \sum_{k=m}^l (c_1^{-1} \zeta_N^k + 4 \Delta t c_4 (1 + |U_N^k|^2)) + 2 M_N^{m,l} + 2 Q_N^{m,l}. \end{aligned} \tag{3.30}$$

Since $\{U_N^n\}_{n=0}^N \subset L^2(\Omega; H)$ is adapted to $\mathcal{F}_n := \mathcal{F}_{t^n}$, it is easy to see that $\{M_N^{m,n}\}_{n=m}^N$ is a martingale relative to $\{\mathcal{F}_n\}_{n=m}^N$ with $M_N^{m,m} \equiv 0$. We would like to apply a discrete version of the Burkholder-Davis-Gundy inequality, recalled here as in Lemma 3.2 to obtain estimates for $\mathbb{E} \max_{m \leq l \leq n} |M_N^{m,l}|$. Unfortunately, it is not clear that $\{M_N^{m,n}\}_{n=m}^N$ is square integrable, so we have to apply a localization argument to make proper use of this inequality. For any $K > 0$, we define the stopping times

$$\widetilde{n}_K = \min_{l \geq m} \{ |U_N^{l-1}| \geq K \} \wedge N.$$

Since

$$\{U_N^n\}_{n=0}^N \subset L^2(\Omega; H),$$

we have that $\widetilde{n}_K \uparrow N$ almost surely as $K \uparrow \infty$. Clearly $\{M_N^{m,n \wedge \widetilde{n}_K}\}_{n=m}^N$ is a square-integrable martingale. For the moment, let us recall a discrete analogue of the Burkholder-Davis-Gundy inequality. This result and other related martingale inequalities can be found in e.g. [22].

Lemma 3.2 *Assume that $\{M^n\}_{n \geq 0}$ is a (discrete) martingale on a Hilbert space $\mathcal{H}$ (with norm $|\cdot|$), relative to a given filtration $\{\mathcal{F}_n\}_{n \geq 0}$. We assume, additionally that $M_0 \equiv 0$ and that $\mathbb{E}|M_n|^2 < \infty$, for all $n \geq 0$. Then, for any $q \geq 1$ and any $n \geq 1$,*

$$\mathbb{E} \max_{1 \leq m \leq n} |M^n|^q \leq c_q \mathbb{E} (A^n)^{\frac{q}{2}}, \tag{3.31}$$

where c_q is a universal positive constant depending only on q[9] *(which is independent of n and $\{M^m\}_{m\geq 0}$), and A^n is the quadratic variation defined by*

$$A^n = \sum_{m=1}^{n} \mathbb{E}(|M^m - M^{m-1}|^2 | \mathcal{F}_{m-1}). \tag{3.32}$$

Hence with the observation that $\mathbb{1}_{\widetilde{n}_K \geq k}$ is $\mathcal{F}_{k-1}$-measurable, we compute the quadratic variation of $\{M_N^{m,n\wedge\widetilde{n}_K}\}_{n=m}^N$ in view of (3.32) as follows:

$$\begin{aligned}
A_N^{m,n} &= \sum_{k=m}^{n} \mathbb{E}(|M_N^{m,k\wedge\widetilde{n}_K} - M_N^{m,(k-1)\wedge\widetilde{n}_K}|^2 | \mathcal{F}_{k-1}) \\
&= \sum_{k=m}^{n} \mathbb{E}(\mathbb{1}_{\widetilde{n}_K \geq k} |g_N(U_N^{k-1}, U_N^{k-1})\eta_N^k|^2 | \mathcal{F}_{k-1}) \\
&= \sum_{k=m}^{\widetilde{n}_K \wedge n} |g_N(U_N^{k-1}, U_N^{k-1})|^2 \Delta t.
\end{aligned}$$

Thus, by Lemma 3.2, (2.10) and (3.6), we infer

$$\begin{aligned}
\mathbb{E} \max_{m\leq l\leq n} |M_N^{m,l\wedge\widetilde{n}_K}| &\leq 3\mathbb{E}\Big(\sum_{k=m}^{n\wedge\widetilde{n}_K} |g_N(U_N^{k-1}, U_N^{k-1})|_{L_2}^2 \Delta t\Big)^{\frac{1}{2}} \\
&\leq 3\mathbb{E}\Big(\sum_{k=m}^{n} |\sigma_N(U_N^{k-1})|_{L_2(\mathfrak{U},H)}^2 |U_N^{k-1}|^2 \Delta t\Big)^{\frac{1}{2}} \\
&\leq 3c_3\mathbb{E}\Big(\sum_{k=m}^{n} 2(1 + |U_N^{k-1}|^2)|U_N^{k-1}|^2 \Delta t\Big)^{\frac{1}{2}} \\
&\leq \frac{1}{4}\mathbb{E} \max_{m\leq k\leq n} |U_N^{k-1}|^2 + 18c_3^2 \mathbb{E} \sum_{k=m}^{n} (1 + |U_N^{k-1}|^2)\Delta t.
\end{aligned}$$

Hence, letting $K \uparrow \infty$, we have, by the monotone convergence theorem,

$$\mathbb{E} \max_{m\leq l\leq n} |M_N^{m,l}| \leq \frac{1}{4}\mathbb{E} \max_{m\leq k\leq n} |U_N^{k-1}|^2 + 18c_3^2 \Delta t \mathbb{E} \sum_{k=m}^{n} (1 + |U_N^{k-1}|^2). \tag{3.33}$$

On the other hand, since U_N^n is adapted to $\mathcal{F}_n$, given the condition (2.10) on σ and (3.6), we infer that

$$\mathbb{E}Q_N^{m,n} = \sum_{k=m}^{n} \mathbb{E}|\sigma_N(U_N^{k-1})|_{L_2(\mathfrak{U},H)}^2 \Delta t \leq 2c_3^2 \Delta t \mathbb{E} \sum_{k=m}^{n} (1 + |U_N^{k-1}|^2). \tag{3.34}$$

We now use (3.33)–(3.34) with (3.30) and infer that

$$\begin{aligned}
\mathbb{E} \max_{m\leq l\leq n} |U_N^l|^2 &\leq \mathbb{E}\Big(2|U_N^{m-1}|^2 + \sum_{k=m}^{n} (c_1^{-1}\zeta_N^k + 4c_4\Delta t(1 + |U_N^k|^2)) + 2 \max_{m\leq l\leq n} |M_N^{m,l}| + 2Q_N^{m,n}\Big) \\
&\leq \mathbb{E}\Big(2|U_N^{m-1}|^2 + \sum_{k=m}^{n} (c_1^{-1}\zeta_N^k + 4c_4\Delta t(1 + |U_N^k|^2)) + 40c_3^2 \Delta t \sum_{k=m}^{n} (1 + |U_N^{k-1}|^2)\Big) \\
&\quad + \frac{1}{2}\mathbb{E} \max_{m\leq k\leq n} |U_N^{k-1}|^2.
\end{aligned}$$

[9]We may often determine c_q in (3.31) explicitly, and in particular, we have that $c_1 = 3$.

Rearranging it, we find that

$$\begin{aligned}\mathbb{E}\max_{m\le l\le n}|U_N^l|^2 &\le \mathbb{E}\Big(2|U_N^{m-1}|^2 + 2c_1^{-1}\sum_{k=m}^{n}\zeta_N^k + c_5\Delta t\,\mathbb{E}\sum_{k=m}^{n+1}(1+|U_N^{k-1}|^2)\Big)\\ &\le \mathbb{E}\Big(2|U_N^{m-1}|^2 + 2c_1^{-1}\sum_{k=m}^{n}\zeta_N^k + c_5\Delta t(n-m+2)\\ &\quad\cdot(1+\mathbb{E}\max_{m\le k\le n+1}|U_N^{k-1}|^2)\Big) \qquad (3.35)\end{aligned}$$

for the constant $c_5 = 8c_4 + 80c_3^2$ which in particular depends only on c_3, c_4. Thus, subject to the condition

$$c_5\Delta t(n-m+2) \le \frac{1}{2}, \quad \text{i.e.,} \quad \frac{n-m+2}{N} \le \frac{1}{2c_5T}, \qquad (3.36)$$

we have

$$\mathbb{E}\max_{m\le l\le n}|U_N^l|^2 \le c_6\mathbb{E}\Big(|U_N^{m-1}|^2 + \sum_{k=m}^{n}\zeta_n^k + 1\Big), \qquad (3.37)$$

where $c_6 = \max\{4c_1^{-1}, 7\}$. Thus, by iterating this inequality and noting from (3.21) that

$$\sum_{k=1}^{N}\zeta_n^k = \|\ell\|^2_{L^2(0,T;V')},$$

we finally conclude that

$$\mathbb{E}\max_{1\le l\le N}|U_N^l|^2 \le c_7\mathbb{E}(|U_N^0|^2 + \|\ell\|^2_{L^2(0,T;V')} + 1) \quad \text{for all } N \ge N_1. \qquad (3.38)$$

Note carefully that, in view of (3.36), we need not iterate (3.37) more than, say, $\ulcorner 16c_5T\urcorner$ times to obtain (3.38).[10] As such, we may take $c_7 = (1+c_6)^{16c_5T} = (1+\max\{4c_1^{-1}, 7\})^{16T(8c_4+80c_3^2)}$, which, crucially, is independent of N.

We now return to (3.30). With (3.34), we infer

$$\begin{aligned}\mathbb{E}\sum_{k=1}^{N}(|U_N^k - U_N^{k-1}|^2 + 2c_1\Delta t\|U_N^k\|^2) &\le \mathbb{E}\Big(|U_N^0|^2 + \sum_{k=1}^{N}(c_1^{-1}\zeta_N^k + 4c_4\Delta t(1+|U_N^k|^2))\\ &\quad + 4c_3^2\Delta t\sum_{k=1}^{N}(1+|U_N^{k-1}|^2)\Big)\\ &\le c_8\mathbb{E}\Big(|U_N^0|^2 + \max_{1\le l\le N}|U_N^l|^2 + \|\ell\|^2_{L^2(0,T;V')} + 1\Big), \qquad (3.39)\end{aligned}$$

where we can take $c_8 = \max\{1, c_1^{-1}, 4T(c_3^2 + c_4)\}$. As such, (3.38)–(3.39) with (3.27) imply (3.28), completing the proof of Proposition 3.2.

[10] Indeed, for $N \ge N_1$, let $\mathfrak{N}(N)$ be the minimum number of iterations of (3.37), subject to the constraint (3.36), which are needed to establish (3.38). Take $\mathfrak{F}(N)$ to be the "fraction of the time interval that can be covered at each step", namely,

$$\mathfrak{F}(N) := \max_{n\in\mathbb{N}}\Big\{\frac{n}{N} : n+2 \le \frac{N}{2c_5T}\Big\} > \frac{1}{2c_5T} - \frac{3}{N} \ge \frac{1}{4c_5T},$$

where the last inequality follows from the standing assumption (3.26). Since $\mathfrak{N}(N)\mathfrak{F}(N) \le 2$, we finally estimate

$$\mathfrak{N}(N) \le \frac{2}{\mathfrak{F}(N)} \le 16c_5T.$$

Here $\ulcorner p\urcorner$ = the smallest integer that is larger than or equal to p.

4 Continuous Time Approximations and Uniform Bounds

In this section, we detail how the sequences $\{U_N^n\}_{n=0}^N$ defined in the sense of Definition 3.1 may be used to define continuous time processes that approximate (2.13). The details of establishing the compactness of the associated sequences of probability laws and of the passage to the limit are given further on in Section 5.

We now fix sequences $\{U_N^n\}_{n=0}^N$ satisfying (3.3) in the sense of Definition 3.1. For $N \geq N_1$, with N_1 as in (3.26), let

$$U_N(t) = \begin{cases} U_N^0 & \text{for } t \in [0, t^1], \\ U_N^n & \text{for } t \in (t^n, t^{n+1}],\ n = 1, \cdots, N-1. \end{cases} \tag{4.1}$$

Of course, we do not have any time derivatives of the U_N's (even fractional in time) as are typically needed for compactness. Furthermore, we would like to be able to associate an approximate stochastic equation for (2.13) with these $\{U_N^n\}_{n=0}^N$'s. For these dual concerns, we introduce further stochastic processes and consider

$$\overline{U}_N(t) = \begin{cases} U_N^0 & \text{for } t \in [0, t^1] \\ U_N^{n-1} + \frac{U_N^n - U_N^{n-1}}{\Delta t}(t - t^n) & \text{for } t \in (t^n, t^{n+1}],\ n = 1, \cdots, N-1. \end{cases} \tag{4.2}$$

Remark 4.1 The processes U_N and $\overline{U}_N$ are slightly different than those typically used in the deterministic case (see, e.g., [71]). Actually, these processes are essentially their deterministic analogues evaluated at time t by their value at time $t - \Delta t$. With this choice, we crucially obtain processes which are adapted to $\{\mathcal{F}_t\}_{t\geq 0}$. Not surprisingly however the present definitions of $U_N, \overline{U}_N$ leads to bothersome error terms in (4.6) below. In turn these error terms dictate the additional convergences in σ and U^0 when we initially defined the discrete scheme (3.3) (see (3.5)–(3.7) and Remark 3.1). These error terms also complicate compactness arguments further in Section 5 (see Remark 4.2).

The rest of this section is now devoted to proving the following desirable properties of U_N and $\overline{U}_N$:

Proposition 4.1 *Let $\mathcal{S} = (\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}, \{W^k\}_{k\geq 1})$ be a stochastic basis, and let N_1 be as in* (3.26) *in Proposition* 3.2. *Consider a sequence $\{U_N^0\}_{N\geq N_1}$ bounded in $L^2(\Omega, H)$ independent of N, with U_N^0 $\mathcal{F}_0$-measurable for each N and such that*

$$\mathbb{E}((1 + \|U_N^0\|^2)(1 + \|U_N^0\|_{V_{(2)}}^2)) \leq c\Delta t^{-1} = cN \tag{4.3}$$

for a constant $c > 0$, independent of N.[11] *Suppose that we also have defined a process $\ell = \ell(t) \in L^2(\Omega; L^2(0, T; V'))$ adapted to $\{\mathcal{F}_t\}_{t\geq 0}$.*

For each $N \geq N_1$, we consider sequences $\{U_N^n\}_{n=1}^N$ which satisfy (3.3) *starting from U_N^0 in the sense of Definition* 3.1. *Once these sequences $\{U_N^n\}_{n=0}^N$ exist, then we define the continuous time processes $\{U_N\}_{N\geq 1}$ and $\{\overline{U}_N\}_{N\geq 1}$ according to* (4.1) *and* (4.2)*, respectively. Then,*

[11] The constraint (4.3) is necessary for (4.4)–(4.7). This is not a serious restriction when we pass to the limit in Section 5. As we described above in Remark 3.1, for any given $U^0 \in L^2(\Omega; H)$ we may obtain a sequence U_N^0 approximating U^0 which maintains (4.3).

(i) *for each* $N \geq N_1$, U_N *and* $\overline{U}_N$ *are* $\{\mathcal{F}_t\}_{t\geq 0}$*-adapted and*

$$\{U_N\}_{N\geq N_1} \text{ and } \{\overline{U}_N\}_{N\geq N_1} \text{ are bounded in } L^2(\Omega; L^2(0,T;V)\cap L^\infty(0,T;H)). \tag{4.4}$$

Moreover, we have that

$$\lim_{N\uparrow\infty} \mathbb{E}\int_0^T |U_N - \overline{U}_N|^2 \mathrm{d}t = 0. \tag{4.5}$$

(ii) U_N *and* $\overline{U}_N$ *satisfy a.s. and for every* $t \geq 0$,

$$\overline{U}_N(t) = U_N^0 + \int_0^t (\mathcal{N}(U_N) + \ell_N)\mathrm{d}s + \int_0^t \sigma_N(U_N)\mathrm{d}W + \mathcal{E}_N^D(t) + \mathcal{E}_N^S(t), \tag{4.6}$$

subject to error terms $\mathcal{E}_N^D(t) \in L^2(\Omega; L^2(0,T;V'))$, $\mathcal{E}_N^S(t) \in L^2(\Omega; L^2(0,T;H))$ *which are defined explicitly in* (4.15)–(4.16) *below.*

(iii) *These error terms* $\mathcal{E}_N^D(t)$, $\mathcal{E}_N^S(t)$ *satisfy*

$$\lim_{N\uparrow\infty} \mathbb{E}\|\mathcal{E}_N^D\|^2_{L^2(0,T;V')} = 0, \tag{4.7}$$

$$\lim_{N\uparrow\infty} \mathbb{E}\|\mathcal{E}_N^S\|^2_{L^2(0,T;H)} = 0, \tag{4.8}$$

respectively, and moreover,

$$\sup_{N\geq N_1} \mathbb{E}\|\mathcal{E}_N^S\|^2_{L^\infty(0,T;H)\cap L^2(0,T;V)} < \infty. \tag{4.9}$$

We proceed to prove Proposition 4.1 in a series of subsections below. The proof of (i) is essentially a direct application of Proposition 3.2, and we provide the details in the subsection immediately following. In Subsection 4.2, we provide the details of the derivation of (4.6) and in particular explain the origin of the error terms $\mathcal{E}_N^D, \mathcal{E}_N^S$. The final Subsection 4.3 provides details of the estimates for these error terms which lead to (4.7)–(4.9).

Remark 4.2 It is not straightforward to obtain fractional in time estimates for $\overline{U}_N$ from (4.6) in view of the error terms which have a rather complicated structure (see (4.15)–(4.16)). As such, we cannot establish sufficient compactness for the sequence $\overline{U}_N$ directly to facilitate the passage to the limit. For this reason, we choose to introduce additional continuous time processes in Section 5 below. An alternate approach will be presented later on in the related work [33].

4.1 Uniform bounds and clustering

It is clear from (4.1) that U_N is $\{\mathcal{F}_t\}_{t\geq 0}$-adapted and that

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |U_N|^2 + \int_0^T \|U_N\|^2 \mathrm{d}t\Big) = \mathbb{E}\Big(\max_{0\leq m\leq N-1} |U_N^m|^2 + \sum_{m=0}^{N-1} \Delta t \|U_N^m\|^2\Big).$$

Thus, since (3.27) holds, we have the uniform bound (3.28) from Proposition 3.2, and we immediately infer that

$$\sup_{N\geq N_1} \mathbb{E}\Big(\sup_{t\in[0,T]} |U_N|^2 + \int_0^T \|U_N\|^2 \mathrm{d}t\Big) < \infty \tag{4.10}$$

with the integer N_1 appearing in (3.26).

As the U_N above, it is easy to see from (4.2) that $\overline{U}_N$ is adapted to $\{\mathcal{F}_t\}_{t\geq 0}$, and that $\{U_N^n\}_{n=1}^N$ is adapted to $\mathcal{F}_n(=\mathcal{F}_{t^n})$. Furthermore, direct calculations show that

$$U_N - \overline{U}_N(t) = \begin{cases} 0 & \text{for } t \in [0, t^1], \\ \dfrac{U_N^n - U_N^{n-1}}{\Delta t}(t^{n+1} - t) & \text{for } t \in (t^n, t^{n+1}],\ n = 1, \cdots, N-1. \end{cases} \tag{4.11}$$

Using (4.11), similarly to [71], we compute that

$$\mathbb{E}\int_0^T |U_N - \overline{U}_N|^2 \mathrm{d}t = \sum_{n=1}^{N-1} \mathbb{E}|U_N^n - U_N^{n-1}|^2 \int_{t^n}^{t^{n+1}} \Big(\frac{t^{n+1} - t}{\Delta t}\Big)^2 \mathrm{d}t = \frac{\Delta t}{3}\mathbb{E}\sum_{n=1}^{N} |U_N^n - U_N^{n-1}|^2.$$

We thus infer (4.5) directly from this observation and (3.28). Based on similar considerations, we also have

$$\mathbb{E}\int_0^T \|\overline{U}_N\|^2 \leq c\Delta t \mathbb{E}\sum_{n=0}^{N} \|U_N^n\|^2 = c\Delta t\mathbb{E}\|U_N^0\|^2 + \Delta t\mathbb{E}\sum_{n=1}^{N}\|U_N^n\|^2.$$

Thus, once again due to (4.3) and (3.28), we finally have

$$\sup_{N\geq N_1} \mathbb{E}\Big(\sup_{t\in[0,T]} |\overline{U}_N|^2 + \int_0^T \|\overline{U}_N\|^2 \mathrm{d}t\Big) < \infty. \tag{4.12}$$

With (4.10) and (4.12), we have now established the first item in Proposition 4.1.

4.2 The approximate stochastic evolution systems

We next derive the equation (4.6) relating U_N and $\overline{U}_N$ giving explicit expressions for $\mathcal{E}_N^D, \mathcal{E}_N^S$. We observe that, almost surely and for almost every $t \geq 0$ (in fact for every $t \notin \{t_0, t_1, \cdots, t_N\}$)

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{U}_N(t) = \sum_{n=1}^{N-1} \frac{U_N^n - U_N^{n-1}}{\Delta t}\chi_{(t^n, t^{n+1})}(t), \tag{4.13}$$

where $\chi(t_1, t_2)$ denotes the indicator function of (t_1, t_2). Recall that $\eta_N^n = W(t^n) - W(t^{n-1})$ and let $N_*^t := \min\{n : t^n \geq t\}$ in other words, we take N_*^t such that

$$N_*^t \Delta t \leq t < (N_*^t + 1)\Delta t.$$

Working from (4.13) and (3.3), we therefore compute

$$\begin{aligned} \overline{U}_N(t) &= U_N^0 + \int_0^t \sum_{n=1}^{N-1} \frac{U_N^n - U_N^{n-1}}{\Delta t}\chi_{(t^n, t^{n+1})}(s)\mathrm{d}s \\ &= U_N^0 + \int_0^t \sum_{n=1}^{N-1} (\mathcal{N}(U_N^n) + \ell_N^n)\chi_{(t^n, t^{n+1})}(s)\mathrm{d}s + \int_0^t \sum_{n=1}^{N-1} \sigma_N(U_N^{n-1})\frac{\eta_N^n}{\Delta t}\chi_{(t_n, t_{n+1})}(s)\mathrm{d}s \\ &= U_N^0 + \int_0^t (\mathcal{N}(U_N) + \ell_N)\mathrm{d}s + \int_0^t \sigma_N(U_N)\mathrm{d}W + \mathcal{E}_N^D(t) + \mathcal{E}_N^S(t), \end{aligned} \tag{4.14}$$

where the "error terms", $\mathcal{E}_N^D(t)$ and $\mathcal{E}_N^S(t)$, are defined as

$$\mathcal{E}_N^D(t) := -\mathcal{N}(U_N^0)\Delta t \wedge t - \Big(\int_{t^{N_*^t-1}}^t \ell_N \mathrm{d}s + \ell_N^{N_*^t-1}(t^{N_*^t} - t)\chi_{t>t^1} \Big) = \mathcal{E}_N^{D,1}(t) + \mathcal{E}_N^{D,2}(t), \quad (4.15)$$

and

$$\mathcal{E}_N^S(t) := -\sigma_N(U_N^{N_*^t-2}) \frac{\eta_N^{N_*^t-1}}{\Delta t}(t^{N_*^t} - t)\chi_{t>t^1} - \int_{t^{N_*^t-1}}^t \sigma_N(U_N)\mathrm{d}W := \mathcal{E}_N^{S,1}(t) + \mathcal{E}_N^{S,2}(t), \quad (4.16)$$

respectively. To understand the origin of these error terms, we observe that

$$\begin{aligned}\int_0^t \sum_{n=1}^{N-1} \mathcal{N}(U_N^n)\chi_{(t^n,t^{n+1})}(s)\mathrm{d}s &= \int_0^t \sum_{n=0}^{N-1} \mathcal{N}(U_N^n)\chi_{(t^n,t^{n+1})}(s)\mathrm{d}s - \mathcal{N}(U_N^0)\Delta t \wedge t \\ &= \int_0^t \mathcal{N}(U_N)\mathrm{d}s + \mathcal{E}_N^{D,1}(t).\end{aligned}$$

Moreover, using the definition of the ℓ_N^n's in (3.4), we have

$$\begin{aligned}\int_0^t \sum_{n=1}^{N-1} \ell_N^n \chi_{(t^n,t^{n+1})}(s)\mathrm{d}s &= \int_0^{t^{N_*^t}} \sum_{n=1}^{N-1} \ell_N^n \chi_{(t^n,t^{n+1})}(s)\mathrm{d}s + \Big(\int_{t^{N_*^t}}^t \ell_N^{N_*^t-1}\mathrm{d}s\Big)\chi_{t>t^1} \\ &= \sum_{n=1}^{N_*^t-1} \ell_N^n \Delta t + \ell_N^{N_*^t-1}(t - t^{N_*^t})\chi_{t>t^1} \\ &= \int_0^{t^{N_*^t-1}} \ell \mathrm{d}s + \ell_N^{N_*^t-1}(t - t^{N_*^t})\chi_{t>t^1} \\ &= \int_0^t \ell \mathrm{d}s - \int_{t^{N_*^t-1}}^t \ell \mathrm{d}s + \ell_N^{N_*^t-1}(t - t^{N_*^t})\chi_{t>t^1}.\end{aligned}$$

On the other hand, for the error terms $\mathcal{E}_N^S(t)$ involving σ_N in (4.16), we compute

$$\begin{aligned}\int_0^t \sum_{n=1}^{N-1} \sigma_N(U_N^{n-1})\frac{\eta_N^n}{\Delta t}\chi_{(t^n,t^{n+1})}(s)\mathrm{d}s &= \int_0^{t^{N_*^t}} \sum_{n=1}^{N_*^t-1} \sigma_N(U_N^{n-1})\frac{\eta_N^n}{\Delta t}\chi_{(t^n,t^{n+1})}(s)\mathrm{d}s \\ &\quad - \int_t^{t^{N_*^t}} \sigma_N(U_N^{N_*^t-2})\frac{\eta_N^{N_*^t-1}}{\Delta t}\mathrm{d}s\chi_{t>t^1} \\ &= \sum_{n=1}^{N_*^t-1} \sigma_N(U_N^{n-1})\eta_N^n - \sigma_N(U_N^{N_*^t-2})\frac{\eta_N^{N_*^t-1}}{\Delta t}(t^{N_*^t} - t)\chi_{t>t^1} \\ &= \int_0^{t^{N_*^t-1}} \sigma_N(U_N)\mathrm{d}W - \sigma_N(U_N^{N_*^t-2})\frac{\eta_N^{N_*^t-1}}{\Delta t}(t^{N_*^t} - t)\chi_{t>t^1} \\ &= \int_0^t \sigma_N(U_N)\mathrm{d}W + \mathcal{E}_N^S(t).\end{aligned}$$

4.3 The estimates for the error terms

We next proceed to make estimates on the error terms $\mathcal{E}_N^D$ and $\mathcal{E}_N^S$ as desired in (4.7), (4.9). Perusing (4.15), we begin with estimates for $\mathcal{E}_N^{D,1}$. Invoking the bounds provided by (2.7) along with the continuity properties of the other operators making up $\mathcal{N}$ in (2.14) defined in Subsection 2.1, we have

$$\mathbb{E}\sup_{t\in[0,T]}\|\mathcal{E}_N^{D,1}(t)\|_{V'}^2 \le \Delta t^2\mathbb{E}\|\mathcal{N}(U_N^0)\|_{V'}^2 \le c\Delta t^2\mathbb{E}((1+\|U_N^0\|^2)(1+\|U_N^0\|_{V_{(2)}}^2)).$$

As such, in view of the standing condition (4.3) (see Remark 3.1), we conclude that

$$\lim_{N\uparrow\infty}\mathbb{E}\|\mathcal{E}_N^{D,1}\|_{L^2(0,T;V')}^2 = \lim_{N\uparrow\infty}\mathbb{E}\|\mathcal{E}_N^{D,1}\|_{L^\infty(0,T;V')}^2 = 0. \tag{4.17}$$

For $\mathcal{E}_N^{D,2}$, we estimate in $L^2(0,T;V')$

$$\begin{aligned}\int_0^T\Big\|\int_{t^{N_*^t-1}}^t \ell\mathrm{d}s\Big\|_{V'}^2\mathrm{d}t &\le \int_0^T\int_{t^{N_*^t-1}}^t\|\ell\|_{V'}^2\mathrm{d}s(t-t^{N_*^t-1})\mathrm{d}t\\ &= \sum_{k=1}^{N-1}\int_{t^{k-1}}^{t^k}\int_{t^{k-1}}^t\|\ell\|_{V'}^2\mathrm{d}s(t-t^{k-1})\mathrm{d}t\\ &\le c\Delta t^2\int_0^T\|\ell\|_{V'}^2\mathrm{d}t,\\ \int_0^T\|\ell_N^{N_*^t-1}(t^{N_*^t}-t)\chi_{t>t^1}\|_{V'}^2\mathrm{d}t &= \sum_{k=1}^{N-1}\|\ell_N^k\|_{V'}^2\int_{t^k}^{t^{k+1}}(t^{k+1}-t)^2\mathrm{d}t\\ &\le \frac{\Delta t}{3}\sum_{k=1}^N\Big\|\int_{t^k}^{t^{k+1}}\ell\mathrm{d}s\Big\|_{V'}^2 \le \frac{\Delta t^2}{3}\int_0^T\|\ell\|_{V'}^2\mathrm{d}t.\end{aligned}$$

In summary, we have

$$\lim_{N\uparrow\infty}\mathbb{E}\|\mathcal{E}_N^{D,2}\|_{L^2(0,T;V')}^2 = 0 \tag{4.18}$$

and so we conclude (4.7) from (4.17)–(4.18).

We next turn to make estimates for $\mathcal{E}_N^S$. We begin with estimates in $L^2(0,T;H)$. For $\mathcal{E}_N^{S,1}$, we observe with (2.10) and (3.6) (see (3.34)) that

$$\begin{aligned}\mathbb{E}\int_0^T|\mathcal{E}_N^{S,1}|^2\mathrm{d}t &= \sum_{k=1}^{N-1}\mathbb{E}\Big|\sigma_N(U_N^{k-1})\frac{\eta_N^k}{\Delta t}\Big|^2\int_{t^k}^{t^{k+1}}(t^{k+1}-t)^2\mathrm{d}t = \frac{\Delta t}{3}\sum_{k=1}^{N-1}\mathbb{E}|\sigma_N(U_N^{k-1})\eta_N^k|^2\\ &= \frac{\Delta t}{3}\sum_{k=1}^{N-1}\mathbb{E}|\sigma_N(U_N^{k-1})|_{L^2(\mathfrak{U},H)}^2\Delta t \le c\Delta t\sum_{k=1}^{N-1}\mathbb{E}(1+|U_N^{k-1}|^2)\Delta t,\end{aligned}$$

and infer from (3.28) in Proposition 3.2 that

$$\lim_{N\uparrow\infty}\mathbb{E}\|\mathcal{E}_N^{S,1}\|_{L^2(0,T;H)}^2 = 0. \tag{4.19}$$

On the other hand, with the Itô isometry and another application of (2.10) and (3.6), we have

$$\begin{aligned}\mathbb{E}\int_0^T|\mathcal{E}_N^{S,2}|^2\mathrm{d}t &= \sum_{k=1}^{N-1}\mathbb{E}\int_{t^k}^{t^{k+1}}\Big|\int_{t^k}^t\sigma_N(U_N)\mathrm{d}W\Big|^2\mathrm{d}t = \sum_{k=1}^{N-1}\mathbb{E}\int_{t^k}^{t^{k+1}}\int_{t^k}^t|\sigma_N(U_N)|_{L^2(\mathfrak{U},H)}^2\mathrm{d}s\mathrm{d}t\\ &= \sum_{k=1}^{N-1}\mathbb{E}|\sigma_N(U_N^{k-1})|_{L^2(\mathfrak{U},H)}^2\int_{t^k}^{t^{k+1}}(t-t_k)\mathrm{d}t \le c\Delta t\sum_{k=1}^{N-1}\mathbb{E}(1+|U_N^{k-1}|^2)\Delta t,\end{aligned}$$

so that

$$\lim_{N\uparrow\infty} \mathbb{E}\|\mathcal{E}_N^{S,2}\|^2_{L^2(0,T;H)} = 0. \tag{4.20}$$

By combining (4.19)–(4.20), we obtain (4.8).

We turn now to establishing the uniform bounds announced in (4.9). Estimates similar to those leading to (4.19)–(4.20), but which instead make use of the condition (3.5), yield bounds in $L^2(0,T;V)$, namely,

$$\begin{aligned}\mathbb{E}\int_0^T \|\mathcal{E}_N^{S,1}\|^2 \mathrm{d}t &= \frac{\Delta t}{3}\sum_{k=1}^{N-1} \mathbb{E}\|\sigma_N(U_N^{k-1})\|^2_{L^2(\mathfrak{U},V)}\Delta t \\ &\le \frac{T}{3}\sum_{k=1}^{N-1} \mathbb{E}|\sigma(U_N^{k-1})|^2_{L^2(\mathfrak{U},V)}\Delta t \\ &\le c\sum_{k=1}^{N-1} \mathbb{E}(1+|U_N^{k-1}|^2)\Delta t,\end{aligned}$$

and similarly

$$\mathbb{E}\int_0^T \|\mathcal{E}_N^{S,2}\|^2 \mathrm{d}t = \sum_{k=1}^{N-1} \mathbb{E}\|\sigma_N(U_N^{k-1})\|^2_{L^2(\mathfrak{U},V)}\int_{t^k}^{t^{k+1}}(t-t_k)\mathrm{d}t \le c\sum_{k=1}^{N-1}\mathbb{E}(1+|U_N^{k-1}|^2)\Delta t,$$

so that, taken together we infer that

$$\sup_{N\ge N_1} \mathbb{E}\|\mathcal{E}_N^{S}\|_{L^2(0,T;V)} < \infty. \tag{4.21}$$

Finally, we supply a bound for $\mathcal{E}_N^S$ in $L^\infty(0,T;H)$. For $\mathcal{E}_N^{S,1}$, we observe with (2.10), (3.6) that

$$\mathbb{E}\sup_{t\in[0,T]}|\mathcal{E}_N^{S,1}|^2 \le \sum_{k=1}^{N-1}\mathbb{E}\sup_{t\in[t^k,t^{k+1}]}|\mathcal{E}_N^{S,1}|^2 \le \sum_{k=1}^{N-1}\mathbb{E}|\sigma_N(U_N^{k-1})\eta^k|^2 \le c\sum_{k=1}^{N-1}\mathbb{E}(1+|U_N^{k-1}|^2)\Delta t.$$

To estimate $\mathcal{E}_N^{S,2}$, we use Doob's inequality and (2.10) to infer

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}|\mathcal{E}_N^{S,2}|^2 &\le \sum_{k=1}^{N-1}\mathbb{E}\sup_{t\in[t^k,t^{k+1}]}|\mathcal{E}_N^{S,2}|^2 = \sum_{k=1}^{N-1}\mathbb{E}\sup_{t\in[t^k,t^{k+1}]}\left|\int_{t^k}^t \sigma_N(U_N)\mathrm{d}W\right|^2 \\ &\le \sum_{k=1}^{N-1}\mathbb{E}\int_{t^k}^{t^{k+1}}|\sigma_N(U_N)|^2_{L^2(\mathfrak{U},H)}\mathrm{d}s \le c\sum_{k=1}^{N-1}\mathbb{E}(1+|U_N^{k-1}|^2)\Delta t.\end{aligned}$$

With these bounds and (3.28), we conclude that

$$\sup_{N\ge N_1}\mathbb{E}\|\mathcal{E}_N^S\|^2_{L^\infty(0,T;H)} < \infty. \tag{4.22}$$

In turn, (4.21)–(4.22) directly imply (4.9), and so the proof of Proposition 4.1 is now complete.

5 Compactness and the Passage to the Limit

In this section, we detail the compactness arguments that we use to prove the existence of martingale solutions of (2.13) using the processes U_N and $\overline{U}_N$ defined in the previous section. As it is not clear how to obtain compactness directly from $\overline{U}_N$ (see Remark 4.2), we must introduce further processes to achieve this end.

Recalling (4.1)–(4.2), (4.15)–(4.16), we define

$$U_N^* = \overline{U}_N - \mathcal{E}_N^S, \quad U_N^{**} = U_N^* - \mathcal{E}_N^D, \tag{5.1}$$

and then consider the associated probability measures

$$\mu_N(\cdot) := \mathbb{P}(U_N \in \cdot), \quad \mu_N^*(\cdot) := \mathbb{P}(U_N^* \in \cdot), \quad \mu_N^{**}(\cdot) := \mathbb{P}(U_N^{**} \in \cdot). \tag{5.2}$$

Notice that, due to Proposition 4.1, μ_N, μ_N^* are defined on the space $\mathcal{X} := L^2(0,T;H)$. Regarding the elements μ_N^{**}, we observe that, as a consequence of (4.6),

$$U_N^{**}(t) = U_N^0 + \int_0^t (\mathcal{N}(U_N) + \ell)\mathrm{d}t + \int_0^t \sigma_N(U_N)\mathrm{d}W. \tag{5.3}$$

As a result of this identity and Proposition 4.1, the elements μ_N^{**} may be regarded as measures on the space $\mathcal{Y} := L^2(0,T;V') \cap \mathcal{C}([0,T];V'_{(3)})$.

We will show below that μ_N and μ_N^{**} converge weakly to a common measure μ and then make careful usage of the Skorohod embedding theorem to pass to the limit in (5.3) on a new stochastic basis. The former compactness arguments, which rely on the intermediate measures μ_N^*, will be carried out in the next subsection and the details of the Skorohod embedding will be discussed in Subsection 5.2 further on.

5.1 Tightness arguments

In this section, we will establish the following compactness properties of the $\{\mu_N\}_{N \geq N_1}$ and $\{\mu_N^{**}\}_{N \geq N_1}$.

Proposition 5.1 *The assumptions are precisely those in Proposition* 4.1. *Define* $\{U_N\}_{N\geq N_1}$ *and* $\{U_N^{**}\}_{N\geq N_1}$ *according to* (4.1) *and* (5.1) *and where* N_1 *is as in* (3.26). *Let* $\{\mu_N\}_{N\geq N_1}$, $\{\mu_N^{**}\}_{N\geq N_1}$ *be the associated Borel measures on*

$$\mathcal{X} := L^2(0,T;H),$$
$$\mathcal{Y} := L^2(0,T;V') \cap \mathcal{C}([0,T];V'_{(3)}),$$

defined according to (5.2). *Then, there exists a Borel measure* μ *on* $L^2(0,T;H) \cap \mathcal{C}([0,T];V'_{(3)})$ *such that, up to a subsequence,*[12]

$$\mu_N \rightharpoonup \mu \quad (\textit{weakly}) \ \textit{on}\ \mathcal{X}, \tag{5.4}$$

$$\mu_N^{**} \rightharpoonup \mu \quad (\textit{weakly}) \ \textit{on}\ \mathcal{Y}. \tag{5.5}$$

[12]We recall the notion of weak compactness of probability measures along with the equivalent notion of tightness in Appendix (see Section A.1).

The rest of this subsection is devoted to the proof of Proposition 5.1. We proceed as follows: First we show that $\{\mu_N^*\}_{N\geq N_1}$ is tight (see Appendix A.1) in $L^2(0,T;H)$ by employing a suitable variant of the Aubin-Lions compactness theorem which we establish in Proposition A.4 below. We next show that $\{\mu_N^{**}\}_{N\geq N_1}$ is tight in $\mathcal{C}([0,T];V'_{(3)})$ via an Arzelá-Ascoli type compact embedding from [25, 70]. We finally employ the estimates (4.5), (4.7) along with the general convergence results recalled in Lemma A.1 to finally infer (5.4)–(5.5).

5.1.1 Tightness for μ_N^* in $L^2(0,T;H)$

With the aid of Proposition A.4, we identify some compact subsets of $\mathcal{X} = L^2(0,T;H)$ that, in conjunction with suitable estimates (see (5.10)–(5.13) immediately below) are used to establish the tightness of $\{\mu_N^*\}_{N\geq N_1}$ in $\mathcal{X}$. For $U \in \mathcal{X}$, $n > 0$, define

$$[U]_j := \Big(j \sup_{0\leq\theta\leq j^{-6}} \int_0^{T-\theta} \|U(t+\theta) - U(t)\|_{V'_{(2)}}^{\frac{4}{3}} \mathrm{d}t\Big)^{\frac{3}{4}}, \tag{5.6}$$

and, for each $R > 0$, consider

$$B_R := \Big\{U \in \mathcal{X} : \|U\|_{L^2(0,T,V)} + \|U\|_{L^\infty(0,T,H)} + \sup_{j\geq 1}[U]_j \leq R\Big\}. \tag{5.7}$$

It is not hard to show that each set B_R is a closed subset of $\mathcal{X}$. Perusing (5.6), it is clear that the condition (A.4) holds uniformly for elements in B_R. Thus, as a consequence of Proposition A.4(ii), these sets B_R are compact in $\mathcal{X} = L^2(0,T;H)$ for each $R > 0$.

Now, for each $R > 0$, we have

$$\mu_N^*(B_R^c) \leq \mathbb{P}\Big(\|U_N^*\|_{L^2(0,T,V)} + \|U_N^*\|_{L^\infty(0,T,H)} > \frac{R}{2}\Big) + \mathbb{P}\Big(\sup_{j\geq 1}[U_N^*]_j > \frac{R}{2}\Big). \tag{5.8}$$

As a consequence of (4.4), (4.9) and (5.1), we have

$$\mathbb{P}\Big(\|U_N^*\|_{L^2(0,T,V)} + \|U_N^*\|_{L^\infty(0,T,H)} > \frac{R}{2}\Big) \leq \frac{c}{R^2} \tag{5.9}$$

for some constant c independent of N.

Next we need to establish suitable uniform estimates for $\sup_{j\geq 1}[U_N^*]_j$ (see (5.6)). To this end, we observe with (4.14) and (5.1) that for any $\theta > 0$,

$$\int_0^{T-\theta} \|U_N^*(t+\theta) - U_N^*(t)\|_{V'_{(2)}}^{\frac{4}{3}} \mathrm{d}t \leq I_N^D(\theta) + I_N^S(\theta), \tag{5.10}$$

with

$$I_N^D(\theta) = c\int_0^{T-\theta} \Big\| \int_t^{t+\theta} \sum_{n=1}^{N-1} (\mathcal{N}(U_N^n) + \ell_N^n)\chi_{(t^n,t^{n+1})}(s)\mathrm{d}s\Big\|_{V'_{(2)}}^{\frac{4}{3}} \mathrm{d}t,$$

$$I_N^S(\theta) = c\int_0^{T-\theta} \Big\| \int_t^{t+\theta} \sigma_N(U_N)\mathrm{d}W\Big\|_{V'_{(2)}}^{\frac{4}{3}} \mathrm{d}t.$$

To address $I_N^D(\theta)$, we observe, with (2.6) and the standing assumptions on the operators that make up $\mathcal{N}$ in (2.14), that for any $U \in V$,

$$\|\mathcal{N}(U)\|_{V'_{(2)}}^{\frac{4}{3}} \leq c(|U|^{\frac{2}{3}} + 1)(\|U\|^2 + 1). \tag{5.11}$$

Furthermore, it is clear from (3.4) and Hölder's inequality that, a.s.

$$\int_0^T \sum_{n=1}^{N-1} \|\ell_N^n\|_{V'_{(2)}}^{\frac{4}{3}} \chi_{(t^n,t^{n+1})}(s)\mathrm{d}t \leq \int_0^T \sum_{n=1}^{N-1} \left(\frac{1}{\Delta t}\int_{(n-1)\Delta t}^{n\Delta t} \|\ell\|_{V'_{(2)}}^{\frac{4}{3}} \mathrm{d}s\right)\chi_{(t^n,t^{n+1})}(s)\mathrm{d}t \leq \int_0^T \|\ell\|_{V'_{(2)}}^{\frac{4}{3}} \mathrm{d}t.$$

Combining these observations, we infer that, a.s.

$$\begin{aligned} I_N^D(\theta) &\leq c\theta^{\frac{1}{3}} \int_0^{T-\theta}\int_t^{t+\theta} \sum_{n=1}^{N-1} \|\mathcal{N}(U_N^n) + \ell_N^n\|_{V'_{(2)}}^{\frac{4}{3}} \chi_{(t^n,t^{n+1})}(s)\mathrm{d}s\mathrm{d}t \\ &\leq c\theta^{\frac{1}{3}} \int_0^T \sum_{n=1}^{N-1} ((|U_N^n|^{\frac{2}{3}}+1)(\|U_N^n\|^2+1) + \|\ell\|_{V'_{(2)}}^{\frac{4}{3}})\chi_{(t^n,t^{n+1})}(s)\mathrm{d}s \\ &\leq c\theta^{\frac{1}{3}}\left(\max_{0\leq l\leq N}(1+|U_N^l|^{\frac{2}{3}})\sum_{j=1}^N \Delta t(\|U_N^j\|^2+1) + \int_0^T \|\ell\|_{V'_{(2)}}^{\frac{4}{3}}\mathrm{d}t\right) \\ &\leq c\theta^{\frac{1}{3}}\left(\max_{0\leq l\leq N}(1+|U_N^l|^2)\sum_{j=1}^N \Delta t(\|U_N^j\|^2+1) + \int_0^T (1+\|\ell\|_{V'}^2)\mathrm{d}t\right). \end{aligned} \tag{5.12}$$

For the term I_N^S, we estimate, for $0 \leq \theta \leq \delta$,

$$\begin{aligned} \mathbb{E}\left(\sup_{0\leq\theta\leq\delta} I_N^S(\theta)\right) &\leq c\int_0^T \left(\mathbb{E}\sup_{0\leq\theta\leq\delta}\left\|\int_t^{(t+\theta)\wedge T}\sigma_N(U_N)\mathrm{d}W\right\|_{V'_{(2)}}^2\right)^{\frac{2}{3}}\mathrm{d}t \\ &\leq c\delta^{\frac{2}{3}}\mathbb{E}\sup_{t\in[0,T]}(1+|U_N|^2), \end{aligned} \tag{5.13}$$

where the second line follows from Doob's inequality and the standing assumptions (2.10) on σ and (3.6) on σ_N:

$$\begin{aligned} \mathbb{E}\sup_{0\leq\theta\leq\delta}\left\|\int_t^{(t+\theta)\wedge T}\sigma_N(U_N)\mathrm{d}W\right\|_{V'_{(2)}}^2 &\leq c\mathbb{E}\int_t^{(t+\delta)\wedge T}\|\sigma_N(U_N)\|_{V'_{(2)}}^2\mathrm{d}s \\ &\leq c\delta\mathbb{E}\sup_{t\in[0,T]}(1+|U_N|^2). \end{aligned}$$

The estimates (5.12)–(5.13) allow the second term in (5.8) to be treated as follows. Observe that according to (5.6) and (5.10), we have

$$\sup_{j\geq 1}[U_N^*]_j^{\frac{4}{3}} \leq \sup_{j\geq 1}\left(j\sup_{|\theta|\leq j^{-6}} I_N^D(\theta)\right) + \sup_{j\geq 1}\left(j\sup_{|\theta|\leq j^{-6}} I_N^S(\theta)\right).$$

For the first term, we observe with (5.12) that

$$\begin{aligned} \sup_{j\geq 1}\left(j\sup_{|\theta|\leq j^{-6}} I_N^D(\theta)\right) &\leq c\left(\max_{0\leq l\leq N}(1+|U_N^l|^2)\sum_{r=1}^N \Delta t(\|U_N^r\|^2+1) + \int_0^T (\|\ell\|_{V'}^2+1)\mathrm{d}t\right) \\ &:= c(T_1^N T_2^N + T_3^N). \end{aligned} \tag{5.14}$$

Regarding the second term, we simply bound

$$\sup_{j\geq 1}\left(j\sup_{|\theta|\leq j^{-6}} I_N^S(\theta)\right) \leq \sum_{j\geq 1} j\sup_{|\theta|\leq j^{-6}} I_N^S(\theta) := T_4^N,$$

so that for $\rho > 0$, sufficiently large,

$$\begin{aligned}\mathbb{P}\Big(\sup_{j\geq 1}[U_N^*]_j^{\frac{4}{3}} > \rho\Big) &\leq \mathbb{P}(c(T_1^N T_2^N + T_3^N) + T_4^N > \rho)\\ &\leq \mathbb{P}\Big(cT_1^N T_2^N > \frac{\rho}{2}\Big) + \mathbb{P}\Big(cT_3^N + T_4^N > \frac{\rho}{2}\Big)\\ &\leq \mathbb{P}\Big(\Big\{T_1^N > \sqrt{\frac{\rho}{2c}}\Big\} \cup \Big\{T_2^N > \sqrt{\frac{\rho}{2c}}\Big\}\Big) + \mathbb{P}\Big(cT_3^N + T_4^N > \frac{\rho}{2}\Big)\\ &\leq \mathbb{P}\Big(T_1^N > \sqrt{\frac{\rho}{2c}}\Big) + \mathbb{P}\Big(T_2^N > \sqrt{\frac{\rho}{2c}}\Big) + \mathbb{P}\Big(cT_3^N + T_4^N > \frac{\rho}{2}\Big)\\ &\leq \frac{c}{\sqrt{\rho}}\mathbb{E}(T_1^N + T_2^N + T_3^N + T_4^N). \end{aligned} \tag{5.15}$$

In view of the uniform bound (3.28) established in Proposition 3.2, $\sup_N \mathbb{E}T_1^N$ and $\sup_N \mathbb{E}T_2^N$ are both finite. The term $\sup_N \mathbb{E}T_3^N$, which is independent of N, is finite due to the standing assumption on ℓ (see (2.17)). For T_4^N, we refer back to (5.13) and apply the monotone convergence theorem to infer

$$\mathbb{E}T_4^N \leq \sum_{n\geq 1} n^{-3}\mathbb{E}\sup_{t\in[0,T]}(1+|U_N|^2) < \infty.$$

We finally conclude that

$$\mathbb{P}\Big(\sup_{j\geq 1}[U_N^*]_j > \frac{R}{2}\Big) = \mathbb{P}\Big(\sup_{j\geq 1}[U_N^*]_j^{\frac{4}{3}} > \Big(\frac{R}{2}\Big)^{\frac{4}{3}}\Big) \leq \frac{c}{R^{\frac{2}{3}}}. \tag{5.16}$$

Combining (5.8)–(5.9) and (5.16), we now conclude that (see Appendix A.1)

$$\{\mu_N^*\}_{N\geq 1} \text{ is tight in } \mathcal{X} = L^2(0,T;H). \tag{5.17}$$

5.1.2 Tightness for μ_N^{**} in $\mathcal{C}([0,T];V'_{(3)})$

We next show that μ_N^{**} is tight in $\mathcal{C}([0,T],V'_{(3)})$. For this purpose, we make appropriate usage of a compact embedding from [25] (see also [70]). Let us fix any $p \in (2,\infty)$, $\alpha \in (0,\frac{1}{2})$ such that $\alpha p > 1$. According to [25],

$$W^{1,\frac{4}{3}}(0,T;V'_{(2)}) \subset\subset \mathcal{C}([0,T];V'_{(3)}), \quad W^{\alpha,p}(0,T;V'_{(2)}) \subset\subset \mathcal{C}([0,T];V'_{(3)}), \tag{5.18}$$

that is, the embeddings are continuous and compact. We now define

$$\begin{aligned} B_R &:= \{X \in \mathcal{C}([0,T];V'_{(3)}) : \|X\|_{W^{1,\frac{4}{3}}(0,T;V'_{(2)})} \leq R\}\\ &\quad + \{Y \in \mathcal{C}([0,T];V'_{(3)}) : \|Y\|_{W^{\alpha,p}(0,T;V'_{(2)})} \leq R\}\\ &:= B_R^D + B_R^S \end{aligned}$$

for any $R > 0$. With (5.18), it is clear that B_R is compact in $\mathcal{C}([0,T];V'_{(3)})$ for every $R > 0$. Observe moreover that, in view of (5.3),

$$\{U_N^{**} \in B_R\} \supseteq \Big\{U_N^0 + \int_0^{\cdot} (\mathcal{N}(U_N) + \ell)\mathrm{d}s \in B_R^D\Big\} \cap \Big\{\int_0^{\cdot} \sigma_N(U_N)\mathrm{d}W \in B_R^S\Big\},$$

and thus that

$$\begin{aligned}\mu_N^{**}(B_R^C) &\leq \mathbb{P}\Big(\Big\|U_N^0 + \int_0^\cdot (\mathcal{N}(U_N)+\ell)\mathrm{d}s\Big\|_{W^{1,\frac{4}{3}}(0,T;V'_{(2)})} > R\Big) \\ &\quad + \mathbb{P}\Big(\Big\|\int_0^\cdot \sigma_N(U_N)\mathrm{d}W\Big\|_{W^{\alpha,p}(0,T;V'_{(2)})} > R\Big) \\ &:= S_N^R + T_N^R. \end{aligned} \tag{5.19}$$

Hence we will infer that $\{\mu_N^{**}\}$ is tight in $\mathcal{C}([0,T],V'_{(3)})$ if we can show that T_N^R, S_N^R converge uniformly in N to zero as $R\uparrow\infty$.

For T_N^R, we estimate, with (5.11),

$$\begin{aligned}&\Big\|U_N^0 + \int_0^\cdot (\mathcal{N}(U_N)+\ell)\mathrm{d}s\Big\|^{\frac{4}{3}}_{W^{1,\frac{4}{3}}(0,T;V'_{(2)})} \\ &\leq c(1+|U_N^0|^2) + c\int_0^T \|\mathcal{N}(U_N)+\ell\|^{\frac{4}{3}}_{V'_{(2)}}\mathrm{d}t \\ &\leq c(1+|U_N^0|^2) + c\int_0^T ((|U_N|^{\frac{2}{3}}+1)(\|U_N\|^2+1)+\|\ell\|^2_{V'})\mathrm{d}t, \\ &\leq c\sup_{t\in[0,T]}(1+|U_N|^2)\cdot\Big(\int_0^T (1+\|U_N\|^2+\|\ell\|^2_{V'})\mathrm{d}t+1\Big).\end{aligned}$$

Thus we find (see (5.15))

$$\begin{aligned}T_N^R &\leq \mathbb{P}\Big(c\sup_{t\in[0,T]}(1+|U_N|^2)\cdot\Big(\int_0^T(1+\|U_N\|^2+\|\ell\|^2_{V'})\mathrm{d}t+1\Big)>R\Big) \\ &\leq \mathbb{P}\Big(c\sup_{t\in[0,T]}(1+|U_N|^2)>R^{\frac{1}{2}}\Big)+\mathbb{P}\Big(\int_0^T(1+\|U_N\|^2+\|\ell\|^2_{V'})\mathrm{d}t+1>R^{\frac{1}{2}}\Big) \\ &\leq \frac{c}{R^{\frac{1}{2}}}\mathbb{E}\sup_{t\in[0,T]}(1+|U_N|^2)+\frac{1}{R^{\frac{1}{2}}}\mathbb{E}\Big(\int_0^T(1+\|U_N\|^2+\|\ell\|^2_{V'})\mathrm{d}t+1\Big).\end{aligned} \tag{5.20}$$

We turn to S_N^R. For this purpose, let us define for any $R>0$ the stopping times

$$\tau_R := \inf_{t\geq 0}\Big\{\sup_{s\in[0,t]}|U_N|\geq R\Big\}\wedge T = \sup_{t\geq 0}\Big\{\sup_{s\in[0,t]}|U_N|<R\Big\}\wedge T.$$

Using τ_R, we now estimate with the Chebyshev inequality that

$$\begin{aligned}S_N^R &\leq \mathbb{P}\Big(\Big\|\int_0^{\cdot\wedge\tau_R}\sigma_N(U_N)\mathrm{d}W\Big\|_{W^{\alpha,p}(0,T;V'_{(2)})}>R,\tau_R\geq T\Big)+\mathbb{P}(\tau_R<T) \\ &\leq \mathbb{P}\Big(\Big\|\int_0^{\cdot\wedge\tau_R}\sigma_N(U_N)\mathrm{d}W\Big\|_{W^{\alpha,p}(0,T;V'_{(2)})}>R\Big)+\mathbb{P}\Big(\sup_{s\in[0,T]}|U_N|\geq R\Big) \\ &\leq \frac{1}{R^p}\mathbb{E}\Big(\Big\|\int_0^{\cdot\wedge\tau_R}\sigma_N(U_N)\mathrm{d}W\Big\|^p_{W^{\alpha,p}(0,T;V'_{(2)})}\Big)+\frac{1}{R^2}\mathbb{E}\sup_{s\in[0,T]}|U_N|^2.\end{aligned} \tag{5.21}$$

Now in order to treat this final stochastic integral term, we recall the following generalization of the Burkholder-Davis-Gundy inequality from [25]: For a given Hilbert space X, $p\geq 2$ and $\alpha\in[0,\frac{1}{2})$, we have for all X-valued predictable $G\in L^p(\Omega;L^p_{\mathrm{loc}}(0,\infty,L_2(\mathfrak{U},X)))$,

$$\mathbb{E}\Big(\Big\|\int_0^\cdot G\mathrm{d}W\Big\|^p_{W^{\alpha,p}(0,T;X)}\Big)\leq c\mathbb{E}\Big(\int_0^T |G|^p_{L_2(\mathfrak{U},X)}\mathrm{d}t\Big),$$

which holds with a constant c depending only on α and p. Continuing now from (5.21), we have

$$\begin{aligned} S_N^R &\le \frac{c}{R^p}\mathbb{E}\int_0^{T\wedge\tau_R} |\sigma_N(U_N)|^p_{L_2(\mathfrak{U},H)}\mathrm{d}t + \frac{1}{R^2}\mathbb{E}\sup_{s\in[0,T]}|U_N|^2 \\ &\le \frac{c}{R^p}\mathbb{E}\sup_{s\in[0,T\wedge\tau_R]}(1+|U_N|^p) + \frac{1}{R^2}\mathbb{E}\sup_{s\in[0,T]}|U_N|^2 \\ &\le \frac{c(1+R^{p-2})}{R^p}\mathbb{E}\sup_{s\in[0,T]}(1+|U_N|^2) + \frac{1}{R^2}\mathbb{E}\sup_{s\in[0,T]}|U_N|^2 \le \frac{c}{R^2}\mathbb{E}\sup_{s\in[0,T]}(1+|U_N|^2). \end{aligned} \tag{5.22}$$

Combining the estimates (5.20), (5.22) with (4.4), we finally conclude

$$\sup_{N\ge N_1}\mu_N^{**}(B_R) \ge 1 - \frac{c}{R^{\frac{1}{2}}},$$

and hence infer that

$$\{\mu_N^{**}\}_{N\ge N_1} \text{ is tight in } \mathcal{C}([0,T];V'_{(3)}). \tag{5.23}$$

Remark 5.1 Let us observe that the tightness bounds for μ_N^{**} and μ_N^* could be carried out differently if we had available, for example, the uniform bounds on "higher moments" like

$$\sup_{N\ge1}\mathbb{E}\Big(\max_{0\le k\le N}|U_N^k|^4 + \Big(\sum_{k=1}^N \Delta t\|U_N^k\|^2\Big)^2\Big) < \infty \tag{5.24}$$

or equivalently that

$$\sup_{N\ge1}\mathbb{E}\Big(\sup_{t\in[0,T]}|U_N|^4 + \Big(\int_0^T \|U_N\|^2\mathrm{d}t\Big)^2\Big) < \infty. \tag{5.25}$$

Indeed, in numerous other previous works related to stochastic fluids equations (see, e.g., [4, 15, 25, 34, 51]) estimates analogous to (5.25) are established essentially via Itô's lemma in order to achieve tightness in the probability laws associated to a regularization scheme.

In the current situation, instead due to the way we carry out the estimates in (5.15), (5.21)–(5.22), we have adopted a different approach, namely, we establish tightness (compactness) estimates without recourse to such higher moment estimates.

A different method using higher moments will be shown in the related work [33].

5.1.3 Cauchy arguments and conclusions

With (5.17) and (5.23) now in hand, it is then simply a matter of collecting the various convergences above to complete the proof of Proposition 5.1.

By making use of Prohorov's theorem (see Section A.1) with (5.17), we infer the existence of a probability measure μ such that, up to a subsequence,

$$\mu_N^* \rightharpoonup \mu \quad \text{on } \mathcal{X} = L^2(0,T;H) \ \ (\text{and also on } L^2(0,T;V')).$$

Due to (5.1) with (4.5) and (4.8), it is clear that $U_N^* - U_N$ converges to zero in $\mathcal{X} = L^2(0,T;H)$ and hence in $L^2(0,T;V')$ a.s. Hence, by now invoking (4.7) and referring back once more to

(5.1), we have that $U_N^* - U_N^{**}$ converges to zero in $L^2(0,T;V')$ a.s. Thus, invoking Lemma A.1, again up to a subsequence, we conclude that

$$\mu_N^{**} \rightharpoonup \mu \quad \text{on } L^2(0,T;V') \quad \text{and} \quad \mu_N \rightharpoonup \mu \quad \text{on } \mathcal{X} = L^2(0,T;H). \tag{5.26}$$

In particular, this is the first desired convergence for $\{\mu_N\}_{N \geq N_1}$, (5.4). On the other hand, invoking Prohorov's theorem with (5.23) and the convergence just established for $\{\mu_N^{**}\}_{N \geq N_1}$ in $L^2(0,T;V')$, we see that μ_N^{**} is tight in $\mathcal{Y} = L^2(0,T;V') \cap \mathcal{C}([0,T];V'_{(3)})$. By Prohorov's theorem in the other direction and passing to a further subsequence as needed, we have

$$\mu_N^{**} \rightharpoonup \widetilde{\mu} \text{ on } \mathcal{Y} = L^2(0,T;V') \cap \mathcal{C}([0,T],V'_{(3)}).$$

Since, clearly, $\widetilde{\mu} = \mu$, this yields the second desired item (5.5). The proof of Proposition 5.1 is therefore complete.

5.2 Proof of Theorem 2.1 conclusion: Almost sure convergence and the passage to the limit on the Skorokhod basis

We now have all of the ingredients to finally prove the main results of this article, namely Theorem 2.1. Suppose that we are given $\mu_{U_0} \in \Pr(H)$ and $\mu_\ell \in \Pr(L^2_{\text{loc}}(0,\infty;V'))$ according to the conditions specified in Definition 2.1. As mentioned in Remark 3.2, now it is necessary to introduce the stochastic basis $\mathcal{S}_\mathcal{G}$ (defined as in Subsection 3.2), an element U^0 which is $\mathcal{G}_0$ measurable and a process $\ell = \ell(t)$ measurable with respect to the sigma algebra generated by the $W(s)$ for $s \in [0,t]$,[13] whose laws coincide with those of μ_{U_0}, μ_ℓ. Thus Proposition 3.1 applies, and we obtain the existence of the U_N^n's adapted to $\mathcal{G}_{t_n}$.

We then approximate $U^0 \in L^2(\Omega;H)$ with a sequence of elements $\{U_N^0\}_{N \geq 1} \subseteq L^2(\Omega, V_{(2)})$, which maintains the bound (4.3) as described in Remark 3.1 above. Proposition 4.1 applies, and hence we can use this sequence $\{U_N^0\}_{N \geq N_1}$, the process ℓ, and the sequence U_N^n to define processes $\{U_N\}_{N \geq N_1}$, $\{U_N^{**}\}_{N \geq N_1}$ according to (4.1) and (5.1), respectively (N_1 is given by (3.26)). In order to pass to the limit in the associated evolution equation (5.3), we consider the product measures

$$\nu_N(\cdot) := \mathbb{P}((U_N^{**}, U_N, \ell, W) \in \cdot)$$

which are defined on the space

$$\mathcal{Z} = \mathcal{Y} \times \mathcal{X} \times L^2(0,T;V') \times \mathcal{C}([0,T];\mathfrak{U}_0), \tag{5.27}$$

where, as above, $\mathcal{Y} = L^2(0,T;V') \cap \mathcal{C}([0,T],V'_{(3)})$, $\mathcal{X} = L^2(0,T;H)$, and $\mathfrak{U}_0$ is defined as in Subsection 2.2, (2.15). By invoking Proposition 5.1, we have that (passing to a subsequences as needed) $\mu_N \rightharpoonup \mu$ on $\mathcal{X}$ and $\mu_N^{**} \rightharpoonup \mu$ on $\mathcal{Y}$, where μ_N and μ_N^{**} are defined as in (5.2). It follows, again up to passing to a subsequence, that ν_N converges weakly to a measure ν on $\mathcal{Z}$ (defined in (5.27)). Furthermore, recalling (5.1) and making use of (4.5), (4.7)–(4.8), it is not hard to see that

$$\nu(\{(U^{**}, U, \ell, W) \in \mathcal{Z} : U \neq U^{**}\}) = 0.$$

[13]Note that since the sigma algebra generated by the $W(s)$ for $s \in [0,t]$) is the smallest respect to which $W(t)$ is measurable, $\ell(t)$ is adapted to $\{\mathcal{F}_t\}_{t \geq 0}$, and hence all the previous results apply.

Thus, by making use of the Skorokhod embedding theorem (see Section A.1), we obtain, relative to a new probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$, a sequence of random variables

$$(\widetilde{U}_N^{**}, \widetilde{U}_N, \widetilde{\ell}_N, \widetilde{W}_N) \to (\widetilde{U}, \widetilde{U}, \widetilde{\ell}, \widetilde{W}), \quad \widetilde{\Omega} \text{ a.s. in } \mathcal{Z}. \tag{5.28}$$

Moreover, the uniform bounds for $\{U_N\}_{N \geq N_1}$ in $L^2(\Omega; L^2(0,T;V) \cap L^\infty(0,T;H))$ from Proposition 4.1, (4.4) imply that in addition to (5.28), we also have

$$\widetilde{U}_N \rightharpoonup \widetilde{U} \quad \text{weakly in } L^2(\Omega; L^2(0,T;V)) \text{ and weakly-star in } L^2(\Omega; L^\infty(0,T;H)). \tag{5.29}$$

Following a procedure very similar to [4], we may now show that $\widetilde{W}_N$ is a cylindrical Brownian motion relative to the filtration $\widetilde{\mathcal{F}}_t^N$ defined as the sigma algebra generated by $(\widetilde{U}_N^{**}(s), \widetilde{U}_N(s), \widetilde{\ell}_N(s), \widetilde{W}_N(s))$ for $s \leq t$ and that $(\widetilde{U}_N^{**}, \widetilde{U}_N, \widetilde{\ell}_N, \widetilde{W}_N)$ satisfies (5.3) on the "Skorokhod space" $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ viz.

$$\widetilde{U}_N^{**}(t) = \widetilde{U}_N^{**}(0) + \int_0^t (\mathcal{N}(\widetilde{U}_N) + \widetilde{\ell}_N)\mathrm{d}s + \int_0^t \sigma_N(\widetilde{U}_N)\mathrm{d}\widetilde{W}_N. \tag{5.30}$$

Using the convergences in (5.28)–(5.29) with (5.30) it is standard[14] to show that $\widetilde{U}$ satisfies (2.18)–(2.20) relative to the stochastic basis $\widetilde{\mathcal{S}} := (\widetilde{\Omega}, \widetilde{\mathcal{F}}, \{\widetilde{\mathcal{F}}_t\}_{t\geq 0}, \widetilde{\mathbb{P}}, \{\widetilde{W}_k\}_{k\geq 1})$, where $\{\widetilde{\mathcal{F}}_t\}_{t\geq 0}$ is defined as the sigma algebra generated by the $(\widetilde{U}(s), \widetilde{\ell}(s), \widetilde{W}(s))$ for $s \leq t$ and $\widetilde{W}_k = (\widetilde{W}, e_k)_{\mathfrak{U}}$. Therefore, $(\widetilde{S}, \widetilde{U}, \widetilde{\ell})$ is a martingale solution to (2.13) relative to μ_{U_0}, μ_ℓ in the sense of Definition 2.1, and the proof of Theorem 2.1 is complete.

6 Convergence of the Euler Scheme

We conclude by reinterpreting from the point of view of numerical analysis, the study above as a result of convergence for the Euler scheme (3.3).

Theorem 6.1 *We assume given $\mu_{U_0} \in \Pr(H)$ and $\mu_\ell \in \Pr(L^2_{\mathrm{loc}}(0,\infty;V'))$ according to Definition* 2.1. *We also assume given the stochastic basis $\mathcal{S}_{\mathcal{G}}$ (defined as in Subsection* 3.2*), an element U^0 which is $\mathcal{G}_0$ measurable and a process $\ell = \ell(t)$ measurable with respect to the sigma algebra generated by the $W(s)$ for $s \in [0,t]$, whose laws coincide with those of μ_{U_0}, μ_ℓ. Let a sequences of elements $\{U_N^0\}_{N\geq 1} \subseteq L^2(\Omega, V_{(2)})$ approximate $U^0 \in L^2(\Omega; H)$ as described in Remark* 3.1. *Then the processes $\{U_N\}_{N\geq N_1}$ defined according to* (4.1) *(N_1 is given by* (3.26)*) adapted to $\{\mathcal{G}_t\}_{t\geq 0}$ exist.*

Moreover, the family $\{\mu_N\}$ of probability laws of $\{U_N\}$, is weakly compact over the phase space $L^2(0,T;\, H) \cap \mathcal{C}([0,T], V'_{(3)})$, and hence converges weakly to a probability measure μ on the same phase space up to a subsequence. Furthermore, there exists a probability space $(\widetilde{\Omega}, \widetilde{F}, \widetilde{P})$ and a subsequence of random vectors $(\widetilde{U}_{N_k}, \widetilde{\ell}_{N_k}, \widetilde{W}_{N_k})$ with values in $\mathcal{Z}_1 := L^2(0,T;\, H) \cap \mathcal{C}([0,T], V'_{(3)}) \times L^2(0,T;V') \times C([0,T];\mathfrak{U}_0)$ such that

(i) *$(\widetilde{U}_{N_k}, \widetilde{\ell}_{N_k}, \widetilde{W}_{N_k})$ have the same probability distribution as (U_{N_k}, ℓ, W).*

(ii) *$(\widetilde{U}_{N_k}, \widetilde{\ell}_{N_k}, \widetilde{W}_{N_k})$ converges almost surely as $N_k \to \infty$, in the topology of $\mathcal{Z}_1$, to an element $(\widetilde{U}, \widetilde{\ell}, \widetilde{W})$. Particularly,*

$$\widetilde{U}_{N_k} \to \widetilde{U} \quad \textit{strongly in } L^2(0,T;\, H) \cap \mathcal{C}([0,T], V'_{(3)}) \textit{ a.s.}, \tag{6.1}$$

[14] Note that, in particular, the stochastic terms involving $\sigma_N(U_N)$ converge due to (3.7).

where $\widetilde{U}$ has the probability distribution μ.

Proof The existence of $\{U_{N_k}\}_{N\geq N_1}$ follows directly from the existence of the U_N^n's proven in Proposition 3.1. (i) and (ii) follow from the Skorokhod embedding theorem (see Section A.1) as shown in Subsection 5.2.

7 Applications for Equations in Geophysical Fluid Dynamics

In this section, we apply the above framework culminating in Theorems 2.1 and 6.1 to a stochastic version of the primitive equations. Our presentation here will focus on the case of the equations of the oceans. Note however that the abstract setting introduced above is equally well suited to derive results for analogous systems for the atmosphere or for the coupled oceanic-atmospheric system (COA for short).[15] We refer the interested reader to [60] for further details on these other interesting situations.

7.1 The oceans equations

The stochastic primitive equations of the oceans take the form

$$\partial_t \mathbf{v} + \nabla_{\mathbf{v}}\mathbf{v} + w\partial_z\mathbf{v} + \frac{1}{\rho_0}\nabla p + f\mathbf{k}\times\mathbf{v} - \mu_{\mathbf{v}}\Delta\mathbf{v} - \nu_{\mathbf{v}}\partial_{zz}\mathbf{v} = F_{\mathbf{v}} + \sigma_{\mathbf{v}}(\mathbf{v},T,S)\dot{W}_1, \tag{7.1a}$$

$$\partial_z p = -\rho g, \tag{7.1b}$$

$$\nabla\cdot\mathbf{v} + \partial_z w = 0, \tag{7.1c}$$

$$\partial_t T + \nabla_{\mathbf{v}}T + w\partial_z T - \mu_T\Delta T - \nu_T\partial_{zz}T = F_T + \sigma_T(\mathbf{v},T,S)\dot{W}_2, \tag{7.1d}$$

$$\partial_t S + \nabla_{\mathbf{v}}S + w\partial_z S - \mu_S\Delta S - \nu_S\partial_{zz}S = F_S + \sigma_S(\mathbf{v},T,S)\dot{W}_3, \tag{7.1e}$$

$$\rho = \rho_0(1 + \beta_T(T - T_r) + \beta_S(S - S_r)). \tag{7.1f}$$

Here, $U := (\mathbf{v},T,S) = (u,v,T,S)$; $\mathbf{v},T,S,p,\rho$ represent the horizontal velocity, temperature, salinity, pressure and density of the fluid under consideration, respectively; $\mu_{\mathbf{v}}$, $\nu_{\mathbf{v}}$, μ_T, ν_T, μ_S, ν_S are positive coefficients which account for the eddy and molecular diffusivities (viscosity) in the equations for $\mathbf{v}$, T and S. The terms $F_{\mathbf{v}}, F_T, F_S$ are volumic sources of momentum, heat and salt which are zero in idealized situations but which we consider to be random in general.

The state dependent stochastic terms are driven by independent Gaussian white noise processes $\dot{W}_j$, $j = 1,2,3$ which are formally delta correlated in time. The stochastic terms may be written in the expansion

$$\sigma_U(U)\dot{W} = \begin{pmatrix} \sigma_{\mathbf{v}}(U)\dot{W}_1(t,\mathbf{x}) \\ \sigma_T(U)\dot{W}_2(t,\mathbf{x}) \\ \sigma_S(U)\dot{W}_3(t,\mathbf{x}) \end{pmatrix} = \sum_{k\geq 1} \begin{pmatrix} \sigma_{\mathbf{v}}^k(U)(t,\mathbf{x})\dot{W}_1^k(t) \\ \sigma_T^k(U)(t,\mathbf{x})\dot{W}_2^k(t) \\ \sigma_S^k(U)(t,\mathbf{x})\dot{W}_3^k(t) \end{pmatrix}, \tag{7.2}$$

where the elements $\dot{W}_j^k$ are independent 1D white (in time) noise processes. We may interpret the multiplication in (7.2) in either the Itô or the Stratonovich sense; as we detail in one example below that the classical correspondence between the Itô and Stratonovich systems

[15] Via a suitable change of variables, the dynamical equations for the compressible gases which constitute the earth's atmosphere may be shown to take a mathematical form essentially similar to the incompressible equations for the oceans.

allows us to treat both situations within the framework of the Itô evolution (2.13). We will describe some physically interesting configurations of these "stochastic terms" in detail below in Subsection 7.3.

The operators $\Delta = \partial_{xx} + \partial_{yy}$ and $\nabla = (\partial_x, \partial_y)$ are the horizontal Laplacian and the gradient operator, respectively. Here the operator $\nabla_{\mathbf{v}}$ captures part of the convective (material) derivative and is defined according to

$$\nabla_{\mathbf{v}} := \mathbf{v} \cdot \nabla = u\partial_x + v\partial_y. \tag{7.3}$$

Remark 7.1 As given, the model (7.1), expresses the equations for oceanic flows in the "beta-plane approximation", that is to say we make use of the fact that the earth is locally flat. This setting is suitable for regional studies, and we will focus on this case for the simplicity of presentation. With suitable adjustments to the definition of the operators Δ, ∇, $\nabla_{\mathbf{v}}$ and to the domain introduced below we could consider the evolutions in the full spherical geometry of the earth. We refer to [47] (and also to [60]) for further details on how to cast a global circulation model in the form of (2.13).

7.1.1 Domain and boundary conditions

The evolution (7.1) takes place on a bounded domain $\mathcal{M} \subset \mathbb{R}^3$ which we define as follows. Fix a bounded, open domain $\Gamma_i \subset \mathbb{R}^2$ with sufficiently smooth boundary ($\mathcal{C}^3$, say); Γ_i represents the surface of the ocean in the region under consideration. We suppose that we have defined a "depth" function $h = h(x, y) : \Gamma_i \to \mathbb{R}$ which is at least $\mathcal{C}^2$ and is subject to the restriction $0 < \underline{h} \leq h(x, y) \leq \overline{h}$. With these ingredients, we then let

$$\mathcal{M} := \{\mathbf{x} := (x, y, z) \in \mathbb{R}^3 : (x, y) \in \Gamma_i, z \in (-h(x, y), 0)\}.$$

The boundary $\partial\mathcal{M}$ of $\mathcal{M}$, is divided into its top Γ_i, lateral Γ_l and bottom Γ_b boundaries. We denote the outward unit normal to $\partial\mathcal{M}$ by $\mathbf{n}$ and the normal to Γ_l in $\mathbb{R}^2$ by $\mathbf{n}_H$.

We next prescribe the following, physically realistic boundary conditions for (7.1) considered in $\mathcal{M}$ (see, e.g., [60] for further details). On Γ_i we suppose

$$\partial_z \mathbf{v} + \alpha_{\mathbf{v}}(\mathbf{v} - \mathbf{v}^a) = \tau_{\mathbf{v}}, \quad w = 0, \quad \partial_z T + \alpha_T (T - T^a) = 0, \quad \partial_z S = 0, \tag{7.4}$$

where $\alpha_{\mathbf{v}}$, α_T are fixed positive constants, and $\tau_{\mathbf{v}}$, $\mathbf{v}^a$, T^a are in general random and non-constant in space and time. Physically speaking, the first two equations in (7.4) account for a boundary layer model, where $\mathbf{v}^a$, T^a represent the values for velocity and temperature of the atmosphere at the surface of the oceans, respectively; $\tau_{\mathbf{v}}$ accounts for the shear of the wind.

At the bottom of the ocean Γ_b, we take

$$\mathbf{v} = 0, \quad w = 0, \quad \partial_{\mathbf{n}} T = 0, \quad \partial_{\mathbf{n}} S = 0. \tag{7.5}$$

Finally for the lateral boundary Γ_l,

$$\mathbf{v} = 0, \quad \partial_{\mathbf{n}} T = 0, \quad \partial_{\mathbf{n}} S = 0. \tag{7.6}$$

Note that, in view of the Neumann (no-flux) boundary conditions imposed on S in (7.4)–(7.6), there is no loss in generality in assuming

$$\int_{\mathcal{M}} S \mathrm{d}\mathcal{M} = 0 = \int_{\mathcal{M}} F_S \mathrm{d}\mathcal{M} \tag{7.7}$$

(see [60] for further details). Finally, (7.1)–(7.7) are supplemented with initial conditions for $\mathbf{v}$, T and S, that is,

$$\mathbf{v} = \mathbf{v}_0, \quad T = T_0, \quad S = S_0, \quad \text{at } t = 0. \tag{7.8}$$

7.1.2 A reformulation of the equations

Starting from the incompressibility condition, (7.1c) and the hydrostatic equation (7.1b), we may derive an equivalent form for (7.1) as follows:

$$\begin{aligned}
&\partial_t \mathbf{v} + \nabla_{\mathbf{v}} \mathbf{v} + w(\mathbf{v})\partial_z \mathbf{v} + \frac{1}{\rho_0}\nabla p_s + f\mathbf{k} \times \mathbf{v} \\
&- \mu_{\mathbf{v}} \Delta \mathbf{v} - \nu_{\mathbf{v}} \partial_{zz} \mathbf{v} = F_{\mathbf{v}} - \nabla P + \sigma_{\mathbf{v}}(\mathbf{v}, T, S)\dot{W}_1, &&\text{(7.9a)}\\
&\partial_t T + \nabla_{\mathbf{v}} T + w(\mathbf{v})\partial_z T - \mu_T \Delta T - \nu_T \partial_{zz} T = F_T + \sigma_T(\mathbf{v}, T, S)\dot{W}_2, &&\text{(7.9b)}\\
&\partial_t S + \nabla_{\mathbf{v}} S + w(\mathbf{v})\partial_z S - \mu_S \Delta S - \nu_S \partial_{zz} S = F_S + \sigma_S(\mathbf{v}, T, S)\dot{W}_3, &&\text{(7.9c)}\\
&\rho = \rho_0(1 - \beta_T(T - T_r) + \beta_S(S - S_r)), \quad P = P(S,T) = g\int_z^0 \rho \mathrm{d}\overline{z}, &&\text{(7.9d)}\\
&w(\mathbf{v})(\cdot, z) = \int_z^0 \nabla \cdot \mathbf{v} \mathrm{d}\overline{z}, \quad \nabla \cdot \int_{-h}^0 \mathbf{v} \mathrm{d}\overline{z} = 0. &&\text{(7.9e)}
\end{aligned}$$

This reformulation is desirable as, in particular, it is more suitable for the typical functional setting of the equations which we describe next. The unknowns and parameters in the equations are precisely those given above immediately after (7.1). Of course, (7.9) is subject to the same initial and boundary conditions as in (7.1), namely (7.4)–(7.8). For further details concerning the equivalence of (7.9) and (7.1) (see [60]).

7.2 The functional setting and connections with the abstract framework

We now proceed to introduce the basic function spaces associated with the primitive equations (7.9) (equivalently (7.1)), and then introduce and explain the variational formulation of the various terms in equation connecting them with the abstract assumptions laid out above in Section 2.

7.2.1 Basic function spaces

To begin, we define the smooth test functions

$$\begin{aligned}
\mathcal{V} := \mathcal{V}_1 \times \mathcal{V}_2 = &\left\{ \mathbf{v} \in \mathcal{C}^\infty(\overline{\mathcal{M}})^2 : \nabla \cdot \int_{-h}^0 \mathbf{v} \mathrm{d}z = 0, \mathbf{v}_{|\Gamma_l \cap \Gamma_b} = 0 \right\} \\
&\times \left\{ (T, S) \in \mathcal{C}^\infty(\overline{\mathcal{M}})^2 : \int_{\mathcal{M}} S \mathrm{d}\mathcal{M} = 0 \right\}.
\end{aligned}$$

We now take H to be the closure of $\mathcal{V}$ in $L^2(\mathcal{M})^4$ or, equivalently, $H := H_1 \times H_2$, which is

$$\Big\{\mathbf{v} \in L^2(\mathcal{M})^2 : \nabla \cdot \int_{-h}^{0} \mathbf{v}\,\mathrm{d}z = 0, n_H \cdot \int_{-h}^{0} \mathbf{v}\,\mathrm{d}z = 0 \text{ on } \partial\Gamma_i\Big\}$$
$$\times\Big\{(T,S) \in L^2(\mathcal{M})^2 : \int_{\mathcal{M}} S\,\mathrm{d}\mathcal{M} = 0\Big\}. \tag{7.10}$$

On H, it is convenient to define the inner product and norm according to

$$(U,\widetilde{U})_H := \int_{\mathcal{M}} (\mathbf{v}\cdot\widetilde{\mathbf{v}} + K_T T\widetilde{T} + K_S S\widetilde{S})\mathrm{d}\mathcal{M}, \quad |U| := (U,U)_H^{\frac{1}{2}}.$$

The constants $K_T, K_S > 0$, which are introduced for coercivity in the principal linear terms in the equations, are chosen in order to fulfill (2.2) for (7.14) below. We define Π to be the orthogonal (Leray-type) projection from $L^2(\mathcal{M})^4$ onto H.

We shall next define the H^1 type space $V = V_1 \times V_2$, which is

$$\Big\{\mathbf{v} \in H^1(\mathcal{M})^2 : \int_{-h}^{0} \nabla\cdot\mathbf{v}\,\mathrm{d}z = 0,\ \mathbf{v} = 0 \text{ on } \Gamma_l \cup \Gamma_b\Big\}$$
$$\times\Big\{(T,S) \in H^1(\mathcal{M})^2 : \int_{\mathcal{M}} S\,\mathrm{d}\mathcal{M} = 0\Big\}. \tag{7.11}$$

We endow V with the inner product and norm

$$((U,\widetilde{U}))_V := ((U,\widetilde{U}))_{\mathbf{v}} + K_T((U,\widetilde{U}))_T + K_S((U,\widetilde{U}))_S, \quad \|U\| := ((U,U))^{\frac{1}{2}}, \tag{7.12}$$

where

$$((U,\widetilde{U}))_{\mathbf{v}} := \int_{\mathcal{M}} (\mu_{\mathbf{v}}\nabla\mathbf{v}\cdot\nabla\widetilde{\mathbf{v}} + \nu_{\mathbf{v}}\partial_z\mathbf{v}\cdot\partial_z\widetilde{\mathbf{v}})\,\mathrm{d}\mathcal{M} + \alpha_{\mathbf{v}}\int_{\Gamma_i} \mathbf{v}\cdot\widetilde{\mathbf{v}}\mathrm{d}\Gamma_i,$$
$$((U,\widetilde{U}))_T := \int_{\mathcal{M}} (\mu_T\nabla T\cdot\nabla\widetilde{T} + \nu_T\partial_z T\cdot\partial_z\widetilde{T})\,\mathrm{d}\mathcal{M} + \alpha_T\int_{\Gamma_i} T\widetilde{T}\mathrm{d}\Gamma_i,$$
$$((U,\widetilde{U}))_S := \int_{\mathcal{M}} (\mu_S\nabla S\cdot\nabla\widetilde{S} + \nu_S\partial_z S\cdot\partial_z\widetilde{S})\,\mathrm{d}\mathcal{M}.$$

From (7.11)–(7.12), we may deduce the Poincaré type inequality $|U| \le c\|U\|$ for every $U \in V$. This justifies taking $\|\cdot\|$ as the norm for V (which is equivalent to the H^1 norm). Finally, we define

$$V_{(2)}, V_{(3)} \text{ are the closures of } \mathcal{V} \text{ in } H^2(\mathcal{M})^4, H^3(\mathcal{M})^4 \text{ norms respectively} \tag{7.13}$$

and simply endow $V_{(2)}$ and $V_{(3)}$ with the $H^2(\mathcal{M})$ and $H^3(\mathcal{M})$ norms, respectively. Let V' (resp. $V'_{(2)}$, $V'_{(3)}$) be the dual of V (resp. $V_{(2)}$, $V_{(3)}$) relative to the H inner product.

It is clear with the Rellich-Kondrachov theorem and standard facts about Hilbert spaces that the spaces introduced in (7.10)–(7.13) provide a suitable Gelfand-Lions inclusion as desired for (2.1). On this functional basis we now turn to describe the variational form of (7.9).

7.2.2 The variational form of the equations

To capture most of the linear structure in (7.9), we define the operator A as a continuous linear map from V to V' via the bilinear form

$$a(U,\widetilde{U}) := ((U,\widetilde{U}))_V - \int_{\mathcal{M}} \Big(g\int_z^0 (\beta_S S - \beta_T T)d\overline{z}\Big)\nabla\cdot\widetilde{\mathbf{v}}\mathrm{d}\mathcal{M}. \tag{7.14}$$

We observe that if K_T, K_S in (7.12) are chosen sufficiently large, then a is coercive, namely, it satisfies the condition required by (2.2).

We next define the main nonlinear portion of (7.9). Motivated by (7.9e), we take $w = w(U) := \int_z^0 \nabla \cdot \mathbf{v}\, d\overline{z}$ and then define a bilinear form $B : V \times V \to V'_{(2)}$ via the trilinear form

$$b(U, \widetilde{U}, U^*) := b_{\mathbf{v}}(U, \widetilde{U}, U^*) + K_T \cdot b_T(U, \widetilde{U}, U^*) + K_S \cdot b_S(U, \widetilde{U}, U^*), \tag{7.15}$$

where

$$\begin{aligned} b_{\mathbf{v}}(U, \widetilde{U}, U^*) &:= \int_{\mathcal{M}} ((\mathbf{v} \cdot \nabla_2)\widetilde{\mathbf{v}} + w(U)\partial_z \widetilde{\mathbf{v}}) \cdot \mathbf{v}^* \, \mathrm{d}\mathcal{M}, \\ b_T(U, \widetilde{U}, U^*) &:= \int_{\mathcal{M}} ((\mathbf{v} \cdot \nabla)\widetilde{T} + w(U)\partial_z \widetilde{T}) T^* \, \mathrm{d}\mathcal{M}, \\ b_S(U, \widetilde{U}, U^*) &:= \int_{\mathcal{M}} ((\mathbf{v} \cdot \nabla)\widetilde{S} + w(U)\partial_z \widetilde{S}) S^* \, \mathrm{d}\mathcal{M}. \end{aligned}$$

To capture the rotation (Coriolis) term in (7.9a), we define $E : H \to H$ via

$$e(U, \widetilde{U}) = \int_{\mathcal{M}} (2f\mathbf{k} \times \mathbf{v}) \cdot \widetilde{\mathbf{v}} \mathrm{d}\mathcal{M}. \tag{7.16}$$

Note carefully that a, e and b satisfy the conditions imposed in Subsection 2.1 which we used in the abstract result Theorem 2.1. The inhomogenous terms in (7.9) are given by the element ℓ defined according to

$$\begin{aligned} \ell(\widetilde{U}) = &\int_{\mathcal{M}} (F_{\mathbf{v}} \widetilde{v} + K_T F_T \widetilde{T} + K_S F_S \widetilde{S}) \mathrm{d}\mathcal{M} + \int_{\mathcal{M}} \Big(g \int_z^0 (1 + \beta_T T_r - \beta_S S_r) \mathrm{d}z \Big) \nabla \cdot \widetilde{\mathbf{v}} \mathrm{d}\mathcal{M} \\ &+ \int_{\Gamma_i} [(\tau_{\mathbf{v}} + \alpha_{\mathbf{v}} \mathbf{v}^a) \cdot \widetilde{v} + \alpha_T T^a \widetilde{T}] \mathrm{d}\Gamma_i. \end{aligned} \tag{7.17}$$

Note that $\mathbf{v}^a, \tau_{\mathbf{v}}, T^a$, which represent the velocity, shear force of the wind and the temperature at the surface of ocean, have significant uncertainties and should thus be considered to have a random component in practice.

7.3 Some stochastic forcing regimes

It remains to complete the connection between (7.1) and (2.13) by describing various physically interesting scenarios for $\sigma(U)\dot{W}$. We connect these "concrete descriptions" with the terms σ and ξ in the abstract equation (2.13) (or equivalently to g, s in (2.19)). We consider three situations in detail below. In each case, we describe how to define σ_U appearing in (7.9), and we then take $\sigma(\cdot) = \Pi\sigma_U(\cdot)$.

7.3.1 Additive noise

The most classical case is to consider an additive noise, where we suppose that σ_U is independent of $U = (\mathbf{v}, T, S)$. In other words $\sigma_U : [0, \infty) \times \mathcal{M} \to (L_2(\mathfrak{U}, L^2(\mathcal{M})))^4$. For (2.10) to be satisfied, we would require that

$$\sup_{t \geq 0} \sum_{k \geq 1} |\sigma_U^k(t)|^2 = \sup_{t \geq 0} |\sigma_U(t)|^2_{L_2(\mathfrak{U}, H)} < \infty. \tag{7.18}$$

Note that since the Itô and Stratonovich interpretations of (7.2) coincide in the additive case, we may take $\xi \equiv 0$ so that (2.11) is automatically satisfied.

We also observe that in this case we may give an explicit (if formal) characterization of the space-time correlation structure of the noise

$$\mathbb{E}[\sigma_U(t,\mathbf{x})\dot{W}(t,\mathbf{x}) \cdot \sigma_U(s,\mathbf{y})\dot{W}(s,\mathbf{y})] = K(t,s,\mathbf{x},\mathbf{y})\delta_{t-s}, \tag{7.19}$$

where the correlation kernel K is given by

$$K(t,s,\mathbf{x},\mathbf{y}) = \sum_{k\geq 1} \sigma_U^k(t,\mathbf{x}) \cdot \sigma_U^k(s,\mathbf{y}).$$

Remark 7.2 Given the condition (7.18), the case of space-time white noise is ruled out under our framework. Of course such a space-time white noise is very degenerate in space (not even defined in L^2_x) and so such a situation is far from reach due to the highly nonlinear character of the PEs. Similar remarks apply to the 3D stochastic Navier-Stokes equations, but see [18] for the 2-D case.

7.3.2 Nemytskii type operators

We next consider stochastic forcings of transformations of the unknown U as follows. Let $\Psi = (\Psi_{\mathbf{v}}, \Psi_T, \Psi_S) : \mathbb{R}^4 \to \mathbb{R}^4$ and suppose, for simplicity, that Ψ is smooth. We denote the partial derivatives of Ψ with respect to the $\mathbf{v}$, T, S variables by $\partial_{\mathbf{v}}\Psi$, $\partial_T\Psi$, $\partial_S\Psi$ respectively and the gradient by $\nabla_U\Psi$. Take a sequence of smooth functions $\alpha^k = \alpha^k(\mathbf{x}) : \mathcal{M} \to \mathbb{R}$ and define

$$\sigma_U^k(U,t,\mathbf{x}) = \Psi(U)\alpha^k(\mathbf{x}). \tag{7.20}$$

We may formally interpret $\sigma_U(U)\dot{W} = \Psi(U)\dot{\eta}$, where

(1) $\dot{\eta}$ is a white in time Gaussian process with the spatial-temporal correction structure $\mathbb{E}(\dot{\eta}(t,\mathbf{x})\dot{\eta}(s,\mathbf{y})) = K(\mathbf{x},\mathbf{y})\delta_{t-s}$, where $K(\mathbf{x},\mathbf{y}) = \sum_{k\geq 1} \alpha^k(\mathbf{x}) \cdot \alpha^k(\mathbf{y})$.

(2) The “multiplication” $\Psi(U)$ and $\dot{\eta}$ may be taken in either the Itô or the Stratonovich sense.

We now connect (7.20) to (2.13) in the Itô or the Stratonovich situations in turn illustrating conditions on Ψ and the α_k’s guarantee that (2.10) holds and in the Stratonovich case that (2.11) holds.

The Itô case Suppose that

$$|\Psi(U)|^2 \leq c_\Psi(1 + |U|^2) \quad \text{ for all } U \in \mathbb{R}^4, \tag{7.21}$$

and for the elements α^k, we suppose that

$$\sum_{k\geq 1} \|\alpha^k\|^2_{V_{(2)}} < \infty. \tag{7.22}$$

Under (7.21)–(7.22), we have

$$|\sigma(U)|^2_{L_2(\mathfrak{U},H)} \leq \sum_{k\geq 1} |\Psi(U)\alpha^k|^2_{L^2} \leq c_\Psi \sum_{k\geq 1} \|\alpha^k\|^2_{L^\infty(\mathcal{M})}(1 + |U|^2_H) \leq c \sum_{k\geq 1} \|\alpha^k\|^2_{V_{(2)}}(1 + |U|^2_H),$$

so that (2.10) holds for the constant c_3 that depends on c_Ψ, $\sum_{k\geq 1}\|\alpha^k\|^2_{V_{(2)}}$ and the constant in Agmon's inequality. Note that, since we are considering the case of an Itô noise, $\xi \equiv 0$.

The Stratonovich case If we understand the multiplication $\Psi(U)\dot{\eta}$ in the Strantonovich sense, then we may convert back to an Itô type evolution according to

$$\Psi(U)\dot{\eta} = \sum_{k\geq 1}\Psi(U)\alpha^k \circ \mathrm{d}W^k = \xi_U(U) + \sum_{k\geq 1}\Psi(U)\alpha^k\mathrm{d}W^k, \tag{7.23}$$

where

$$\xi_U(U,x) = \Psi(U)\cdot\nabla_U\Psi(U)\sum_{k=1}\alpha^k(x)^2.$$

One can refer to, e.g., [1, 40] for further details on this conversion formula. Under the additional assumption

$$|\nabla_U\Psi(U)| \leq c < \infty \quad \text{ for all } U \in \mathbb{R}^4, \tag{7.24}$$

we define $\xi_U(U) := \Pi\xi(U)$ for any $U \in H$. It is clear that ξ satisfies (2.11).

Remark 7.3 We note here that the relationship (7.23) is, for now, only formal; we prove the existence of martingale solutions to the system that results from a formal application of this conversion formula (see, e.g., [1, 40]). We leave the rigorous justification of (7.23) and the related issues of an approximation of Wong-Zakai type (see [74]) of (2.13) for future work. Note however that (7.23) has already been explored in [12, 29, 72] in an infinite dimensional fluids context for pathwise solutions and in [73] for martingale solutions to a class of abstract, nonlinear, stochastic PDEs.

7.3.3 Stochastic forcing of functionals

Finally, we examine the case when we stochastically force functionals of the unknown, i.e., terms which have a non-local dependence on the solution U. For example consider, for $k \geq 1$ continuous (not necessarily linear) $\phi^k := \phi^k(U) : H \to \mathbb{R}$, and sufficiently smooth $\alpha^k = \alpha^k(t,\mathbf{x}) : [0,\infty)\times\mathcal{M} \to \mathbb{R}^4$. We define

$$\sigma_U^k(U,t,\mathbf{x}) = \phi^k(U)\alpha^k(t,\mathbf{x}). \tag{7.25}$$

Here, we interpret $\sigma_U(U)\dot{W}$ in the Itô sense. Subject to, for example,

$$\sup_k|\phi^k(U)|^2 \leq c(1+|U|^2), \quad \sup_{t\geq 0}\sum_{k\geq 1}\|\alpha^k(t)\|^2 < \infty, \tag{7.26}$$

we obtain a σ from (7.25) which satisfies (2.10). For a "concrete example" of a σ of the form (7.25) which satisfies (7.26), let $\{\psi^k\}_{k\geq 1}$ be a sequence of elements in $L^2(\mathcal{M})^2$ with $\sup_k|\psi^k|_{L^2(\mathcal{M})} < \infty$ and let $\alpha^k \in V$ satisfying the sumability condition in (7.26). We take $\phi^k(U) = \int_M \mathbf{v}(\mathbf{x})\cdot\psi^k(\mathbf{x})\mathrm{d}\mathcal{M}$ and obtain

$$\sigma(U)\dot{W} = \sum_{k\geq 1}\Big(\int_{\mathcal{M}}\mathbf{v}(\mathbf{x})\cdot\psi^k(\mathbf{x})\mathrm{d}\mathcal{M}\Big)\alpha^k(t,\mathbf{x})\mathrm{d}W^k(t). \tag{7.27}$$

A Appendix: Technical Complements

We collect here, for the convenience of the reader, various technical results which have been used in the course of the analysis above. While some of the material may be considered to be somewhat "classical" by specialists, we believe that the stochastic type results will be useful to the non-probabilists and that the deterministic results will be helpful for the probabilists.

A.1 Some convergence properties of measures

We next briefly review some basic notations of convergence for collections of Borel probability measures. In particular, we highlight a certain abstract convergence lemma that has been used in a crucial way in the passage to the limit several times above. For further details concerning the general theory of convergence in spaces of probability measures, one can refer to, e.g., [5, 65].

Let $(\mathcal{H}, \rho)$ be a complete metric space and denote by $\Pr(\mathcal{H})$ the collection of Borel probability measures on $\mathcal{H}$. We recall that a sequence $\{\mu_n\}_{n\geq 1} \subset \Pr(\mathcal{H})$ is said to converge weakly to a measure μ on $\mathcal{H}$ (denoted by $\mu_n \rightharpoonup \mu$) if and only if

$$\lim_{n\to\infty} \int f(x)\mathrm{d}\mu_n(x) = \int f(x)\mathrm{d}\mu(x) \text{ for every bounded continuous function } f : \mathcal{H} \to \mathbb{R}. \quad \text{(A.1)}$$

We recall that a collection $\Lambda \subset \Pr(\mathcal{H})$ is said to be weakly relatively compact if every sequence $\{\mu_n\}_{n\geq 1} \subset \Lambda$ possesses a weakly convergent subsequence. On the other hand, we say that $\Lambda \subset \Pr(\mathcal{H})$ is tight if, for every $\epsilon > 0$ there exists a compact set $K_\epsilon \subset \mathcal{H}$ such that $\mu(K_\epsilon) \geq 1-\epsilon$ for each $\mu \in \Lambda$. The Prokhorov theorem asserts that these two notions, namely tightness and weak compactness of probability measures, are equivalent.

We also make use of the Skorokhod embedding theorem which states that, whenever $\mu_n \rightharpoonup \mu$ on $\mathcal{H}$, then there exists a probability space $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}})$ and a sequence of random variables $X_n : \widetilde{\Omega} \to \mathcal{H}$ such that $\widetilde{\mathbb{P}}(X_n \in \cdot) = \mu_n(\cdot)$ and which converges a.s. to a random variable $X : \widetilde{\Omega} \to \mathcal{H}$ with $\widetilde{\mathbb{P}}(X \in \cdot) = \mu(\cdot)$.

The following convergence result, found in e.g. [5], relates roughly speaking weak convergence and clustering in probability, and was used to facilitate the proof of (5.26) in Subsection 5.1.3.

Lemma A.1 *Let $(\mathcal{H}, \rho)$ be an arbitrary metric space. Suppose that X_n and Y_n are $\mathcal{H}$-valued random variables, and let $\mu_n(\cdot) = \mathbb{P}(X_n \in \cdot)$ and $\nu_n(\cdot) = \mathbb{P}(Y_n \in \cdot)$ be the associated sequences of the probability laws. If the sequence $\{\mu_n\}_{n\geq 0}$ converges weakly to a probability measure μ and if, for all $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(\rho(X_n, Y_n) \geq \epsilon) = 0,$$

then ν_n also converges weakly to μ.

A.2 An extension of the Doob-Dynkin lemma

We extend the Doob-Dynkin lemma (see, e.g., [54]) to the case where the image space of the measurable functions are complete separable metric spaces. In order to achieve this goal, let us recall the following notions and results from [21].

If $(\Omega, \mathcal{F})$ is a measure space and $E \subset \Omega$, let $\mathcal{F}_E := \{B \cap E : B \in \mathcal{F}\}$. Then $\mathcal{F}_E$ is a sigma algebra of subsets of E, and $\mathcal{F}_E$ will be called the relative sigma algebra (of $\mathcal{F}$ on E).

Proposition A.1 *Let $(\Omega, \mathcal{F})$ be any measurable space and E be any subset of Ω (not necessarily in $\mathcal{F}$). Let f be a function on E with values in a Polish space $\mathcal{H}$ and measurable with respect to $\mathcal{F}_E$. Then f can be extended to a function on all of Ω, measurable with respect to $\mathcal{F}$.*

Proof The proof is direct combining Theorem 4.2.5 and Proposition 4.2.6 in [21].

Now let $(\mathcal{Y}, \mathcal{M})$ be a measure space, $\mathcal{X}$ be any set, and ψ be a function from $\mathcal{X}$ into $\mathcal{Y}$. Let $\psi^{-1}[\mathcal{M}] := \{\psi^{-1}(M) : M \in \mathcal{M}\}$. Then $\psi^{-1}[\mathcal{M}]$ is a sigma algebra of subsets of $\mathcal{X}$.

Theorem A.1 *Given a set $\mathcal{X}$, a measure space $(\mathcal{Y}, \mathcal{M})$, and a function ψ from $\mathcal{X}$ into $\mathcal{Y}$, if a function ℓ on $\mathcal{X}$ with values in a Polish space $\mathcal{H}$ is $\psi^{-1}[\mathcal{M}]$ measurable, then there exists an $\mathcal{M}$-measurable function L on $\mathcal{Y}$ such that $\ell = L \circ \psi$.*

Proof Whenever $\psi(u) = \psi(v)$, we have $\ell(u) = \ell(v)$. Otherwise, let B be a Borel set in $\mathcal{H}$ with $\ell(u) \in B$ but $\ell(v) \notin B$. Then $\ell^{-1}(B) = \psi^{-1}(C)$ for some $C \in \mathcal{M}$, with $\psi(u) \in C$ but $\psi(v) \notin C$, a contradiction. Thus, $\ell = L \circ \psi$ for some function L from $D :=$ range ψ into $\mathcal{H}$. For any Borel set $E \subset \mathcal{H}$, $\psi^{-1}(L^{-1}(E)) = \ell^{-1}(E) = \psi^{-1}(F)$ for some $F \in \mathcal{M}$, so $F \cap D = L^{-1}(E)$ and L is $\mathcal{M}_D$ measurable. By Proposition A.1, L has an $\mathcal{M}$-measurable extension to all of $\mathcal{Y}$.

A.3 A measurable selection theorem

We turn now to restate the measurable selection theorem which was proven in [10] and is based on the earlier works (see [11, 41]). We employed this result above to establish the existence of adapted solutions to (3.8) in Proposition 3.1.

Firstly, we recall the definition of a Radon measure. Let X be a locally compact Hausdorff spaces and $\mathcal{B}(X)$ be the Borel sigma algebra on X. A Radon measure on X is a measure defined on $\mathcal{B}(X)$ that is finite on all compact sets, outer regular on all Borel sets, and inner regular on all open sets (see [26, p. 212]).

Theorem A.2 *Let X and Y be separable Banach spaces and suppose that Λ is a "multivalued map" from X into Y, i.e., a map from X into the subsets of Y. We assume that Λ takes values in closed, non-empty subsets of Y and that its graph is closed viz.,*

$$\text{if } x_n \to x \text{ in } X, \text{ and } y_n \to y \text{ in } Y, \text{ with } y_n \in \Lambda x_n, \text{ then } y \in \Lambda x.$$

Then, Λ admits a universal Radon measurable section, Γ, that is there exists a map $\Gamma : X \to Y$ such that $\Gamma x \in \Lambda x$ for every x, and such that Γ is Radon measurable for every Radon measure on X.

Remark A.1 Note that since X is a separable Banach space, any probability measure on X is Radon; this is because any separable Banach space is a Polish space (separable and complete metric space) and that every Polish space is a Radon space (a Hausdorff space X is called a Radon space if every finite Borel measure on X is a Radon measure, i.e., is inner regular (see [66]).

The following results are from [20, 66]. The final goal is to establish Corollary A.1 below, which we have employed in the article to prove that the map χ defined in (3.17) (see Subsection 3.2) is universally Radon measurable. For that purpose, we need to introduce the following results (from Proposition A.2 to Theorem A.3).

Definition A.1 (Lusin μ-Measurable) *Let X be a topological space. Let μ be a Radon measure on X and let h map X into Y, where Y is a Hausdorff topological space. Then the mapping h is said to be Lusin μ-measurable if, for every compact set $K \subset X$ and every $\delta > 0$, there exists a compact set $K_\delta \subset K$ with $\mu(K - K_\delta) \leq \delta$ such that h restricted to K_δ is continuous.*

Proposition A.2 *A function whose restriction to every compact set is continuous, is Lusin measurable for every Radon measure* (*see* [66, p. 25]).

Proposition A.3 *The assumptions are the same as in Definition A.1. If $h : X \to Y$ is Lusin μ-measurable, then h is μ-measurable, and conversely, if Y is metrizable and separable, then every μ-measurable function is also Lusin μ-measurable* (*see* [66, p. 26]).

Theorem A.3 *Let X, Y and Z be separable Banach spaces and μ be a Radon measure on X. Let $\varphi : X \to Y$ be a μ-measurable mapping. Let $\Gamma : Y \to Z$ be universally Radon measurable. Then $G := \Gamma \circ \varphi$ is μ-measurable on X.*

Proof From Proposition A.3, φ is Lusin μ-measurable. Then Theorem A.3 follows from the proof of Theorem 3.2 in [10].

Corollary A.1 *Let X, Y and Z be separable Banach spaces and $\varphi : X \to Y$ be a continuous mapping. Let $\Gamma : Y \to Z$ be universally Radon measurable. Then $G := \Gamma \circ \varphi$ is universally Radon measurable.*

Proof This can be directly deduced from Proposition A.2 and Theorem A.3.

A.4 Compact embedding results

In order to establish the compactness of a sequence of probability measures associated with the solutions to (3.3), we made use of the following compact embedding theorem which is close to that found in [69] and of course generalizes the classical Aubin-Lions compactness theorem (see [2]).

Proposition A.4 *Let $Z \subset\subset Y \subset X$ be a collection of three Banach spaces with Z compactly embedded in Y and Y continuously embedded in X.*

(i) *Suppose that $\mathfrak{G}$ is a bounded subset of $L^p(\mathbb{R}, Z) \cap L^\infty(\mathbb{R}, Y)$, where $1 < p \leq \infty$, and assume that for some $1 < q < \infty$,*

$$\int_{-\infty}^{\infty} |g(t+s) - g(s)|_X^q \mathrm{d}s \to 0 \quad \text{as } t \to 0 \tag{A.2}$$

uniformly for $g \in \mathfrak{G}$ and that there exists $L > 0$ such that

$$\mathrm{supp}\{g\} \subset [-L, L] \quad \text{for every } g \in \mathfrak{G}. \tag{A.3}$$

Then, the set $\mathfrak{G}$ is relatively compact in $L^p(\mathbb{R}, Y)$.

(ii) *For $T > 0$, if $\mathfrak{G}$ is a bounded subset of $L^p(0,T,Z) \cap L^\infty(0,T,Y)$ and*

$$\int_0^{T-a} |g(t+s) - g(s)|_X^q \mathrm{d}s \to 0 \quad \text{as } t \to 0 \tag{A.4}$$

uniformly for elements in $\mathfrak{G}$, then $\mathfrak{G}$ is relatively compact in $L^p(0,T,Y)$.

Proof The proof is a fairly straightforward generalization of [70, Theorem 13.2]. Observe that if $q > p$, then (A.2) and (A.3) together imply that

$$\int_{-\infty}^{\infty} |g(t+s) - g(s)|_X^p \mathrm{d}s \to 0 \quad \text{as } t \to 0$$

uniformly for $g \in \mathfrak{G}$. Therefore there is no loss of generality in supposing that $q \leq p$ in what follows.

For $a > 0$, define the averaging operator J_a according to

$$(J_a f)(s) = \frac{1}{2a} \int_{s-a}^{s+a} f(t)\mathrm{d}t = \frac{1}{2a} \int_{-a}^{a} f(s+t)\mathrm{d}t.$$

We take $\mathfrak{G}_a = \{J_a g : g \in \mathfrak{G}\}$. Arguing exactly as in [69], we have, for $a > 0$, that $\mathfrak{G}_a$ is relatively compact in $L^p(\mathbb{R}; Y)$.

To show that $\mathfrak{G}$ is itself relatively compact in $L^p(\mathbb{R}; Y)$, we prove that it is a totally bounded subset of $L^p(\mathbb{R}; Y)$; in other words, we prove that, for every $\epsilon > 0$, there exists finitely many elements $g_1, \cdots, g_N$ in $L^p(\mathbb{R}, Y)$ such that $\mathfrak{G}$ is contained in the union of the ϵ balls centered at these points.

Again, arguing exactly as in [69], we have that, as a consequence of (A.2), for every $\delta > 0$ there exists $a = a(\delta) > 0$ such that

$$|J_a g - g|_{L^q(\mathbb{R},X)} \leq \delta, \quad \text{for every } g \in \mathfrak{G}. \tag{A.5}$$

On the other hand, from [71, Chapter 3, Lemma 2.1], we infer that, for every $\eta > 0$, there exists $C_\eta > 0$ such that, for every $g \in L^p(\mathbb{R}, Z)$,

$$\begin{aligned} |J_a g - g|_{L^p(\mathbb{R},Y)} &\leq C_\eta |J_a g - g|_{L^p(\mathbb{R},X)} + \eta |J_a g - g|_{L^p(\mathbb{R},Z)} \\ &\leq C_\eta |J_a g - g|_{L^p(\mathbb{R},X)} + 2\eta |g|_{L^p(\mathbb{R},Z)}. \end{aligned} \tag{A.6}$$

The last inequality follows from the fact that $|J_a f|_{L^p(\mathbb{R},Z)} \leq |f|_{L^p(\mathbb{R},Z)}$ for all $f \in L^p(\mathbb{R}, Z)$. Now, on the other hand, we have

$$|J_a g - g|_{L^p(\mathbb{R},X)} \leq |J_a g - g|_{L^\infty(\mathbb{R},X)}^{\frac{p-q}{p}} |J_a g - g|_{L^q(\mathbb{R},X)}^{\frac{q}{p}} \leq (2|g|_{L^\infty(\mathbb{R},X)})^{\frac{p-q}{p}} |J_a g - g|_{L^q(\mathbb{R},X)}^{\frac{q}{p}}.$$

So, for a constant κ depending only on $\sup_{g \in \mathfrak{G}} |g|_{L^\infty(\mathbb{R},Y)}$, p, q and the constant associated with the continuous embedding of Y into X, we find

$$|J_a g - g|_{L^p(\mathbb{R},X)} \leq \kappa |J_a g - g|_{L^q(\mathbb{R},X)}^{\frac{q}{p}}. \tag{A.7}$$

Fix $\epsilon > 0$. Let $|f|_{L^p(\mathbb{R},Z)} \leq \kappa$, $\forall f \in \mathfrak{G}$ and let $\eta = \frac{\epsilon}{6\kappa}$, and pick $a > 0$, sufficiently small, so that (A.5) holds for $\delta := \left(\frac{\epsilon}{3C_\eta \kappa}\right)^{\frac{p}{q}}$, where C_η is the constant corresponding to η in (A.6). Using that $\mathfrak{G}_a$ is precompact in $L^p(\mathbb{R}, Y)$, we next choose a finite collection $\mathfrak{F} = \{g_1, \cdots, g_n\} \subset \mathfrak{G}$, such that the $L^p(\mathbb{R}, Y)$ $\frac{\epsilon}{3}$-balls centered at $J_a g_k$ cover $\mathfrak{G}_a$. Now, with these various choices, we have that for any $g \in \mathfrak{G}$, there exists $g_k \in \mathfrak{F}$ such that $|J_a g_k - J_a g|_{L^p(\mathbb{R},Y)} \le \frac{\epsilon}{3}$. As such, we employ (A.6) with $\eta = \frac{\epsilon}{6\kappa}$, followed by (A.7) and estimate

$$\begin{aligned} |J_a g_k - g|_{L^p(\mathbb{R},Y)} &\le |J_a g_k - J_a g|_{L^p(\mathbb{R},Y)} + |J_a g - g|_{L^p(\mathbb{R},Y)} \\ &\le \frac{\epsilon}{3} + C_\eta |J_a g - g|_{L^p(\mathbb{R},X)} + 2\eta \sup_{f \in \mathfrak{G}} |f|_{L^p(\mathbb{R},Z)} \\ &\le \frac{2\epsilon}{3} + C_\eta \kappa |J_a g - g|^{\frac{q}{p}}_{L^q(\mathbb{R},X)} \le \epsilon. \end{aligned}$$

Since, $\epsilon > 0$ is arbitrary to begin with, this shows that $\mathfrak{G}$ is a totally bounded subset of $L^p(\mathbb{R}; Y)$, and we thus infer (i). The second item (ii) follows directly from (i) as in [69]. The proof of Proposition A.4 is therefore complete.

Acknowledgements The authors have benefited from the hospitality of the department of mathematics Virginia Tech and from the Newton institute for mathematical sciences, University of Cambridge where the final stage of the writing was completed. The authors wish to thank Arnaud Debussche for his help with the use of the universally Radon measurable selection theorem. The authors also thank the referee(s) for bringing their attention to the article [3] which they regrettably overlooked.

References

[1] Arnold, L., Stochastic differential equations: Theory and applications, translated from the German, Wiley-Interscience, John Wiley Sons, New York, 1974.

[2] Aubin J.-P., Approximation of elliptic boundary-value problems, Wiley-Interscience, Pure and Applied Mathematics, Vol. XXVI, A Division of John Wiley Sons, Inc., New York-London-Sydney, 1972.

[3] Brzeźniak, Z., Carelli, E. and Prohl, A., Finite-element-based discretizations of the incompressible Navier-Stokes equations with multiplicative random forcing, *IMA J. Numer. Anal.*, **33**(3), 2013, 771–824.

[4] Bensoussan, A., Stochastic Navier-Stokes equations, *Acta Appl. Math.*, **38**(3), 1995, 267–304.

[5] Billingsley, P., Convergence of probability measures, 2nd edition, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley Sons Inc., New York, 1999.

[6] Bjerknes, V., Das problem der wettervorhersage, betrachtet vom standpunkte der mechanik und der physik, *Meteorol. Z.*, **21**, 1904, 1–7.

[7] Brézis, H., Problèmes unilatéraux, *J. Math. Pures Appl.*, **51**(9), 1972, 1–168.

[8] Browder, F. E., Nonlinear monotone and accretive operators in Banach spaces, *Proc. Nat. Acad. Sci. U.S.A.,* **61**, 1968, 388–393.

[9] Berner, J., Shutts, G. J., Leutbecher, M. and Palmer, T. N., A spectral stochastic kinetic energy backscatter scheme and its impact on ow-dependent predictability in the ecmwf ensemble prediction system, *J. Atmospheric Sci.*, **66**(3), 2009, 603–626.

[10] Bensoussan, A. and Temam, R., Équations stochastiques du type Navier-Stokes, *J. Functional Analysis*, **13**, 1973, 195–222.

[11] Castaing, C., Sur les multi-applications measurables, *Rev. Française Informat. Recherche Opérationnell*, **1**(1), 1967, 91–126.

[12] Chueshov, I. and Millet, A., Stochastic two-dimensional hydrodynamical systems: Wong-Zakai approximation and support theorem, *Stoch. Anal. Appl.*, **29**(4), 2011, 570–611.

[13] Cao, C. and Titi, E., Global well-posedness of the three-dimensional viscous primitive equations of large scale ocean and atmosphere dynamics, *Ann. Math.* (2), **166**(1), 2007, 245–267.

[14] De Bouard, A. and Debussche, A., A semi-discrete scheme for the stochastic nonlinear Schrödinger equation, *Numer. Math.*, **96**(4), 2004, 733–770.

[15] Debussche, A., Glatt-Holtz, N. and Temam, R., Local martingale and pathwise solutions for an abstract uids model, *Phys. D*, **240**(14–15), 2011, 1123–1144.

[16] Debussche, A., Glatt-Holtz, N., Temam, R. and Ziane, M., Global existence and regularity for the 3D stochastic primitive equations of the ocean and atmosphere with multiplicative white noise, *Nonlinearity*, **25**(7), 2012, 2093–2118.

[17] Debussche, A. and Printems, J., Convergence of a semi-discrete scheme for the stochastic Korteweg-de Vries equation, *Discrete Contin. Dyn. Syst. Ser. B*, **6**(4), 2006, 761–781.

[18] Da Prato, G. and Debussche, A., Two-dimensional Navier-Stokes equations driven by a space-time white noise, *J. Funct. Anal.*, **196**(1), 2002, 180–210.

[19] Da Prato, G. and Zabczyk, J., Stochastic equations in infinite dimensions, Encyclopedia of Mathematics and Its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.

[20] Dunford, N. and Schwartz, J. T., Linear Operators, Part I, Wiley Classics Library, John Wiley Sons Inc., New York, 1988, General theory, with the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[21] Dudley, R. M., Real analysis and probability, Cambridge Studies in Advanced Mathematics, vol. 74, Cambridge University Press, Cambridge, 2002, Revised reprint of the 1989 original.

[22] Durrett, R., Probability: Theory and examples, 4th edition, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge, 2010.

[23] Ewald, B. and Penland, C., Numerical generation of stochastic differential equations in climate models, Special Volume on Computational Methods for the Atmosphere and the Oceans, Handbook of Numerical Analysis, vol. 14, Elsevier/North-Holland, Amsterdam, 2009, 279–306.

[24] Ewald, B., Petcu, M. and Temam, R., Stochastic solutions of the two-dimensional primitive equations of the ocean and atmosphere with an additive noise, *Anal. Appl.* (*Singap.*), **5**(2), 2007, 183–198.

[25] Flandoli, F. and Gatarek, D., Martingale and stationary solutions for stochastic Navier-Stokes equations, *Probab. Theory Related Fields*, **102**(3), 1995, 367–391.

[26] Folland, G. B., Real analysis, 2nd edition, Pure and Applied Mathematics (New York), John Wiley Sons Inc., New York, 1999, Modern Techniques and Their Applications, A Wiley-Interscience Publication.

[27] Guo, B. and Huang, D., 3D stochastic primitive equations of the large-scale ocean: Global well-posedness and attractors, *Comm. Math. Phys.*, **286**(2), 2009, 697–723.

[28] Gyöongy, I. and Millet, A., On discretization schemes for stochastic evolution equations, *Potential Anal.,* **23**(2), 2005, 99–134.

[29] Grecksch, W. and Schmalfuß, B., Approximation of the stochastic Navier-Stokes equation, *Mat. Apl. Comput.*, **15**(3), 1996, 227–239.

[30] Glatt-Holtz, N. and Temam, R., Cauchy convergence schemes for some nonlinear partial differential equations, *Applicable Analysis*, **90**(1), 2011, 85–102.

[31] Glatt-Holtz, N. and Temam, R., Pathwise solutions of the 2-D stochastic primitive equations, *Applied Mathematics and Optimization*, **63**(3), 2011, 401–433.

[32] Glatt-Holtz, N., Temam, R. and Tribbia, J., Some remarks on the role of stochastic parameterization in the equations of the ocean and atmosphere, in preparation.

[33] Glatt-Holtz, N., Temam, R. and C. Wang, Numerical analysis of the stochastic Navier-Stokes equations: Stability and convergence results, in preparation.

[34] Glatt-Holtz, N. and Vicol, V. C., Local and global existence of smooth solutions for the stochastic Euler equations with multiplicative noise, *Ann. Probab.*, **42**(1), 2014, 80–145.

[35] Glatt-Holtz, N. and Ziane, M., The stochastic primitive equations in two space dimensions with multiplicative noise, Discrete Contin, *Dyn. Syst. Ser. B*, **10**(4), 2008, 801–822.

[36] Hasselmann, K., Stochastic climate models, part I: Theory, *Tellus*, **28**, 1976, 474–485.

[37] Horsthemke, W. and Lefever, R., Noise-induced transitions: Theory and applications in physics, chemistry and biology, Springer-Verlag, New York, 1984.

[38] Kobelkov, G. M., Existence of a solution "in the large" for the 3D large-scale ocean dynamics equations, *C. R. Math. Acad. Sci. Paris*, **343**(4), 2006, 283–286.

[39] Kobelkov, G. M., Existence of a solution "in the large" for ocean dynamics equations, *J. Math. Fluid Mech.*, **9**(4), 2007, 588–610.

[40] Kloeden, P. and Platen, E., Numerical solution of stochastic differential equations, Applications of Mathematics (New York), vol. 23, Springer-Verlag, Berlin, 1992.

[41] Kuratowski, K. and Ryll-Nardzewski, C., A general theorem on selectors, *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, **13**, 1965, 397–403.

[42] Karatzas, I. and Shreve, S. E., Brownian motion and stochastic calculus, 2nd edition, Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991.

[43] Kukavica, I. and Ziane, M., On the regularity of the primitive equations of the ocean, *Nonlinearity*, **20**(12), 2007, 2739–2753.

[44] Leray, J. and Lions, J.-L., Quelques résulatats de Višik sur les problèmes elliptiques nonlinéaires par les méthodes de Minty-Browder (in France), *Bull. Soc. Math.*, **93**, 1965, 97–107.

[45] Leslie, D. C. and Quarini, G. L., The application of turbulence theory to the formulation of subgrid modelling procedures, *Journal of Fluid Mechanics*, **91**, 1979, 65–91.

[46] Lions, J.-L., Temam, R. and Wang, S. H., New formulations of the primitive equations of atmosphere and applications, *Nonlinearity*, **5**(2), 1992, 237–288.

[47] Lions, J.-L., Temam, R. and Wang, S. H., On the equations of the large-scale ocean, *Nonlinearity*, **5**(5), 1992, 1007–1053.

[48] Lions, J.-L., Temam, R. and Wang, S., Models for the coupled atmosphere and ocean (CAO I, II), *Comput. Mech. Adv.*, **1**(1), 1993, 120.

[49] Minty, G. J., Monotone (nonlinear) operators in Hilbert space, *Duke Math. J.*, **29**, 1962, 341–346.

[50] Manna, U, Menaldi, J. L. and Sritharan, S. S., Stochastic 2-D Navier-Stokes equation with artificial compressibility, *Commun. Stoch. Anal.*, **1**(1), 2007, 123–139.

[51] Menaldi, J. L. and Sritharan, S. S., Stochastic 2-D Navier-Stokes equation, *Appl. Math. Optim.*, **46**(1), 2002, 31–53.

[52] Mason, P. J. and Thomson, D. J., Stochastic backscatter in large-eddy simulations of boundary layers, *Journal of Fluid Mechanics*, **242**, 1992, 51–78.

[53] Marion, M. and Temam, R., Navier-Stokes equations: theory and approximation, Handbook of numerical analysis, Vol. VI, *Handb. Numer. Anal.*, VI, North-Holland, Amsterdam, 1998, 503–688. MR 1665429 (2000a:76002)

[54] Øksendal, B., Stochastic Differential Equations, An Introduction with Applications, 6th edition, Universitext, Springer-Verlag, Berlin, 2003.

[55] Penland, C. and Ewald, B. D., On modelling physical systems with stochastic models: Diffusion versus Lévy processes, *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **366**(1875), 2008, 2457–2476.

[56] Pedlosky, J., Geophysical uid Dynamics, Springer-Verlag, New York, 1982.

[57] Penland, C., A stochastic approach to nonlinear dynamics: A review, *Bulletin of the American Meteorological Society*, **84**, 2003, ES43–ES51.

[58] Prévôt, C. and Röockner, M., A concise course on stochastic partial differential equations, Lecture Notes in Mathematics, **1905**, Springer-Verlag, Berlin, 2007.

[59] Penland, C. and Sardeshmukh, P. D., The optimal growth of tropical sea surface temperature anomalies, *Journal of Climate*, **8**(8), 1995, 1999–2024.

[60] Petcu, M., Temam, R. and Ziane, M., Some mathematical problems in geophysical fluid dynamics, Special Volume on Computational Methods for the Atmosphere and the Oceans, Handbook of Numerical Analysis, **14**, Elsevier, 2008, 577–750.

[61] Richardson, L. F., Weather prediction by numerical process, with a foreword by Peter Lynch, 2nd edition, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2007.

[62] Rose, H. A., Eddy diffusivity, eddy noise and subgrid-scale modelling, *Journal of Fluid Mechanics*, **81**, 1977, 719–734.

[63] Rozovskii, B., Temam, R. and Tribbia, J., AIM Workshop: Mathematical and Geophysical Fluid Dynamics, Analytical and Stochastic Methods, Palo Alto, 2006.

[64] Rousseau, A, Temam, R. M. and Tribbia, J. J., Boundary value problems for the inviscid primitive equations in limited domains, Handbook of numerical analysis, Vol. XIV, Special volume: Computational methods for the atmosphere and the oceans, *Handb. Numer. Anal.*, **14**, Elsevier/North-Holland, Amsterdam, 2009, 481–575.

[65] Revuz, D. and Yor, M., Continuous martingales and Brownian motion, 3rd edition, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **293**, Springer-Verlag, Berlin, 1999.

[66] Schwartz, L, Radon measures on arbitrary topological spaces and cylindrical measures, Published for the Tata Institute of Fundamental Research, Bombay by Oxford University Press, London, 1973, Tata Institute of Fundamental Research Studies in Mathematics, No. 6.

[67] Temam, R., Sur l'approximation des solutions des équations de Navier-Stokes, *C. R. Acad. Sci. Paris Sér. A-B*, **262**, 1966, A219–A221.

[68] Temam, R., Une méthode d'approximation de la solution des équations de Navier-Stokes (in France), *Bull. Soc. Math.*, **96** 1968, 115–152.

[69] Temam, R., Navier-Stokes equations and nonlinear functional analysis, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 41, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1983.

[70] Temam, R., Navier-Stokes equations and nonlinear functional analysis, 2nd edition, CBMS-NSF Regional Conference Series in Applied Mathematics, **66**, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995.

[71] Temam, R., Navier-Stokes Equations: Theory and Numerical Analysis, Reprint of the 1984 edition, A. M. S., Providence, RI, 2001.

[72] Twardowska, K., An approximation theorem of Wong-Zakai type for stochastic Navier-Stokes equations, *Rend. Sem. Mat. Univ. Padova*, **96**, 1996, 15–36.

[73] Tessitore, G. and Zabczyk, J., Wong-Zakai approximations of stochastic evolution equations, *J. Evol. Equ.*, **6**(4), 2006, 621–655.

[74] Wong, E. and Zakai, M., On the relation between ordinary and stochastic differential equations, *Internat. J. Engrg. Sci.*, **3**, 1965, 213–229.

[75] Zidikheri, M. J. and Frederiksen, J. S., Stochastic subgrid-scale modelling for non-equilibrium geophysical ows, Philosophical Transactions of the Royal Society A: Mathematical, *Physical and Engineering Sciences*, **368**(1910), 2010, 145–160.