Discrete and Continuous Dynamical Systems
Vol. 43, No. 2, February 2023, pp. 568-618
doi:10.3934/dcds.2022163

GLOBAL WELL-POSEDNESS OF THE VISCOUS CAMASSA–HOLM EQUATION WITH GRADIENT NOISE

Helge Holden*[1], Kenneth H. Karlsen[2] and Peter H. C. Pang[1,2]

[1]Department of Mathematical Sciences,
NTNU Norwegian University of Science and Technology, NO-7491 Trondheim, Norway

[2] Department of Mathematics, University of Oslo, NO-0316 Oslo, Norway

(Communicated by Adrian Constantin)

Abstract. We analyse a nonlinear stochastic partial differential equation that corresponds to a viscous shallow water equation (of the Camassa–Holm type) perturbed by a convective, position-dependent noise term. We establish the existence of weak solutions in H^m ($m \in \mathbb{N}$) using Galerkin approximations and the stochastic compactness method. We derive a series of a priori estimates that combine a model-specific energy law with non-standard regularity estimates. We make systematic use of a stochastic Gronwall inequality and also stopping time techniques. The proof of convergence to a solution argues via tightness of the laws of the Galerkin solutions, and Skorokhod–Jakubowski a.s. representations of random variables in quasi-Polish spaces. The spatially dependent noise function constitutes a complication throughout the analysis, repeatedly giving rise to nonlinear terms that "balance" the martingale part of the equation against the second-order Stratonovich-to-Itô correction term. Finally, via pathwise uniqueness, we conclude that the constructed solutions are probabilistically strong. The uniqueness proof is based on a finite-dimensional Itô formula and a DiPerna–Lions type regularisation procedure, where the regularisation errors are controlled by first and second order commutators.

Contents

1. Introduction and main results 569
2. Solution concepts 574
3. The Galerkin approximation 575
4. A priori estimates 576
5. Tightness of probability laws 582
6. Weak (martingale) solutions 585
7. Pathwise uniqueness and proof of Theorem 1.1 591
8. Higher regularity solutions (Theorem 1.2) 605

2020 *Mathematics Subject Classification.* Primary: 35R60, 35G25, 60H15; Secondary: 35A01, 35A02.

Key words and phrases. Shallow water equation, viscous Camassa–Holm equation, stochastic perturbation, convective noise, existence, Faedo–Galerkin method, compactness, tightness, Skorokhod–Jakubowski representation, uniqueness, commutator estimate.

This research was partially supported by the Research Council of Norway Toppforsk project *Waves and Nonlinear Phenomena* (250070). The final author is also partially supported by the Research Council of Norway project *INICE* (301538).

*Corresponding author: Helge Holden (helge.holden@ntnu.no).

Acknowledgments 612
Appendix A. Stochastic toolbox 612
REFERENCES 617

1. **Introduction and main results.**

1.1. **Background.**
We are interested in the initial-value problem for the stochastic parabolic-elliptic system

$$\begin{aligned} 0 &= \mathrm{d}u + \left[u\,\partial_x u + \partial_x P - \varepsilon \partial_x^2 u\right]\,\mathrm{d}t \\ &\quad - \frac{1}{2}\sigma(x)\partial_x\left(\sigma(x)\partial_x u\right)\,\mathrm{d}t + \sigma(x)\partial_x u\,\mathrm{d}W, \\ -\partial_x^2 P + P &= u^2 + \frac{1}{2}\left(\partial_x u\right)^2, \quad \text{for } (t,x) \in (0,T)\times\mathbb{S}^1, \end{aligned} \tag{1.1}$$

where $\mathbb{S}^1 = \mathbb{R}/\mathbb{Z}$ is the 1D torus (circle), ε and T are positive numbers, $\sigma = \sigma(x) \in W^{2,\infty}(\mathbb{S}^1)$ is a position-dependent noise amplitude, and W is a 1D Brownian motion defined on a probability space and adapted to some filtration (further details will be given later). Formally, by the Itô–Stratonovich conversion formula, the two σ terms in (1.1) can be combined into the simple looking Stratonovich differential $\sigma\,\partial_x u\circ\mathrm{d}W$, which in the literature is referred to as a gradient, transport or convection noise term. The elliptic equation for P can be "solved" to give

$$P = P[u] := K * \left(u^2 + \frac{1}{2}\left(\partial_x u\right)^2\right), \tag{1.2}$$

where K denotes the Green's function of the operator $1-\partial_x^2$ on $\mathbb{S}^1$, which can be given in explicit form, and $*$ denotes convolution in x. Consequently, (1.1) takes the form of a nonlinear, nonlocal stochastic partial differential equation (SPDE).

If $\varepsilon = 0$ and $\sigma \equiv 0$, then (1.1) becomes the classical (deterministic) Camassa–Holm (CH) equation [13, 28], which is a nonlinear dispersive PDE that models shallow water waves. Besides, it is nonlocal, completely integrable and may be written in (bi-)Hamiltonian form in terms of the so-called momentum variable $m := \left(1-\partial_{xx}^2\right)u$. The inclusion of gradient type noise is natural in that the perturbation can be thought of as one on the transporting velocity field, i.e., $u\,\partial_x u$ is replaced by $(u + \sigma\circ\dot{W})\circ\partial_x u$, and hence as an additive perturbation of the underlying Lagrangian dynamics; see Section 1.2 for more details.

The CH equation is a supercritical PDE in the sense that the competition between dispersion, which tends to spread out a wave, and nonlinearity, which causes a wave to concentrate, leads to the development of singularities in finite time (wave breaking). The well-posedness of the CH wave equation, in different classes of weak solutions for general finite-energy initial data $u|_{t=0} = u_0 \in H^1$, has been widely studied, see for example [7, 8, 19, 31, 51] (and the references therein). The relevance of the Sobolev space H^1 is that its norm is preserved (up to an inequality) by the solution operator, and H^1 regularity is needed to make distributional sense to the equation. This space is consistent with wave breaking, i.e., a solution u remains bounded while its x-derivative $\partial_x u$ becomes (negatively) unbounded [13] (this is rigorously demonstrated in [20, 21]).

Random effects are important when developing good mathematical models of complex phenomena, with carefully crafted SPDEs providing tools for modelling, analysis, and prediction. Randomness can enter models differently, such as through stochastic transport, stochastic forcing, or uncertain system parameters like random initial and boundary data. The work [34] proposes a general approach to deriving SPDEs for fluid dynamics from a stochastic variational principle. This approach constitutes a stochastic extension of the classical variational derivation of Eulerian fluid dynamics. The corresponding stochastic perturbation of the CH equation leads to an SPDE similar to (1.1) (with $\varepsilon = 0$), see [22] and also [3]. For the related stochastic Hunter–Saxton equation, see [33].

1.2. **Stochastic CH equation.** Let us discuss the derivation of (1.1) (with $\varepsilon = 0$) in more detail. Denote by M_m the multiplication operator by $m = \big(1 - \partial_x^2\big)u$, i.e., $M_m[v] = mv$, and by D the (spatial) differentiation operator on $\mathbb{S}^1$. As is well known, the deterministic CH equation can be written in a bi-Hamiltonian form as

$$0 = \partial_t m + M_m D \frac{\delta \tilde{h}[m]}{\delta m} + D M_m \frac{\delta \tilde{h}[m]}{\delta m}, \tag{1.3}$$

where the Hamiltonian is

$$\tilde{h}[m] = \frac{1}{2} \int_{\mathbb{S}^1} m(t,x)\,(K * m)(t,x)\,\mathrm{d}x, \tag{1.4}$$

and the kernel

$$K(x) := \frac{\cosh(x - 2\pi\left[\frac{x}{2\pi}\right] - \pi)}{2\sinh(\pi)} \tag{1.5}$$

is the Green's function for the operator $1 - \partial_x^2$ on $\mathbb{S}^1$. One can formally convert the bi-Hamiltonian equation (1.3) into the "transport" system

$$0 = \partial_t u + u\,\partial_x u + \partial_x P, \qquad \text{with } P = P[u] \text{ defined in (1.2)},$$

which is a popular formulation of the CH equation. It was suggested in [34, 22] that the Hamiltonian ought to be perturbed by noise directly, so that a physically significant stochastic analogue of the CH equation should be based on the integrated Hamiltonian

$$H[m] = \int_{\mathbb{S}^1} \int_0^t \frac{1}{2} m(s,x)(K * m)(t,x)\,\mathrm{d}s + \int_0^t \big(m(s,x)\sigma(x)\big) \circ \mathrm{d}W(s)\,\mathrm{d}x.$$

With $\sigma \equiv 0$, we identify $\tilde{h}$, cf. (1.4), with $\mathrm{d}H/\mathrm{d}t$. The first variation of $H[m]$ is

$$\frac{\delta H[m]}{\delta m} = u + \sigma(x)\,\dot{W},$$

This expression is of class $C^{-1/2-0}$ in time, and it is far from being a time-continuous object. However, at the formal level, compared with (1.3), the analogous stochastic CH equation becomes

$$0 = \mathrm{d}m + M_m D\big(u\,\mathrm{d}t + \sigma(x)\,\mathrm{d}W\big) + D M_m\big(u\,\mathrm{d}t + \sigma(x)\,\mathrm{d}W\big),$$

where the multiplication operator M_m here uses the Stratonovich product $\circ$; written out more explicitly, we have

$$0 = \mathrm{d}m + \big(m\,\partial_x u + \partial_x(mu)\big)\,\mathrm{d}t + m\,\partial_x\sigma(x) \circ \mathrm{d}W + \partial_x\big(m\sigma(x)\big) \circ \mathrm{d}W. \tag{1.6}$$

We can derive an equation for u that is heuristically equivalent to (1.6). Under the assumption that the functions u, $m = u - \partial_x^2 u$ and σ are sufficiently regular, we can convolve (1.6) by K to obtain

$$0 = \mathrm{d}\Big(K * \big(u - \partial_x^2 u\big)\Big) + K * \Big(3u\,\partial_x u - 2\partial_x u\,\partial_x^2 u - u\partial_x^3 u\Big)\,\mathrm{d}t$$

$$+ K * \Big(\partial_x\sigma\, u - 2\partial_x\sigma\, \partial_x^2 u + \partial_x\left(\sigma u\right) - \sigma\partial_x^3 u\Big) \circ \mathrm{d}W.$$

Recalling the definition of K, cf. (1.5), we obtain

$$\begin{aligned} 0 &= \mathrm{d}u + u\,\partial_x u\,\mathrm{d}t + K * \left(2u\,\partial_x u + \partial_x u\,\partial_x^2 u\right)\,\mathrm{d}t \\ &\quad + \left[\sigma\partial_x u + K * \left(\partial_x^2\sigma\,\partial_x u + 2\partial_x\sigma\, u\right)\right] \circ \mathrm{d}W. \end{aligned}$$

Setting $P = P[u]$, cf. (1.2), we arrive at the final form

$$0 = \mathrm{d}u + \left[u\,\partial_x u + \partial_x P\right]\,\mathrm{d}t + \left[\sigma\partial_x u + K * \left(2\partial_x\sigma\, u + \partial_x^2\sigma\,\partial_x u\right)\right] \circ \mathrm{d}W. \tag{1.7}$$

Mathematically, as explained in Remark 4.4, the convolution part of the noise term offers no new (essential) difficulties compared to $\sigma\partial_x u \circ \mathrm{d}W$. For the sake of clarity, we will therefore focus on the equation

$$0 = \mathrm{d}u + \left[u\,\partial_x u + \partial_x P\right]\,\mathrm{d}t + \sigma(x)\partial_x u \circ \mathrm{d}W.$$

By the Stratonovich–Itô conversion formula, the foregoing equation takes the operational form

$$0 = \mathrm{d}u + \left[u\,\partial_x u + \partial_x P\right]\,\mathrm{d}t - \frac{1}{2}\sigma\partial_x\left(\sigma(x)\partial_x u\right)\,\mathrm{d}t + \sigma\partial_x u\,\mathrm{d}W. \tag{1.8}$$

Regarding the analysis of the stochastic CH equation (1.8), there are few results available at the moment. To better describe the situation, let us note that the equations discussed so far are all nonlinear SPDEs of the form

$$0 = \partial_t u + u\partial_x u + S\big(u, \partial_x u, \partial_x^2 u\big) + \Gamma\big(x, u, \partial_x u\big)\,\mathrm{d}W.$$

Depending on the specification of the functions S and Γ, randomness enters the equation in different ways, including stochastic forcing (noise through a lower order "source term") or gradient-dependent noise (noise through a transport/convection operator). Examples of stochastic forcing arise if $\Gamma\big(x, u, \partial_x u\big) = \beta(x, u)$, for some function β. A typical gradient noise example is $\Gamma\big(x, u, \partial_x u\big) = \sigma(x)\partial_x u$, for some function σ, like in (1.8) or (1.1). Now most of the results in the literature concern the "stochastic forcing" case, either via additive ($\beta = \beta(x)$) or multiplicative ($\beta = \beta(u)$) noise, see the works [16, 17, 18, 32, 52, 38, 42, 47, 48, 53]. For gradient noise, we refer to [1] for a local well-posedness result (up to wave-breaking) for (1.6). The idea in [1] is to transform the equation into a PDE with random coefficients, and apply Kato's operator theory. The work [2] extends this result to a stochastic two-component CH system with gradient noise $\sigma(x)\partial_x \circ \mathrm{d}W$, for smooth ($C^\infty$) noise functions $\sigma(x)$. The approach [2] is based on an abstract SDE framework *à la* [37], but one that is adapted to handle gradient-dependent noise operators (the original framework [37] applies to stochastic forcing operators). The global-in-time existence of properly defined weak solutions for the stochastic CH equation (1.8) is an open problem, but see [15] for some partial results if σ is a constant.

1.3. **Main results.** In this paper, we study a regularised version of (1.8), namely the SPDE (1.1), which contains a viscous dissipation term $\varepsilon\partial_x^2 u$, $\varepsilon > 0$. The second-order operator in (1.8) involving σ is not a regularising (parabolic) operator. The technical reason for this is that the quadratic variation of the martingale part of the equation coincides with the dissipation generated by the second-order operator. The difference between these two terms arises when computing the nonlinear composition $S(u)$ using Itô's formula. There are several reasons why we focus on (1.1) instead of (1.8). First of all, with a few exceptions (discussed above), a general (global) well-posedness theory for (1.8) is missing, and (1.1) appears to be a natural place to start.

Apart from that, in ongoing work, we are investigating the existence of dissipative weak solutions for (1.8). This class of global weak solutions is strongly linked to the well-posedness of (1.1). Indeed, in the deterministic literature there are two natural classes of weak solutions, "dissipative" and "conservative", which differ in how they continue the solution past the blow-up time. Conservative solutions ask that the PDE holds weakly and that the total energy is preserved. In contrast, dissipative solutions are characterized by a drop in the total energy (at the time of blow up). To demonstrate the existence of an appropriately defined dissipative solution to (1.8), one starts from the well-posedness of the viscous SPDE (1.1) to construct an approximate solution sequence $\{u_\varepsilon\}_{\varepsilon>0}$, exhibiting good regularity properties and a priori estimates, and then attempt to pass to the limit $\varepsilon \to 0$ to produce a solution of the inviscid equation (1.8), making use of subtle weak convergence and propagation of compactness techniques (the details will be presented in an upcoming work).

In the present paper, as a first step towards global existence for (1.8), we will develop a rather complete (global) well-posedness theory for (1.1), which allows for general "non-smooth" noise functions $\sigma(x)$. Roughly speaking, by a solution to (1.1) we mean a stochastic process $(\omega, t) \mapsto u(\omega, t, \cdot)$ that takes values in $H^1(\mathbb{S}^1)$ and satisfies the SPDE in the weak sense in x. These solutions are strong (or pathwise) in the probabilistic sense, i.e., they are adapted to an underlying fixed filtration. For a detailed description of the concept of solution, see Definitions 2.1 and 2.2. The first main theorem of the paper is the following result.

Theorem 1.1 (Well-posedness in H^1)**.** *Suppose $\sigma \in W^{2,\infty}(\mathbb{S}^1)$, $p_0 > 4$, and $u_0 \in L^{p_0}(\Omega; H^1(\mathbb{S}^1))$. There exists a unique strong H^1 solution to* (1.1) *with initial condition $u|_{t=0} = u_0$.*

There is a sense in which the "natural" energy space given by the structure of the CH equation is $L^\infty_t H^1_x$; see beginning of Section 8.1 where H^m_x estimates are discussed and compared against estimates in H^1_x. There is also a slight difference in the definitions of H^m solutions pertaining to the function spaces in which they are required to inhabit (see (c) of Definition 2.1). So we record as our second main result a separate theorem on well-posedness in higher-regularity classes:

Theorem 1.2 (Well-posedness in H^m)**.** *Fix $m \geq 2$ and $p_0 > 4$. Suppose $\sigma \in W^{m+1,\infty}(\mathbb{S}^1)$, and $u_0 \in L^{p_0}(\Omega; H^m(\mathbb{S}^1))$. There exists a unique strong H^m solution to* (1.1) *with initial condition $u|_{t=0} = u_0$.*

The H^m_x estimates ($m \geq 2$) do not follow from standard parabolic regularity theory because of nonlinear factors of cubic type, and the fundamental lack of $L^1_t L^\infty_{\omega,x}$ estimates on u and $\partial_x u$ (this is in contrast to the deterministic equation [51]). We cope with these problems using stopping time arguments and a stochastic Gronwall inequality. The moment requirement $p_0 > 4$ on the initial condition in H^m_x comes from technical lemmas (Lemmas 7.1, 7.2 and Proposition 7.4, and see also Remark 8.2).

1.4. **Organisation of paper.** We bring this introduction to an end by outlining the organization of the paper, along with a quick exposition of ideas behind the proofs.

First, in Section 2, we state precisely the different solution concepts used throughout the paper. The existence parts of Theorems 1.1 and 1.2 are based on weak solutions and the introduction of suitable Faedo–Galerkin approximations $\{u_n\}$, where the epithet "weak" refers to probabilistic weak and so-called martingale solutions.

A refined stochastic compactness method [39] is used to conclude convergence $\{u_n\}$ towards a weak solution. In the context of SPDEs, the stochastic compactness method goes back to [4] and it was subsequently used in numerous works, see for example [23, 27, 29] and the references therein. In Section 3, we define and establish the well-posedness of the Faedo–Galerkin approximations. A priori estimates and tightness properties of the approximations $\{u_n\}$ are proved in Sections 4 and 5. More precisely, in Proposition 4.2 and Lemma 5.1 we supply several n-uniform (and ε-uniform) bounds that imply

$$\{u_n\} \subseteq_b L^p\big(\Omega; L^\infty([0,T]; H^1(\mathbb{S}^1))\big) \cap L^p\big(\Omega; C^\theta([0,T]; L^2(\mathbb{S}^1))\big),$$

for appropriate ranges of p and θ (where $\subseteq_b$ means "bounded inclusion", i.e., $A \subseteq_b X$ if $A \subseteq X$ and $\sup_{a\in A} \|a\|_X < \infty$). We use this and the compact inclusion

$$L^\infty([0,T]; H^1(\mathbb{S}^1)) \cap C^\theta([0,T]; L^2(\mathbb{S}^1)) \hookrightarrow C([0,T]; H^1_w(\mathbb{S}^1))$$

to deduce the tightness of the probability laws of the Faedo–Galerkin solutions in the quasi-Polish space $C([0,T]; H^1_w(\mathbb{S}^1))$. Here $H^1_w(\mathbb{S}^1)$ denotes the space $H^1(\mathbb{S}^1)$ with the weak topology. Because of a uniform-in-n bound on $\mathbb{E}\,\|u_n\|^2_{L^2([0,T];H^2(\mathbb{S}^1))}$, arising from the ε-dissipation operator in (1.1), we also obtain the uniform stochastic boundedness of $\{u_n\}$ in the space $L^2([0,T]; H^2(\mathbb{S}^1)) \cap W^{\theta',2}([0,T]; L^2(\mathbb{S}^1))$, with $\theta' < \theta$. Hence, it follows that the probability laws of $\{u_n\}$ are tight on $L^2([0,T]; H^1(\mathbb{S}^1))$, cf. Lemma 5.5. Using the Skorokhod–Jakubowski theorem [35] of almost sure representations of random variables in quasi-Polish spaces (see Appendix A), we deduce in Section 6 the existence of weak (martingale) solutions to the viscous stochastic CH equation (1.1). In Sections 7.1–7.2, we prove pathwise uniqueness for (1.1) by a renormalisation procedure, bypassing the need for an infinite dimensional Itô formula. The uniquenss proof requires some non-standard first- and second-order commutator estimates (that extend beyond the standard DiPerna–Lions estimates), which are established in Lemmas 7.1, 7.2 and Proposition 7.4. Pathwise uniqueness, along with the weak existence result and also the Gyöngy–Krylov characterization of convergence in probability, allows us to conclude in Section 7.3 the existence of a unique strong (pathwise) H^1 solution to (1.1), thus concluding the proof of Theorem 1.1.

One-sided strong temporal continuity characterises *dissipative* weak solutions in the inviscid $\varepsilon \downarrow 0$ limit. For fixed positive viscosity, solutions satisfy (two-sided) strong temporal continuity. This is demonstrated afterwards in Section 7.4.

In Section 8, we turn to Theorem 1.2 and solutions with higher regularity. In Section 8.1, we fix $m \geq 2$ and prove n-uniform bounds in $L^p\big(\Omega; L^\infty([0,\tau]; H^m(\mathbb{S}^1))\big)$, for $p \in [1,\infty)$, up to a suitable stopping time τ (Proposition 8.1). Using this we conclude the stochastic boundedness (see (A.1)) in the higher regularity space $L^2([0,T]; H^{m+1}(\mathbb{S}^1)) \cap W^{\theta,2}([0,T]; L^2(\mathbb{S}^1))$, for some $\theta < 1$, as long as the initial condition u_0 belongs to $L^p(\Omega; H^1(\mathbb{S}^1)) \cap L^2(\Omega; H^m(\mathbb{S}^1))$. By some additional stopping time arguments, this implies that the laws of $\{u_n\}$ are tight on $L^2([0,T]; H^m(\mathbb{S}^1))$, (see Lemma 8.4), and by a Skorokhod–Jakubowski procedure (as in Section 6) we extract a weak solution in $L^2([0,T]; H^m(\mathbb{S}^1))$. The key difference between the H^1 and H^m cases lies in the lack of a bound on $\mathbb{E}\,\|u_n\|^2_{L^2([0,T];H^m(\mathbb{S}^1))}$, that is, if $m \geq 2$, then Lemma 8.4 is available for H^m only up to some stopping time $\tau < T$ but not on $[0,T]$. This obstacle, which is peculiar to the stochastic problem, makes it necessary to argue along several layers of stopping times, see Lemma 8.4 and its proof. Finally, in Section 8.2, we establish the pathwise uniqueness in $L^1(\Omega; L^\infty([0,T]; H^m(\mathbb{S}^1)))$ and conclude the well-posedness of strong H^m solutions.

In Appendix A, we record some results of stochastic analysis frequently deployed in this paper.

2. **Solution concepts.** In this section, we present the solution concept used in Theorems 1.1 and 1.2. We denote by $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ a complete *probability space* with (countably generated) σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$. We consider filtrations $\{\mathcal{F}_t\}_{t\in[0,T]}$ that satisfy the "usual conditions" of being complete and right-continuous. We refer to

$$\mathcal{S} := \big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\big) \tag{2.1}$$

as a *stochastic basis* (sometimes called a filtered probability space).

Theorems 1.1 and 1.2 speak of *strong* H^m solutions. These are weak (distributional) solutions in the PDE sense in the Sobolev space H^m. From the probabilistic point of view, however, we will have to consider first so-called *martingale* solutions, which are also referred to as *weak* solutions. The notions of weak/strong probabilistic solutions have a different meaning from weak/strong solutions in the PDE literature. If the stochastic basis $\mathcal{S}$ (2.1) and the Wiener process W are fixed in advance, we speak of a strong (or pathwise) solution. If $\big(\mathcal{S}, W\big)$ is a part of the unknown solution, the relevant notion is a martingale solution. In what follows, "weak H^m solutions" refer to solutions that are probabilistic weak and weak in the PDE sense, whereas "strong H^m solutions" refer to solutions that are pathwise and weak in the PDE sense.

In view of Theorems 1.1 and 1.2, the H^1 well-posedness theory deviates slightly from the H^m theory for $m \geq 2$. The corresponding solution concepts differ in their requirement on the initial condition and the condition $u \in L^2\left(\Omega; L^2([0,T]; H^2(\mathbb{S}^1))\right)$ if $m = 1$ versus the weaker stochastic boundedness condition in $L^2([0,T]; H^{m+1}(\mathbb{S}^1))$ if $m \geq 2$, as is seen in the next definition.

Definition 2.1 (Weak H^m solution)**.** Fix $m \in \mathbb{N}$ and $p_0 > 4$. Let Λ be a probability measure on $H^m(\mathbb{S}^1)$ satisfying

$$\int_{H^m(\mathbb{S}^1)} \|v\|_{H^m(\mathbb{S}^1)}^{p_0} \, \Lambda(\mathrm{d}v) < \infty. \tag{2.2}$$

The triple $\big(\mathcal{S}, u, W\big)$ is a weak (or martingale) H^m solution to (1.1) with initial distribution Λ if the following conditions hold:

(a) $\mathcal{S} = \big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\big)$ is stochastic basis, cf. (2.1);
(b) W is a standard Wiener process on $\mathcal{S}$;
(c) $u\colon \Omega \times [0,T] \to H^1(\mathbb{S}^1)$ is adapted, with $u \in L^{p_0}\left(\Omega; C([0,T]; H^1(\mathbb{S}^1)\right)$ and $u \in L^2([0,T]; H^m(\mathbb{S}^1))$, $\mathbb{P}$-almost surely. Moreover,

$$\begin{cases} u \in L^2\left(\Omega; L^2([0,T]; H^2(\mathbb{S}^1)\right), & \textit{if } m = 1, \\ u \in_{\mathrm{sb}} L^2([0,T]; H^{m+1}(\mathbb{S}^1)) \cap L^\infty([0,T]; H^m(\mathbb{S}^1)), & \textit{if } m \geq 2, \end{cases}$$

where $\in_{\mathrm{sb}}$ means stochastically bounded (see (A.1));
(d) the law of the initial data $u_0 := u(0)$ on $H^m(\mathbb{S}^1)$ is Λ, i.e., $(u(0))_* \mathbb{P} = \Lambda$, or $\Lambda(A) = \mathbb{P}(u(0)^{-1}(A))$ for measurable sets A;

(e) for all $t \in [0,T]$ and all $\varphi \in C^1(\mathbb{S}^1)$ the following equation holds $\mathbb{P}$-almost surely (in the sense of Itô):

$$\begin{aligned}\int_{\mathbb{S}^1} u(t)\varphi\,\mathrm{d}x - \int_{\mathbb{S}^1} u_0\varphi\,\mathrm{d}x \\ &= \int_0^t \int_{\mathbb{S}^1} \left(-u\,\partial_x u\,\varphi + (P - \varepsilon\partial_x u)\,\partial_x\varphi\right)\mathrm{d}x\,\mathrm{d}s \\ &\quad - \frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \sigma\partial_x u\,\partial_x(\sigma\varphi)\,\mathrm{d}x\,\mathrm{d}s - \int_0^t \int_{\mathbb{S}^1} \varphi\sigma\,\partial_x u\,\mathrm{d}x\,\mathrm{d}W(s),\end{aligned}$$

where $P = P[u] := K * \left(u^2 + \frac{1}{2}(\partial_x u)^2\right)$.

Finally, we introduce the notion of strong (pathwse) H^m solution.

Definition 2.2 (Strong H^m solution)**.** Fix a stochastic basis $\mathcal{S}$, cf. (2.1), and a Wiener process W defined on $\mathcal{S}$. Fix $m \in \mathbb{N}$ and $p_0 > 4$, and consider a random variable $u_0 \in L^{p_0}(\Omega; H^1(\mathbb{S}^1))$. A process u, defined relative to $\mathcal{S}$, is a strong H^m solution to (1.1) if $\left(\mathcal{S}, u, W\right)$ is a weak H^m solution to (1.1) with initial law $\Lambda := (u_0)_*\,\mathbb{P}$, i.e., Λ obeys (2.2) and $\left(\mathcal{S}, u, W\right)$ satisfies (a)–(e) in Definition 2.1.

3. The Galerkin approximation.

We now specify our Galerkin scheme for constructing approximate solutions. Let $\{e_1, e_2, \dots\} \subseteq H^1(\mathbb{S}^1)$ be an orthonormal basis of $L^2(\mathbb{S}^1)$ that is dense in $H^1(\mathbb{S}^1)$ and set $H_n = \operatorname{span}\{e_1, \dots, e_n\}$. In particular, we take $\{e_i\}_{i\in\mathbb{N}}$ to be the eigenfunctions of ∂_x^2 on the circle $\mathbb{S}^1$, i.e., $e_{2j}(x) = \cos(2\pi jx)$ and $e_{2j+1}(x) = \sin(2\pi jx)$, $x \in [0,1]$, for concreteness. Let $\mathbf{\Pi}_n : \left(H^1(\mathbb{S}^1)\right)^* \to H_n$ be defined by

$$\mathbf{\Pi}_n u := \sum_{i=1}^n \left\langle u, e_i \right\rangle_{L^2(\mathbb{S}^1)} e_i,$$

so that, restricted to $L^2(\mathbb{S}^1)$, $\mathbf{\Pi}_n$ is the orthogonal projection onto H_n.

For each $n \in \mathbb{N}$, we consider the Galerkin approximation of (1.1) on H_n, that is, we seek a function

$$u_n(\omega, t, x) = \sum_{i=1}^n w_i(\omega, t) e_i(x),$$

where the unknown coefficients $\{w_i = w_i(\omega, t)\}_{i=1}^n$ are determined by requiring that

$$\begin{aligned} 0 &= \mathrm{d}u_n - \varepsilon\partial_x^2 u_n\,\mathrm{d}t + \mathbf{\Pi}_n\left(u_n\partial_x u_n + \partial_x P[u_n]\right)\mathrm{d}t \\ &\quad - \frac{1}{2}\mathbf{\Pi}_n\left(\sigma\partial_x\left(\sigma\partial_x u_n\right)\right)\mathrm{d}t + \mathbf{\Pi}_n\left(\sigma\partial_x u_n\right)\mathrm{d}W, \\ u_n(0) &= \mathbf{\Pi}_n u_0. \end{aligned} \tag{3.1}$$

Here, u_0 is a random variable $\Omega \to H^1(\mathbb{S}^1)$ with law Λ and a bounded second moment, i.e., $\mathbb{E}\left\|u_0\right\|_{H^1(\mathbb{S}^1)}^2 < \infty$.

Theorem 3.1. *For any fixed n, there exists a unique $C([0,T]; H_n)$-valued adapted process u_n that is a strong solution to* (3.1)*.*

Proof. The proof consists of noting that (3.1) is a SDE system with coefficients that are locally Lipschitz continuous in $w = \{w_i\}_{i\in\mathbb{N}}$. By a standard well-posedness theorem for SDEs [41, Thm. IX.2.1], this immediately implies the existence and uniqueness of a continuous (strong) solution of (3.1) on $[0,T]$.

It remains to argue that the Galerkin equation (3.1) can be viewed as a SDE system in w. First, by properties of the basis functions,

$$\partial_x u_n = \sum_{i=1}^n C_i\, w_i\, e_i, \quad \partial_x^2 u_n = \sum_{i=1}^n -C_i^2\, w_i\, e_i,$$

where C_i are constants depending only on i. Next, the nonlinear term

$$\begin{aligned}\mathbf{\Pi}_n\big(u_n\partial_x u_n\big) &= \sum_{i,j=1}^n C_j\, w_i w_j\, \mathbf{\Pi}_n\big(e_i e_j\big)\\ &= \sum_{i,j,k=1}^n C_j\, w_i w_j \int_{\mathbb{S}^1} e_i(y)e_j(y)e_k(y)\,\mathrm{d}y\, e_k,\end{aligned}$$

is locally Lipschitz in w. Regarding the nonlocal operator, we can calculate thus:

$$\begin{aligned}&\mathbf{\Pi}_n\partial_x P[u_n]\\ &= \sum_{k=1}^n \int_{\mathbb{S}^1} \partial_y K * \left(u_n^2 + \frac{1}{2}\left(\partial_y u_n\right)^2\right) e_k(y)\,\mathrm{d}y\, e_k\\ &= \sum_{i,j,k=1}^n \int_{\mathbb{S}^1}\int_{\mathbb{S}^1} \partial_z K(z-y)\Big(w_i w_j e_i(y) e_j(y)\\ &\qquad\qquad + \frac{1}{2} C_i C_j w_i w_j e_i(y) e_j(y)\Big)\, e_k(z)\,\mathrm{d}y\,\mathrm{d}z\, e_k\\ &= \sum_{i,j,k=1}^n w_i w_j \left(\int_{\mathbb{S}^1}\int_{\mathbb{S}^1} \partial_z K(z-y)\left(1+\frac{1}{2}C_i C_j\right) e_i(y) e_j(y)\, e_k(z)\,\mathrm{d}y\,\mathrm{d}z\right)\; e_k,\end{aligned}$$

which is then seen also to be locally Lipschitz in w. Similarly, we can show that the linear terms $\mathbf{\Pi}_n\left(\sigma\partial_x\left(\sigma\partial_x u_n\right)\right)$ and $\mathbf{\Pi}_n\left(\sigma\partial_x u_n\right)$ are locally Lipschitz in w. □

4. **A priori estimates.** Our first result is a fundamental model-specific energy estimate that we will refer to repeatedly throughout this work.

We use frequently the fact that for any function $f \in L^2(\mathbb{S}^1)$,

$$\int_{\mathbb{S}^1} u_n \mathbf{\Pi}_n f\,\mathrm{d}x = \int_{\mathbb{S}^1} u_n f\,\mathrm{d}x, \tag{4.1}$$

because $\mathbf{\Pi}_n$ is self-adjoint and idempotent. For any $f \in H^1(\mathbb{S}^1)$ and $\frac{n-1}{2} \in \mathbb{N}$, we compute the spatial derivative of $\mathbf{\Pi}_n f$ as follows:

$$\begin{aligned}\partial_x\left(\mathbf{\Pi}_n f\right) &= \sum_{2j\le n}\langle f, e_{2j}\rangle_{L^2(\mathbb{S}^1)}\partial_x e_{2j} + \sum_{2j+1\le n}\langle f, e_{2j+1}\rangle_{L^2(\mathbb{S}^1)}\partial_x e_{2j+1}\\ &= 2\pi\Big(\sum_{2j\le n} j\langle f, e_{2j}\rangle_{L^2(\mathbb{S}^1)} e_{2j+1} - \sum_{2j+1\le n} j\langle f, e_{2j+1}\rangle_{L^2(\mathbb{S}^1)} e_{2j}\Big)\\ &= -\sum_{2j\le n}\langle f, \partial_x e_{2j+1}\rangle_{L^2(\mathbb{S}^1)} e_{2j+1} - \sum_{2j+1\le n}\langle f, \partial_x e_{2j}\rangle_{L^2(\mathbb{S}^1)} e_{2j}\\ &= \sum_{2j+1\le n}\langle \partial_x f, e_{2j+1}\rangle_{L^2(\mathbb{S}^1)} e_{2j+1} + \sum_{2j\le n}\langle \partial_x f, e_{2j}\rangle_{L^2(\mathbb{S}^1)} e_{2j}\\ &= \mathbf{\Pi}_n\left(\partial_x f\right).\end{aligned} \tag{4.2}$$

Proposition 4.1 (Energy estimate)**.** *For each* $n \in \mathbb{N}$*, let* u_n *be a solution to* (3.1) *with* $\mathbb{E}\left\|u_0\right\|_{H^1(\mathbb{S}^1)}^2 < \infty$*. There exists a constant*

$$C = C\left(T, \mathbb{E}\left\|u_0\right\|_{H^1(\mathbb{S}^1)}^2, \|\sigma\|_{W^{2,\infty}(\mathbb{S}^1)}\right),$$

independent of n *and* ε*, such that*

$$\mathbb{E}\left\|u_n\right\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^2 + \varepsilon\mathbb{E}\int_0^T \left\|\partial_x u_n(t)\right\|_{H^1(\mathbb{S}^1)}^2 \, \mathrm{d}t \le C. \tag{4.3}$$

Proof. We multiply (via Itô's formula) the SDE (3.1) against u_n and then integrate in $x \in \mathbb{S}^1$. Using (4.1) and (4.2), we obtain the SDE

$$\begin{aligned}
\frac{1}{2}\mathrm{d}\int_{\mathbb{S}^1} |u_n|^2 \, \mathrm{d}x + \varepsilon \int_{\mathbb{S}^1} |\partial_x u_n|^2 \, \mathrm{d}x \, \mathrm{d}t \\
&= -\int_{\mathbb{S}^1} \left(u_n^2 \partial_x u_n + u_n \partial_x K * \left(u_n^2 + \frac{1}{2}(\partial_x u_n)^2\right)\right) \mathrm{d}x \, \mathrm{d}t \\
&\quad + \frac{1}{2}\int_{\mathbb{S}^1} \left(\sigma u_n \partial_x(\sigma \partial_x u_n) + \left|\mathbf{\Pi}_n(\sigma \partial_x u_n)\right|^2\right) \mathrm{d}x \, \mathrm{d}t \\
&\quad - \int_{\mathbb{S}^1} \sigma u_n \partial_x u_n \, \mathrm{d}x \, \mathrm{d}W.
\end{aligned}$$

Differentiating (3.1) and multiplying through by $\partial_x u_n = \mathbf{\Pi}_n \partial_x u_n$ (via Itô's formula), using again (4.1) and (4.2), yields

$$\begin{aligned}
\frac{1}{2}\mathrm{d}\int_{\mathbb{S}^1} |\partial_x u_n|^2 \, \mathrm{d}x + \varepsilon \int_{\mathbb{S}^1} \left|\partial_x^2 u_n\right|^2 \, \mathrm{d}x \, \mathrm{d}t \\
&= \int_{\mathbb{S}^1} \left(u_n \partial_x u_n \partial_x^2 u_n + \partial_x^2 u_n \partial_x K * \left(u_n^2 + \frac{1}{2}(\partial_x u_n)^2\right)\right) \mathrm{d}x \, \mathrm{d}t \\
&\quad - \frac{1}{2}\int_{\mathbb{S}^1} \left(\partial_x^2 u_n \sigma \partial_x(\sigma \partial_x u_n) - \left|\partial_x\left(\mathbf{\Pi}_n(\sigma \partial_x u_n)\right)\right|^2\right) \mathrm{d}x \, \mathrm{d}t \\
&\quad + \int_{\mathbb{S}^1} \sigma \partial_x u_n \, \partial_x^2 u_n \, \mathrm{d}x \, \mathrm{d}W.
\end{aligned}$$

Adding the previous two equations, we arrive at

$$\frac{1}{2}\mathrm{d}\left\|u_n\right\|_{H^1(\mathbb{S}^1)}^2 + \varepsilon \int_{\mathbb{S}^1} \left(|\partial_x u_n|^2 + \left|\partial_x^2 u_n\right|^2\right) \mathrm{d}x \, \mathrm{d}t = I_1^n \, \mathrm{d}t + I_2^n \, \mathrm{d}t + I_3^n \, \mathrm{d}W, \tag{4.4}$$

where

$$\begin{aligned}
I_1^n &:= -\int_{\mathbb{S}^1} \left(u_n^2 \partial_x u_n + u_n \partial_x K * \left(u_n^2 + \frac{1}{2}(\partial_x u_n)^2\right)\right) \mathrm{d}x \\
&\qquad + \int_{\mathbb{S}^1} \left(u_n \partial_x u_n \partial_x^2 u_n + \partial_x^2 u_n \, \partial_x K * \left(u_n^2 + \frac{1}{2}(\partial_x u_n)^2\right)\right) \mathrm{d}x, \\
I_2^n &:= \frac{1}{2}\int_{\mathbb{S}^1} \left(\sigma u_n \partial_x(\sigma \partial_x u_n) + \left|\mathbf{\Pi}_n(\sigma \partial_x u_n)\right|^2\right) \mathrm{d}x \\
&\qquad - \frac{1}{2}\int_{\mathbb{S}^1} \left(\partial_x^2 u_n \sigma \partial_x(\sigma \partial_x u_n) - \left|\partial_x\left(\mathbf{\Pi}_n(\sigma \partial_x u_n)\right)\right|^2\right) \mathrm{d}x \\
&=: \frac{1}{2} I_{2,1}^n - \frac{1}{2} I_{2,2}^n, \\
I_3^n &:= -\int_{\mathbb{S}^1} \left(\sigma u_n \partial_x u_n - \sigma \partial_x u_n \partial_x^2 u_n\right) \mathrm{d}x.
\end{aligned} \tag{4.5}$$

1. Estimate of I_1^n.

Using integration by parts and the kernel property of K that $K - \partial_x^2 K = \boldsymbol{\delta}$, the Dirac mass,

$$\begin{aligned} I_1^n = &-\int_{\mathbb{S}^1} \left(u_n^2 \partial_x u_n - \partial_x u_n K * \left(u_n^2 + \frac{1}{2}\left(\partial_x u_n\right)^2\right)\right) \mathrm{d}x \\ &+ \int_{\mathbb{S}^1} \left(u_n \partial_x u_n \partial_x^2 u_n - \partial_x u_n \partial_x^2 K * \left(u_n^2 + \frac{1}{2}\left(\partial_x u_n\right)^2\right)\right) \mathrm{d}x = 0. \end{aligned} \tag{4.6}$$

2. Estimate of I_2^n.

By Bessel's inequality,

$$\int_{\mathbb{S}^1} \left|\left(\mathbf{\Pi}_n\left(\sigma \partial_x u_n\right)\right)\right|^2 \mathrm{d}x \leq \int_{\mathbb{S}^1} \left|\sigma \partial_x u_n\right|^2 \mathrm{d}x.$$

Combining this with an integration by parts in the $I_{2,1}^n$-term $\sigma u_n \partial_x\left(\sigma \partial_x u_n\right)$, and then expanding out $\partial_x\left(\sigma u_n\right)$ followed by another integration by parts, yields

$$I_{2,1}^n \leq -\int_{\mathbb{S}^1} \sigma \partial_x \sigma u_n \partial_x u_n \,\mathrm{d}x = -\frac{1}{4}\int_{\mathbb{S}^1} \partial_x \sigma^2 \partial_x u_n^2 \,\mathrm{d}x.$$

Similarly, by (4.2) and Bessel's inequality,

$$\begin{aligned} \int_{\mathbb{S}^1} \left|\partial_x\big(\mathbf{\Pi}_n\left(\sigma \partial_x u_n\right)\big)\right|^2 \mathrm{d}x &= \int_{\mathbb{S}^1} \left|\mathbf{\Pi}_n\big(\partial_x\left(\sigma \partial_x u_n\right)\big)\right|^2 \mathrm{d}x \\ &\leq \int_{\mathbb{S}^1} \left|\partial_x\left(\sigma \partial_x u_n\right)\right|^2 \mathrm{d}x \\ &= \int_{\mathbb{S}^1} \left(\left|\partial_x \sigma \partial_x u\right|^2 + \frac{1}{2}\partial_x \sigma^2 \partial_x\left|\partial_x u_n\right|^2 + \left|\sigma \partial_x^2 u_n\right|^2\right) \mathrm{d}x \\ &= \int_{\mathbb{S}^1} \left(\left|\partial_x \sigma \partial_x u\right|^2 - \frac{1}{2}\partial_x^2 \sigma^2 \left|\partial_x u_n\right|^2 + \left|\sigma \partial_x^2 u_n\right|^2\right) \mathrm{d}x. \end{aligned}$$

We combine this with an expansion of the $I_{2,2}^n$-term $\partial_x^2 u_n \sigma \partial_x\left(\sigma \partial_x u_n\right)$ into the sum $\sigma^2\left|\partial_x^2 u_n\right|^2 + \frac{1}{4}\partial_x \sigma^2 \partial_x\left(\partial_x u_n\right)^2$, along with an integration by parts in the latter term:

$$\begin{aligned} I_{2,2}^n \geq &\int_{\mathbb{S}^1} \Big(\sigma^2\left|\partial_x^2 u_n\right|^2 + \frac{1}{4}\partial_x \sigma^2 \partial_x\left(\partial_x u_n\right)^2 \\ &\qquad - \left|\partial_x \sigma \partial_x u\right|^2 + \frac{1}{2}\partial_x^2 \sigma^2\left|\partial_x u_n\right|^2 - \left|\sigma \partial_x^2 u_n\right|^2\Big) \mathrm{d}x \\ = &\int_{\mathbb{S}^1} \left(\frac{1}{4}\partial_x^2 \sigma^2\left|\partial_x u_n\right|^2 - \left|\partial_x \sigma \partial_x u\right|^2\right) \mathrm{d}x. \end{aligned}$$

Hence

$$\begin{aligned} 2 I_2^n &= I_{2,1}^n - I_{2,2}^n \\ &\leq \int_{\mathbb{S}^1} \left(\frac{1}{4}\partial_x \sigma^2 u_n^2 - \frac{1}{4}\partial_x^2 \sigma^2\left|\partial_x u_n\right|^2 + \left|\partial_x \sigma \partial_x u\right|^2\right) \mathrm{d}x \\ &\leq C_\sigma \left\|u_n\right\|_{H^1(\mathbb{S}^1)}^2. \end{aligned} \tag{4.7}$$

3. Estimate of martingale term I_3^n.

First, since $I_3^n = \frac{1}{2}\int_{\mathbb{S}^1} \partial_x \sigma\left(\left|u_n\right|^2 - \left|\partial_x u_n\right|^2\right) \mathrm{d}x$, we have the estimate

$$\left|I_3^n\right| \leq \frac{1}{2}\left\|\partial_x \sigma\right\|_{L^\infty(\mathbb{S}^1)} \left\|u_n\right\|_{H^1(\mathbb{S}^1)}^2.$$

By the Burkholder–Davis–Gundy inequality,

$$\mathbb{E} \sup_{s\in[0,t]} \left| \int_0^s I_3^n \, \mathrm{d}W \right| \le \mathbb{E} \left(\int_0^t |I_3^n|^2 \, \mathrm{d}s \right)^{1/2}$$
$$\le \tilde{C}_\sigma \, \mathbb{E} \left(\int_0^t \|u_n(s)\|_{H^1(\mathbb{S}^1)}^4 \, \mathrm{d}s \right)^{1/2} =: \tilde{I}_3^n.$$

By the Hölder and Young inequalities, we can further estimate the above by

$$\tilde{I}_3^n \le \tilde{C}_\sigma \, \mathbb{E} \left(\int_0^t \|u_n(s)\|_{H^1(\mathbb{S}^1)}^2 \, \mathrm{d}s \sup_{s\in[0,t]} \|u_n(s)\|_{H^1(\mathbb{S}^1)}^2 \right)^{1/2}$$
$$\le C_\sigma \mathbb{E} \int_0^t \|u_n(s)\|_{H^1(\mathbb{S}^1)}^2 \, \mathrm{d}s + \frac{1}{2} \mathbb{E} \sup_{s\in[0,t]} \|u_n(s)\|_{H^1(\mathbb{S}^1)}^2 . \tag{4.8}$$

4. Conclusion.
Gathering the estimates (4.6), (4.7), and (4.8), we conclude that there exists a constant C, independent of n and ε, such that

$$\frac{1}{2}\mathbb{E} \sup_{s\in[0,t]} \|u_n(s)\|_{H^1(\mathbb{S}^1)}^2 + \varepsilon \int_0^t \int_{\mathbb{S}^1} \left(|\partial_x u_n|^2 + \left|\partial_x^2 u_n\right|^2 \right) \mathrm{d}x \, \mathrm{d}s$$
$$\le \mathbb{E} \|u_n(0)\|_{H^1(\mathbb{S}^1)}^2 + C\mathbb{E} \int_0^t \|u_n(s)\|_{H^1(\mathbb{S}^1)}^2 \, \mathrm{d}s, \quad t \in [0,T],$$

which implies (4.3) by Gronwall's inequality. □

The previous lemma supplies control of the second moment of the $H^1(\mathbb{S}^1)$-norm. This effectively guarantees that higher moments are bounded as well.

Lemma 4.2 (Higher moment bounds for the $H^1(\mathbb{S}^1)$-norm)**.** *Fix $p \in (4,\infty)$, and let u_n be a solution to (3.1) with $\mathbb{E}\|u_0\|_{H^1(\mathbb{S}^1)}^p < \infty$. There exists a constant*

$$C = C\left(p, T, \|\sigma\|_{W^{2,\infty}(\mathbb{S}^1)}\right),$$

independent of n (and ε), such that

$$\mathbb{E} \sup_{t\in[0,T]} \|u_n\|_{H^1(\mathbb{S}^1)}^p + \varepsilon^{p/2}\mathbb{E} \left(\int_0^T \int_{\mathbb{S}^1} \left(|\partial_x u_n|^2 + \left|\partial_x^2 u_n\right|^2 \right) \mathrm{d}x \, \mathrm{d}t \right)^{p/2} \le C\mathbb{E} \|u_0\|_{H^1(\mathbb{S}^1)}^p . \tag{4.9}$$

Remark 4.3. Insofar as the bound $\mathbb{E}\|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^p \le C$ is concerned, since $(\Omega, \mathbb{P})$ is a finite measure space, the bound with the same constant holds for any $r \in [1,p)$ in place of p, though the theorem is stated for $p > 4$. By (4.9) and the one-dimensional embedding $H^1(\mathbb{S}^1) \hookrightarrow L^\infty(\mathbb{S}^1)$, we have also that

$$\mathbb{E}\|u_n\|_{L^\infty([0,T]\times\mathbb{S}^1)}^p \lesssim_p 1, \quad p \in [1,\infty), \tag{4.10}$$

but u_n is not uniformly bounded in $L^\infty_{\omega,t,x}$.

Proof. By (4.4) and parts 1, 2 of the proof of Proposition 4.1, we have

$$\frac{1}{2}\mathrm{d}\,\|u_n\|_{H^1(\mathbb{S}^1)}^2 + \varepsilon \int_{\mathbb{S}^1} \left(|\partial_x u_n|^2 + \left|\partial_x^2 u_n\right|^2 \right) \mathrm{d}x \, \mathrm{d}t$$

$$
\begin{aligned}
&= \frac{1}{2}\int_{\mathbb{S}^1} \left(\sigma u_n \partial_x\left(\sigma \partial_x u_n\right) + \left|\mathbf{\Pi}_n\left(\sigma\partial_x u_n\right)\right|^2\right) \mathrm{d}x\,\mathrm{d}t \\
&\quad - \frac{1}{2}\int_{\mathbb{S}^1} \left(\partial_x^2 u_n \sigma \partial_x\left(\sigma\partial_x u_n\right) - \left|\partial_x\left(\mathbf{\Pi}_n\left(\sigma\partial_x u_n\right)\right)\right|^2\right) \mathrm{d}x\,\mathrm{d}t \\
&\quad - \int_{\mathbb{S}^1} \left(\sigma\, u_n\partial_x u_n - \sigma\, \partial_x u_n \partial_x^2 u_n\right) \mathrm{d}x\,\mathrm{d}W.
\end{aligned}
$$

We again use Bessel's inequality to eliminate the two projection operators (remembering that projection commutes with differentiation). Then, integrating in time, rasing both sides to the power $p/2$, and taking expectation, we find

$$
\begin{aligned}
&\frac{1}{2^{p/2}}\mathbb{E}\sup_{t\in[0,T]}\|u_n\|_{H^1(\mathbb{S}^1)}^p + \varepsilon^{p/2}\mathbb{E}\left(\int_0^T\int_{\mathbb{S}^1}\left|\partial_x u_n\right|^2 + \left|\partial_x^2 u_n\right|^2\,\mathrm{d}x\,\mathrm{d}t\right)^{p/2} \\
&\lesssim_p \mathbb{E}\sup_{t\in[0,T]}\Bigg|\int_0^t\int_{\mathbb{S}^1}\sigma u_n\partial_x\left(\sigma\partial_x u_n\right) + \left|\left(\sigma\partial_x u_n\right)\right|^2\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad - \int_0^t\int_{\mathbb{S}^1}\partial_x^2 u_n\sigma\partial_x\left(\sigma\partial_x u_n\right) - \left|\partial_x\left(\sigma\partial_x u_n\right)\right|^2\,\mathrm{d}x\,\mathrm{d}t\Bigg|^{p/2} \\
&\quad + \mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t\int_{\mathbb{S}^1}\sigma\, u_n\partial_x u_n - \sigma\,\partial_x u_n\partial_x^2 u_n\,\mathrm{d}x\,\mathrm{d}W\right|^{p/2} =: I_1 + I_2.
\end{aligned}
$$

For I_1, we find:

$$
\begin{aligned}
&\int_{\mathbb{S}^1}\sigma u_n\partial_x\left(\sigma\partial_x u_n\right) + \left|\left(\sigma\partial_x u_n\right)\right|^2\,\mathrm{d}x \\
&= \int_{\mathbb{S}^1} - \left|\sigma\,\partial_x u_n\right|^2 + \sigma u_n\,\partial_x\sigma\,\partial_x u_n + \left|\left(\sigma\partial_x u_n\right)\right|^2\,\mathrm{d}x \\
&= \int_{\mathbb{S}^1} -\frac{1}{2}\partial_x\left(\sigma\,\partial_x\sigma\right)\,u_n^2\,\mathrm{d}x,
\end{aligned}
$$

and, similarly,

$$
\begin{aligned}
&\int_{\mathbb{S}^1}\partial_x^2 u_n\sigma\partial_x\left(\sigma\partial_x u_n\right) - \left|\partial_x\left(\sigma\partial_x u_n\right)\right|^2\,\mathrm{d}x \\
&= \int_{\mathbb{S}^1}\left|\partial_x\left(\sigma\partial_x u_n\right)\right|^2 - \partial_x\sigma\,\partial_x u_n\,\partial_x\left(\sigma\,\partial_x u_n\right) - \left|\partial_x\left(\sigma\partial_x u_n\right)\right|^2\,\mathrm{d}x \\
&= \int_{\mathbb{S}^1}\left(\frac{1}{2}\partial_x\left(\sigma\,\partial_x\sigma\right) - \left|\partial_x\sigma\right|^2\right)\left|\partial_x u_n\right|^2\,\mathrm{d}x.
\end{aligned}
$$

This give us

$$
I_1 \lesssim_\sigma \int_0^T \mathbb{E}\,\|u_n\|_{H^1(\mathbb{S}^1)}^p\,\mathrm{d}t,
$$

where the implied constant depends on $\|\sigma\|_{W^{2,\infty}(\mathbb{S}^1)}$.

For I_2, by the Burkholder–Davis–Gundy inequality, we have

$$
\begin{aligned}
I_2 &\le \mathbb{E}\left(\int_0^T\left|\int_{\mathbb{S}^1}\sigma\,u_n\partial_x u_n - \sigma\,\partial_x u_n\partial_x^2 u_n\,\mathrm{d}x\right|^2\mathrm{d}t\right)^{p/4} \\
&\le \mathbb{E}\left(\int_0^T\left|\int_{\mathbb{S}^1}\frac{1}{2}\partial_x\sigma\,\left(\left|u_n\right|^2 - \left|\partial_x u_n\right|^2\right)\,\mathrm{d}x\right|^2\mathrm{d}t\right)^{p/4}
\end{aligned}
$$

$$\lesssim_\sigma \int_0^T \mathbb{E} \|u_n\|_{H^1(\mathbb{S}^1)}^p \, \mathrm{d}t,$$

where we used the convexity of $x \mapsto x^{p/4}$ in the final inequality, provided by the assumption $p > 4$.

The estimates on I_1 and I_2 then allow us to derive the stated bound in the Lemma statement by a standard application of Gronwall's inequality. □

Remark 4.4 (Full Euler–Poincaré structure in the noise)**.** It can be verified that there is no additional difficulty with the incorporation of full Euler–Poincaré noise of the form

$$\begin{aligned} \mathcal{B}(u) \circ \mathrm{d}W &= (\sigma\, \partial_x u + \mathcal{J}_1(u)) \circ \mathrm{d}W, \\ \mathcal{J}_1(u) &:= K * \tilde{B}(u) := K * \left(2\partial_x \sigma\, u + \partial_x^2 \sigma\, \partial_x u\right), \end{aligned}$$

in place of $\sigma \partial_x u \circ \mathrm{d}W$ in (1.1), see (1.7).

Written out in Itô form, the noise $\mathcal{B}(u) \circ \mathrm{d}W = \mathcal{B}(u)\,\mathrm{d}W + \frac{1}{2}\mathcal{C}(u)\,\mathrm{d}t$ gives a Stratonovich–Itô correction $\mathcal{C}$ of the form

$$\begin{aligned} 2\mathcal{C}(u) &= \left\langle \mathcal{B}(u), W \right\rangle = -\mathcal{B}\left(\mathcal{B}(u)\right) = -\sigma\, \partial_x \left(\sigma\, \partial_x u\right) + 2\mathcal{J}_2(u), \\ 2\mathcal{J}_2(u) &= -\sigma \partial_x K * \tilde{B}(u) - 2K * \left(\partial_x \sigma \left(\sigma \partial_x u + K * \tilde{B}(u)\right)\right) \\ &\quad - K * \left(\partial_x^2 \sigma \partial_x \left(\sigma \partial_x u + K * \tilde{B}(u)\right)\right). \end{aligned}$$

Since the transformation $\sigma \partial_x u \mapsto \mathcal{B}(u) = \sigma \partial_x u + K * \tilde{B}(u)$ does not introduce higher-order derivatives on u, but it does on σ, the only extra requirement for bounds on $\mathcal{B}$ or $\mathcal{C}$ is that $\sigma \in W^{3,\infty}(\mathbb{S}^1)$ instead of $W^{2,\infty}(\mathbb{S}^1)$.

The corresponding energy balance is

$$\begin{aligned} \frac{1}{2}\mathrm{d} \|u\|_{H^1(\mathbb{S}^1)}^2 &+ \varepsilon \|\partial_x u\|_{H^1(\mathbb{S}^1)}^2 \,\mathrm{d}t \\ &= I_2 \,\mathrm{d}t + \int_{\mathbb{S}^1} \left(u \mathcal{J}_2(u) + \partial_x u \partial_x \mathcal{J}_2(u)\right) \mathrm{d}x \,\mathrm{d}t \\ &\quad + \frac{1}{2} \int_{\mathbb{S}^1} \left(|\mathcal{J}_1(u)|^2 + |\partial_x \mathcal{J}_1(u)|^2 \right) \mathrm{d}x \,\mathrm{d}t \\ &\quad + \int_{\mathbb{S}^1} \left(\mathcal{J}_1(u) \sigma \partial_x u + \partial_x \mathcal{J}_1(u) \partial_x \left(\sigma\, \partial_x u\right) \right) \mathrm{d}x \,\mathrm{d}t \\ &\quad + \left(I_3 - \int_{\mathbb{S}^1} u \left(2\partial_x \sigma u + \partial_x^2 \sigma \partial_x u\right) \,\mathrm{d}x \right) \mathrm{d}W, \end{aligned}$$

where

$$\begin{aligned} I_2 &:= \frac{1}{2} \int_{\mathbb{S}^1} \left(\sigma u \partial_x \left(\sigma \partial_x u\right) + |\sigma \partial_x u|^2 \right) \mathrm{d}x \\ &\quad - \frac{1}{2} \int_{\mathbb{S}^1} \left(\partial_x^2 u\, \sigma \partial_x \left(\sigma \partial_x u\right) - |\partial_x \left(\sigma \partial_x u\right)|^2 \right) \mathrm{d}x, \\ I_3 &:= - \int_{\mathbb{S}^1} \left(\sigma u \partial_x u - \sigma \partial_x u \partial_x^2 u\right) \mathrm{d}x \end{aligned}$$

are as in (4.5) for ready comparison.

5. Tightness of probability laws.

We will prove that the probability laws $\{(u_n)_*\mathbb{P}\}$ of the Galerkin approximations $\{u_n\}$ are (n-uniformly) tight, on suitable Polish and quasi-Polish spaces. We will later construct weak solutions by applying a stochastic compactness argument. In one step of the argument, one makes use of tightness, which is linked to weak compactness of the laws. In contrast to the results in Section 4, tightness results are not uniform-in-ε.

5.1. Tightness on $L^2([0,T];H^1(\mathbb{S}^1))$ and on $C([0,T];H^1_w(\mathbb{S}^1))$.

We will first show improved temporal regularity in $L^2(\mathbb{S}^1)$. This will be used to establish the tightness of laws of $\{u_n\}$ on the Polish space $L^2([0,T];H^1(\mathbb{S}^1))$ and on the quasi-Polish space $C([0,T];H^1_w(\mathbb{S}^1))$.

Lemma 5.1 (Temporal L^2 continuity)**.** *For each $n \in \mathbb{N}$, let u_n be a solution to (3.1) with $\mathbb{E}\left\|u_0\right\|^p_{H^1(\mathbb{S}^1)} < \infty$ for $p > 2$. For any $\theta \in [0,(p-2)/4p)$, there exists a constant*

$$C = C\left(T, p, \theta, \mathbb{E}\left\|u_0\right\|^{4p}_{H^1(\mathbb{S}^1)}, \left\|\sigma\right\|_{W^{2,\infty}(\mathbb{S}^1)}\right),$$

independent of n and ε, such that

$$\mathbb{E}\left\|u_n\right\|^{2p}_{C^\theta([0,T];L^2(\mathbb{S}^1))} \le C. \tag{5.1}$$

Proof. We shall estimate $\mathbb{E}\left\|u_n(t)-u_n(s)\right\|^{2p}_{L^2(\mathbb{S}^1)}$ in terms of $|t-s|^{1+\gamma}$ for some $\gamma > 1$, and then appeal to Kolmogorov's continuity criterion.

First, we separate the spatial integral as

$$\int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big)^2 \,\mathrm{d}x = \int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big) \int_s^t \mathrm{d}u_n(r)\,\mathrm{d}x = \sum_{i=1}^5 I_i^n,$$

where

$$\begin{aligned}
I_1^n &:= -\int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big) \int_s^t u_n(r)\partial_x u_n(r)\,\mathrm{d}r\,\mathrm{d}x,\\
I_2^n &:= -\int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big) \int_s^t \partial_x P[u_n](r)\,\mathrm{d}r\,\mathrm{d}x,\\
I_3^n &:= \varepsilon\int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big) \int_s^t \partial_x^2 u_n(r)\,\mathrm{d}r\,\mathrm{d}x,\\
I_4^n &:= \frac{1}{2}\int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big) \int_s^t \sigma(r)\partial_x\left(\sigma\partial_x u\right)(r)\,\mathrm{d}r\,\mathrm{d}x,\\
I_5^n &:= \int_{\mathbb{S}^1} \big(u_n(t)-u_n(s)\big) \int_s^t \sigma(r)\partial_x u_n(r)\,\mathrm{d}W(r)\,\mathrm{d}x.
\end{aligned}$$

After an integration by parts involving $u_n\partial_x u_n = \partial_x\left(u_n^2/2\right)$, and using the bound $\left\|\partial_x u_n\right\|_{L^\infty([0,T];L^1(\mathbb{S}^1))} \le C\left\|u_n\right\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}$,

$$\begin{aligned}
\mathbb{E}\left|I_1^n\right|^p &\le \mathbb{E}\left(\left\|u_n\right\|^{2p}_{L^\infty([0,T]\times\mathbb{S}^1)}\left\|\partial_x u_n\right\|^p_{L^\infty([0,T];L^1(\mathbb{S}^1))}\right)|t-s|^p\\
&\le \left(\mathbb{E}\left\|u_n\right\|^{4p}_{L^\infty([0,T]\times\mathbb{S}^1)}\right)^{1/2}\left(\mathbb{E}\left\|u_n\right\|^{2p}_{L^\infty([0,T];H^1(\mathbb{S}^1))}\right)^{1/2}|t-s|^p.
\end{aligned}$$

We conclude, given the higher moment bounds (4.9) and (4.10), that $\mathbb{E}\left|I_1^n\right|^p \le C\left|t-s\right|^p$.

Recalling that $\partial_x P[u_n] = \partial_x K * \left(u_n^2 + \frac{1}{2}\left(\partial_x u_n\right)^2\right)$ and using Young's convolution inequality, we obtain

$$\begin{aligned}\|\partial_x P[u_n]\|_{L^2(\mathbb{S}^1)} &\le \|\partial_x K\|_{L^2(\mathbb{S}^1)} \left\| u_n^2 + \frac{1}{2}\left(\partial_x u_n\right)^2 \right\|_{L^1(\mathbb{S}^1)} \\ &\le \|\partial_x K\|_{L^2(\mathbb{S}^1)} \|u_n\|_{H^1(\mathbb{S}^1)}^2 ,\end{aligned}$$

and therefore

$$\begin{aligned}\mathbb{E}\,|I_2^n|^p &\le \left(2\,\|\partial_x K\|_{L^2(\mathbb{S}^1)}\right)^p \mathbb{E}\left(\|u_n\|_{L^\infty([0,T]\times\mathbb{S}^1)}^p \|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^{2p}\right)|t-s|^p \\ &\le C\left(\mathbb{E}\,\|u_n\|_{L^\infty([0,T]\times\mathbb{S}^1)}^{2p}\right)^{1/2}\left(\mathbb{E}\,\|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^{4p}\right)^{1/2}|t-s|^p .\end{aligned}$$

Again making use of (4.9) and (4.10), we arrive at $\mathbb{E}\,|I_2^n|^p \le C\,|t-s|^p$.

Similarly, after integration by parts, we obtain

$$\mathbb{E}\,|I_3^n|^p \le (2\varepsilon)^p\,\mathbb{E}\,\|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^{2p}\,|t-s|^p \le C\,|t-s|^p$$

and, noting that

$$\begin{aligned}&\partial_x\big(\sigma\left(u_n(t)-u_n(s)\right)\big)\sigma\partial_x u(r) \\ &\quad = \sigma\partial_x\sigma\left(u_n(t)-u_n(s)\right)\partial_x u(r) + \sigma^2\partial_x\left(u_n(t)-u_n(s)\right)\partial_x u(r)\end{aligned}$$

generates terms of the type handled before, $\mathbb{E}\,|I_4^n|^p \le C\,|t-s|^p$.

Finally, we estimate the stochastic term I_5^n. We cannot exchange the temporal and spatial integrals because the stochastic process

$$(\omega,t)\mapsto \int_{\mathbb{S}^1}\sigma\left(u_n(t)-u_n(s)\right)\partial_x u_n(r)\,\mathrm{d}x$$

is not $\mathcal{F}_r$-measurable, so we will instead estimate using the Cauchy–Schwarz inequality repeatedly and then the Burkholder–Davis–Gundy inequality. We have that $\left\|\partial_x\big(\sigma\left(u_n(t)-u_n(s)\right)\big)\right\|_{L^2(\mathbb{S}^1)} \le C_\sigma\,\|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1)))}$. After an integration by parts,

$$\begin{aligned}\mathbb{E}\,|I_5^n|^p &\le \mathbb{E}\left|\left\|\partial_x\big(\sigma\left(u_n(t)-u_n(s)\right)\big)\right\|_{L^2(\mathbb{S}^1)}\left(\int_{\mathbb{S}^1}\left|\int_s^t u_n\,\mathrm{d}W(r)\right|^2\mathrm{d}x\right)^{1/2}\right|^p \\ &\le \tilde{C}_{\sigma,p}\mathbb{E}\left(\|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^p\left(\int_{\mathbb{S}^1}\left|\int_s^t u_n\,\mathrm{d}W(r)\right|^2\mathrm{d}x\right)^{p/2}\right) \\ &\le \tilde{C}_{\sigma,p}\left(\mathbb{E}\,\|u_n\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}^{2p}\right)^{1/2}\left(\mathbb{E}\left(\int_{\mathbb{S}^1}\left|\int_s^t u_n\,\mathrm{d}W(r)\right|^2\mathrm{d}x\right)^p\right)^{1/2} \\ &\le C_{\sigma,p}\left(\int_{\mathbb{S}^1}\mathbb{E}\left|\int_s^t u_n\,\mathrm{d}W(r)\right|^{2p}\mathrm{d}x\right)^{1/2},\end{aligned}$$

where the final inequality is the result of (4.9) and Jensen's inequality. Finally, by (4.9) and the Burkholder–Davis–Gundy inequality, pointwise in x,

$$\left(\int_{\mathbb{S}^1}\mathbb{E}\left|\int_s^t u_n\,\mathrm{d}W(r)\right|^{2p}\mathrm{d}x\right)^{1/2} \le C_p\left(\mathbb{E}\int_{\mathbb{S}^1}\left(\int_s^t u_n^2\,\mathrm{d}r\right)^p\mathrm{d}x\right)^{1/2}$$

$$\le C\left(\mathbb{E}\left\|u_n\right\|_{L^\infty([0,T]\times\mathbb{S}^1)}^{2p}\right)^{1/2}|t-s|^{p/2}$$
$$\overset{(4.10)}{\le} C\,|t-s|^{p/2}.$$

Summarising, we have obtained

$$\mathbb{E}\left\|u_n(t)-u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p} \le C\,|t-s|^{p/2} = C\,|t-s|^{1+(p-2)/2},$$

where the constant C is independent of n and ε. By Kolmogorov's continuity criterion, there is a version of u_n in $C^\theta([0,T];L^2(\mathbb{S}^1))$, for any $\theta\in[0,(2-p)/4p)$, and a bound of the form (5.1). ☐

Remark 5.2. The temporal continuity bound (5.1) can also be carried out with respect to a fractional Sobolev norm via a computation following, e.g., [27, Lemma 2.1].

Lemma 5.3 (Tightness on $C([0,T];H^1_w(\mathbb{S}^1))$)**.** *For each $n\in\mathbb{N}$, let u_n be a solution to (3.1) with $\mathbb{E}\|u_0\|_{H^1(\mathbb{S}^1)}^p<\infty$ for $p>2$. The laws of $\{u_n\}$ are tight on $C([0,T];H^1_w(\mathbb{S}^1))$.*

Proof. Choose $\theta\in(0,(2-p)/4p)$. Given the compact embedding [39, Cor. B.2]

$$L^\infty([0,T];H^1(\mathbb{S}^1))\cap C^\theta([0,T];L^2(\mathbb{S}^1))\hookrightarrow C([0,T];H^1_w(\mathbb{S}^1)),$$

the laws of $\{u_n\}$ are tight on $C([0,T];H^1_w(\mathbb{S}^1))$ via the following standard computation: By the compact embedding, the sets

$$\mathcal{K}_R := \left\{u\in C([0,T];H^1_w(\mathbb{S}^1)) : \|u\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}+\|u\|_{C^\theta([0,T];L^2(\mathbb{S}^1))}\le R\right\}$$

are compact in $\mathcal{X}:=C([0,T];H^1_w(\mathbb{S}^1))$. Therefore, by Markov's inequality,

$$(u_n)_*\,\mathbb{P}\big(\mathcal{X}\backslash\mathcal{K}_R\big)\le\frac{1}{R}\mathbb{E}\left\|u_n\right\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}+\frac{1}{R}\mathbb{E}\left\|u_n\right\|_{C^\theta([0,T];L^2(\mathbb{S}^1))}.$$

By (4.3) and (5.1), the right-hand side tends to zero as $R\to\infty$. ☐

We need the following variant of the Aubin–Lions lemma [27, Thm. 2.1] (see also [49, Sec. 13.3]) to establish tightness on $L^2([0,T];H^1(\mathbb{S}^1))$.

Lemma 5.4. *Let $B_0\subseteq B\subseteq B_1$ be Banach spaces, B_0 and B_1 reflexive, with compact embedding of B_0 in B. Fix $p\in(1,\infty)$ and $\alpha\in(0,1)$. Let*

$$\mathcal{Y}=L^p([0,T];B_0)\cap W^{\alpha,p}([0,T];B_1),$$

be endowed with the natural norm. The embedding of $\mathcal{Y}$ in $L^p([0,T];B)$ is compact.

Tightness of probability measures is related to the stochastic boundedness of random variables. Below we prove that $u_n\in_{\mathrm{sb}} L^2_tH^2_x\cap W^{\theta,2}_tL^2_x$, uniformly in n, for some $\theta<(2-p)/4p$, making essential use of the dissipation part of (4.3).

Lemma 5.5 (Tightness on $L^2([0,T];H^1(\mathbb{S}^1))$)**.** *For each $n\in\mathbb{N}$, let u_n be a solution to (3.1) with $\mathbb{E}\|u_0\|_{H^1(\mathbb{S}^1)}^p<\infty$ for $p>2$. Let $\theta'\in(0,(2-p)/4p)$. The following stochastic boundedness estimate holds uniformly in n:*

$$\lim_{M\to\infty}\mathbb{P}\left(\|u_n\|_{L^2([0,T];H^2(\mathbb{S}^1))\cap W^{\theta',2}([0,T];L^2(\mathbb{S}^1))}>M\right)=0. \tag{5.2}$$

Moreover, the laws of $\{u_n\}$ are tight on $L^2([0,T];H^1(\mathbb{S}^1))$.

Proof. A natural norm on $L^2([0,T];H^2(\mathbb{S}^1))\cap W^{\theta',2}([0,T];L^2(\mathbb{S}^1))$ is

$$\|u_n\|_{L^2([0,T];H^2(\mathbb{S}^1))\cap W^{\theta',2}([0,T];L^2)} = \|u_n\|_{L^2([0,T];H^2(\mathbb{S}^1))} + \|u_n\|_{W^{\theta',2}([0,T];L^2(\mathbb{S}^1))}.$$

For $\theta \in (\theta',(2-p)/4p)$, the embeddings $C^\theta([0,T];B_1) \hookrightarrow C^{\theta'}([0,T];B_1) \hookrightarrow W^{\theta',2}([0,T];B_1)$ are continuous. Using Markov's inequality and (5.1) with $\theta = 1/5$,

$$\mathbb{P}\left(\left\{\|u_n\|_{W^{\theta',2}([0,T];L^2(\mathbb{S}^1))} > M\right\}\right) \le \frac{1}{M}\,\mathbb{E}\,\|u_n\|_{C^\theta([0,T];L^2(\mathbb{S}^1))} \lesssim \frac{1}{M}.$$

Next, using the energy estimate of Proposition 4.1, we obtain

$$\mathbb{P}\left(\left\{\|u_n\|_{L^2([0,T];H^2(\mathbb{S}^1))} > M\right\}\right) \le \frac{1}{M^2}\,\mathbb{E}\,\|u_n\|^2_{L^2([0,T];H^2(\mathbb{S}^1))} \lesssim \frac{1}{\varepsilon M^2}.$$

In view of the natural norm for the intersection space, this implies (5.2).

Tightness of the laws on $L^2([0,T];H^1(\mathbb{S}^1))$ now follows from Lemma 5.4 and the stochastic boundedness estimate (5.2). In particular, for each $\delta > 0$, there exists a number $M > 0$ and a compact set

$$\mathcal{A}_M = \Big\{v \in L^2([0,T];H^2(\mathbb{S}^1))\cap W^{\theta',2}([0,T];L^2(\mathbb{S}^1)) : \\ \|v\|_{L^2([0,T];H^2(\mathbb{S}^1))} + \|v\|_{W^{\theta',2}([0,T];L^2(\mathbb{S}^1))} \le M\Big\},$$

such that the complement $\mathcal{A}_M^c$ satisfies $(u_n)_*\,\mathbb{P}\big(\mathcal{A}_M^c\big) < \delta$. □

6. **Weak (martingale) solutions.** To be able to pass to the limit in the nonlinear terms in the SPDE (1.1), we must show that the Galerkin approximations $\{u_n\}$ converge strongly in (ω,t,x). Setting aside the probability variable ω, strong (t,x) convergence is linked to the spatial and temporal a priori estimates established in Section 4 and Section 5. On the other hand, the available estimates only ensure weak convergence in ω. To rectify this unfortunate (but typical) situation, we will replace the random variables $\{u_n\}$ by Skorokhod–Jakubowski a.s. representations $\{\tilde{u}_n\}$, which are defined on a new stochastic basis and will converge almost surely. The existence of $\{\tilde{u}_n\}$ will follow from the tightness estimates established in Section 5. Finally, we will show that the strong limit of $\{\tilde{u}_n\}$ constitutes a weak (martingale) solution according to Definition 2.1.

6.1. **Skorokhod–Jakubowski a.s. representations.** Introduce the path spaces

$$\begin{aligned} \mathcal{X}_{u,s} &:= L^2([0,T];H^1(\mathbb{S}^1)) \\ \mathcal{X}_{u,w} &:= C([0,T];H^1_w(\mathbb{S}^1)), \\ \mathcal{X}_W &:= C([0,T]), \\ \mathcal{X}_0 &:= H^1(\mathbb{S}^1), \end{aligned} \tag{6.1}$$

and set $\mathcal{X} := \mathcal{X}_{u,s}\times\mathcal{X}_{u,w}\times\mathcal{X}_W\times\mathcal{X}_0$. Denote by μ^n the (joint) law of the $\big(\mathcal{X},\mathcal{B}(\mathcal{X})\big)$-valued random variable $\big(u_{n,s},u_{n,w},W,\mathbf{\Pi}_n u_0\big)$. We denote by $\mu^n_{u,s}$, $\mu^n_{u,w}$, μ^n_W and μ^n_0 the laws of u_n, u_n, W and $\mathbf{\Pi}_n u_0$ on $\mathcal{X}_{u,s}$, $\mathcal{X}_{u,w}$, $\mathcal{X}_W$ and $\mathcal{X}_0$, respectively. (The subscripts "s" and "w" refer to the "strong" and "weak" topologies used in the subscripted path spaces and laws defined on them.)

Note carefully that we have used two copies of u_n in separate spaces $\mathcal{X}_{u,s}$ and $\mathcal{X}_{u,w}$ that do not inject continuously into one another. The aim of this manoeuvre is to ensure convergence in two separate topologies of the Skorokhod–Jakubowski representations of u_n. The rationale for this is explained in the appendix following Theorem A.5. The two variables are identified *post hoc* in Lemma 6.4.

Lemma 6.1 (Tightness of Galerkin approximations)**.** *The laws $\{\mu^n\}$ are tight.*

Proof. The tightness on $\mathcal{X}$ of the product measures $\{\mu^n_{u,s}\otimes\mu^n_{u,w}\otimes\mu^n_W\otimes\mu^n_0\}$ implies the tightness of the joint laws $\{\mu^n\}$ on $\mathcal{X}$. The tightness of $\{\mu^n_{u,i}\}$ on $\mathcal{X}_{u,i}$ for $i=1,2$ are stated in Lemmas 5.3 and 5.5. Since $\mathbf{\Pi}_n u_0\to u_0$ in $H^1(\mathbb{S}^1)$, the laws $\{\mu^n_0\}$ are tight on $H^1(\mathbb{S}^1)$. The elements of $\{\mu^n_W\}$ do not change with n, each μ^n_W is equal to the law of the Wiener process W (which is tight on $\mathcal{X}_W$). Hence, the tightness of the product measures $\{\mu^n_{u,s}\otimes\mu^n_{u,w}\otimes\mu^n_W\otimes\mu^n_0\}$ follows. □

Theorem 6.2 (Skorokhod–Jakubowski representations)**.** *There exist a probability space $(\tilde\Omega,\tilde{\mathscr{F}},\tilde{\mathbb{P}})$ and $\mathcal{X}$-valued variables $\{(\tilde u_{n,s},\tilde u_{n,w},\tilde W_n,\tilde u_{0,n})\}_{n\in\mathbb{N}}$, $(\tilde u_s,\tilde u_w,\tilde W,\tilde u_0)$, defined on $(\tilde\Omega,\tilde{\mathscr{F}},\tilde{\mathbb{P}})$, such that along a subsequence (not relabelled),*

$$\tilde u_{n,s}\sim u_n,\quad \tilde u_{n,w}\sim u_n,\quad \tilde W_n\sim W,\quad \tilde u_{0,n}\sim\mathbf{\Pi}_n u_0 \tag{6.2}$$

and, $\tilde{\mathbb{P}}$-almost surely,

$$(\tilde u_{n,s},\tilde u_{n,w},\tilde W_n,\tilde u_{0,n})\overset{n\uparrow\infty}{\longrightarrow}(\tilde u_s,\tilde u_w,\tilde W,\tilde u_0)\quad \text{in } \mathcal{X}.$$

Proof. Apply Theorem A.5. □

Remark 6.3. We need Jakubowski's version [35] of the Skorokhod representation theorem because our path space $\mathcal{X}$ contains the non-metrisable (but quasi-Polish) space $C([0,T];H^1_w(\mathbb{S}^1))$.

Lemma 6.4 (identification of doubled variables)**.** *For the sequence of variables defined in (6.2), $\tilde u_{n,s}=\tilde u_{n,w}$, $\tilde{\mathbb{P}}\otimes \mathrm{d}t\otimes\mathrm{d}x$–a.e. Moreover, $\tilde u_s=\tilde u_w$, $\tilde{\mathbb{P}}\otimes \mathrm{d}t\otimes\mathrm{d}x$–a.e.*

Remark 6.5. It is then henceforward sufficient to speak of $\tilde u_n:=\tilde u_{n,s}=\tilde u_{n,w}$ and $\tilde u:=\tilde u_s=\tilde u_w$.

Proof. For a fixed n, this follows directly from [35, Lemma 1], where an identification was made for variables in two Polish spaces. However, the completeness and separability of the path spaces were not used in the proof, and the lemma can be proven unchanged for quasi-Polish spaces.

For any $\varphi\in C^\infty([0,T]\times\mathbb{S}^1)$, $\eta\in L^\infty(\tilde\Omega)$, as $n\to\infty$,

$$\tilde{\mathbb{E}}\int_0^T\int_{\mathbb{S}^1}\eta\varphi\tilde u_{n,s}\,\mathrm{d}x\,\mathrm{d}t\to\tilde{\mathbb{E}}\int_0^T\int_{\mathbb{S}^1}\eta\varphi\tilde u_s\,\mathrm{d}x\,dt,$$
$$\tilde{\mathbb{E}}\int_0^T\int_{\mathbb{S}^1}\eta\varphi\tilde u_{n,w}\,\mathrm{d}x\,\mathrm{d}t\to\tilde{\mathbb{E}}\int_0^T\int_{\mathbb{S}^1}\eta\varphi\tilde u_w\,\mathrm{d}x\,dt,$$

and, since $\tilde u_{n,s}=\tilde u_{n,w}$,

$$\tilde{\mathbb{E}}\left(\eta\int_0^T\int_{\mathbb{S}^1}\varphi\tilde u_s\,\mathrm{d}x\,dt\right)=\tilde{\mathbb{E}}\left(\eta\int_0^T\int_{\mathbb{S}^1}\varphi\tilde u_w\,\mathrm{d}x\,dt\right).$$

From this it easily follows that $\tilde u_s=\tilde u_w$, $\tilde{\mathbb{P}}\otimes\mathrm{d}t\otimes\mathrm{d}x$-a.e. □

With $t\in[0,T]$ and X denoting $L^2([0,T];H^1(\mathbb{S}^1))$, $C([0,T];H^1_w(\mathbb{S}^1))$ or $C([0,T])$, let $f\mapsto f|_{[0,t]}:X\to X|_{[0,t]}$ denote the restriction operator to $[0,t]$. We define $\{\tilde{\mathcal{F}}_t\}_{t\ge0}$ to be the $\tilde{\mathbb{P}}$-augmented canonical filtration of $(\tilde u,\tilde W,\tilde u_0)$, i.e.,

$$\tilde{\mathcal{F}}_t:=\Sigma\left(\Sigma\big(\tilde u\big|_{[0,t]},\tilde W\big|_{[0,t]},\tilde u_0\big)\cup\{N\in\tilde{\mathcal{F}}:\tilde{\mathbb{P}}(N)=0\}\right).$$

where, for a collection E of subsets of $\tilde{\Omega}$, $\Sigma(E)$ denotes the smallest sigma algebra containing E. Denote by $\tilde{\mathcal{S}}$ the corresponding stochastic basis, that is,

$$\tilde{\mathcal{S}} := \big(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\geq 0}, \tilde{\mathbb{P}}\big). \tag{6.3}$$

Similarly, based on $\big(\tilde{u}_n, \tilde{W}_n, \tilde{u}_{0,n}\big)$, we define $\tilde{\mathcal{S}}_n = \big(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t^n\}_{t\geq 0}, \tilde{\mathbb{P}}\big)$. Then $\tilde{u}$, $\tilde{W}$ and $\tilde{u}_n$, $\tilde{W}_n$ are adapted relative to the stochastic bases $\tilde{\mathcal{S}}$, $\tilde{\mathcal{S}}_n$, respectively. Besides, by the equality of laws, $\tilde{W}_n$ is a Brownian motion on $\tilde{\mathcal{S}}_n$, and we have the following result.

Lemma 6.6 (Brownian motion)**.** *The process $\tilde{W}$ is a Brownian motion on $\tilde{\mathcal{S}}$.*

Proof. The proof is standard (see, e.g., [24, Lemma 4.8]), relying on Lévy's characterisation theorem (e.g., [41, Thm. IV.3.6]) and the equality of laws. The claim follows if we establishes that $\tilde{W}$ is a martingale relative to $\tilde{\mathcal{S}}$.

By the equivalence of laws, for $0 \leq s \leq t \leq T$,

$$\begin{aligned}
&\tilde{\mathbb{E}}\Big(\big(\tilde{W}_n(t) - \tilde{W}_n(s)\big)\,\gamma\big(\tilde{u}_n\big|_{[0,s]}, \tilde{u}_n\big|_{[0,s]}, \tilde{W}_n\big|_{[0,s]}\big)\Big)\\
&\qquad = \mathbb{E}\Big(\big(W(t) - W(s)\big)\,\gamma\big(u_n\big|_{[0,s]}, \tilde{u}_n\big|_{[0,s]}, W\big|_{[0,s]}\big)\Big) = 0,
\end{aligned}$$

because W is a martingale relative to $\mathcal{S} = \big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\big)$, for any continuous function $\gamma\colon L^2([0,s]; H^1(\mathbb{S}^1)) \times C([0,s]; H^1_w(\mathbb{S}^1)) \times C([0,s]) \to \mathbb{R}$. Moreover,

$$\sup_{n\in\mathbb{N}} \tilde{\mathbb{E}}\left|\tilde{W}_n(t)\right|^2 = \mathbb{E}\left|W(t)\right|^2 = t < \infty.$$

Therefore, by the Vitali convergence theorem,

$$\tilde{\mathbb{E}}\Big(\big(\tilde{W}(t) - \tilde{W}(s)\big)\gamma\big(\tilde{u}\big|_{[0,s]}, \tilde{u}\big|_{[0,s]}, \tilde{W}\big|_{[0,s]}\big)\Big) = 0,$$

so that $\tilde{W}$ is a martingale (and hence a Brownian motion) on $\tilde{\mathcal{S}}$. □

Next, we collect the convergence and continuity properties that are needed later to prove that the limit $\big(\tilde{\mathcal{S}}, \tilde{u}, \tilde{W}\big)$ is a weak H^m solution.

Lemma 6.7 (Convergence)**.** *Let u_n, $\tilde{u}_n$ and $\tilde{u}$ be defined as in Theorem 6.2 and Remark 6.5, and set $q_n := \partial_x u_n$, $\tilde{q}_n := \partial_x \tilde{u}_n$ and $\tilde{q} := \partial_x \tilde{u}$. Then $q_n \sim \tilde{q}_n$, and the following convergences hold $\tilde{\mathbb{P}}$-almost surely:*

$$\tilde{u}_n \overset{n\uparrow\infty}{\longrightarrow} \tilde{u}, \quad \text{in } C([0,T]; H^1_w(\mathbb{S}^1)), \tag{6.4a}$$

$$\tilde{q}_n \overset{n\uparrow\infty}{\longrightarrow} \tilde{q} \quad \text{in } L^2([0,T]\times\mathbb{S}^1), \tag{6.4b}$$

$$\tilde{u}_n^2 \overset{n\uparrow\infty}{\longrightarrow} \tilde{u}^2 \quad \text{in } L^1([0,T]; W^{1,1}(\mathbb{S}^1)), \tag{6.4c}$$

$$\tilde{q}_n^2 \overset{n\uparrow\infty}{\longrightarrow} \tilde{q}^2 \quad \text{in } L^1([0,T]\times\mathbb{S}^1), \tag{6.4d}$$

$$\tilde{u}_n\tilde{q}_n \to \tilde{u}\tilde{q} \quad \text{in } L^2([0,T]\times\mathbb{S}^1). \tag{6.4e}$$

Proof. Since $u_n \sim \tilde{u}_n$ on $L^2([0,T]; H^1(\mathbb{S}^1))$, we have $\partial_x u_n \sim \partial_x \tilde{u}_n$ on $L^2([0,T]\times\mathbb{S}^1)$. Next, regarding the convergence claims (6.4), the limits (6.4a) and (6.4b) follow directly from Theorem 6.2.

By the standard calculus inequality

$$\|fg\|_{W^{1,1}} \leq \|f\|_{H^1}\|g\|_{L^2} + \|g\|_{H^1}\|f\|_{L^2},$$

we also have (with $f = \tilde{u}_n - \tilde{u}$ and $g = \tilde{u}_n + \tilde{u}$),

$$\begin{aligned}\left\|\tilde{u}_n^2 - \tilde{u}^2\right\|_{L^1([0,T];W^{1,1}(\mathbb{S}^1))} \\ \le \left\|\tilde{u}_n + \tilde{u}\right\|_{L^2([0,T];H^1(\mathbb{S}^1))} \left\|\tilde{u}_n - \tilde{u}\right\|_{L^2([0,T];H^1(\mathbb{S}^1))} \stackrel{n\uparrow\infty}{\longrightarrow} 0,\end{aligned}$$

$\tilde{\mathbb{P}}$-almost surely. Finally, by (4.3), the embedding $H^1(\mathbb{S}^1) \hookrightarrow L^\infty(\mathbb{S}^1)$ and the equivalence of laws,

$$\begin{aligned}\left\|\tilde{u}_n\tilde{q}_n - \tilde{u}\tilde{q}\right\|_{L^2([0,T]\times\mathbb{S}^1)} &\le \left\|\tilde{u}_n - \tilde{u}\right\|_{L^2([0,T];L^\infty(\mathbb{S}^1))} \left\|\tilde{q}_n\right\|_{L^\infty([0,T];L^2(\mathbb{S}^1))} \\ &\quad + \left\|\tilde{q}_n - \tilde{q}\right\|_{L^2([0,T]\times\mathbb{S}^1)} \left\|\tilde{u}\right\|_{L^\infty([0,T]\times\mathbb{S}^1)} \stackrel{n\uparrow\infty}{\longrightarrow} 0,\end{aligned}$$

$\tilde{\mathbb{P}}$-almost surely. This establishes (6.4c)–(6.4e). □

Theorem 6.8 (Weak H^1 solution)**.** *Suppose $\sigma \in W^{2,\infty}(\mathbb{S}^1)$, $p > 2$, and that $\mathbb{E}\left\|u_0\right\|^p_{H^1(\mathbb{S}^1)} < \infty$. Let $\tilde{u}$, $\tilde{W}$, $\tilde{u}_0$ be the Skorokhod–Jakubowski a.s. representations from Theorem 6.2 (and Remark 6.5), and let $\tilde{\mathcal{S}}$ be the corresponding stochastic basis (6.3). Then $\left(\tilde{\mathcal{S}}, \tilde{u}, \tilde{W}\right)$ is a weak H^1 solution of (1.1) with initial law $\tilde{\Lambda} := (\tilde{u}_0)_*\,\tilde{\mathbb{P}}$, substituting for (c) in Definition 2.1 the following (here, $m = 1$):*

(c') $\tilde{u} : \Omega \times [0,T] \to H^1(\mathbb{S}^1)$ is

$$\tilde{u}(\omega, \cdot\,) \in C([0,T]; H^1_w(\mathbb{S}^1))$$

for $\tilde{\mathbb{P}}$-almost every $\omega \in \Omega$. Moreover, $\tilde{u} \in L^p(\tilde{\Omega}; L^\infty([0,T]; H^1(\mathbb{S}^1))) \cap L^2(\tilde{\Omega} \times [0,T]; H^2(\mathbb{S}^1))$.

Remark 6.9. The difference between (c') and (c) of Definition 2.1 is the weakened $\tilde{\mathbb{P}}$-almost sure inclusion $\tilde{u} \in C([0,T]; H^1_w(\mathbb{S}^1))$ along with the lack of temporal continuity requirement in $\tilde{u} \in L^p(\Omega; L^\infty([0,T]; H^1(\mathbb{S}^1)))$ in (c') in place of $\tilde{u} \in L^p(\Omega; C([0,T]; H^1(\mathbb{S}^1)))$ of (c). This continuity in $H^1(\mathbb{S}^1)$ ("strong temporal continuity") is not necessary for establishing pathwise uniqueness in Section 7.2 below, and subsequent (stochastic) strong existence (Section 7.3). We therefore relegate the proof of strong temporal continuity to Proposition 7.8.

Proof. We continue to use the notations from Lemma 6.7. By the equality of joint laws on $\mathcal{X}$, see (6.2), we also have

$$\left(u_n, u_n, W, \mathbf{\Pi}_n u_0, q_n\right) \sim \left(\tilde{u}_n, \tilde{u}_n, \tilde{W}_n, \tilde{u}_{0,n}, \tilde{q}_n\right) \quad \text{on } \mathcal{X} \times L^2([0,T]\times\mathbb{S}^1), \tag{6.5}$$

because ∂_x is a bounded operator from $H^1(\mathbb{S}^1)$ to $L^2(\mathbb{S}^1)$. For each fixed $n \in \mathbb{N}$ and $\varphi \in C^1(\mathbb{S}^1)$, consider the function $F_{\varphi,n} : \mathcal{X} \times L^2([0,T]\times\mathbb{S}^1) \to \mathbb{R}$ defined by

$$F_{\varphi,n}\left[\left(\tilde{u}_n(s), \tilde{W}_n(s), \tilde{u}_{0,n}, \tilde{q}_n(s), s \in [0,t]\right)\right] = I_1^n(t) + I_2^n(t) + I_3^n(t) + I_4^n(t),$$

where

$$\begin{aligned}I_1^n(t) &:= \int_{\mathbb{S}^1} \tilde{u}_n(t)\varphi\,\mathrm{d}x - \int_{\mathbb{S}^1} \tilde{u}_{0,n}\varphi\,\mathrm{d}x - \int_0^t\int_{\mathbb{S}^1} \varepsilon\,\tilde{q}_n\partial_x\varphi\,\mathrm{d}x\,\mathrm{d}s, \\ I_2^n(t) &:= \int_0^t\int_{\mathbb{S}^1}\left[\tilde{u}_n\,\tilde{q}_n + \mathbf{\Pi}_n\partial_x P\left[\tilde{u}_n\right]\right]\varphi\,\mathrm{d}x\,\mathrm{d}s, \\ I_3^n(t) &:= -\frac{1}{2}\int_0^t\int_{\mathbb{S}^1} \sigma\tilde{q}_n\,\partial_x\left(\sigma\mathbf{\Pi}_n\varphi\right)\,\mathrm{d}x\,\mathrm{d}s, \\ I_4^n(t) &:= \int_0^t\int_{\mathbb{S}^1} \sigma\partial_x\tilde{u}_n\,\mathbf{\Pi}_n\varphi\,\mathrm{d}x\,\mathrm{d}\tilde{W}_n.\end{aligned}$$

Recall that a Baire function of class κ, where κ is an ordinal number, is a function that is the pointwise limit of Baire functions of class $\kappa-1$, and class 0 Baire functions are the continuous functions. We have that $F_{\varphi,n}$ is a Baire function of class 1. In particular, the inclusion of the stochastic integral in this class can be seen by it being the pointwise limit of temporally mollified approximations along the lines of Benssousan [4, Sec. 4.3.5] or [23, Lemma 2.1] (see Lemma A.3). Hence, by the equivalence of joint laws (6.5), we have [46, p. 105]

$$\begin{aligned}
&\tilde{\mathbb{P}}\left(\left\{F_{\varphi,n}\big[\big(\tilde{u}_n(s), \tilde{W}_n(s), \tilde{u}_{0,n}, \tilde{q}_n(s), s\in[0,t]\big)\big]=0\right\}\right)\\
&\qquad=\mathbb{P}\Big(\big\{F_{\varphi,n}\big[\big(u_n(s), W(s), \mathbf{\Pi}_n u_0, q_n(s), s\in[0,t]\big)\big]=0\big\}\Big)\overset{(3.1)}{=}1.
\end{aligned}$$

We now establish the convergence of $I_1^n,\ldots,I_4^n$ separately.

1. Convergence of I_2^n.
We estimate as follows:

$$\begin{aligned}
&\Big\|\tilde{u}\,\tilde{q}+\partial_x P[\tilde{u}]-\mathbf{\Pi}_n\Big(\tilde{u}_n\,\tilde{q}_n+\partial_x P\big[\tilde{u}_n\big]\Big)\Big\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\quad\le\|\tilde{u}\tilde{q}-\mathbf{\Pi}_n\left(\tilde{u}_n\tilde{q}_n\right)\|_{L^2([0,T]\times\mathbb{S}^1)}+\big\|\partial_x P\big[\tilde{u}\big]-\mathbf{\Pi}_n\partial_x P\big[\tilde{u}_n\big]\big\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\quad\le\|\tilde{u}\tilde{q}-\mathbf{\Pi}_n\left(\tilde{u}\tilde{q}\right)\|_{L^2([0,T]\times\mathbb{S}^1)}+\|\mathbf{\Pi}_n\left(\tilde{u}\tilde{q}-\tilde{u}_n\tilde{q}_n\right)\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\qquad+\big\|\partial_x P\big[\tilde{u}\big]-\mathbf{\Pi}_n\partial_x P\big[\tilde{u}\big]\big\|_{L^2([0,T]\times\mathbb{S}^1)}+\big\|\mathbf{\Pi}_n\left(\partial_x P\big[\tilde{u}\big]-\partial_x P\big[\tilde{u}_n\big]\right)\big\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\quad\le\|\tilde{u}\tilde{q}-\mathbf{\Pi}_n\left(\tilde{u}\tilde{q}\right)\|_{L^2([0,T]\times\mathbb{S}^1)}+\big\|\partial_x P\big[\tilde{u}\big]-\mathbf{\Pi}_n\partial_x P\big[\tilde{u}\big]\big\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\qquad+\|\tilde{u}\tilde{q}-\tilde{u}_n\tilde{q}_n\|_{L^2([0,T]\times\mathbb{S}^1)}+\|\partial_x P[\tilde{u}]-\partial_x P[\tilde{u}_n]\|_{L^2([0,T]\times\mathbb{S}^1)}\ \text{(Bessel's ineq.)}\\
&\quad\le\|\left(1-\mathbf{\Pi}_n\right)\left(\tilde{u}\tilde{q}\right)\|_{L^2([0,T]\times\mathbb{S}^1)}+\|\left(1-\mathbf{\Pi}_n\right)\partial_x P[\tilde{u}]\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\qquad+\|\tilde{u}\tilde{q}-\tilde{u}_n\tilde{q}_n\|_{L^2([0,T]\times\mathbb{S}^1)}+\|\partial_x K\|_{L^2(\mathbb{S}^1)}\left\|\tilde{u}^2-\tilde{u}_n^2\right\|_{L^1([0,T]\times\mathbb{S}^1)}\\
&\qquad+\frac{1}{2}\|\partial_x K\|_{L^2(\mathbb{S}^1)}\left\|\tilde{q}^2-\tilde{q}_n^2\right\|_{L^1([0,T]\times\mathbb{S}^1)}\overset{n\uparrow\infty}{\longrightarrow}0,\quad\tilde{\mathbb{P}}\text{-a.s.},\quad\text{(by Lemma 6.7)}
\end{aligned}$$

using the convergence $\mathbf{\Pi}_n\to 1$ in the operator norm on $L^2(\mathbb{S}^1)\to L^2(\mathbb{S}^1)$. This implies that

$$I_2^n\overset{n\uparrow\infty}{\longrightarrow}\int_0^t\int_{\mathbb{S}^1}\Big[\tilde{u}\,\partial_x\tilde{u}+\partial_x P\big[\tilde{u}\big]\Big]\varphi\,\mathrm{d}x\,\mathrm{d}s,\quad\tilde{\mathbb{P}}\text{-a.s.}\tag{6.6}$$

2. Convergence of I_1^n and I_3^n.
For any $\varphi\in C^1(\mathbb{S}^1)$, $\tilde{\mathbb{P}}$-almost surely,

$$\begin{aligned}
&\left|-\frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\sigma\tilde{q}\,\partial_x\left(\sigma\varphi\right)\,\mathrm{d}x\,\mathrm{d}s-I_3^n\right|\\
&\le\frac{1}{2}\int_0^t\left|\int_{\mathbb{S}^1}\sigma\big(\tilde{q}-\tilde{q}_n\big)\partial_x\left(\sigma\varphi\right)\,\mathrm{d}x\right|\,\mathrm{d}s\\
&\quad+\frac{1}{2}\int_0^t\left|\int_{\mathbb{S}^1}\sigma\tilde{q}_n\,\partial_x\big(\sigma\left(1-\mathbf{\Pi}_n\right)\varphi\big)\,\mathrm{d}x\right|\,\mathrm{d}s\\
&\le C_{\sigma,\varphi}\left\|\tilde{q}-\tilde{q}_n\right\|_{L^2([0,T]\times\mathbb{S}^1)}\\
&\qquad+C_\sigma\left\|\tilde{q}_n\right\|_{L^2([0,T]\times\mathbb{S}^1)}\left\|\left(1-\mathbf{\Pi}_n\right)\varphi\right\|_{H^1(\mathbb{S}^1)}\overset{n\uparrow\infty}{\longrightarrow}0.
\end{aligned}$$

Similarly, by the $\tilde{\mathbb{P}}$-a.s. $L^2_{t,x}$ convergence of $\tilde{q}_n$, cf. (6.4),

$$\left|\int_0^t \int_{\mathbb{S}^1} \left(\varepsilon\tilde{q} - \varepsilon\tilde{q}_n\right) \partial_x\varphi \, \mathrm{d}x \, \mathrm{d}s\right| \overset{n\uparrow\infty}{\longrightarrow} 0, \quad \tilde{\mathbb{P}}\text{-a.s.}$$

The convergence

$$\int_{\mathbb{S}^1} \tilde{u}_n(t)\varphi \, \mathrm{d}x \to \int_{\mathbb{S}^1} \tilde{u}(t)\varphi \, \mathrm{d}x, \qquad \tilde{\mathbb{P}}\text{-a.s.}$$

follows from the $\tilde{\mathbb{P}}$-a.s. convergence $\tilde{u}_n \to \tilde{u}$ in $C([0,T]; H^1_w(\mathbb{S}^1))$, see (6.4), noting that $\varphi \in C^1(\mathbb{S}^1) \subseteq H^1(\mathbb{S}^1)$. Finally, the convergence

$$\int_{\mathbb{S}^1} \tilde{u}_{0,n}\varphi \, \mathrm{d}x \overset{n\uparrow\infty}{\longrightarrow} \int_{\mathbb{S}^1} \tilde{u}_0\varphi \, \mathrm{d}x, \quad \tilde{\mathbb{P}}\text{-a.s.}$$

is a direct consequence of Theorem 6.2 and (6.1).

Combining these results we find that

$$\begin{aligned} I_1^n + I_3^n \overset{n\uparrow\infty}{\longrightarrow} & \int_{\mathbb{S}^1} \tilde{u}(t)\varphi \, \mathrm{d}x - \int_{\mathbb{S}^1} \tilde{u}_0\varphi \, \mathrm{d}x - \int_0^t \int_{\mathbb{S}^1} \varepsilon\partial_x\tilde{u} \, \partial_x\varphi \, \mathrm{d}x \, \mathrm{d}s \\ & - \frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \sigma \, \partial_x\tilde{u} \, \partial_x\left(\sigma\varphi\right) \, \mathrm{d}x \, \mathrm{d}s, \quad \tilde{\mathbb{P}}\text{-a.s.} \end{aligned} \tag{6.7}$$

3. Convergence of I_4^n.

First, $\tilde{\mathbb{P}}$-almost surely., we have

$$\begin{aligned} \left\|\sigma\tilde{q} - \mathbf{\Pi}_n\left(\sigma\tilde{q}_n\right)\right\|_{L^2([0,T]\times\mathbb{S}^1)} \le & \left\|\sigma\left(\tilde{q} - \tilde{q}_n\right)\right\|_{L^2([0,T]\times\mathbb{S}^1)} \\ & + \left\|\sigma\tilde{q}_n - \mathbf{\Pi}_n\left(\sigma\tilde{q}_n\right)\right\|_{L^2([0,T]\times\mathbb{S}^1)} \overset{n\uparrow\infty}{\longrightarrow} 0, \end{aligned}$$

and so $\mathbf{\Pi}_n\left(\sigma\,\tilde{q}_n\right) \overset{n\uparrow\infty}{\longrightarrow} \sigma\tilde{q}$ in $L^2([0,T]\times\mathbb{S}^1)$ in probability. Besides, $\tilde{W}_n \overset{n\uparrow\infty}{\longrightarrow} \tilde{W}$ in $C([0,T])$ $\tilde{\mathbb{P}}$-almost surely, and thus in probability. Therefore, by Lemma A.3,

$$I_4^n \overset{n\uparrow\infty}{\longrightarrow} \int_0^t \sigma\tilde{q} \, \mathrm{d}\tilde{W}, \quad \text{in } L^2([0,T]\times\mathbb{S}^1), \tag{6.8}$$

in probability and hence $\tilde{\mathbb{P}}$-almost surely along a subsequence.

4. Weak formulation.

Gathering (6.6), (6.7), and (6.8), we have shown that $\tilde{u}$, $\tilde{W}$, $\tilde{u}_0$ satisfy, for any $\varphi \in C^1_c(\mathbb{S}^1)$,

$$\begin{aligned} 0 = & \int_{\mathbb{S}^1} \varphi\tilde{u} \, \mathrm{d}x\Big|_0^T + \int_0^T \int_{\mathbb{S}^1} \varphi \, \tilde{u} \, \partial_x\tilde{u} \, \mathrm{d}x \, \mathrm{d}t - \varepsilon\int_0^T \int_{\mathbb{S}^1} \varphi\partial_x^2\tilde{u} \, \mathrm{d}x \, \mathrm{d}t \\ & + \int_0^T \int_{\mathbb{S}^1} \varphi\partial_x P\left[\tilde{u}\right] \mathrm{d}x \, \mathrm{d}t - \frac{1}{2}\int_0^T \int_{\mathbb{S}^1} \sigma \, \partial_x\left(\sigma\partial_x\tilde{u}\right) \, \mathrm{d}x \, \mathrm{d}t \\ & + \int_0^T \int_{\mathbb{S}^1} \sigma\partial_x\tilde{u} \, \mathrm{d}x \, \mathrm{d}\tilde{W}, \\ \tilde{u}(0) = & \ \tilde{u}_0 \end{aligned} \tag{6.9}$$

as in Definition 2.1(e).

5. Appropriate inclusions.

The $\tilde{\mathbb{P}}$-almost sure inclusion $\tilde{u} \in C([0,T]; H^1_w(\mathbb{S}^1))$ follows directly from the Skorokhod argument of Theorem 6.2. Therefore, we are left to show that $\tilde{u} \in L^p(\Omega; L^\infty([0,T]; H^1(\mathbb{S}^1))) \cap L^2(\Omega\times[0,T]; H^2(\mathbb{S}^1))$.

By the Lusin–Souslin theorem [36, Thm. 15.1], the inclusion $L^2([0,T];H^2(\mathbb{S}^1)) \hookrightarrow L^2([0,T];H^1(\mathbb{S}^1))$ is Borel. We can then invoke the equality of laws to obtain

$$\tilde{\mathbb{E}}\,\|\tilde{u}_n\|^2_{L^2([0,T];H^2(\mathbb{S}^1))} = \mathbb{E}\,\|u_n\|^2_{L^2([0,T];H^2(\mathbb{S}^1))} < C,$$

where C is independent of n, by Theorem 4.1. This implies that $\tilde{q}_n$ are uniformly bounded in $L^2(\Omega\times[0,T]\times\mathbb{S}^1)$. Therefore, by reflexivity, any weak limit is also in $L^2(\Omega\times[0,T]\times\mathbb{S}^1)$.

The inclusion $L^\infty([0,T];H^1(\mathbb{S}^1)) \hookrightarrow C([0,T];H^1_w(\mathbb{S}^1))$ is continuous because for any $\varphi\in H^1(\mathbb{S}^1)^* = H^{-1}(\mathbb{S}^1)$,

$$\sup_{t\in[0,T]} \left|\langle u,\varphi\rangle_{H^{-1},H^1,}\right| \le \|\varphi\|_{H^{-1}(\mathbb{S}^1)} \sup_{t\in[0,T]} \|u\|_{H^1(\mathbb{S}^1)}\,.$$

Therefore, by the quasi-Polish version of the Lusin–Souslin theorem [39, Cor. A.2], we maintain as before the higher moment bound

$$\tilde{\mathbb{E}}\,\|\tilde{u}_n\|^p_{L^\infty([0,T];H^1(\mathbb{S}^1))} = \mathbb{E}\,\|u_n\|^p_{L^\infty([0,T];H^1(\mathbb{S}^1))} < C.$$

□

7. **Pathwise uniqueness and proof of Theorem 1.1.** In this section, we will show pathwise uniqueness and, consequently, the existence of strong solutions in the energy space $L^2\left(\Omega;L^\infty([0,T];H^1(\mathbb{S}^1))\right)$ (Theorem 7.6). This will involve estimates similar to the energy inequality in Proposition 4.1. However, as we are dealing with solutions a.s. in $L^\infty([0,T];H^1(\mathbb{S}^1))$, calculations using smooth Galerkin approximations cannot be reproduced here. To keep using the standard (finite-dimensional) Itô formula, we convolve the SPDE against a standard Friedrichs mollifier J_δ, making it possible to interpret the SPDE pointwise in x. Mollification introduces error terms to the equation, (see (7.16)). We will first state and prove convergence results for these error terms.

7.1. **Regularisation errors.** We begin this subsection by proving first order commutator estimates in the stochastic setting. Notice that the fourth moment assumption is made. This assumption is the reason that a bounded $p>4$ moment is needed on the initial condition (e.g., in Theorem 1.1). The assumption itself arises from (7.3), where the $L^\infty_t L^2_x$ boudedness of $\partial_x u$ is exploited in applying Young's convolution inequality. It is true that $\partial_x u$ is in $L^{3-}_{\omega,t,x}$ uniformly in ε, but because of the extra square in the exponent, this is difficult to exploit. Higher integrability bounds for fixed $\varepsilon>0$ exist but may only hold up to stopping time.

Throughout the paper we let J_δ be a standard Friedrichs (spatial) mollifier, and set $u_\delta := u * J_\delta$, and use δ as subscript to denote a mollified function.

Lemma 7.1 (Commutator estimates)**.** *Let $u,v,w\in L^4(\Omega;L^\infty([0,T];H^1(\mathbb{S}^1)))$, and suppose $\sigma\in W^{1,\infty}(\mathbb{S}^1)$. Finally, let $K\in W^{1,\infty}(\mathbb{S}^1)$ be a given kernel function. Define the commutator functions:*

$$\begin{aligned} E^1_\delta = E^1_\delta(u,v) &:= \Big(u\,\partial_x u - v\,\partial_x v\Big) * J_\delta - (u_\delta\,\partial_x u_\delta - v_\delta\,\partial_x v_\delta) \\ &\qquad + \partial_x K * \left(u^2 - v^2 + \frac{1}{2}\Big((\partial_x u)^2 - (\partial_x v)^2\Big)\right) * J_\delta \\ &\qquad - \partial_x K * \left(u_\delta^2 - v_\delta^2 + \frac{1}{2}\Big((\partial_x u_\delta)^2 - (\partial_x v_\delta)^2\Big)\right), \end{aligned} \tag{7.1}$$

$$E^2_\delta = E^2_\delta(w) := \big(\sigma\,\partial_x w\big) * J_\delta - \sigma\,\partial_x w_\delta,$$

$$E^3_\delta = E^3_\delta(w) := -\frac{1}{2}\big(\sigma\,\partial_x\left(\sigma\,\partial_x w\right)\big) * J_\delta + \frac{1}{2}\sigma\,\partial_x\left(\sigma\,\partial_x w_\delta\right).$$

The following convergences hold:

$$\mathbb{E}\left\|E_\delta^1\right\|_{L^2([0,T]\times\mathbb{S}^1)}^2, \mathbb{E}\left\|E_\delta^2\right\|_{L^2([0,T];H^1(\mathbb{S}^1))}^2, \mathbb{E}\left\|E_\delta^3\right\|_{L^2([0,T]\times\mathbb{S}^1)}^2 \xrightarrow{\delta\downarrow 0} 0. \tag{7.2}$$

Proof. Whilst these commutator estimates are similar to the classical ones of [25], we prove them here both because we are in the stochastic setting, with an extra integral in $\mathrm{d}\mathbb{P}$, and also because the extra temporal integrability on $\|u\|_{H^1(\mathbb{S}^1)}$ permits for the slightly stronger results that we shall be using. In particular, we have bounds in L^2_x, and not only L^1_x, for E^1_δ.

1. Convergence of E^1_δ.

For the transport terms we have

$$\begin{aligned}
&\|(u\,\partial_x u) * J_\delta - u_\delta\,\partial_x u_\delta\|_{L^2([0,T]\times\mathbb{S}^1)}^2 \\
&\quad\lesssim \|(u\,\partial_x u) * J_\delta - u\,\partial_x u_\delta\|_{L^2([0,T]\times\mathbb{S}^1)}^2 + \|u\,\partial_x u_\delta - u_\delta\,\partial_x u_\delta\|_{L^2([0,T]\times\mathbb{S}^1)}^2 \\
&\quad= \left\|\int_{\mathbb{S}^1} \frac{u(\,\cdot\,)-u(y)}{\cdot - y}\partial_y u(y)\,(\,\cdot\, - y) J_\delta(\,\cdot\, - y)\,\mathrm{d}y\right\|_{L^2([0,T]\times\mathbb{S}^1)}^2 \\
&\qquad+ \left\|\int_{\mathbb{S}^1} \frac{u(\,\cdot\,)-u(y)}{\cdot - y}\partial_x u_\delta(\,\cdot\,)\,(\,\cdot\, - y)\,J_\delta(\,\cdot\, - y)\,\mathrm{d}y\right\|_{L^2([0,T]\times\mathbb{S}^1)}^2 \\
&\quad=: I_1^\delta + I_2^\delta.
\end{aligned}$$

By Young's convolution inequality, and the fact that $\delta\,\|J_\delta\|_{L^2(\mathbb{S}^1)} \lesssim \sqrt{\delta}$,

$$\begin{aligned}
\mathbb{E}\left|I_1^\delta\right| &\lesssim \mathbb{E}\int_0^T \left\||\partial_x u| \sup_{|h|\le\delta}\left|\frac{u(\,\cdot\,+h)-u(\,\cdot\,)}{h}\right|\right\|_{L^1(\mathbb{S}^1)}^2 \delta^2\,\|J_\delta\|_{L^2(\mathbb{S}^1)}^2\,\mathrm{d}s \\
&\lesssim \mathbb{E}\,\delta\int_0^T \|\partial_x u\|_{L^2(\mathbb{S}^1)}^4\,\mathrm{d}s \lesssim_T \delta.
\end{aligned} \tag{7.3}$$

Similarly,

$$\begin{aligned}
\mathbb{E}\left|I_2^\delta\right| &\lesssim \delta\,\mathbb{E}\left|\int_0^T \|\partial_x u\|_{L^2}^2\,\|\partial_x u_\delta\|_{L^2(\mathbb{S}^1)}^2\,\mathrm{d}s\right| \\
&\lesssim \delta\,\mathbb{E}\int_0^T \|\partial_x u\|_{L^2(\mathbb{S}^1)}^4\,\mathrm{d}s \lesssim_T \delta.
\end{aligned}$$

Therefore,

$$\mathbb{E}\,\|(u\,\partial_x u) * J_\delta - u_\delta\,\partial_x u_\delta\|_{L^2([0,t]\times\mathbb{S}^1)}^2 \le \mathbb{E}\big(I_1^\delta + I_2^\delta\big) \xrightarrow{\delta\downarrow 0} 0.$$

Consider the terms in E^1_δ involving the kernel K, for which $\|\partial_x K\|_{L^2(\mathbb{S}^1)} \lesssim 1$. For any $\xi \in L^4\big(\Omega; L^\infty([0,T]; L^2(\mathbb{S}^1))\big)$, we find

$$\begin{aligned}
&\left\|\partial_x K * \xi^2 * J_\delta - \partial_x K * \xi_\delta^2\right\|_{L^2(\mathbb{S}^1)}^2 \\
&\qquad\le \|\partial_x K\|_{L^2(\mathbb{S}^1)}^2 \left\|\xi^2 * J_\delta - \xi_\delta^2\right\|_{L^1(\mathbb{S}^1)}^2 \\
&\qquad\lesssim \left\|\xi^2 * J_\delta - \xi^2\right\|_{L^1(\mathbb{S}^1)}^2 + \left\|\xi^2 - \xi\xi_\delta\right\|_{L^1(\mathbb{S}^1)}^2 + \left\|\xi\xi_\delta - \xi_\delta^2\right\|_{L^1(\mathbb{S}^1)}^2.
\end{aligned}$$

By standard properties of Friedrichs mollifiers, the terms on the right-hand side all tend to zero as $\delta \to 0$. We take $\xi = u, v$ or $\xi = \partial_x u, \partial_x v$ in the calculation above.

Combining the foregoing calculations, we arrive at

$$\mathbb{E}\left\|E_\delta^1\right\|^2_{L^2([0,T]\times\mathbb{S}^1)} \xrightarrow{\delta\downarrow 0} 0.$$

2. Convergence of E_δ^2.

For any $\xi \in L^2(\Omega\times[0,T]\times\mathbb{S}^1)$ (such as $\xi = u$ or $\xi = \partial_x u$), the convergence

$$\big(\sigma\xi\big) * J_\delta - (\sigma\,\xi_\delta) \xrightarrow{\delta\downarrow 0} 0 \quad \text{in } L^2(\Omega\times[0,T]\times\mathbb{S}^1).$$

is a direct result of the dominated convergence theorem.

The convergence

$$\partial_x\big(\sigma\xi\big) * J_\delta - \partial_x\left(\sigma\,\xi_\delta\right) \xrightarrow{\delta\downarrow 0} 0 \quad \text{in } L^2([0,T]\times\mathbb{S}^1)$$

follows directly from [25, Lemma II.1], where it was shown that

$$\left\|\partial_x\big(\sigma\xi\big) * J_\delta - \partial_x\left(\sigma\,\xi_\delta\right)\right\|_{L^2([0,T]\times\mathbb{S}^1)} \le C_{\delta,\sigma}\left\|\xi\right\|_{L^2([0,T]\times\mathbb{S}^1)},$$

where $C_{\delta,\sigma} \xrightarrow{\delta\downarrow 0} 0$ and is independent of ξ (and therefore deterministic). This gives

$$\mathbb{E}\left\|E_\delta^2\right\|^2_{L^2([0,T];H^1(\mathbb{S}^1))} \xrightarrow{\delta\downarrow 0} 0.$$

3. Convergence of E_δ^3.

Since $w \in L^4\big(\Omega; L^\infty([0,T];H^1(\mathbb{S}^1))\big)$, again setting $\xi = \partial_x w$, so that ξ belongs to $L^4\big(\Omega; L^\infty([0,T];L^2(\mathbb{S}^1))\big)$, the commutator can be written as

$$\begin{aligned}
-2E_\delta^3 &= \Big(\sigma\,\partial_x\left(\sigma\xi\right)\Big) * J_\delta - \sigma\,\partial_x\left(\sigma\xi_\delta\right) \\
&= \Big[\Big(\sigma\,\partial_x\sigma\xi\Big) * J_\delta - \sigma\,\partial_x\sigma\xi_\delta\Big] + \Big[\Big(\sigma^2\,\partial_x\xi\Big) * J_\delta - \sigma^2\,\partial_x\xi_\delta\Big].
\end{aligned}$$

We can then apply step 2 of the proof to conclude that (7.2) holds.

We point out that we needed only L^2-integrability in ω in *Steps 2* and *3*. ☐

Next, we introduce an operator notation that will be indispensable in the next two results. For $f \in L^p(\mathbb{S}^1)$, $1 \le p \le \infty$, set

$$\mathbf{j}_\delta f := f_\delta = f * J_\delta, \quad \mathbf{\Sigma} f := \partial_x\left(\sigma f\right) \tag{7.4}$$

where J_δ is the mollifier used in the definition of E_δ^3. Finally, we define

$$\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](f) := \mathbf{\Sigma}\mathbf{j}_\delta f - \mathbf{j}_\delta\mathbf{\Sigma} f = \partial_x\big(\sigma f_\delta\big) - \partial_x(\sigma f) * J_\delta.$$

Lemma 7.2 (Double commutator estimate)**.** *Let $\xi \in L^2(\Omega\times[0,T]\times\mathbb{S}^1)$, and suppose $\sigma \in W^{2,\infty}(\mathbb{S}^1)$. Then*

$$R_\delta := J_\delta * \partial_x\big(\sigma\,\partial_x\left(\sigma\xi\right)\big) - 2\partial_x\big(\sigma\,J_\delta * \partial_x\left(\sigma\xi\right)\big) + \partial_x\big(\sigma\,\partial_x\left(\sigma\xi_\delta\right)\big) \xrightarrow{\delta\downarrow 0} 0$$

in $L^2(\Omega\times[0,T]\times\mathbb{R})$.

Remark 7.3. An almost sure version of this lemma (instead of convergence in $L^2(\Omega)$) was stated with a similar proof in [33, Lemma B.3]. Furthermore, we allow for a more general σ here than in [40]; the work [40] imposes a divergence-free condition on σ (with $\mathbb{S}^1$ replaced by $\mathbb{R}^d$).

Proof. Using the operator notation defined in (7.4), we can write $-R_\delta$ as a double commutator:

$$-R_\delta = \Big[\big[\mathbf{\Sigma}, \mathbf{j}_\delta\big], \mathbf{\Sigma}\Big](\xi) = \big[\mathbf{\Sigma}, \mathbf{j}_\delta\big](\mathbf{\Sigma}\xi) - \mathbf{\Sigma}\big[\mathbf{\Sigma}, \mathbf{j}_\delta\big](\xi)$$
$$= 2\mathbf{\Sigma}\mathbf{j}_\delta\mathbf{\Sigma}\xi - \mathbf{j}_\delta\mathbf{\Sigma}\mathbf{\Sigma}\xi - \mathbf{\Sigma}\mathbf{\Sigma}\mathbf{j}_\delta\xi. \tag{7.5}$$

Term-by-term we have

$$2\mathbf{\Sigma}\mathbf{j}_\delta\mathbf{\Sigma}\xi(x) = 2\int_{\mathbb{R}} \partial_{xx}^2 J_\delta(x-y)\sigma(x)\sigma(y)\xi(y)\,\mathrm{d}y \tag{7.6}$$
$$+\, 2\int_{\mathbb{R}} \partial_x J_\delta(x-y)\partial_x\sigma(x)\sigma(y)\xi(y)\,\mathrm{d}y, \tag{7.7}$$
$$\mathbf{j}_\delta\mathbf{\Sigma}\mathbf{\Sigma}\xi(x) = \int_{\mathbb{R}} \partial_{xx}^2 J_\delta(x-y)\sigma^2(y)\xi(y)\,\mathrm{d}y \tag{7.8}$$
$$-\int_{\mathbb{R}} \partial_x J_\delta(x-y)\sigma(y)\partial_y\sigma(y)\xi(y)\,\mathrm{d}y, \tag{7.9}$$

and

$$\mathbf{\Sigma}\mathbf{\Sigma}\mathbf{j}_\delta\xi(x) = \int_{\mathbb{R}} J_\delta(x-y)\partial_x\big(\sigma(x)\partial_x\sigma(x)\big)\xi(y)\,\mathrm{d}y \tag{7.10}$$
$$+\, 3\int_{\mathbb{R}} \partial_x J_\delta(x-y)\sigma(x)\partial_x\sigma(x)\xi(y)\,\mathrm{d}y \tag{7.11}$$
$$+\int_{\mathbb{R}} \partial_{xx}^2 J_\delta(x-y)\sigma^2(x)\xi(y)\,\mathrm{d}y. \tag{7.12}$$

We will estimate (7.6) to (7.12) by considering the sums

$$\mathfrak{I}_1 := (7.7) - (7.9) - (7.11), \quad \mathfrak{I}_2 := (7.6) - (7.8) - (7.12),$$

and the stand-alone integral (7.10), where, from (7.5), we see that

$$-R_\delta = \Big[\big[\mathbf{\Sigma}, \mathbf{j}_\delta\big], \mathbf{\Sigma}\Big](\xi) = \mathfrak{I}_1 + \mathfrak{I}_2 - (7.10). \tag{7.13}$$

We will use [25, Lemma II.1] to establish that (7.13) tends to zero in an appropriate sense. Estimating the terms in (7.13) separately, we have

$$\begin{aligned}
\|\mathfrak{I}_1\|_{L^2(\mathbb{R})} &= \Big\|\int_{\mathbb{R}} \partial_x J_\delta(\cdot - y) \\
&\quad \times \Big(2\sigma(y)\partial_x\sigma(\cdot) + \sigma(y)\partial_y\sigma(y) - 3\sigma(\cdot)\partial_x\sigma(\cdot)\Big)\xi(y)\,\mathrm{d}y\Big\|_{L^2(\mathbb{R})} \\
&= \Big\|\int_{\mathbb{R}} \partial_x J_\delta(\cdot - y) \\
&\quad \times \Big(2\big(\sigma(y) - \sigma(\cdot)\big)\partial_x\sigma(\cdot) + \big(\sigma(y)\partial_y\sigma(y) - \sigma(\cdot)\partial_x\sigma(\cdot)\big)\Big)\xi(y)\,\mathrm{d}y\Big\|_{L^2(\mathbb{R})} \\
&\le \Big\|\int_{\mathbb{R}} |\partial_x J_\delta(\cdot - y)| \\
&\quad \times \Big(2\,|\sigma(y) - \sigma(\cdot)|\;|\partial_x\sigma(\cdot)| + |\sigma(y)\partial_y\sigma(y) - \sigma(\cdot)\partial_x\sigma(\cdot)|\Big)\;|\xi(y)|\;\mathrm{d}y\Big\|_{L^2(\mathbb{R})} \\
&\le C\Big\|\int_{\mathbb{R}} |\cdot - y|\,|\partial_x J_\delta(\cdot - y)| \\
&\quad \times \left(2\left|\frac{\sigma(y) - \sigma(\cdot)}{y - \cdot}\partial_x\sigma(\cdot)\right| + \left|\frac{\sigma(y)\partial_y\sigma(y) - \sigma(\cdot)\partial_x\sigma(\cdot)}{y - \cdot}\right|\right)|\xi(y)|\;\mathrm{d}y\Big\|_{L^2(\mathbb{R})}
\end{aligned}$$

$$
\begin{aligned}
&\le C\left(\|\partial_x\sigma\|_{L^\infty(\mathbb{R})}^2 + \left\|\partial_x\left(\sigma\partial_x\sigma\right)\right\|_{L^\infty(\mathbb{R})}\right) \\
&\qquad \times \left\|\int_{\mathbb{R}} |\cdot - y|\,|\partial_x J_\delta(\cdot - y)|\,|\xi(y)|\,\mathrm{d}y\right\|_{L^2(\mathbb{R})} \\
&\le C\Big(\|\partial_x\sigma\|_{L^\infty(\mathbb{R})}^2 + \left\|\partial_x\left(\sigma\partial_x\sigma\right)\right\|_{L^\infty(\mathbb{R})}\Big)\Big\||\cdot|\,\partial_x J_\delta(\cdot)\Big\|_{L^1(\mathbb{R})}\|\xi\|_{L^2(\mathbb{R})} \\
&\le C\Big(\|\partial_x\sigma\|_{L^\infty(\mathbb{R})}^2 + \left\|\partial_x\left(\sigma\partial_x\sigma\right)\right\|_{L^\infty(\mathbb{R})}\Big)\|\xi\|_{L^2(\mathbb{R})},
\end{aligned}
$$

where we have used Young's convolution inequality and subsequently the basic estimate $\big\||\cdot|\,\partial_x J_\delta(\cdot)\big\|_{L^1(\mathbb{R})} \lesssim 1$. Similarly,

$$
\begin{aligned}
\|\mathfrak{I}_2\|_{L^2(\mathbb{R})} &= \left\|\int_{\mathbb{R}} \partial_{xx}^2 J_\delta(\cdot - y)\big(2\sigma(\cdot)\sigma(y) - \sigma^2(\cdot) - \sigma^2(y)\big)\xi(y)\,\mathrm{d}y\right\|_{L^2(\mathbb{R})} \\
&= \left\|\int_{\mathbb{R}} \partial_{xx}^2 J_\delta(\cdot - y)\big(\sigma(\cdot) - \sigma(y)\big)^2\xi(y)\,\mathrm{d}y\right\|_{L^2(\mathbb{R})} \\
&\le \left\|\int_{\mathbb{R}} \left|\partial_{xx}^2 J_\delta(\cdot - y)\right|\left|2\sigma(\cdot)\sigma(y) - \sigma^2(\cdot) - \sigma^2(y)\right|\,|\xi(y)|\,\mathrm{d}y\right\|_{L^2(\mathbb{R})} \\
&\le C\left\|\int_{\mathbb{R}} (\cdot - y)^2\left|\partial_{xx}^2 J_\delta(\cdot - y)\right|\left|\frac{\sigma(\cdot) - \sigma(y)}{\cdot - y}\right|^2|\xi(y)|\,\mathrm{d}y\right\|_{L^2(\mathbb{R})} \\
&\le C\,\|\partial_x\sigma\|_{L^\infty(\mathbb{R})}^2\left\|\int_{\mathbb{R}} (\cdot - y)^2\left|\partial_{xx}^2 J_\delta(\cdot - y)\right||\xi(y)|\,\mathrm{d}y\right\|_{L^2(\mathbb{R})} \\
&\le C\,\|\partial_x\sigma\|_{L^\infty(\mathbb{R})}^2\left\|(\cdot)^2\partial_{xx}^2 J_\delta(\cdot)\right\|_{L^1(\mathbb{R})}\|\xi\|_{L^2(\mathbb{R})} \\
&\le C\,\|\partial_x\sigma\|_{L^\infty(\mathbb{R})}^2\|\xi\|_{L^2(\mathbb{R})}.
\end{aligned}
$$

We also have

$$
\|(7.10)\|_{L^2(\mathbb{R})} \le C\,\|J_\delta\|_{L^1(\mathbb{R})}\,\|\partial_x(\sigma\partial_x\sigma)\|_{L^\infty(\mathbb{R})}\,\|\xi(t)\|_{L^2(\mathbb{R})}.
$$

Given the last three (δ-independent) bounds, it is sufficient to establish convergence of (7.5) under the assumption that σ, ξ are smooth (in x). The general case follows by density using the established bounds. Under this assumption, we have

$$
\begin{aligned}
\mathfrak{I}_2 &= \int_{\mathbb{R}} \partial_{xx}^2 J_\delta(x - y)\big(2\sigma(x)\sigma(y) - \sigma^2(x) - \sigma^2(y)\big)\xi(y)\,\mathrm{d}y \\
&= -2\int_{\mathbb{R}} \partial_{xx}^2 J_\delta(x - y)\frac{(x-y)^2}{2}\left(\frac{\sigma(y) - \sigma(x)}{y - x}\right)^2\xi(y)\,\mathrm{d}y \\
&= -2\big(\partial_x\sigma(x)\big)^2\xi(x)\int_{\mathbb{R}} \frac{z^2}{2}\partial_{zz}^2 J_\delta(z)\,\mathrm{d}z + o_\delta(1),
\end{aligned}
$$

where $\int_{\mathbb{R}} \frac{z^2}{2}\partial_{zz}^2 J_\delta(z)\,\mathrm{d}z = 1$. A similar calculation can be done for $\mathfrak{I}_1$, in which case there is only one derivative on the mollifier and then the calculation can be found in the proof of [25, Lemma II.1]. The limit of (7.10) is standard. Reasoning as in the proof of [25, Lemma II.1], we arrive at

$$
\begin{gathered}
\mathfrak{I}_1 \overset{\delta\downarrow 0}{\to} \partial_x\big(\sigma\partial_x\sigma\big)\xi + 2\big(\partial_x\sigma\big)^2\xi, \quad \mathfrak{I}_2 \overset{\delta\downarrow 0}{\to} -2\big(\partial_x\sigma\big)^2\xi, \\
-(7.10) \overset{\delta\downarrow 0}{\to} -\partial_x\big(\sigma\partial_x\sigma\big)\xi \quad \text{in } L^2(\mathbb{R}), \text{ for } \mathrm{d}\mathbb{P}\otimes\mathrm{d}t\text{-a.e.}
\end{gathered}
$$

Adding these terms together, with reference to (7.13), and using the dominated convergence theorem, we conclude that $-R_\delta \overset{\delta\downarrow 0}{\longrightarrow} 0$ in $L^2(\Omega\times[0,T]\times\mathbb{R})$. □

Proposition 7.4 (Itô–Stratonovich conversion terms and regularisation errors)**.** *Let $S \in C^1(\mathbb{R}) \cap \dot{W}^{2,\infty}(\mathbb{R})$ satisfy $S'(r) = O(r)$ and $\sup_r |S''(r)| < \infty$. Let $\varphi \in C^\infty([0,T] \times \mathbb{S}^1)$. With w, w_δ, E_δ^2, and E_δ^3 defined as in (7.1) of Lemma 7.1, we have*

$$\begin{aligned}\mathbb{E} \int_0^T \bigg| \int_{\mathbb{S}^1} & - \varphi S'(\partial_x w_\delta) \partial_x E_\delta^3 \\ & + \varphi S''(\partial_x w_\delta) \left(\frac{1}{2} \left| \partial_x E_\delta^2 \right|^2 + \partial_x \left(\sigma \, \partial_x w_\delta \right) \, \partial_x E_\delta^2 \right) \, \mathrm{d}x \bigg| \, \mathrm{d}t \xrightarrow{\delta \downarrow 0} 0.\end{aligned} \tag{7.14}$$

Remark 7.5. An almost sure version of this proposition, instead of convergence in $L^1(\Omega)$, and with $\varphi \equiv 1$, was stated with a similar proof in [33, Lemma B.3]. Moreover, the result in [33] was stated with slightly more stringent conditions on S, requiring $|S'(r)| = O(1)$ instead of $|S'(r)| = O(r)$. For the remainder of this paper, we shall use $\varphi \equiv 1$.

Proof. In the following, we continue using the operator notation consisting of $\mathbf{j}_\delta$ and $\mathbf{\Sigma}$ defined in (7.4). The estimate (7.14) takes inspiration from the proof of [40, Prop. 3.4]. However, whereas they considered the commutator between the operators $\tilde{\mathbf{\Sigma}} f := \sigma \partial_x f$ and $\mathbf{j}_\delta f$, we have to consider the analogous question for $\mathbf{\Sigma} f = \partial_x(\sigma f)$ and $\mathbf{j}_\delta$. Insofar as $\partial_x w$ can be any element $\xi \in L^2(\Omega; L^\infty([0,T]; L^2(\mathbb{S}^1)))$ (in fact, even just $\xi \in L^2(\Omega \times [0,T] \times \mathbb{S}^1)$!) for the purpose of the convergence, denote $\partial_x w$ by ξ, and $\partial_x w_\delta$ by ξ_δ, since mollification commutes with (weak) differentiation.

We can express $\partial_x E_3^\delta$ in terms of commutator brackets as follows:

$$\partial_x E_\delta^3(\xi) = \frac{1}{2}\Big(\mathbf{\Sigma}\mathbf{\Sigma}\mathbf{j}_\delta \xi - \mathbf{j}_\delta \mathbf{\Sigma}\mathbf{\Sigma}\xi \Big) = \frac{1}{2}\Big(\mathbf{\Sigma}\big[\mathbf{\Sigma}, \mathbf{j}_\delta\big](\xi) + \big[\mathbf{\Sigma}, \mathbf{j}_\delta\big]\mathbf{\Sigma}(\xi) \Big). \tag{7.15}$$

Similarly, we can write the remaining part of the integrand of (7.14) in the form $\frac{1}{2} S''(\xi_\delta) E_\delta^4$, where

$$E_\delta^4(\xi) := \big(\partial_x(\sigma\xi) * J_\delta\big)^2 - \big(\partial_x(\sigma\xi_\delta)\big)^2 = \big(\mathbf{j}_\delta \mathbf{\Sigma}\xi\big)^2 - \big(\mathbf{\Sigma}\mathbf{j}_\delta \xi\big)^2.$$

Therefore, following the calculations in [40, p. 655],

$$\begin{aligned} -\frac{1}{2} S''(\xi_\delta) E_\delta^4 &= \frac{1}{2} S''(\xi_\delta) \big(\mathbf{\Sigma}\mathbf{j}_\delta \xi - \mathbf{j}_\delta \mathbf{\Sigma}\xi\big)\big(\mathbf{\Sigma}\mathbf{j}_\delta\xi + \mathbf{j}_\delta\mathbf{\Sigma}\xi\big) \\ &= -\frac{1}{2} S''(\xi_\delta)\big(\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\big(\mathbf{\Sigma}\mathbf{j}_\delta\xi\big)\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi) \\ &= -\frac{1}{2} S''(\xi_\delta)\big(\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\partial_x\sigma\xi_\delta\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi) \\ &\quad + \sigma\partial_x(S'(\xi_\delta))\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi) \\ &= -\frac{1}{2} S''(\xi_\delta)\big(\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\partial_x\sigma\xi_\delta\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi) \\ &\quad + \partial_x\big(\sigma S'(\xi_\delta)\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big) - S'(\xi_\delta)\partial_x\big(\sigma\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big) \\ &= -\frac{1}{2} S''(\xi_\delta)\big(\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\partial_x\sigma\xi_\delta\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi) \\ &\quad + \partial_x\big(\sigma S'(\xi_\delta)\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big) - S'(\xi_\delta)\mathbf{\Sigma}\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi), \end{aligned}$$

by invoking the definition of $\mathbf{\Sigma}$. Adding this to (7.15), we find that

$$\begin{aligned} &-\frac{1}{2} S''(\xi_\delta) E_\delta^4 + S'(\xi_\delta)\,\partial_x E_\delta^3 \\ &\quad = -\frac{1}{2} S''(\xi_\delta)\big(\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\partial_x\sigma\xi_\delta\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi) + \partial_x\big(\sigma S'(\xi_\delta)\big[\mathbf{\Sigma},\mathbf{j}_\delta\big](\xi)\big) \end{aligned}$$

$$
\begin{aligned}
&- S'(\xi_\delta)\boldsymbol{\Sigma}\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi) + \frac{1}{2}S'(\xi_\delta)\big(\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\boldsymbol{\Sigma}\xi) - \boldsymbol{\Sigma}\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big)\\
&= -\frac{1}{2}S''(\xi_\delta)\big(\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\partial_x\sigma\xi_\delta\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\\
&\quad + \partial_x\big(\sigma S'(\xi_\delta)\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big) + \frac{1}{2}S'(\xi_\delta)\big(\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\boldsymbol{\Sigma}\xi) - \boldsymbol{\Sigma}\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big)\\
&= -\frac{1}{2}S''(\xi_\delta)\big(\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big)^2 + S''(\xi_\delta)\partial_x\sigma\xi_\delta\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\\
&\quad + \partial_x\big(\sigma S'(\xi_\delta)\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big) + \frac{1}{2}S'(\xi_\delta)\Big[\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big],\boldsymbol{\Sigma}\Big](\xi).
\end{aligned}
$$

For the term $\partial_x\big(\sigma S'(\xi_\delta)\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi)\big)$, we integrate- by-parts in x against φ. We know already that $\big[\boldsymbol{\Sigma},\mathbf{j}_\delta\big](\xi) = \partial_x E_\delta^2 \xrightarrow{\delta\downarrow 0} 0$ in $L^2(\Omega\times[0,T]\times\mathbb{S}^1)$, cf. Lemma 7.1. (So in fact cancellation only occurs between the $S'(\xi_\delta)\partial_x E_\delta^3$ and the $S''(\xi_\delta)\partial_x(\sigma\xi_\delta)\,\partial_x E_\delta^2$ terms.) Convergence of the double commutator bracket is established in Lemma 7.2. Now the entire claim (7.14) follows, as $S\in C^1(\mathbb{R})\cap \dot W^{2,\infty}(\mathbb{R})$ and $S'(r)=O(r)$, so that $S'(\xi_\delta)\in L^2(\Omega\times[0,T]\times\mathbb{S}^1)$ and $S''(\xi_\delta)\in L^\infty(\Omega\times[0,T]\times\mathbb{S}^1)$. ☐

7.2. **Pathwise uniqueness.** To quickly establish pathwise uniqueness of solutions in the energy space $L^2([0,T];H^1(\mathbb{S}^1))$, we would need bounds on the solution in $L^2\left([0,T];L^\infty(\Omega;W^{1,\infty}(\mathbb{S}^1))\right)$ to control exponential moments of cubic terms that appear in the exponent resulting from a Gronwall inequality. Unfortunately, such bounds are not available unless T is replaced by a stopping time $\eta_R<T$ that converges a.s. to T as $R\to\infty$. However, integrating up to a stopping time, it is not possible to interchange the expectation (integral w.r.t. $\mathrm{d}\mathbb{P}$) with the temporal integral and appeal to a standard Gronwall inequality. We will therefore rely on the stochastic Gronwall inequalities, see Lemmas A.1 and A.2. Having shown uniqueness on $\big[0,\eta_R\big]$, we send $R\to\infty$ to conclude uniqueness on $[0,T]$.

Theorem 7.6 (Pathwise uniqueness in H^1)**.** *Let u, v be strong H^1 solutions to the viscous stochastic Camassa–Holm equation* (1.1) *with $\sigma\in W^{2,\infty}(\mathbb{S}^1)$ and intial condition $u_0\in L^{p_0}(\Omega;H^1(\mathbb{S}^1))$ for some $p_0>4$. Then $\mathbb{E}\,\|u-v\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}=0$.*

Proof. Suppose u and v are strong solutions defined relative to the (same) stochastic basis $\big(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\ge 0},\mathbb{P}\big)$ and Brownian motion W. The difference $w=u-v$ obeys

$$
\begin{aligned}
0 &= \mathrm{d}w - \varepsilon\partial_x^2 w\,\mathrm{d}t + (u\,\partial_x u - v\,\partial_x v)\;\mathrm{d}t\\
&\quad + \partial_x K * \left(u^2 - v^2 + \frac{1}{2}\left((\partial_x u)^2 - (\partial_x v)^2\right)\right)\,\mathrm{d}t\\
&\quad - \frac{1}{2}\sigma\,\partial_x\left(\sigma\,\partial_x w\right)\,\mathrm{d}t + \sigma\,\partial_x w\,\mathrm{d}W.
\end{aligned}
$$

The spatial derivative satisfies

$$
\begin{aligned}
0 &= \mathrm{d}\partial_x w - \varepsilon\partial_x^3 w\,\mathrm{d}t + \partial_x\left(u\,\partial_x u - v\,\partial_x v\right)\,\mathrm{d}t\\
&\quad + \partial_x^2 K * \left(u^2 - v^2 + \frac{1}{2}\left((\partial_x u)^2 - (\partial_x v)^2\right)\right)\,\mathrm{d}t\\
&\quad - \frac{1}{2}\partial_x\left(\sigma\,\partial_x\left(\sigma\,\partial_x w\right)\right)\,\mathrm{d}t + \partial_x\left(\sigma\,\partial_x w\right)\,\mathrm{d}W.
\end{aligned}
$$

Recall that J_δ is a standard Friedrichs mollifier and $w_\delta = w * J_\delta$. We convolve both the foregoing equations against J_δ in order to obtain SPDEs that can be understood in the pointwise sense (in x):

$$
\begin{aligned}
0 &= \mathrm{d}w_\delta - \varepsilon\partial_x^2 w_\delta \,\mathrm{d}t + (u_\delta\,\partial_x u_\delta - v_\delta\,\partial_x v_\delta)\,\mathrm{d}t \\
&\quad + \partial_x K * \left(u_\delta^2 - v_\delta^2 + \frac{1}{2}\left((\partial_x u_\delta)^2 - (\partial_x v_\delta)^2\right)\right)\mathrm{d}t \\
&\quad - \frac{1}{2}\sigma\,\partial_x(\sigma\,\partial_x w_\delta)\,\mathrm{d}t + \sigma\,\partial_x w_\delta\,\mathrm{d}W + E_\delta^1\,\mathrm{d}t + E_\delta^2\,\mathrm{d}W + E_\delta^3\,\mathrm{d}t, \\
0 &= \mathrm{d}\partial_x w_\delta - \varepsilon\partial_x^3 w_\delta\,\mathrm{d}t + \partial_x(u_\delta\,\partial_x u_\delta - v_\delta\,\partial_x v_\delta)\,\mathrm{d}t \\
&\quad + \partial_x^2 K * \left(u_\delta^2 - v_\delta^2 + \frac{1}{2}\left((\partial_x u_\delta)^2 - (\partial_x v_\delta)^2\right)\right)\mathrm{d}t \\
&\quad - \frac{1}{2}\partial_x(\sigma\,\partial_x(\sigma\,\partial_x w_\delta))\,\mathrm{d}t + \partial_x(\sigma\,\partial_x w_\delta)\,\mathrm{d}W \\
&\quad + \partial_x E_\delta^1\,\mathrm{d}t + \partial_x E_\delta^2\,\mathrm{d}W + \partial_x E_\delta^3\,\mathrm{d}t,
\end{aligned}
\tag{7.16}
$$

where E_δ^1, E_δ^2 and E_δ^3 are as in (7.1).

Apart from the technical addition of the commutator terms E_δ^i, $i = 1,2,3$, uniqueness follows from a straightforward calculation. The quantities u_δ, v_δ, and w_δ are necessarily $\tilde{\mathbb{P}}$-almost surely in $C([0,T];H^1(\mathbb{S}^1))$ by the inclusion of u, v, and consequently w in $C([0,T];H^1_w(\mathbb{S}^1))$. As in deriving the energy inequality of Proposition 4.1 for the Galerkin approximations, repeated applications of the (finite-dimensional) Itô formula gives

$$
\begin{aligned}
&\frac{1}{2}\|w_\delta(t)\|_{H^1(\mathbb{S}^1)}^2 + \varepsilon\int_0^t \|\partial_x w_\delta(s)\|_{H^1(\mathbb{S}^1)}^2\,\mathrm{d}s \\
&= -\int_0^t\int_{\mathbb{S}^1}\left[w_\delta(u_\delta\,\partial_x u_\delta - v_\delta\,\partial_x v_\delta) + \partial_x w_\delta\,\partial_x(u_\delta\,\partial_x u_\delta - v_\delta\,\partial_x v_\delta)\right]\mathrm{d}x\,\mathrm{d}s \\
&\quad - \int_0^t\int_{\mathbb{S}^1}\left[w_\delta\,\partial_x K * \left(u_\delta^2 - v_\delta^2 + \frac{1}{2}\left((\partial_x u_\delta)^2 - (\partial_x v_\delta)^2\right)\right)\right. \\
&\qquad\qquad \left. + \partial_x w_\delta\,\partial_x^2 K * \left(u_\delta^2 - v_\delta^2 + \frac{1}{2}\left((\partial_x u_\delta)^2 - (\partial_x v_\delta)^2\right)\right)\right]\mathrm{d}x\,\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\left[\sigma w_\delta\,\partial_x(\sigma\,\partial_x w_\delta) + \partial_x w_\delta\,\partial_x(\sigma\,\partial_x(\sigma\,\partial_x w_\delta))\right]\mathrm{d}x\,\mathrm{d}s \\
&\quad - \int_0^t\int_{\mathbb{S}^1}\left[w_\delta\big(E_\delta^1 + E_\delta^3\big) + \partial_x w_\delta\left(\partial_x E_\delta^1 + \partial_x E_\delta^3\right)\right]\mathrm{d}x\,\mathrm{d}s \\
&\quad + \frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\left[\left(\sigma\,\partial_x w_\delta + E_\delta^2\right)^2 + \left(\partial_x(\sigma\,\partial_x w_\delta) + \partial_x E_\delta^2\right)^2\right]\mathrm{d}x\,\mathrm{d}s \\
&\quad + \int_0^t\int_{\mathbb{S}^1}\left[\sigma\,w_\delta\,\partial_x w_\delta + \partial_x w_\delta\,\partial_x(\sigma\,\partial_x w_\delta)\right]\mathrm{d}x\,\mathrm{d}W \\
&\quad + \int_0^t\int_{\mathbb{S}^1}\left[w_\delta E_\delta^2 + \partial_x w_\delta\,\partial_x E_\delta^2\right]\mathrm{d}x\,\mathrm{d}W \\
&=: I_1^\delta + I_2^\delta + I_3^\delta + I_4^\delta + I_5^\delta + M_1^\delta + M_2^\delta,
\end{aligned}
\tag{7.17}
$$

recalling that $\|w_\delta(0)\|_{H^1(\mathbb{S}^1)}^2 = 0$. We split the remaining analysis into two parts — one for I_1^δ and I_2^δ, consisting of the mollified terms from the "deterministic" part of the equation, and another for the remaining integrals consisting of the effects of the convective noise and all the mollification error terms.

1. Estimating I_1^δ and I_2^δ.

For I_1^δ we have

$$\begin{aligned}\left|I_1^\delta\right| &= \frac{1}{2}\left|\int_0^t \int_{\mathbb{S}^1} \left[w_\delta\, \partial_x \left(u_\delta^2 - v_\delta^2\right) + \partial_x w_\delta\, \partial_x^2 \left(u_\delta^2 - v_\delta^2\right)\right] \mathrm{d}x\, \mathrm{d}s\right| \\ &= \frac{1}{2}\left|\int_0^t \int_{\mathbb{S}^1} \left[\partial_x w_\delta\, w_\delta \left(u_\delta + v_\delta\right) + \partial_x^2 w_\delta\, \partial_x \left(w_\delta \left(u_\delta + v_\delta\right)\right)\right] \mathrm{d}x\, \mathrm{d}s\right| \\ &\le \frac{1}{2}\int_0^t \Big[\left\|\partial_x w_\delta\right\|_{L^2(\mathbb{S}^1)} \left\|w_\delta\right\|_{L^2(\mathbb{S}^1)} \left\|u_\delta + v_\delta\right\|_{L^\infty(\mathbb{S}^1)} \\ &\qquad\qquad + \left\|\partial_x^2 w_\delta\right\|_{L^2(\mathbb{S}^1)} \left\|w_\delta\right\|_{H^1(\mathbb{S}^1)} \left\|u_\delta + v_\delta\right\|_{W^{1,\infty}(\mathbb{S}^1)} \Big] \mathrm{d}s \\ &\le \frac{1}{2}\int_0^t \Big[\left\|w_\delta\right\|_{H^1(\mathbb{S}^1)}^2 \left\|u_\delta + v_\delta\right\|_{L^\infty(\mathbb{S}^1)} \\ &\qquad\qquad + \frac{\varepsilon}{2} \left\|\partial_x^2 w_\delta\right\|_{L^2(\mathbb{S}^1)}^2 + \frac{1}{2\varepsilon} \left\|w_\delta\right\|_{H^1(\mathbb{S}^1)}^2 \left\|u_\delta + v_\delta\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2 \Big] \mathrm{d}s.\end{aligned}$$

Using the identity $\left(K - \partial_x^2 K\right) * f = f$,

$$\begin{aligned}\left|I_2^\delta\right| &= \left|\int_0^t \int_{\mathbb{S}^1} \partial_x w_\delta \left(u_\delta^2 - v_\delta^2 + \frac{1}{2}\left(\left(\partial_x u_\delta\right)^2 - \left(\partial_x v_\delta\right)^2\right)\right) \mathrm{d}x\, \mathrm{d}s\right| \\ &= \left|\int_0^t \int_{\mathbb{S}^1} \partial_x w_\delta \left(w_\delta \left(u_\delta + v_\delta\right) + \frac{1}{2}\partial_x w_\delta\, \partial_x \left(u_\delta + v_\delta\right)\right) \mathrm{d}x\, \mathrm{d}s\right| \\ &= \left|\int_0^t \int_{\mathbb{S}^1} \left[w_\delta\, \partial_x w_\delta \left(u_\delta + v_\delta\right) - \partial_x w_\delta\, \partial_x^2 w_\delta \left(u_\delta + v_\delta\right)\right] \mathrm{d}x\, \mathrm{d}s\right| \\ &\le \int_0^t \Big[\left\|w_\delta\right\|_{H^1(\mathbb{S}^1)}^2 \left\|u_\delta + v_\delta\right\|_{L^\infty(\mathbb{S}^1)} \\ &\qquad\qquad + \frac{\varepsilon}{4} \left\|\partial_x^2 w_\delta\right\|_{L^2(\mathbb{S}^1)}^2 + \frac{1}{4\varepsilon} \left\|\partial_x w_\delta\right\|_{L^2(\mathbb{S}^1)}^2 \left\|u_\delta + v_\delta\right\|_{L^\infty(\mathbb{S}^1)}^2 \Big] \mathrm{d}s.\end{aligned}$$

In the right-hand sides of $\left|I_1^\delta\right|$ and $\left|I_2^\delta\right|$, the terms involving $\varepsilon \int_0^t \left\|\partial_x^2 w_\delta\right\|_{L^2(\mathbb{S}^1)}^2 \mathrm{d}s$ can be absorbed into (7.17).

2. Estimating $I_3^\delta + I_4^\delta + I_5^\delta$.

Next, we have

$$\begin{aligned}I_3^\delta &= -\frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \left|\sigma\, \partial_x w_\delta\right|^2 \mathrm{d}x\, \mathrm{d}s - \frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \sigma\, \partial_x \sigma\, w_\delta\, \partial_x w_\delta\, \mathrm{d}x\, \mathrm{d}s \\ &\quad - \frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \left|\sigma \partial_x^2 w_\delta\right|^2 \mathrm{d}x\, \mathrm{d}s - \frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \sigma\, \partial_x \sigma\, \partial_x w_\delta\, \partial_x^2 w_\delta\, \mathrm{d}x\, \mathrm{d}s \\ &= -\frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \left[\left|\sigma\, \partial_x w_\delta\right|^2 + \left|\sigma \partial_x^2 w_\delta\right|^2\right] \mathrm{d}x\, \mathrm{d}s - \frac{1}{2}\int_0^t \int_{\mathbb{S}^1} \sigma\, \partial_x \sigma\, w_\delta\, \partial_x w_\delta\, \mathrm{d}x\, \mathrm{d}s \\ &\quad + \frac{1}{4}\int_0^t \int_{\mathbb{S}^1} \partial_x \left(\sigma\, \partial_x \sigma\right) \left|\partial_x w_\delta\right|^2 \mathrm{d}x\, \mathrm{d}s.\end{aligned}$$

We add I_4^δ and I_5^δ together and use Lemma 7.1 and Proposition 7.4, yielding

$$I_4^\delta + I_5^\delta = \int_0^t \int_{\mathbb{S}^1} \left[- w_\delta \left(E_\delta^1 + E_\delta^3\right) + \frac{1}{2}\left(\sigma \partial_x w_\delta + E_\delta^2\right)^2 \right] \mathrm{d}x\, \mathrm{d}s$$

$$
\begin{aligned}
&+\int_0^t\int_{\mathbb{S}^1}\left[-\partial_x w_\delta\,\partial_x E_\delta^1+\frac{1}{2}\left|\partial_x\left(\sigma\,\partial_x w_\delta\right)\right|^2\right]\mathrm{d}x\,\mathrm{d}s\\
&+\int_0^t\int_{\mathbb{S}^1}\left[-\partial_x w_\delta\,\partial_x E_\delta^3+\frac{1}{2}\left|\partial_x E_\delta^2\right|^2+\partial_x\left(\sigma\,\partial_x w_\delta\right)\,\partial_x E_\delta^2\right]\mathrm{d}x\,\mathrm{d}s\\
=:{}& I_{4+5,1}^\delta+I_{4+5,2}^\delta+\int_0^t I_{4+5,3}^\delta\,\mathrm{d}s.
\end{aligned}
$$

For $I_{4+5,1}^\delta$, we have

$$
\begin{aligned}
\left|I_{4+5,1}^\delta\right| &\le \|w_\delta\|_{L^2([0,t]\times\mathbb{S}^1)}^2+\left\|E_\delta^1+E_\delta^3\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2\\
&\quad+\|\sigma\|_{L^\infty(\mathbb{S}^1)}^2\|\partial_x w_\delta\|_{L^2([0,t]\times\mathbb{S}^1)}^2+\left\|E_\delta^2\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2\\
&\le C_\sigma\|w_\delta\|_{L^2([0,t];H^1(\mathbb{S}^1))}^2+2\left\|E_\delta^1\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2\\
&\quad+\left\|E_\delta^2\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2+2\left\|E_\delta^3\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2.
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
I_{4+5,2}^\delta &= \int_0^t\int_{\mathbb{S}^1}\left[-\partial_x w_\delta\,\partial_x E_\delta^1+\sigma\partial_x\sigma\,\partial_x w_\delta\,\partial_x^2 w_\delta\right]\mathrm{d}x\,\mathrm{d}s\\
&\quad+\frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\left[\left|\sigma\,\partial_x^2 w_\delta\right|^2+\left|\partial_x\sigma\,\partial_x w_\delta\right|^2\right]\mathrm{d}x\,\mathrm{d}s\\
&= -\int_0^t\int_{\mathbb{S}^1}\left[\partial_x w_\delta\,\partial_x E_\delta^1+\frac{1}{2}\,\partial_x\left(\sigma\partial_x\sigma\right)\,\left|\partial_x w_\delta\right|^2\right]\mathrm{d}x\,\mathrm{d}s\\
&\quad+\frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\left[\left|\sigma\,\partial_x^2 w_\delta\right|^2+\left|\partial_x\sigma\,\partial_x w_\delta\right|^2\right]\mathrm{d}x\,\mathrm{d}s.
\end{aligned}
$$

Adding $I_{4+5,2}^\delta$ to I_3^δ, we obtain

$$
\begin{aligned}
I_3^\delta+I_{4+5,2}^\delta &= -\frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\left|\sigma\,\partial_x w_\delta\right|^2\,\mathrm{d}x\,\mathrm{d}s-\frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\sigma\,\partial_x\sigma\,w_\delta\,\partial_x w_\delta\,\mathrm{d}x\,\mathrm{d}s\\
&\quad+\int_0^t\int_{\mathbb{S}^1}\partial_x^2 w_\delta\,E_\delta^1\,\mathrm{d}x\,\mathrm{d}s+\frac{1}{2}\int_0^t\int_{\mathbb{S}^1}\left|\partial_x\sigma\,\partial_x w_\delta\right|^2\,\mathrm{d}x\,\mathrm{d}s\\
&\quad-\frac{1}{4}\int_0^t\int_{\mathbb{S}^1}\partial_x(\sigma\partial_x\sigma)(\partial_x w_\delta)^2\,\mathrm{d}x\,\mathrm{d}s\\
&\le\left(\left\|\partial_x^2\sigma\right\|_{L^\infty(\mathbb{S}^1)}^2+\|\partial_x\sigma\|_{L^\infty(\mathbb{S}^1)}^2+\|\sigma\|_{L^2(\mathbb{S}^1)}^2+1\right)\|w_\delta\|_{L^2([0,t];H^1(\mathbb{S}^1))}^2\\
&\quad+\frac{\varepsilon}{8}\left\|\partial_x^2 w_\delta\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2+\frac{16}{\varepsilon}\left\|E_\delta^1\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2.
\end{aligned}
$$

The term involving $\varepsilon\left\|\partial_x^2 w_\delta\right\|_{L^2([0,t]\times\mathbb{S}^1)}^2$ can be absorbed into (7.17).

Given Proposition 7.4, taking $S(r)=r^2/2$, we immediately get that

$$
\mathbb{E}\int_0^T\left|I_{4+5,3}^\delta\right|\,\mathrm{d}t\xrightarrow{\delta\downarrow 0}0.
$$

Therefore, by Lemma 7.1,

$$
I_3^\delta+I_4^\delta+I_5^\delta\le C_\sigma\|w_\delta\|_{L^2([0,t];H^1(\mathbb{S}^1))}^2+\varrho_\delta(t),
$$

where $\varrho_\delta(t)\ge 0$ and $\sup_{t\in[0,T]}\mathbb{E}\,\varrho_\delta(t)\xrightarrow{\delta\downarrow 0}0$.

3. Conclusion.

Putting the estimates for I_1^δ through I_5^δ together we arrive at

$$\begin{aligned}\frac{1}{2}\left\|w_\delta(t)\right\|_{H^1(\mathbb{S}^1)}^2+\frac{\varepsilon}{4}\int_0^t\left\|\partial_x w_\delta(s)\right\|_{H^1(\mathbb{S}^1)}^2\,\mathrm{d}s&\\ \le C_{\varepsilon,\sigma}\int_0^t\left\|w_\delta\right\|_{H^1(\mathbb{S}^1)}^2&\left(1+\left\|u_\delta+v_\delta\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\right)\mathrm{d}s\\ &+M_\delta(t)+\varrho_\delta(t),\end{aligned}\tag{7.18}$$

where $M_\delta(t):=M_1^\delta(t)+M_2^\delta(t)$.

The estimates on I_1^δ through I_5^δ also show, by the equality (7.17), that the process $t\mapsto\left\|w_\delta(t)\right\|_{H^1(\mathbb{S}^1)}^2$ is $\mathbb{P}$-almost surely continuous.

By Young's convolution inequality,

$$\int_0^t\left\|u_\delta(s)+v_\delta(s)\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\,\mathrm{d}s\le\int_0^t\left\|J_\delta\right\|_{L^1(\mathbb{S}^1)}^2\left\|u(s)+v(s)\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\,\mathrm{d}s,$$

and $\left\|J_\delta\right\|_{L^1(\mathbb{S}^1)}=1$ by construction. Let us therefore introduce the stopping time

$$\eta_R=\inf\left\{t\in\mathbb{R}_+:\int_0^{t\wedge T}\left\|u(s)+v(s)\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\,\mathrm{d}s>R\right\},$$

with $\eta_R=\infty$ if the set on the right-hand side is empty. We have that $\eta_R\overset{R\uparrow\infty}{\longrightarrow}T$ a.s. (for fixed $\varepsilon>0$). Indeed, from part (c) of Definition 2.1 — which says that $u,v\in L^2_\omega L^2_t H^2_x$ — and the embedding $H^2(\mathbb{S}^1)\hookrightarrow W^{1,\infty}(\mathbb{S}^1)$,

$$\begin{aligned}\mathbb{P}\left(\{\eta_R<T\}\right)&\le\mathbb{P}\left(\left\{\int_0^T\left\|u(t)+v(t)\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\,\mathrm{d}t>R\right\}\right)\\ &\le\frac{1}{R}\mathbb{E}\int_0^T\left\|u(t)+v(t)\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\,\mathrm{d}t\\ &\le\frac{\tilde{C}}{R}\mathbb{E}\int_0^T\left\|u(t)\right\|_{H^2(\mathbb{S}^1)}^2+\left\|v(t)\right\|_{H^2(\mathbb{S}^1)}^2\,\mathrm{d}t\le\frac{C}{R}\overset{R\uparrow\infty}{\longrightarrow}0,\end{aligned}\tag{7.19}$$

where C depends on ε, cf. (4.3).

With this stopping time, $M_\delta(t\wedge\eta_R)$ is a square-integrable martingale term.

Specifying $t=\eta_R$ in (7.18), noting that

$$\int_0^{\eta_R}1+\left\|u_\delta(s)+v_\delta(s)\right\|_{W^{1,\infty}(\mathbb{S}^1)}^2\,\mathrm{d}s\le T+R,$$

we can use the stochastic Gronwall inequality (Lemma A.2 with $\nu=1/2$ and any $1/2<r<1$) to conclude that

$$\lim_{\delta\to0}\left(\mathbb{E}\sup_{t\in[0,\eta_R]}\left\|w_\delta(t)\right\|_{H^1(\mathbb{S}^1)}\right)^2=0.\tag{7.20}$$

Recalling that the stopping times $\eta_R\overset{R\uparrow\infty}{\longrightarrow}T$ are independent of δ and, by the properties of mollification, $w_\delta(t)\to w(t)$ in $H^1(\mathbb{S}^1)$ for $\mathrm{d}\mathbb{P}\otimes\mathrm{d}t$-a.e. $(\omega,t)\in\Omega\times[0,T]$, combining (7.20) with the dominated convergence theorem implies that

$$\mathbb{E}\left\|w\right\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}=0.$$

□

7.3. **Strong H^1-existence.** To establish the existence of strong H^1 solutions, and thereby concluding the proof of Theorem 1.1, we shall use an infinite dimensional version of the Yamada–Watanabe principle, following from Lemma A.4. As the path space $\mathcal{X}$ constructed immediately following the definitions (6.1) is not a Polish space, we provide a slightly refined argument.

Concluding the proof of Theorem 1.1. Recalling (6.1), we consider the extended path space $\mathcal{Y} := (\mathcal{X}_{u,s} \times \mathcal{X}_{u,w}) \times (\mathcal{X}_{u,s} \times \mathcal{X}_{u,w}) \times \mathcal{X}_W \times \mathcal{X}_0$. Let $\{u_n\}$ be the Galerkin solutions with initial conditions $\{\mathbf{\Pi}_n u_0\}$, cf. (3.1). Set

$$\mu^n_{u,s} := (u_n : \Omega \to \mathcal{X}_{u,s})_* \mathbb{P}, \quad \mu^n_{u,w} := (u_n : \Omega \to \mathcal{X}_{u,w})_* \mathbb{P},$$
$$\mu^n_W := (W)_* \mathbb{P}, \quad \mu^n_0 := (\mathbf{\Pi}_n u_0)_* \mathbb{P},$$

as probability measures respectively on $\mathcal{X}_{u,s}$, $\mathcal{X}_{u,w}$, $\mathcal{X}_W$, and $\mathcal{X}_0$. Finally, define on $\mathcal{Y}$ the product measure

$$\mu^{m,n} := \mu^m_{u,s} \otimes \mu^m_{u,w} \otimes \mu^n_{u,s} \otimes \mu^n_{u,w} \otimes \mu^m_W \otimes \mu^m_0.$$

Consider an arbitrary subsequence $\{\mu^{m_k,n_k}\}_{k\in\mathbb{N}}$ so that $\{m_k\}_{k\in\mathbb{N}}$ and $\{n_k\}_{k\in\mathbb{N}}$ are increasing sequences. The tightness of $\{\mu^n_j\}$ in $\mathcal{X}_j$, taking j equal to (u,s), (u,w), W, 0, respectively, see Lemma 6.1, implies the tightness of $\{\mu^{m_k,n_k}\}_{k\in\mathbb{N}}$ on $\mathcal{Y}$. By Prohorov's theorem, this subsequence converges weakly to a probability measure $\tilde{\mu}$ on $\mathcal{Y}$.

By the Skorokhod–Jakubowski representation theorem (cf. Theorem 6.2) (and the identification of Lemma 6.4), there exist a new probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and, passing to a further subsequence (not relabelled), new random variables

$$\big(\tilde{u}_{m_k}, \tilde{u}_{m_k}, \tilde{u}_{n_k}, \tilde{u}_{n_k}, \tilde{W}, \tilde{u}_{0,m_k}\big), \quad \text{with joint laws } \mu^{m_k,n_k}, \tag{7.21}$$

converging in $\mathcal{Y}$ to a limit $\big(\tilde{u}^\alpha, \tilde{u}^\alpha, \tilde{u}^\beta, \tilde{u}^\beta, \tilde{W}, \tilde{u}_0\big)$, $\tilde{\mathbb{P}}$-a.s., whose joint law is $\tilde{\mu}$.

Construct now a (filtered) stochastic basis $\tilde{\mathcal{S}}$ as in the paragraph following Theorem 6.2. It then follows (as in Theorem 6.8) that $\big(\tilde{u}^\alpha, \tilde{W}\big)$ and $\big(\tilde{u}^\beta, \tilde{W}\big)$ are weak (martingale) H^1 solutions with initial condition $\tilde{u}_0$ on $\tilde{\mathcal{S}}$. Therefore, by pathwise uniqueness (cf. Theorem 7.6),

$$\tilde{\mu}\Big((u,u,v,v) \in \mathcal{X}_{u,s} \times \mathcal{X}_{u,w} \times \mathcal{X}_{u,s} \times \mathcal{X}_{u,w} : u = v\Big)$$
$$= \tilde{\mathbb{P}}\Big(\tilde{u}^\alpha = \tilde{u}^\beta \text{ in } \mathcal{X}_{u,s} \text{ and in } \mathcal{X}_{u,w}\Big) = 1.$$

We have constructed a pair of weak H^1 solutions that coincide almost surely. To use the Gyöngy–Krylov theorem to conclude convergence on the initial probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a Polish space is needed, whereas our path space $\mathcal{Y}$ is only quasi-Polish. We can, however, map $\mathcal{Y}$ into the Polish space $[-1,1]^{\mathbb{N}}$. This is the technique Jakubowski used to extend the Skorokhod representation theorem to non-metric (quasi-Polish) spaces [35]. Indeed, because $\mathcal{Y}$ is a quasi-Polish space, there is a countable family of continuous functions $\{f_\ell : \mathcal{Y} \to [-1,1]\}_{\ell\in\mathbb{N}}$ that separate points, see (A.2). Introduce the continuous map $f : \mathcal{Y} \to [-1,1]^{\mathbb{N}}$ (equipped with the product topology) by $f : u \mapsto \{f_\ell(u)\}_{\ell\in\mathbb{N}}$. The map f is a measurable bijective function when restricted to a σ-compact subspace of $\mathcal{Y}$ (i.e., a countable union of compact subspaces) of $\mathcal{Y}$, see [35, Sec. 2] and [26, Cor. 3.1.14, p. 126].

Considering the first and third entries (u_{m_k}, u_{n_k}) of (7.21), also recalling their limits $\big(u^\alpha, u^\alpha\big)$, and using f defined in the foregoing paragraph, we find by continuity of f that $\big(f(u_{m_k}), f(u_{n_k})\big)$ converge in distribution (law) to $(f(u^\alpha), f(u^\alpha))$

as $k \to \infty$. By Lemma A.4, there is a subsequence $\left\{f(u_{n_{k_j}})\right\}_{j\in\mathbb{N}}$ that converges in probability. Since f separates points of $\mathcal{Y}$, we must necessarily have that also $\left\{u_{n_{k_j}}\right\}_{j\in\mathbb{N}}$ converges in probability on $\left(\Omega, \mathcal{F}, \mathbb{P}\right)$, and hence $\mathbb{P}$-almost surely along a further subsequence. □

7.4. **Strong temporal continuity.** In this subsection, we finally establish the strong continuity of solutions as stipulated in (c) of Definition 2.1, completing the omission described in Remark 6.9.

Lemma 7.7 (Energy equality)**.** *Let u be the solution found in Theorem 6.8, and q be the weak x-derivative of u. The following energy equality holds for every $s, t \in [0,T]$:*

$$\begin{aligned}\mathbb{E}\left\|u(r)\right\|_{H^1(\mathbb{S}^1)}^2\Big|_s^t &+ \varepsilon\mathbb{E}\int_s^t \left\|q(r)\right\|_{H^1(\mathbb{S}^1)}^2\,\mathrm{d}r \\ &= -\mathbb{E}\int_s^t\int_{\mathbb{S}^1}\sigma\,\partial_x\sigma\,u\,q\,\mathrm{d}x\,dr + \mathbb{E}\int_s^t\int_{\mathbb{S}^1}\left(\frac{1}{4}\partial_x^2\sigma^2 - \left|\partial_x\sigma\right|^2\right)q^2\,\mathrm{d}x\,\mathrm{d}r.\end{aligned} \tag{7.22}$$

Proof. We can derive the energy inequality by mollification as in (7.16) (e.g., by taking $v_0, v \equiv 0$, which is clearly a solution). That is, we have

$$\begin{aligned}0 = \mathrm{d}u_\delta - \varepsilon\partial_x^2 u_\delta + \left(u_\delta\,\partial_x u_\delta\right)\mathrm{d}t + \partial_x K * \left(u_\delta^2 + \frac{1}{2}\left(\partial_x u_\delta\right)^2\right)\mathrm{d}t \\ - \frac{1}{2}\sigma\,\partial_x\left(\sigma\,\partial_x u_\delta\right)\mathrm{d}t + \left(\sigma\,\partial_x u_\delta + E_\delta^2\right)\mathrm{d}W + E_\delta^1\,\mathrm{d}t + E_\delta^3\,\mathrm{d}t,\end{aligned}$$

$$\begin{aligned}0 = \mathrm{d}q_\delta - \varepsilon\partial_x^2 q_\delta + \partial_x\left(u_\delta\,q_\delta\right)\mathrm{d}t + \partial_x^2 K * \left(u_\delta^2 + \frac{1}{2}\left(q_\delta\right)^2\right)\mathrm{d}t \\ - \frac{1}{2}\partial_x\left(\sigma\,\partial_x\left(\sigma\,q_\delta\right)\right)\mathrm{d}t + \left(\partial_x\left(\sigma\,q_\delta\right) + \partial_x E_\delta^2\right)\mathrm{d}W \\ + \partial_x E_\delta^1\,\mathrm{d}t + \partial_x E_\delta^3\,\mathrm{d}t,\end{aligned}$$

where, again,

$$\begin{aligned}E_\delta^1 &:= \left(u\,\partial_x u\right) * J_\delta - u_\delta\,\partial_x u_\delta \\ &\quad + \partial_x K * \left(u^2 + \frac{1}{2}\left(\partial_x u\right)^2\right) * J_\delta - \partial_x K * \left(u_\delta^2 + \frac{1}{2}\left(\partial_x u_\delta\right)^2\right), \\ E_\delta^2 &:= \left(\sigma\,\partial_x u\right) * J_\delta - \sigma\,\partial_x u_\delta, \\ E_\delta^3 &:= -\frac{1}{2}\left(\sigma\,\partial_x\left(\sigma\,\partial_x u\right)\right) * J_\delta + \frac{1}{2}\sigma\,\partial_x\left(\sigma\,\partial_x u_\delta\right).\end{aligned}$$

We multiply the equation for $\mathrm{d}u_\delta$ by u_δ and the equation for $\mathrm{d}q_\delta$ by q_δ. Manipulations using the pointwise Itô formula as in (4.4) of Proposition 4.1 then leads us upon integration to:

$$\frac{1}{2}\left\|u_\delta\right\|_{H^1(\mathbb{S}^1)}^2\Big|_s^t + \int_s^t \varepsilon\left\|q_\delta\right\|_{H^1(\mathbb{S}^1)}^2\,\mathrm{d}r = \sum_{i=1}^5 I_i^\delta + M^\delta,$$

where

$$\begin{aligned}I_1^\delta &:= \frac{1}{2}\int_s^t\int_{\mathbb{S}^1} u_\delta\,\sigma\,\partial_x\left(\sigma\,\partial_x u_\delta\right)\mathrm{d}x\,\mathrm{d}r + \frac{1}{2}\int_s^t\int_{\mathbb{S}^1} q_\delta\partial_x\left(\sigma\,\partial_x\left(\sigma\,q_\delta\right)\right)\mathrm{d}x\,\mathrm{d}r, \\ I_2^\delta &:= -\int_s^t\int_{\mathbb{S}^1} u_\delta\left(E_\delta^1 + E_\delta^3\right)\mathrm{d}x\,\mathrm{d}r - \int_s^t\int_{\mathbb{S}^1} q_\delta\partial_x E_\delta^1\,\mathrm{d}x\,\mathrm{d}r,\end{aligned}$$

$$
\begin{aligned}
I_3^\delta &:= -\int_s^t \int_{\mathbb{S}^1} q_\delta \, \partial_x E_\delta^3 \,\mathrm{d}x\,\mathrm{d}r, \\
I_4^\delta &:= \int_s^t \int_{\mathbb{S}^1} \left[\frac{1}{2}\left|\partial_x E_\delta^2\right|^2 + \partial_x\left(\sigma q_\delta\right)\,\partial_x E_\delta^2\right] \mathrm{d}x\,\mathrm{d}r, \\
I_5^\delta &:= \int_s^t \int_{\mathbb{S}^1} \left[\frac{1}{2}\left|E_\delta^2\right|^2 + \sigma\, q_\delta E_\delta^2\right] \mathrm{d}x\,\mathrm{d}r, \\
I_6^\delta &:= \frac{1}{2}\int_s^t \int_{\mathbb{S}^1} \left[\left|\sigma\, q_\delta\right|^2 + \left|\partial_x\left(\sigma q_\delta\right)\right|^2\right] \mathrm{d}x\,\mathrm{d}r, \\
M^\delta &:= -\int_s^t \int_{\mathbb{S}^1} \Big[\left(\sigma\,\partial_x u_\delta + E_\delta^2\right) + \left(\partial_x\left(\sigma\, q_\delta\right) + \partial_x E_\delta^2\right)\Big] \mathrm{d}x\,\mathrm{d}W.
\end{aligned}
$$

Terms associated with the deterministic CH equation (where σ does not appear) cancel out due to the structure of the equation as in the proof of Proposition 4.1. I_1^δ to I_3^δ arise from the standard chain rule, and I_4^δ to I_6^δ are Itô correction terms. M^δ is a martingale term.

As in the proof of Theorem 7.6, by Lemma 7.1 and Proposition 7.4, as $\delta \downarrow 0$,

$$
\mathbb{E} I_2^\delta,\ \mathbb{E} I_5^\delta \to 0, \qquad \mathbb{E}\left[I_3^\delta + I_4^\delta\right] \to 0.
$$

Adding I_1^δ to I_6^δ, and performing integration-by-parts multiple times,

$$
\begin{aligned}
-I_1^\delta - I_6^\delta &= \frac{1}{2}\int_s^t \int_{\mathbb{S}^1} \partial_x\left(\sigma\, u_\delta\right)\,\sigma\, q_\delta \,\mathrm{d}x\,\mathrm{d}r - \frac{1}{2}\int_s^t \int_{\mathbb{S}^1} \left|\sigma q_\delta\right|^2 \,\mathrm{d}x\,\mathrm{d}r \\
&\quad + \frac{1}{2}\int_s^t \int_{\mathbb{S}^1} \sigma\,\partial_x q_\delta\,\partial_x\left(\sigma\, q_\delta\right) \mathrm{d}x\,\mathrm{d}r - \frac{1}{2}\int_s^t \int_{\mathbb{S}^1} \left|\partial_x\left(\sigma q_\delta\right)\right|^2 \,\mathrm{d}x\,\mathrm{d}r \\
&= \int_s^t \int_{\mathbb{S}^1} \sigma\,\partial_x\sigma\, u_\delta\, q_\delta \,\mathrm{d}x\,dr + \int_s^t \int_{\mathbb{S}^1} \left(\left|\partial_x\sigma\right|^2 - \frac{1}{4}\partial_x^2\sigma^2\right) q_\delta^2 \,\mathrm{d}x\,\mathrm{d}r
\end{aligned}
$$

We need now to take $\delta \to 0$ in $\mathbb{E}[I_1^\delta + I_6^\delta]$. Since $u \in L^2(\Omega \times [0,T]; H^2(\mathbb{S}^1))$, it holds that $u, q \in L^2(\Omega \times [0,T] \times \mathbb{S}^1)$, so by Young's inequality and the dominated convergene theorem,

$$
-\mathbb{E}\left[I_1^\delta + I_6^\delta\right] \xrightarrow{\delta\downarrow 0} \mathbb{E}\int_s^t \int_{\mathbb{S}^1} \sigma\,\partial_x\sigma\, u\, q \,\mathrm{d}x\,dr + \mathbb{E}\int_s^t \int_{\mathbb{S}^1} \left(\left|\partial_x\sigma\right|^2 - \frac{1}{4}\partial_x^2\sigma^2\right) q^2 \,\mathrm{d}x\,\mathrm{d}r.
$$

Finally, $\mathbb{E}M^\delta = 0$, since its quadratic variation satisfies

$$
\mathbb{E}\int_0^T \left|\int_{\mathbb{S}^1} \Big[\left(\sigma\,\partial_x u_\delta + E_\delta^2\right) + \left(\partial_x\left(\sigma\, q_\delta\right) + \partial_x E_\delta^2\right)\Big] \mathrm{d}x\right|^2 \mathrm{d}r < \infty.
$$

On the left-hand side, we can also pass $\delta \to 0$ at every $t \in [0,T]$ using

$$
\mathbb{E}\lim_{\delta\to 0} \left\|u_\delta(t)\right\|_{H^1(\mathbb{S}^1)}^2 = \lim_{\delta\to 0} \mathbb{E}\left\|u_\delta(t)\right\|_{H^1(\mathbb{S}^1)}^2,
$$

(by the Lebesgue dominated convergence theorem) and the energy bound (4.3). The passage in $\delta \to 0$ for the temporal integral $\varepsilon\mathbb{E}\int_s^t \left\|q_\delta(r)\right\|_{H^1(\mathbb{S}^1)}^2 \,\mathrm{d}r$ is similar. We therefore arrive at (7.22). □

Proposition 7.8. *Let u be the solution of Theorem 6.8. For any $t_0 \in (0,T)$,*

$$
\lim_{t\to t_0} \mathbb{E}\left\|u(t) - u(t_0)\right\|_{H^1(\mathbb{S}^1)}^2 = 0,
$$

with corresponding one-sided limits $t \downarrow 0$ and $t \uparrow T$ at the end-points $t_0 = 0$ and $t_0 = T$, respectively. Moreover, for a $p_0 > 4$, $u \in L^{p_0}(\Omega; C([0,T]; H^1(\mathbb{S}^1)))$.

Proof. We can upgrade the weak H^1_x-continuity into strong continuity using the Brezis–Lieb lemma (for L^2). Since we already know that $u \in C([0,T]; L^2(\mathbb{S}^1))$ (see Lemma 5.1 and argue as in Part 6 of the proof to Theorem 6.8), we need only establish temporal continuity for $q = \partial_x u$ in $L^2(\mathbb{S}^1)$.

From $u \in C([0,T]; H^1_w(\mathbb{S}^1))$, $q(t) \rightharpoonup q(s)$ in $L^2(\mathbb{S}^1)$ as $t \to s$. Since $\mathbb{E}\, \|q\|^2_{H^1(\mathbb{S}^1)} \in L^1([0,T])$ (Theorem 6.8), the map $t \mapsto \int_0^t \mathbb{E}\, \|q(r)\|^2_{H^1(\mathbb{S}^1)}\, \mathrm{d}r$ is absolutely continuous on $[0,T]$, and from the energy equality of Lemma 7.7, we obtain

$$\mathbb{E}\, \|u(t)\|^2_{H^1(\mathbb{S}^1)} \overset{t\to t_0}{\longrightarrow} \mathbb{E}\, \|u(t_0)\|^2_{H^1(\mathbb{S}^1)} \qquad \text{for a.e. } t_0 \in [0,T]. \tag{7.23}$$

Finally,

$$\mathbb{E}\, \|u(t) - u(t_0)\|^2_{H^1(\mathbb{S}^1)} = \mathbb{E}\, \|u(t)\|^2_{H^1(\mathbb{S}^1)} - 2\mathbb{E}\langle u(t), u(t_0)\rangle_{H^1(\mathbb{S}^1)} + \mathbb{E}\, \|u(t_0)\|^2_{H^1(\mathbb{S}^1)}.$$

By weak continuity, the middle term tends to $-2\mathbb{E}\, \|u(t_0)\|^2_{H^1(\mathbb{S}^1)}$; together with (7.23), we attain the lemma statement.

Since $\lim_{t\to t_0} \mathbb{E}\, \|u(t) - u(t_0)\|^2_{H^1(\mathbb{S}^1)} = 0$, by Fatou's lemma, we also have

$$\mathbb{E} \lim_{t\to t_0} \|u(t) - u(t_0)\|^2_{H^1(\mathbb{S}^1)} = 0,$$

and therefore $\lim_{t\to t_0} \|u(t) - u(t_0)\|^2_{H^1(\mathbb{S}^1)}$, $\mathbb{P}$-almost surely.

Since $u \in C([0,T]; H^1(\mathbb{S}^1))$, $\mathbb{P}$-almost surely, and for initial conditions $u_0 \in L^{p_0}(\Omega; H^1(\mathbb{S}^1))$, $u \in L^{p_0}(\Omega; L^\infty([0,T]; H^1(\mathbb{S}^1)))$, we can readily conclude that $u \in L^p(\Omega; C([0,T]; H^1(\mathbb{S}^1)))$. This follows from the fact that the $C([0,T]; H^1(\mathbb{S}^1))$ norm coincides with the $L^\infty([0,T]; H^1(\mathbb{S}^1))$ for any element in $C([0,T]; H^1(\mathbb{S}^1))$. □

8. **Higher regularity solutions (Theorem 1.2).** In this section we fix $m \geq 2$ and consider the well-posedness of strong H^m solutions. We will emphasise the parts that differ from the well-posedness theory for strong H^1 solutions. Throughout this section we will require that the noise function σ belongs to $W^{m+1,\infty}(\mathbb{S}^1)$.

8.1. **Weak existence.** We begin by proving the existence of weak (martingale) H^m solutions. Cubic nonlinearities in the SDE for $\mathrm{d}\, \|u_n(t)\|^2_{H^1(\mathbb{S}^1)}$, which disappear due to the structure of the equation if $m = 1$, are retained at the level of $H^m(\mathbb{S}^1)$. Therefore standard calculations, involving first taking the expectation and then applying a standard Gronwall inequality, oblige us to use Gagliardo–Nirenberg inequalities. These fail to give sufficiently controllable powers on certain norms due to the extra expectation integral under which these norms are bounded. In particular, there are no uniform bounds in $L^\infty(\Omega)$ for any stochastic quantity. As in Theorem 7.6, we introduce stopping times $\eta^n_R < T$ to control exponential moments, so that the estimates derived below hold only on $[0, \eta^n_R]$, where the stopping time η^n_R depend on n and an "auxiliary parameter" R. These stopping times converge a.s. to the final time T as $R \to \infty$. Finally, we use the obtained estimate to conclude uniform-in-n stochastic boundedness in $L^2([0,T]; H^m(\mathbb{S}^1))$, which is precisely the tightness condition required to apply the Skorokhod–Jakubowski procedure.

We make here the technical observation that the projection operator $\mathbf{\Pi}_n$ acting on $f \in H^m(\mathbb{S}^1)$ also satisfies

$$\|\mathbf{\Pi}_n f - f\|_{H^m(\mathbb{S}^1)} \overset{n\uparrow\infty}{\longrightarrow} 0, \tag{8.1}$$

because the basis in $H^1(\mathbb{S}^1)$ of trigonometric functions of integral frequencies forms a basis in $H^m(\mathbb{S}^1)$ as well.

Proposition 8.1 (H^m and H^{m+1} estimates up to a stopping time)**.** *Let u_n be a solution to (3.1) with $\sigma \in W^{m+1,\infty}(\mathbb{S}^1)$ and initial condition $u_0 \in L^{2p}(\Omega; H^m(\mathbb{S}^1))$, for some $p \in [1,\infty)$. For $R > 1$, let η_R^n be the stopping time*

$$\eta_R^n := \inf \left\{ t \in [0,T] : \int_0^t \|u_n(s)\|_{W^{1,\infty}(\mathbb{S}^1)}^2 \, \mathrm{d}s > R \right\}, \tag{8.2}$$

setting $\eta_R^n = T$ if the set on the right-hand side is empty. Then $\eta_R^n \overset{R\uparrow\infty}{\longrightarrow} T$, $\mathbb{P}$-a.s., uniformly in n. Moreover, there exists a constant

$$C = C\left(p, T, R, \varepsilon, \mathbb{E}\,\|u_0\|_{H^m(\mathbb{S}^1)}^{2p}, \|\sigma\|_{W^{m+1,\infty}(\mathbb{S}^1)}\right),$$

independent of $n \in \mathbb{N}$, such that

$$\mathbb{E}\,\|u_n\|_{L^\infty([0,\eta_R^n];H^m(\mathbb{S}^1))}^p \le C. \tag{8.3}$$

Finally, there is a constant $C = C\left(T, R, \varepsilon, \mathbb{E}\,\|u_0\|_{H^m(\mathbb{S}^1)}^2, \|\sigma\|_{W^{m+1,\infty}(\mathbb{S}^1)}\right)$, which is independent of n, such that

$$\mathbb{E}\left(\int_0^{\eta_R^n} \|u_n(t)\|_{H^{m+1}(\mathbb{S}^1)}^2 \, \mathrm{d}t\right)^{1/2} \le C. \tag{8.4}$$

Remark 8.2. Subsequently, we will take $2p = 2$ to show existence, but require $2p = 8$ for uniqueness, to use Lemmas 7.1, 7.2, and Proposition 7.4. From Lemma 7.1, to establish uniqueness, we require 4th moments on $\|u\|_{L^\infty([0,T];H^1(\mathbb{S}^1))}$, which in turn is bounded by the 4th moment of the initial condition. Application of the stochastic Gronwall inequality requires that a strictly higher moment be boundeed. It is possible simply to take $2p > 4$, but it is more convenient simply to take $2p = 8$.

Proof. We divide the proof into several steps.

1. Pointwise limit of the stopping times.

The fact that η_R^n as defined in (8.2) converges a.s. to T as $R \to \infty$, uniformly in n, follows from a calculation similar to that in (7.19). More precisely, given the n-uniform (but ε-dependent) H^2 estimate implied by (4.3) and the embedding $H^2(\mathbb{S}^1) \hookrightarrow W^{1,\infty}(\mathbb{S}^1)$, we have

$$\begin{aligned}\mathbb{P}\left(\{\eta_R^n < T\}\right) &\le \mathbb{P}\left(\left\{\int_0^T \|u_n(t)\|_{W^{1,\infty}(\mathbb{S}^1)}^2 \, \mathrm{d}t > R\right\}\right) \\ &\le \frac{1}{R}\mathbb{E}\int_0^T \|u_n(t)\|_{W^{1,\infty}(\mathbb{S}^1)}^2 \, \mathrm{d}t \le \frac{\tilde{C}}{R}\mathbb{E}\int_0^T \|u_n(t)\|_{H^2(\mathbb{S}^1)}^2 \, \mathrm{d}t \\ &\le \frac{C_\varepsilon}{R} \overset{R\uparrow\infty}{\longrightarrow} 0.\end{aligned} \tag{8.5}$$

2. Bounds on higher regularity norms.

Next we prove the bound (8.3). Taking the ℓth derivative of (3.1), we find that

$$\begin{aligned}0 = \mathrm{d}\partial_x^\ell u_n - \varepsilon \partial_x^{\ell+2} u_n \, \mathrm{d}t + \left[\partial_x^\ell \mathbf{\Pi}_n \left(u_n \, \partial_x u_n\right) + \partial_x^{\ell+1} P[u_n]\right] \, \mathrm{d}t \\ - \frac{1}{2}\partial_x^\ell \mathbf{\Pi}_n \left(\sigma \, \partial_x \left(\sigma \, \partial_x u_n\right)\right) \mathrm{d}t + \partial_x^\ell \mathbf{\Pi}_n \left(\sigma \, \partial_x u_n\right) \, \mathrm{d}W.\end{aligned}$$

First we multiply through by $\partial_x^\ell u_n$, integrate in space, and use the commutativity between the projection and the derivative. Then we apply the Itô formula for

$r \mapsto r^p$. The result is

$$\begin{aligned}
&\frac{1}{2p}\left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p}\Big|_0^t + \varepsilon\int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\left\|\partial_x^{\ell+1} u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2}\,\mathrm{d}s\\
&\quad= \frac{1}{2}\int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\int_{\mathbb{S}^1}\partial_x u_n\left(\partial_x^\ell u_n\right)^2\,\mathrm{d}x\,\mathrm{d}s\\
&\qquad - \int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\int_{\mathbb{S}^1}\partial_x^{\ell+1} P[u_n]\,\partial_x^\ell u_n\,\mathrm{d}x\,\mathrm{d}s\\
&\qquad + \int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\int_{\mathbb{S}^1}\left(u_n\partial_x^{\ell+1}u_n - \partial_x^\ell\left(u_n\,\partial_x u_n\right)\right)\partial_x^\ell u\,\mathrm{d}x\,\mathrm{d}s\\
&\qquad + \frac{1}{2}\int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\int_{\mathbb{S}^1}\partial_x^\ell u_n\,\partial_x^\ell\left(\sigma\,\partial_x\left(\sigma\,\partial_x u_n\right)\right)\,\mathrm{d}x\,\mathrm{d}s\\
&\qquad + \frac{1}{2}\int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-4}\left|\int_{\mathbb{S}^1}\partial_x^\ell u_n\,\partial_x^\ell\left(\sigma\,\partial_x u_n\right)\,\mathrm{d}x\right|^2\,\mathrm{d}s\\
&\qquad + \int_0^t \left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\int_{\mathbb{S}^1}\partial_x^\ell u_n\,\partial_x^\ell\left(\sigma\,\partial_x u_n\right)\,\mathrm{d}x\,\mathrm{d}W\\
&\quad =: \sum_{i=1}^{5}\int_0^t I_i^n\,\mathrm{d}s + \int_0^t I_6^n\,\mathrm{d}W.
\end{aligned}$$

We again estimate I_1^n to I_5^n, leaving the martingale term $\int_0^t I_6^n\,\mathrm{d}W$ to be handled by the stochastic Gronwall inequality. We have readily that

$$|I_1^n| \le \frac{1}{2}\left\|\partial_x u_n\right\|_{L^\infty(\mathbb{S}^1)}\left\|\partial_x^\ell u_n\right\|_{L^2(\mathbb{S}^1)}^{2p}.$$

By the Cauchy–Schwarz and Young's inequalites,

$$|I_2^n| \le C_\varepsilon\left\|\partial_x^\ell u_n\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\left\|\partial_x^\ell P[u_n]\right\|_{L^2(\mathbb{S}^1)}^{2} + \frac{\varepsilon}{2}\left\|\partial_x^\ell u_n\right\|_{L^2(\mathbb{S}^1)}^{2p-2}\left\|\partial_x^{\ell+1} u_n\right\|_{L^2(\mathbb{S}^1)}^{2}.$$

By the Leibniz rule and the Gagliardo–Nirenberg inequality,

$$\begin{aligned}
\left\|\partial_x^\ell u_n^2\right\|_{L^2(\mathbb{S}^1)} &= \left\|u_n\partial_x^\ell u_n + \ell\partial_x u_n\,\partial_x^{\ell-1}u_n + \ldots + \partial_x^\ell u_n\,u_n\right\|_{L^2(\mathbb{S}^1)}\\
&\lesssim \left\|u_n\right\|_{L^\infty(\mathbb{S}^1)}\left\|\partial_x^\ell u_n\right\|_{L^2(\mathbb{S}^1)},
\end{aligned}\tag{8.6}$$

and likewise

$$\left\|\partial_x^{\ell-1}\left(\partial_x u_n\right)^2\right\|_{L^2(\mathbb{S}^1)} \lesssim \left\|\partial_x u_n\right\|_{L^\infty(\mathbb{S}^1)}\left\|\partial_x^\ell u_n\right\|_{L^2(\mathbb{S}^1)}.$$

Remark 8.3. We add some details on the estimate (8.6). It suffices to show

$$\left\|u^{(j)}u^{(\ell-j)}\right\|_{L^2(\mathbb{S}^1)} \lesssim \left\|u\right\|_{L^\infty(\mathbb{S}^1)}\left\|u^{(\ell)}\right\|_{L^2(\mathbb{S}^1)}, \quad j = 0,\ldots,\ell,$$

with $u^{(j)} = \partial_x^j u$. Hölder's inequality gives

$$\left\|u^{(j)}u^{(\ell-j)}\right\|_2 \le \left\|u^{(j)}\right\|_r\left\|u^{(\ell-j)}\right\|_{r'}, \quad \frac{1}{r} + \frac{1}{r'} = \frac{1}{2}.$$

Apply now the Gagliardo–Nirenberg inequality [5, Lemma 2.1] to find

$$\left\|u^{(\beta)}\right\|_r \lesssim \left\|u^{(m)}\right\|_p^\theta\left\|u\right\|_q^{1-\theta},$$

assuming $\int_{\mathbb{S}^1} u \,\mathrm{d}x = 0$. Here

$$\frac{1}{r} = \beta - \theta(m - \frac{1}{p}) + (1-\theta)\frac{1}{q}, \quad m > \beta.$$

If p equals 1 or ∞, then $\theta = \beta/m$. Assume for the moment that $\int_{\mathbb{S}^1} u \,\mathrm{d}x = 0$. We get

$$\left\| u^{(j)} \right\|_r \lesssim \left\| u^{(\ell)} \right\|_2^{j/\ell} \|u\|_\infty^{1-j/\ell},$$

with

$$\frac{1}{r} = j - \frac{j}{\ell}(\ell - \frac{1}{2}).$$

Similarly,

$$\left\| u^{(\ell-j)} \right\|_{r'} \lesssim \left\| u^{(\ell)} \right\|_2^{\ell-j/\ell} \|u\|_\infty^{j/\ell},$$

with

$$\frac{1}{r'} = \ell - j - \frac{\ell - j}{\ell}(\ell - \frac{1}{2}).$$

Note that $1/r + 1/r' = 1/2$. Furthermore,

$$\begin{aligned}\left\| u^{(j)} \right\|_r \left\| u^{(\ell-j)} \right\|_{r'} &\lesssim \left\| u^{(\ell)} \right\|_2^{j/\ell} \|u\|_\infty^{1-j/\ell} \left\| u^{(\ell)} \right\|_2^{\ell-j/\ell} \|u\|_\infty^{j/\ell} \\ &= \left\| u^{(\ell)} \right\|_2 \|u\|_\infty .\end{aligned}$$

In the general case, let $v = u - \left|\mathbb{S}^1\right|^{-1} \int_{\mathbb{S}^1} u \,\mathrm{d}x$. Then we find

$$\begin{aligned}\left\| u^{(j)} \right\|_r = \left\| v^{(j)} \right\|_r &\lesssim \left\| v^{(\ell)} \right\|_2^{j/\ell} \|v\|_\infty^{1-j/\ell} \\ &\lesssim \left\| u^{(\ell)} \right\|_2^{j/\ell} \|u\|_\infty^{1-j/\ell},\end{aligned}$$

and similar for the other estimate. This justifies (8.6).

Writing the term $\partial_x^\ell P$ above as

$$\partial_x^\ell K * \left(u_n^2 + \frac{1}{2} (\partial_x u_n)^2 \right) = K * \partial_x^\ell u_n^2 + \frac{1}{2} \partial_x K * \partial_x^{\ell-1} (\partial_x u_n)^2,$$

the L_x^2 norm can be bounded by the Young convolution and Gagliardo–Nirenberg inequalities, see [51, (2.6), (2.8)]:

$$\begin{aligned}\left\| \partial_x^\ell P[u_n] \right\|_{L^2(\mathbb{S}^1)} &\lesssim \left(\left\| \partial_x^\ell u_n^2 \right\|_{L^2(\mathbb{S}^1)} + \left\| \partial_x^{\ell-1} (\partial_x u_n)^2 \right\|_{L^2(\mathbb{S}^1)} \right) \\ &\lesssim \left(\|u_n\|_{L^\infty(\mathbb{S}^1)} + \|\partial_x u_n\|_{L^\infty} \right) \left\| \partial_x^\ell u_n \right\|_{L^2(\mathbb{S}^1)} .\end{aligned}$$

Because $u_n \partial_x^{\ell+1} u_n - \partial_x^\ell (u_n \,\partial_x u_n)$ precisely removes all instances of the $(\ell+1)$st derivative, by the Gagliardo–Nirenberg inequality, as in [51, (2.7)],

$$\begin{aligned}|I_3^n| &\le \left\| \partial_x^\ell u_n \right\|_{L^2(\mathbb{S}^1)}^{2p-2} \left\| u_n \partial_x^{\ell+1} u_n - \partial_x^\ell (u_n \,\partial_x u_n) \right\|_{L^2(\mathbb{S}^1)} \left\| \partial_x^\ell u_n \right\|_{L^2(\mathbb{S}^1)} \\ &\le C \|\partial_x u_n\|_{L^\infty(\mathbb{S}^1)} \left\| \partial_x^\ell u_n \right\|_{L^2(\mathbb{S}^1)}^{2p} .\end{aligned}$$

In the parts involving σ, for I_4^n we have

$$\begin{aligned}&2 I_4^n \left\| \partial_x^\ell u_n \right\|_{L^2(\mathbb{S}^1)}^{-2p+2} \\ &\quad = - \int_{\mathbb{S}^1} \partial_x^{\ell+1} u_n \, \partial_x^{\ell-1} (\sigma \,\partial_x (\sigma \,\partial_x u_n)) \,\mathrm{d}x\end{aligned}$$

$$
\begin{aligned}
&= -\int_{\mathbb{S}^1} \partial_x^{\ell+1} u_n \, \partial_x^{\ell-1}\left(\sigma^2 \partial_x^2 u_n\right) + \partial_x^{\ell+1} u_n \, \partial_x^{\ell-1}\left(\sigma \, \partial_x \sigma \, \partial_x u_n\right) \mathrm{d}x \\
&= -\int_{\mathbb{S}^1} \sigma^2 \left|\partial_x^{\ell+1} u_n\right|^2 \, \mathrm{d}x - \int_{\mathbb{S}^1} \partial_x^{\ell+1} u_n \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} \partial_x^{\ell-1-k} \sigma^2 \partial_x^{2+k} u_n \, \mathrm{d}x \\
&\quad - \int_{\mathbb{S}^1} \sigma \, \partial_x \sigma \, \partial_x^{\ell+1} u_n \partial_x^{\ell} u_n + \partial_x^{\ell+1} u_n \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} \partial_x^{\ell-1-k} \left(\sigma \, \partial_x \sigma\right) \, \partial_x^{k+1} u_n \, \mathrm{d}x \\
&= -\int_{\mathbb{S}^1} \sigma^2 \left|\partial_x^{\ell+1} u_n\right|^2 \, \mathrm{d}x + \frac{1}{2} \int_{\mathbb{S}^1} \partial_x \left(\sigma \partial_x \sigma\right) \left|\partial_x^{\ell} u_n\right|^2 \, \mathrm{d}x \\
&\quad - \int_{\mathbb{S}^1} \partial_x^{\ell+1} u_n \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} \partial_x^{\ell-1-k} \sigma^2 \partial_x^{2+k} u_n \, \mathrm{d}x \\
&\quad - \int_{\mathbb{S}^1} \partial_x^{\ell+1} u_n \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} \partial_x^{\ell-1-k} \left(\sigma \, \partial_x \sigma\right) \, \partial_x^{k+1} u_n \, \mathrm{d}x,
\end{aligned}
$$

which implies

$$
2 \left|I_4^n\right| \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{-2p+2} \leq C_{\sigma,\ell} \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2},
$$

because the summands do not have derivatives of order higher than ℓ.

Similarly for I_5^n we have

$$
\begin{aligned}
&\int_{\mathbb{S}^1} \partial_x^{\ell+1} u_n \, \partial_x^{\ell-1}\left(\sigma \partial_x u_n\right) \mathrm{d}x \\
&\quad = \int_{\mathbb{S}^1} \left[\sigma \, \partial_x^{\ell+1} u_n \, \partial_x^{\ell} u_n + \partial_x^{\ell} u_n \partial_x \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} \partial_x^{\ell-1-k} \sigma \, \partial_x^{k+1} u_n\right] \mathrm{d}x \\
&\quad = -\int_{\mathbb{S}^1} \left[\frac{1}{2} \partial_x \sigma \left|\partial_x^{\ell} u_n\right|^2 - \partial_x^{\ell} u_n \partial_x \sum_{k=0}^{\ell-2} \binom{\ell-1}{k} \partial_x^{\ell-1-k} \sigma \, \partial_x^{k+1} u_n\right] \mathrm{d}x,
\end{aligned}
$$

and therefore

$$
\left|I_5^n\right| \leq C_{\sigma,\ell} \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2p-4} \left(\left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2}\right)^2 \leq C_{\sigma,\ell} \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2p}.
$$

Gathering the estimates for I_1^n to I_5^n, we find

$$
\begin{aligned}
&\frac{1}{2p} \mathrm{d} \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2p} + \frac{\varepsilon}{2} \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2p-2} \left\|\partial_x^{\ell+1} u_n\right\|_{L^2(\mathbb{S}^1)}^{2} \, \mathrm{d}t \\
&\quad \leq C_{\sigma,\ell,\varepsilon} \left(1 + \left\|u_n\right\|_{L^\infty(\mathbb{S}^1)} + \left\|\partial_x u_n\right\|_{L^\infty(\mathbb{S}^1)}\right) \left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2p} \, \mathrm{d}t + I_6^n \, \mathrm{d}W,
\end{aligned}
$$

where $\int_0^{t\wedge\eta_R^n} I_6^n \, \mathrm{d}W$ is a square-integrable martingale. We can overestimate the right-hand side by adding to "$\left\|\partial_x^{\ell} u_n\right\|_{L^2(\mathbb{S}^1)}^{2p}$" the term "$\int_0^t \frac{\varepsilon}{2} \dots \mathrm{d}s$". Setting

$$
\begin{aligned}
\xi_n(t) &:= \frac{1}{2p} \left\|\partial_x^{\ell} u_n(t)\right\|_{L^2(\mathbb{S}^1)}^{2p} + \frac{\varepsilon}{2} \int_0^t \left\|\partial_x^{\ell} u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p-2} \left\|\partial_x^{\ell+1} u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2} \, \mathrm{d}s, \\
A_n(t) &:= \int_0^t C_{\sigma,\ell,\varepsilon} \left(1 + \left\|u_n(s)\right\|_{L^\infty(\mathbb{S}^1)} + \left\|\partial_x u_n(s)\right\|_{L^\infty(\mathbb{S}^1)}\right) \mathrm{d}s,
\end{aligned}
$$

and $M_n(t) := \int_0^t I_6^n \,\mathrm{d}W(s)$, we obtain $\mathrm{d}\xi_n(t) \le \xi_n(t)\,\mathrm{d}A_n(t) + \mathrm{d}M_n(t)$. Now an application of the stochastic Gronwall inequality (Lemma A.2) gives

$$\begin{aligned}\Bigg(\mathbb{E} \sup_{s\in[0,\eta_R^n]} \Bigg(&\left\|\partial_x^\ell u_n(s)\right\|_{L^2(\mathbb{S}^1)}^{2p} \\ &+ \frac{\varepsilon}{2}\int_0^s \left\|\partial_x^\ell u_n(s')\right\|_{L^2(\mathbb{S}^1)}^{2p-2} \left\|\partial_x^{\ell+1} u_n(s')\right\|_{L^2(\mathbb{S}^1)}^2 \,\mathrm{d}s' \Bigg)^{1/2}\Bigg)^2 \\ &\le C_{p,\sigma,\varepsilon,T,\ell,R}\, \mathbb{E}\left\|\partial_x^\ell u_n(0)\right\|_{L^2(\mathbb{S}^1)}^{2p}, \qquad \ell = 0,\dots,m,\end{aligned} \tag{8.7}$$

from which (8.3) easily follows.

With $p=1$ and $\ell = 1,\dots,m$, it also follows from (8.7) that

$$\mathbb{E}\left(\frac{\varepsilon}{2}\int_0^{\eta_R^n} \left\|\partial_x^{\ell+1} u_n(s')\right\|_{L^2(\mathbb{S}^1)}^2 \,\mathrm{d}s'\right)^{1/2} \le C_{\sigma,\varepsilon,T,\ell,R,u_0},$$

which implies (8.4). □

Next, we establish stochastic boundedness and tightness of laws for $\{u_n\}$. The proof differs from the straightforward deduction leading to Lemma 5.5 ($m=1$), and $u_n \in_{\mathrm{sb}} L^2([0,T];H^2(\mathbb{S}^1)) \cap W^{\theta',2}([0,T];L^2(\mathbb{S}^1))$ and the tightness of laws in $L^2([0,T];H^1(\mathbb{S}^1))$, where $u_n \in_{\mathrm{sb}} L_t^2 H_x^2$ follows trivially from $u_n \in_b L^2(\Omega; L_t^2 H_x^2)$. For $m \ge 2$, we do not have $u_n \in_b L^2\left(\Omega; L_t^2 H_x^{m+1}\right)$ for the entire interval $[0,T]$, but rather only up to a suitable stopping time, cf. (8.4). The next proof develops a refined stopping time argument to deal with this issue, leading to $u_n \in_{\mathrm{sb}} L_t^2 H_x^{m+1}$.

Lemma 8.4. *Let u_n be a solution to (3.1) with $\mathbb{E}\|u_0\|_{H^1(\mathbb{S}^1)}^p, \mathbb{E}\|u_0\|_{H^m(\mathbb{S}^1)}^2 < \infty$, for some $p>2$. For $\theta' < (2-p)/4p$, the laws of $\{u_n\}_{n\in\mathbb{N}}$ are uniformly stochastically bounded in $L^2([0,T];H^{m+1}(\mathbb{S}^1)) \cap W^{\theta',2}([0,T];L^2(\mathbb{S}^1))$, i.e.,*

$$\lim_{M\to\infty} \mathbb{P}\left(\left\{\|u_n\|_{L^2([0,T];H^{m+1}(\mathbb{S}^1))\cap W^{\theta',2}([0,T];L^2(\mathbb{S}^1))} > M\right\}\right) = 0 \tag{8.8}$$

holds, uniformly in n. The laws of $\{u_n\}_{n\in\mathbb{N}}$ are also uniformly stochastically bounded in $L^\infty([0,T];H^m(\mathbb{S}^1))$.

Moreover, the laws of $\{u_n\}$ are tight on $L^2([0,T];H^m(\mathbb{S}^1))$.

Proof. As in the proof of Lemma 5.5, for any $\theta \in (\theta', (2-p)/4p)$,

$$\mathbb{P}\left(\left\{\|u_n\|_{W^{\theta',r}([0,T];L^2(\mathbb{S}^1))} > M\right\}\right) \le \frac{1}{M}\,\mathbb{E}\,\|u_n\|_{C^\theta([0,T];L^2(\mathbb{S}^1))} \lesssim \frac{1}{M},$$

where, in passing, we mention that the requirement $u_0 \in L_\omega^p H_x^1$ is linked to the application of Lemma 5.1, which allows us to arrive at the final $1/M$ estimate.

Following the proof of Lemma 5.5, set

$$X_n(t) := \left(\int_0^t \|u_n(s)\|_{H^{m+1}(\mathbb{S}^1)}^2 \,\mathrm{d}s\right)^{1/2} = \|u_n\|_{L^2([0,t];H^{m+1}(\mathbb{S}^1))},$$

and introduce the stopping time

$$\xi_M^n = \inf\left\{t\in[0,T] : X_n(t) > M\right\},$$

setting $\xi_M^n = T$ if the set is empty. Let η_R^n be the stopping time defined in (8.2). For any fixed R, we have

$$\{X_n(t) > M\} = \{\xi_M^n < t\} = \{\xi_M^n < t, \eta_R^n < t\} \cup \{\xi_M^n < t, \eta_R^n \ge t\}$$

$$\subseteq \{\eta_R^n < t\} \cup \{\xi_M^n < t, \eta_R^n \geq t\}.$$

We can estimate the probability of the last event on the right-hand side separately as follows:

$$\{\xi_M^n < t, \eta_R^n \geq t\} \subseteq \{\xi_M^n < (t \wedge \eta_R^n)\} \subseteq \{X_n(t \wedge \xi_M^n \wedge \eta_R^n) \geq M\},$$

which implies

$$\begin{aligned} \mathbb{P}\big(\{\xi_M^n < t, \eta_R^n \geq t\}\big) &\leq \mathbb{P}\big(\{X_n(t \wedge \xi_M^n \wedge \eta_R^n) \geq M\}\big) \\ &\leq \frac{1}{M}\mathbb{E}\, X_n(t \wedge \xi_M^n \wedge \eta_R^n) \leq \frac{1}{M}\mathbb{E}\, X_n(t \wedge \eta_R^n) \overset{(8.4)}{\leq} \frac{C_{T,R,\varepsilon}}{M}. \end{aligned}$$

Separately, via (8.5),

$$\mathbb{P}\big(\{\eta_R^n < t\}\big) \leq \frac{C_{T,\varepsilon}}{R}.$$

Therefore,

$$\mathbb{P}\big(\{X_n(t) > M\}\big) \leq \frac{C_{T,R,\varepsilon}}{M} + \frac{C_{T,\varepsilon}}{R}.$$

Sending $M \to \infty$, recalling the definition of X_n, we arrive at

$$\lim_{M\to\infty} \mathbb{P}\left(\left\{\left(\int_0^t \|u_n(s)\|_{H^{m+1}(\mathbb{S}^1)}^2 \,\mathrm{d}s\right)^{1/2} > M\right\}\right) \leq \frac{C_{T,\varepsilon}}{R},$$

which can be made arbitrarily small by taking R large, uniformly in $t \in [0,T]$. This implies (8.8). Tightness on $L^2([0,T]; H^m(\mathbb{S}^1))$ follows from this, Lemma 5.4, and the n-uniformity of the limit $M \to \infty$, arguing as in in the proof of Lemma 5.5.

The same argument yields stochastic boundedness in $L^\infty([0,T]; H^m(\mathbb{S}^1))$ for the laws of $\{u_n\}_{n\in\mathbb{N}}$. □

Introducing the path spaces:

$$\begin{aligned} &\mathcal{X}_{u,s}^m := L^2([0,T]; H^m(\mathbb{S}^1)), \quad \mathcal{X}_{u,w}^m := C([0,T]; H_w^1(\mathbb{S}^1)), \\ &\mathcal{X}_W := C([0,T]), \quad \mathcal{X}_0 := H^m(\mathbb{S}^1), \end{aligned}$$

and setting $\mathcal{X}_m := \mathcal{X}_{u,s}^m \times \mathcal{X}_{u,w}^m \times \mathcal{X}_W \times \mathcal{X}_0$, we repeat the procedure in Section 6.

Lemma 8.5. *The joint laws of* $\big(u_n, u_n, W, \mathbf{\Pi}_n u_0\big)$ *are tight on* $\mathcal{X}_m$.

Proof. By Proposition 8.4 and Lemma 5.3, the laws of u_n are tight on $\mathcal{X}_{u,s}^m$ and $\mathcal{X}_{u,w}^m$. Since $\mathbf{\Pi}_n u_0 \to u_0$ in $H^m(\mathbb{S}^1)$, cf. (8.1), the laws of $\mathbf{\Pi}_n u_0$ are tight on $H^m(\mathbb{S}^1)$. As $n \to \infty$, the law of W is stationary on $\mathcal{X}_W$ and therefore tight. □

Theorem 8.6 (Weak H^m solution)**.** *Suppose* $\sigma \in W^{m+1,\infty}(\mathbb{S}^1)$ *and that* u_0 *belongs to* $L^p(\Omega; H^1(\mathbb{S}^1)) \cap L^2(\Omega; H^m(\mathbb{S}^1))$, *for* $p \in [1,\infty)$. *There exists a weak* H^m *solution* $\big((\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\geq 0}, \tilde{\mathbb{P}}), \tilde{u}, \tilde{W}\big)$ *to the viscous stochastic CH equation* (1.1) *with initial condition* $u|_{t=0} = u_0$.

Proof. From the Skorokhod–Jakubowski theorem (Theorem 6.2), we can extract variables $(\tilde{u}_{n,s}, \tilde{u}_{n,w}, \tilde{W}_n, \tilde{u}_{0,n})$ and $(\tilde{u}_s, \tilde{u}_w, \tilde{W}, \tilde{u}_0)$ on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ such that

$$\begin{aligned} &(\tilde{u}_{n,s}, \tilde{u}_{n,w}, \tilde{W}_n, \tilde{u}_{0,n}) \sim (u_n, u_n, W, \mathbf{\Pi}_n u_0) \quad \text{in } \mathcal{X}_m, \\ &(\tilde{u}_{n,s}, \tilde{u}_{n,w}, \tilde{W}_n, \tilde{u}_{0,n}) \overset{n\uparrow\infty}{\longrightarrow} (\tilde{u}_s, \tilde{u}_w, \tilde{W}, \tilde{u}_0) \quad \text{in } \mathcal{X}_m, \quad \mathbb{P}\text{-a.s.}, \end{aligned}$$

along a subsequence that is not relabelled. From Lemma 6.6, $\tilde{W}$ is a Brownian motion on $\big(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\geq 0}, \tilde{\mathbb{P}}\big)$, where $\{\tilde{\mathcal{F}}_t\}_{t\geq 0}$ is the canonical filtration defined by

$$\tilde{\mathcal{F}}_t := \Sigma\left(\Sigma\big(u\big|_{[0,t]}, W\big|_{[0,t]}\big) \cup \left\{N \in \tilde{\mathcal{F}} : \tilde{\mathbb{P}}(N) = 0\right\}\right).$$

As in Lemma 6.4, we can identify $\tilde{u}_n := \tilde{u}_{n,s} = \tilde{u}_{n,w}$, and $\tilde{u} := \tilde{u}_s = \tilde{u}_w$, $\tilde{\mathbb{P}} \otimes \mathrm{d}t \otimes \mathrm{d}x$-a.e.

Following the proof of Theorem 6.8, the Galerkin equation (3.1) holds in the PDE weak sense using the equivalence of laws, for the variables $\big(\tilde{u}_n, \tilde{W}_n, \tilde{u}_{0,n}\big)$ in place of $(u_n, W, \mathbf{\Pi}_n u_0)$, $\tilde{\mathbb{P}}$-almost surely, up to any $t \in [0,T]$. Using the $\tilde{\mathbb{P}}$-almost everywhere convergence of $\big(\tilde{u}_n, \tilde{W}_n, \tilde{u}_{0,n}\big)$ in the joint path space $\mathcal{X}^m$, as in the proof of Theorem 6.8, we can extract the limiting equation for $\big(\tilde{u}, \tilde{W}, \tilde{u}_0\big)$, thereby establishing the existence of a weak (martingale) H^m solution in the $n \to \infty$ limit.

Strong temporal continuity in $H^1(\mathbb{S}^1)$ can be established exactly as in Section 7.4, and stochastic boundedness in $L^2([0,T]; H^{m+1}(\mathbb{S}^1))$ follows from equality of laws and Lemma 8.4 because, by the Lusin–Souslin theorem, $L^2([0,T]; H^{m+1}(\mathbb{S}^1))$ injects continuously into $L^2([0,T]; H^m(\mathbb{S}^1))$ and hence is Borel in the bigger space (see Part 6 of the proof of Theorem 6.8). □

8.2. **Pathwise uniqueness and strong H^m solutions.** In this section, we briefly conclude with pathwise uniqueness in H^m.

Theorem 8.7 (Pathwise uniqueness in H^m)**.** *Let u, v be strong H^m solutions to the viscous stochastic CH equation (1.1), with $\sigma \in W^{m+1,\infty}(\mathbb{S}^1)$ and initial condition $u|_{t=0} = v|_{t=0} = u_0 \in L^8(\Omega; H^m(\mathbb{S}^1))$. Then*

$$\mathbb{E}\,\|u - v\|_{L^\infty([0,T];H^m(\mathbb{S}^1))} = 0.$$

Proof. Having established that $\mathbb{E}\,\|u - v\|_{L^\infty([0,T];H^1(\mathbb{S}^1))} = 0$ in Theorem 7.6, we conclude that $u = v$, $\mathbb{P}\otimes\mathrm{d}t\otimes\mathrm{d}x$-a.e. Then necessarily, $\mathbb{E}\,\|u - v\|_{L^\infty([0,T];H^m(\mathbb{S}^1))} = 0$ also. This is uniqueness in $L^1(\Omega; L^\infty([0,T]; H^m(\mathbb{S}^1)))$. □

With the same argument that was employed in Section 7.3, we can now conclude that the second main theorem of the paper (Theorem 1.2), holds.

Acknowledgments. We thank the anonymous reviewers whose comments helped improve and clarify this manuscript.

Appendix A. **Stochastic toolbox.** In this section, we recall some notations and results from stochastic analysis that are used throughout the paper. We use [14, 37, 41] as general references on stochastic analysis and SPDEs. Recall that $\big(\Omega, \mathcal{F}, \mathbb{P}\big)$ is a complete probability space with a countably generated σ-algebra $\mathcal{F}$ and probability measure $\mathbb{P}$. Let $\mathbb{B}$ be a separable Banach space, equipped with the Borel σ-algebra $\mathcal{B}(\mathbb{B})$. A $\mathbb{B}$-valued random variable v is a measurable mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $\big(\mathbb{B}, \mathcal{B}(\mathbb{B})\big)$, $\omega \mapsto v(\omega)$. The expectation of v is $\mathbb{E}\,v := \int_\Omega v\, d\mathbb{P}$. We often use the abbreviation a.s. or almost surely to mean for $\mathbb{P}$-almost every $\omega \in \Omega$. The collection of $\mathbb{B}$-valued random variables v for which $\mathbb{E}\,|v| < \infty$ is denoted by $L^1(\Omega) = L^1(\Omega, \mathcal{F}, \mathbb{P})$. This is a Banach space with norm $\|v\|_{L^1(\Omega)} = \mathbb{E}\,\|v\|_{\mathbb{B}}$. For $p > 1$, $L^p(\Omega)$ is defined similarly, with $\|v\|_{L^p(\Omega)}$ given by $\big(\mathbb{E}\,\|v\|_{\mathbb{B}}^p\big)^{1/p}$ if $p < \infty$ and $\operatorname{ess\,sup}_{\omega\in\Omega} \|v(\omega)\|_{\mathbb{B}}$ if $p = \infty$.

A stochastic process $v = \{v(t)\}_{t\in[0,T]}$, for $T > 0$, is a collection of $\mathbb{B}$-valued random variables $v(t)$. We say that v is measurable if v is jointly measurable from $\mathcal{F} \times \mathcal{B}([0,T])$ to $\mathcal{B}(\mathbb{B})$. Recall that we consider filtrations $\{\mathcal{F}_t\}_{t\in[0,T]}$ that satisfy the "usual conditions" of being complete and right-continuous, and we refer to $\mathcal{S} := \big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\big)$, see (2.1), as a stochastic basis. A stochastic process v is *adapted* if $v(t)$ is $\mathcal{F}_t$ measurable for all $t \in [0,T]$. When a filtration is involved there are additional notions of measurability (predictable, optional, progressive) that are more convenient to work with. Here we use the (stronger) notion of a predictable process. A *predictable* process v is a $\mathcal{P}_T \times \mathcal{B}([0,T])$-measurable map $\Omega \times [0,T] \to \mathbb{B}$, $(\omega,t) \mapsto v(\omega,t)$, where $\mathcal{P}_T$ denotes the predictable σ-algebra on $\Omega \times [0,T]$ associated with $\{\mathcal{F}_t\}_{t\in[0,T]}$ (the σ-algebra generated by all left-continuous adapted processes). A predictable process is adapted. Although the converse is not true, adaptive processes with regular (e.g., continuous) paths are predictable. To check for continuity, one uses the *Kolmogorov test* [14, p. 7]: suppose there are constants $\kappa > 1$, $\delta > 0$, and $K > 0$ such that

$$\mathbb{E}\,\|v(t) - v(s)\|_{\mathbb{B}}^{\kappa} \leq K\,|t-s|^{1+\delta}, \quad \forall s,t \in [0,T],$$

then there exists a continuous modification of v, still denoted by v, such that $\mathbb{E}\,\|v\|_{C^\gamma([0,T];\mathbb{B})}^{\kappa} \leq K$, where the constant K is independent of v and $\gamma \in \big[0, \frac{\delta}{\kappa}\big)$.

Throughout the work, we repeatedly end up with SDE inequalities of the form $\mathrm{d}\xi \leq \eta\,\mathrm{d}t + L\xi\,\mathrm{d}t + \mathrm{d}M$, for some quantity of interest $\xi = \xi(\omega,t)$ and a zero-mean martingale M. For us $L \geq 0$ is often a stochastic process, so that the standard (deterministic) Gronwall inequality cannot be applied. The following stochastic Gronwall inequality is taken from [50, Lemma 3.8], which is a version of a result proved first in [44, Thm. 4]. The term $L\xi\,\mathrm{d}t$ can be written as $\xi\mathrm{d}\int_0^t L(s)\,\mathrm{d}s = \xi\,\mathrm{d}A(t)$, which is the form used in the lemma. Besides, the inequality provides a bound on the νth moment of ξ that does not depend on the martingale term M. It is this "martingale uniformity" that forces the non-standard condition $\nu \in (0,1)$.

Lemma A.1 (Stochastic Gronwall inequality)**.** *Relative to the stochastic basis $\mathcal{S}$, see (2.1), let $\xi(t)$ and $\eta(t)$ be two non-negative adapted processes, $A(t)$ be an adapted non-decreasing process with $A(0) = 0$, and M a local martingale with $M(0) = 0$. Suppose ξ is càdlàg in time and satisfies the following SDE inequality on $[0,T]$:*

$$\mathrm{d}\xi \leq \eta\,\mathrm{d}t + \xi\,\mathrm{d}A + \mathrm{d}M.$$

For $0 < \nu < r < 1$ and $t \in [0,T]$, we have

$$\Big(\mathbb{E}\sup_{s\in[0,t]}|\xi(s)|^{\nu}\Big)^{1/\nu} \leq \Big(\frac{r}{r-\nu}\Big)^{1/\nu}\left(\mathbb{E}\exp\left(\frac{rA(t)}{1-r}\right)\right)^{(1-r)/r}\mathbb{E}\Big(\xi(0) + \int_0^t \eta(s)\,\mathrm{d}s\Big).$$

This lemma can be formulated for stopping times τ in place of t. For suppose ξ, η, A, and M are as in Lemma A.1, then for any stopping time τ,

$$\mathrm{d}\xi(t\wedge\tau) \leq \eta(t\wedge\tau)\,\mathrm{d}(t\wedge\tau) + \xi(t\wedge\tau)\,\mathrm{d}A(t\wedge\tau) + \mathrm{d}M(t\wedge\tau).$$

Since τ is a stopping time, $M(t\wedge\tau)$ remains a local martingale (see [41, Cor. II.3.6, Def. IV.1.5]), moreover, we can write $\eta(t\wedge\tau)\mathrm{d}(t\wedge\tau)$ as $\mathbb{1}_{\{t\leq\tau\}}\eta(t)\,\mathrm{d}t$, so using the elementary equality

$$\sup_{s\in[0,T]}|\xi(s\wedge\tau)| = \sup_{s\in[0,T\wedge\tau]}|\xi(s)|,$$

Lemma A.1 is readily seen to imply:

Lemma A.2. *Let ξ, η, A and M be as in Lemma A.1. Let τ be a stopping time on the same filtration as M is a martingale. For $0 < \nu < r < 1$, we have*

$$\left(\mathbb{E} \sup_{s\in[0,T\wedge\tau]} |\xi(s)|^{\nu}\right)^{1/\nu} \leq \left(\frac{r}{r-\nu}\right)^{1/\nu} \left(\mathbb{E} \exp\left(\frac{rA(T\wedge\tau)}{1-r}\right)\right)^{(1-r)/r} \mathbb{E}\left(\xi(0) + \int_0^{T\wedge\tau} \eta(s)\,\mathrm{d}s\right).$$

Next, we use on a few occasions the following convergence result for stochastic integrals, which is due to Debussche, Glatt-Holtz, and Temam, see [23, Lemma 2.1].

Lemma A.3 (Convergence of stochastic integrals)**.** *Fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For each $n \in \mathbb{N}$, consider a stochastic basis $\mathcal{S}_n = \big(\Omega, \mathcal{F}, \{\mathcal{F}_t^n\}_{t\in[0,T]}, \mathbb{P}\big)$, a Wiener process W^n on $\mathcal{S}_n$, and a predictable $L^2(\mathbb{S}^1)$-valued process G^n on $\mathcal{S}_n$ satisfying $G^n \in L^2([0,T]; L^2(\mathbb{S}^1))$, $\mathbb{P}$-almost surely. Suppose there is a stochastic basis $\mathcal{S} = \big(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P}\big)$, a Wiener process W on $\mathcal{S}$, and a predictable $L^2(\mathbb{S}^1)$-valued process G on $\mathcal{S}$ with $G \in L^2((0,T); L^2(\mathbb{S}^1))$ $\mathbb{P}$-almost surely, such that*

$$W^n \overset{n\uparrow\infty}{\longrightarrow} W \text{ in } C([0,T]), \quad G^n \overset{n\uparrow\infty}{\longrightarrow} G \text{ in } L^2([0,T]; L^2(\mathbb{S}^1)), \quad \text{in probability.}$$

Then

$$\int_0^t G^n \,\mathrm{d}W^n \overset{n\uparrow\infty}{\longrightarrow} \int_0^t G \,\mathrm{d}W \quad \text{in } L^2([0,T]; L^2(\mathbb{S}^1)), \text{ in probability.}$$

A sequence $\{v_n\}$ of $\mathbb{B}$-valued random variables is *stochastically bounded* (in $\mathbb{B}$) if

$$\mathbb{P}\big(\|v_n\|_{\mathbb{B}} > M\big) \to 0, \text{ as } M \to \infty, \text{ uniformly in } n, \tag{A.1}$$

here written $v_n \in_{\mathrm{sb}} \mathbb{B}$. A simple approach for proving stochastic boundedness is—via Markov's (or Chebychev's) inequality—to bound $\|v_n\|_{\mathbb{B}}$ in $L^p(\Omega)$, uniformly in n. Denote by $\mu_n := (v_n)_*\mathbb{P}$ the probability law of v_n, i.e., for any $A \in \mathcal{B}(\mathbb{B})$, $\mu_n(A) = (v_n)_*\mathbb{P}(A) := \mathbb{P}\big(X_n \in A\big)$. Stochastic boundedness is equivalent to the requirement that $\mu_n\big(\{v \in \mathbb{B} : \|v\|_{\mathbb{B}} > M\}\big) \to 0$ as $M \to \infty$, uniformly in n. If $\mathbb{B}$ is finite dimensional, this condition is that of tightness of the probability laws $\{\mu_n\}$. If $\mathbb{B}$ is infinite dimensional, or more generally for a topological space $\big(\mathcal{X}, \mathcal{B}(\mathcal{X})\big)$, by *tightness* of a sequence of (Borel) probability measures $\{\mu_n\}$ on $\mathcal{X}$, we mean that for any $\delta > 0$, there is a compact set $K_\delta \subset \mathcal{X}$ such that $\mu_n\big(\mathcal{X} \backslash K_\delta\big) < \delta$, uniformly in n. The identification of a suitable compact set relies on Aubin–Lions–Simon type embedding theorems, see for example [45]. In a separable metric (or even a Hausdorff) space $\mathcal{X}$, by the well-known *Prokhorov theorem*, tightness of the laws $\{\mu_n\}$ implies weak compactness of $\{\mu_n\}$, where we recall that $\{\mu_n\}$ is *weakly (or narrowly) convergent* to μ if $\int_{\mathcal{X}} f\,d\mu_n \overset{n\uparrow\infty}{\longrightarrow} \int_{\mathcal{X}} f\,d\mu$, for all $f \in C_b(\mathcal{X})$, the set of bounded continuous functions. If $\mathcal{X}$ is a Polish space, i.e., a separable completely metrisable topological space, then weak compactness implies tightness.

Finally, we will need the Gyöngy-Krylov characterization of convergence in probability [30]. It will be used to upgrade weak (martingale) solutions to pathwise solutions.

Lemma A.4 (Gyöngy–Krylov)**.** *Let $\mathcal{X}$ be a Polish space. For a sequence $\{v_n\}$ of $\mathcal{X}$-valued random variables define the joint probability laws $\{\mu^{m,n}\}_{m,n}$ by setting, for all $A \in \mathcal{B}(\mathcal{X} \times \mathcal{X})$, $\mu^{m,n}(A) := \mathbb{P}\big(\{(v_m, v_n) \in A\}\big)$. Then the sequence $\{v_n\}$ converges in probability if and only if for every subsequence $\{\mu^{m_k,n_k}\}_k$, there exists* *a further subsequence that converges weakly to a probability measure* μ *supported on the diagonal:* $\mu\big(\{(v,w) \in \mathcal{X} \times \mathcal{X} : v = w\}\big) = 1$.

The fact that the support of the limit of the joint laws $\mu^{m,n}$ in Lemma A.4 lies on the diagonal follows from a pathwise uniqueness property, that is, for two solutions v_a and v_b of the same SPDE sharing the same initial condition, one has

$$\mathbb{P}\big(\{\omega \in \Omega : \|v_a(\omega,t) - v_b(\omega,t)\|_{\mathcal{X}} = 0,\ \forall t \in [0,T]\}\big) = 1.$$

We point out that pathwise uniqueness also implies uniqueness in law [41, Thm. IX.1.7], i.e., that for two weak solutions $\big(v_a, W_a, \mathcal{S}_a\big)$ and $\big(v_b, W_b, \mathcal{S}_b\big)$, with their respective Brownian motions W_a, W_b and stochastic bases $\mathcal{S}_a, \mathcal{S}_b$, one has that the laws of v_a and v_b coincide, i.e., $v_a \sim v_b$.

Generally, to ensure convergence of a sequence of approximate solutions towards a solution for a nonlinear SPDE, it is essential that we secure strong compactness in the ω variable (a.s. convergence). To that end, one often relies on the Skorokhod representation theorem for random variables taking values in a Polish space $\mathcal{X}$, delivering a new probability space and new random variables, with the same laws as the original ones, converging almost surely. In this work, we use the spaces $L^2([0,T]; H^1(\mathbb{S}^1))$ and $C([0,T]; H^1_w(\mathbb{S}^1))$. The former is a Polish space, whereas the latter is not. Here $C([0,T]; H^1_w(\mathbb{S}^1))$ refers to the continuous functions from $[0,T]$ to the Hilbert space $H^1(\mathbb{S}^1)$ equipped with the *weak topology*. This is a locally convex space with the weak topology generated by the system of seminorms $\|v\|_\phi = \sup_{t\in[0,T]} \big|\langle v(t), \phi\rangle_X\big|$, for $\phi \in X := H^1(\mathbb{S}^1)$. Since X is separable and reflexive, the unit ball $B_X \subset X$ is a metrisable compact set and one can equip $C([0,T]; B_X)$ with a complete metric topology induced by the above system of seminorms. On $C([0,T]; H^1_w(\mathbb{S}^1))$ we consider the σ-algebra $\mathcal{B}_T$ generated by the mappings $C([0,T]; H^1_w(\mathbb{S}^1)) \ni v \mapsto v(t) \in X$, $t \in [0,T]$.

Weakly continuous functions taking values in a separable Banach space are not Polish but rather quasi-Polish. *Quasi-Polish* refers to a topological space $\big(\mathcal{X}, \tau\big)$ that asks for point-separability by countably many continuous functions, i.e., that there exists a countable family

$$\{f_\ell : \mathcal{X} \to [-1,1]\}_{\ell\in\mathbb{N}} \tag{A.2}$$

of continuous functions that separate points of $\mathcal{X}$ [35]. In other words, $\mathcal{X}$ is quasi-Polish if $\mathcal{X}$ is a Hausdorff space (but need not be regular) that admits a continuous injection $f(v) = \{f_\ell(v)\}_{\ell\in\mathbb{N}}$ to the Polish space $[-1,1]^{\mathbb{N}}$. The idea behind the proof of the theorem below [35] is to transfer the Skorokhod representation problem via homeomorphism methods to a compact subset of $[-1,1]^{\mathbb{N}}$, where the Skorokhod representation theorem is known to hold, and then map back to $\mathcal{X}$ via f^{-1}, noting that every compact set in $\mathcal{X}$ is $\sigma\big(\{f_\ell\}\big)$-measurable and metriseable. Whenever the σ-algebra $\sigma\big(\{f_\ell\}\big)$ is strictly smaller than the Borel σ-algebra $\mathcal{B}_\tau$, it turns out that every *tight* Borel probability measure on $\big(\mathcal{X}, \tau\big)$ is uniquely determined by its values on $\sigma\big(\{f_\ell\}\big)$ and can be uniquely extended to $\mathcal{B}_\tau$. Besides, f has a continuous inverse (f is a homeomorphic embedding) when restricted to a τ-compact subset of $\mathcal{X}$. As in [9, Cor. 3.12] (see also [10, 11, 39]), one can easily prove that $C([0,T]; H^1_w(\mathbb{S}^1))$ is quasi-Polish, and that the separating sequence $\{f_\ell\}_{\ell\in\mathbb{N}}$ generates the σ-algebra $\mathcal{B}_T$. We refer to [12, Sec. 3] for a discussion collecting relevant properties of quasi-Polish spaces, including $C([0,T]; X_w)$ for an arbitrary separable Hilbert space X.

As the original Skorokhod theorem is not applicable in quasi-Polish spaces, we use the more recent version by Jakubowski [35]. The following form of the theorem is taken from [9, 10, 11, 39], which are some of the first works to employ the theorem to construct martingale solutions of nonlinear SPDEs, including stochastic nonlinear wave equations and the stochastic incompressible Navier–Stokes equations, see also [6] for an application to the compressible Navier–Stokes equations.

Theorem A.5 (Skorokhod–Jakubowski a.s. representations)**.** *Let $(\mathcal{X}, \tau, \mathcal{B}_\tau)$ be a quasi-Polish space, and denote by $\Sigma_f \subset \mathcal{B}_\tau$ the σ-algebra generated by the sequence $\{f_\ell\}$ of continuous functions that separate points. Then*

1. *every τ-compact subset of $\mathcal{X}$ is metrisable;*
2. *every Borel subset of a sigma compact set in $\mathcal{X}$ belongs to Σ_f;*
3. *every probability measure supported by a sigma compact set in $\mathcal{X}$ has a unique Radon extension to the Borel σ-algebra $\mathcal{B}_\tau = \mathcal{B}(\mathcal{X})$.*

Moreover, if $\{\mu_n\}$ is a tight sequence of probability measures on $(\mathcal{X}, \Sigma_f)$, then there exist a subsequence $\{n_k\}_k$, a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, and Borel measurable $\mathcal{X}$-valued random variables $\tilde{v}_k$, $\tilde{v}$, such that μ_{n_k} is the law of $\tilde{v}_k$ and $\tilde{v}_k \to \tilde{v}$ $\tilde{\mathbb{P}}$-a.s. in $\mathcal{X}$. Besides, the law μ of $\tilde{v}$ is a Radon measure on $\mathcal{B}_\tau$.

Proof. See [35, pp. 169–173]. □

A path space for a sequence of variables $\{v_n\}$ defines the topology in which we would like the Skorokhod–Jakubowski representations $\{\tilde{v}_k\}$ to converge (a.s.). It is often important that $\tilde{v}_k \to \tilde{v}$ in multiple spaces/topologies (say, $\mathcal{X}_1$ and $\mathcal{X}_2$). When both spaces $\mathcal{X}_1$ and $\mathcal{X}_2$ are normed spaces, or when one space injects continuously into another, it is often possible to set up a topology on the intersection space $\mathcal{Y} := \mathcal{X}_1 \cap \mathcal{X}_2$ directly that meet two criteria:

(i) $\mathcal{Y}$ is quasi-Polish.
(ii) Compact sets on $\mathcal{Y}$ are sufficiently plentiful; in particular, tightness of laws on $\mathcal{Y}$ can be readily deduced by the separate tightness on $\mathcal{X}_1$ and on $\mathcal{X}_2$.

These criteria are opposed in the sense that a topology on the intersection space $\mathcal{Y}$ stronger than (the subspace topology induced by) each of the topologies on $\mathcal{X}_1$ and $\mathcal{X}_2$ makes it easy to show that $\mathcal{Y}$ is quasi-Polish. One such example is the supremum topology. On the other hand, the strength of the topology placed on $\mathcal{Y}$ makes convergence there more difficult and compact sets harder to come by.

Herein, the difficulty of characterising compact sets on any sufficiently strong topology on the intersection space $\mathcal{Y}$ is side-stepped by finding a.s. representations and limits for $\{(v_n, v_n)\}$ on the product space $\mathcal{X}_1 \times \mathcal{X}_2$, and after that identifying their limits as the same process (Lemma 6.4).

(Countable) products of quasi-Polish spaces are quasi-Polish. We shall apply Theorem A.5 in the product space $\mathcal{X} = \mathcal{X}_1 \times \mathcal{X}_2 \times \mathcal{X}_3 \times \mathcal{X}_4$, where $\mathcal{X}_1$ and $\mathcal{X}_2$ are two path spaces for two copies of the same variable. A Cartesian product of topological spaces is always equipped with the product topology and, thus, the Borel σ-algebra generated by the product topology.

On a product space there are two natural σ-algebras: the product of the Borel σ-algebras and the already introduced Borel σ-algebra for the product topology. Although, in general, these two are not the same, they do coincide on a separable metric space. This implies that coordinatewise measurability and tightness is the same as joint measurability and tightness, which is convenient since we would want to use the product of the Borel σ-algebras in computations leading up to joint tightness and weak convergence in the product space. Whilst the setting of Theorem A.5 goes far beyond separable metric spaces, in applications a priori estimates ensure that the involved random variables take values in a compact set, and then we can rely on (1) and (2) of Theorem A.5.

REFERENCES

[1] S. Albeverio, Z. Brzeźniak and A. Daletskii, Stochastic Camassa–Holm equation with convection type noise, *J. Differential Equations*, **276** (2021), 404-432.

[2] D. Alonso-Orán, C. Rohde and H. Tang, A local-in-time theory for singular sdes with applications to fluid models with transport noise, *Journal of Nonlinear Science*, **31** (2021), 98.

[3] T. M. Bendall, C. J. Cotter and D. D. Holm, Perspectives on the formation of peakons in the stochastic Camassa–Holm equation, *Proc. A.*, **477** (2021), Paper No. 20210224, 18 pp.

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

[5] A. Boritchev, Decaying turbulence in the generalised Burgers equation, *Arch. Ration. Mech. Anal.*, **214** (2014), 331-357.

[6] D. Breit and M. Hofmanová, Stochastic Navier–Stokes equations for compressible fluids, *Indiana Univ. Math. J.*, **65** (2016), 1183-1250.

[7] A. Bressan and A. Constantin, Global conservative solutions of the Camassa–Holm equation, *Arch. Ration. Mech. Anal.*, **183** (2007), 215-239.

[8] A. Bressan and A. Constantin, Global dissipative solutions of the Camassa–Holm equation, *Anal. Appl. (Singap.)*, **5** (2007), 1-27.

[9] Z. Brzeźniak and E. Motyl, Existence of a martingale solution of the stochastic Navier–Stokes equations in unbounded 2D and 3D domains, *J. Differential Equations*, **254** (2013), 1627-1685.

[10] Z. Brzeźniak and M. Ondreját, Weak solutions to stochastic wave equations with values in Riemannian manifolds, *Comm. Partial Differential Equations*, **36** (2011), 1624-1653.

[11] Z. Brzeźniak and M. Ondreját, Stochastic geometric wave equations with values in compact Riemannian homogeneous spaces, *Ann. Probab.*, **41** (2013), 1938-1977.

[12] Z. Brzeźniak, M. Ondreját and J. Seidler, Invariant measures for stochastic nonlinear beam and wave equations, *J. Differential Equations*, **260** (2016), 4157-4179.

[13] R. Camassa and D. D. Holm, An integrable shallow water equation with peaked solitons, *Phys. Rev. Lett.*, **71** (1993), 1661-1664.

[14] P.-L. Chow, *Stochastic Partial Differential Equations*, Advances in Applied Mathematics. CRC Press, Boca Raton, FL, second edition, 2015.

[15] Y. Chen, J. Duan and H. Gao, Global well-posedness of the stochastic Camassa–Holm equation, *Commun. Math. Sci.*, **19** (2021), 607-627.

[16] Y. Chen and H. Gao, Well-posedness and large deviations of the stochastic modified Camassa–Holm equation, *Potential Anal.*, **45** (2016), 331-354.

[17] Y. Chen, H. Gao and B. Guo, Well-posedness for stochastic Camassa–Holm equation, *J. Differential Equations*, **253** (2012), 2353-2379.

[18] Y. Chen and L. Ran, The effect of a noise on the stochastic modified Camassa–Holm equation, *J. Math. Phys.*, **61** (2020), 091504, 16 pp.

[19] G. M. Coclite, H. Holden and K. H. Karlsen, Global weak solutions to a generalized hyperelastic-rod wave equation, *SIAM J. Math. Anal.*, **37** (2005), 1044-1069 (electronic).

[20] A. Constantin and J. Escher, Global existence and blow-up for a shallow water equation, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, **26** (1998), 303-328, `http://www.numdam.org/item/?id=ASNSP_1998_4_26_2_303_0`.

[21] A. Constantin and J. Escher, Wave breaking for nonlinear nonlocal shallow water equations, *Acta Math.*, **181** (1998), 229-243.

[22] D. Crisan and D. D. Holm, Wave breaking for the stochastic Camassa–Holm equation, *Phys. D*, **376/377** (2018), 138-143.

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

[24] A. Debussche, M. Hofmanová and J. Vovelle, Degenerate parabolic stochastic partial differential equations: Quasilinear case, *Ann. Probab.*, **44** (2016), 1916-1955.

[25] R. J. DiPerna and P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces, *Invent. Math.*, **98** (1989), 511-547.

[26] R. Engelking, *General Topology*, Heldermann Verlag, Berlin, second edition, 1989.

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

[28] B. Fuchssteiner and A. S. Fokas, Symplectic structures, their Bäcklund transformations and hereditary symmetries, *Phys. D*, **4** (1981/82), 47-66.

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

[30] I. Gyöngy and N. Krylov, Existence of strong solutions for Itô's stochastic equations via approximations, *Probab. Theory Related Fields*, **105** (1996), 143-158.

[31] H. Holden and X. Raynaud, Global conservative solutions of the Camassa–Holm equation—a Lagrangian point of view, *Comm. Partial Differential Equations*, **32** (2007), 1511-1549.

[32] Z. Huang, H. Tang and Z. Liu, Random attractor for a stochastic viscous coupled Camassa–Holm equation, *J. Inequal. Appl.*, **2013** (2013), 201, 30 pp.

[33] H. Holden, K. H. Karlsen and P. H. C. Pang, The Hunter–Saxton equation with noise, *J. Differential Equations*, **270** (2021), 725-786.

[34] D. D. Holm, Variational principles for stochastic fluid dynamics, *Proc. A.*, **471** (2015), 20140963, 19 pp.

[35] A. Jakubowski, The almost sure Skorokhod representation for subsequences in nonmetric spaces, *Theory Probab. Appl.*, **42** (1997), 167-174.

[36] A. S. Kechris, *Classical Descriptive Set Theory*, Springer-Verlag, New York, 1995.

[37] W. Liu and M. Röckner, *Stochastic Partial Differential Equations: An Introduction*, Universitext. Springer, 2015.

[38] W. Lv, P. He and Q. Wang, Well-posedness and blow-up solution for the stochastic Dullin–Gottwald–Holm equation, *J. Math. Phys.*, **60** (2019), 083513, 10 pp.

[39] M. Ondreját, Stochastic nonlinear wave equations in local Sobolev spaces, *Electron. J. Probab.*, **15** (2010), 1041-1091.

[40] S. Punshon-Smith and S. Smith, On the Boltzmann equation with stochastic kinetic transport: Global existence of renormalized martingale solutions, *Arch. Ration. Mech. Anal.*, **229** (2018), 627-708.

[41] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, Springer, Berlin, third edition, 1999.

[42] C. Rohde and H. Tang, On the stochastic Dullin–Gottwald–Holm equation: Global existence and wave-breaking phenomena, *NoDEA Nonlinear Differential Equations Appl.*, **28** (2021), 34 pp.

[43] C. Rohde and H. Tang, On a stochastic Camassa–Holm type equation with higher order nonlinearities, *Journal of Dynamics and Differential Equations*, **33** (2021), 1823-1852.

[44] M. Scheutzow, A stochastic Gronwall lemma, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, **16** (2013), 1350019, 4 pp.

[45] J. Simon, Compact sets in the space $L^p(0,T;B)$, *Ann. Mat. Pura Appl.*, **146** (1987), 65-96.

[46] R. Situ, *Theory of Stochastic Differential Equations with Jumps and Applications*, Springer, New York, 2005.

[47] H. Tang, On the pathwise solutions to the Camassa–Holm equation with multiplicative noise, *SIAM J. Math. Anal.*, **50** (2018), 1322-1366.

[48] H. Tang, Noise effects on dependence on initial data and blow-up for stochastic Euler–Poincaré equations, `arXiv:2002.08719v2`, 2020.

[49] R. Temam, *Navier–Stokes Equations and Nonlinear Functional Analysis*, SIAM, Philadelphia, PA, second edition, 1995.

[50] L. Xie and X. Zhang, Ergodicity of stochastic differential equations with jumps and singular coefficients, *Ann. Inst. Henri Poincaré Probab. Stat.*, **56** (2020), 175-229.

[51] Z. Xin and P. Zhang, On the weak solutions to a shallow water equation, *Comm. Pure Appl. Math.*, **53** (2000), 1411-1433.

[52] L. Zhang, Local and global pathwise solutions for a stochastically perturbed nonlinear dispersive PDE, *Stochastic Process. Appl.*, **130** (2020), 6319-6363.

[53] L. Zhang, Effect of random noise on solutions to the modified two-component Camassa–Holm system on $\mathbb{T}^d$, `arXiv:2107.09603`, 2021.

Received September 2022; revised October 2022; early access November 2022.