The Annals of Probability
2015, Vol. 43, No. 3, 1202–1273
DOI: 10.1214/13-AOP887

SUB AND SUPERCRITICAL STOCHASTIC QUASI-GEOSTROPHIC EQUATION[1]

BY MICHAEL RÖCKNER, RONGCHAN ZHU[2] AND XIANGCHAN ZHU

University of Bielefeld, Beijing Institute of Technology and Beijing Jiaotong University

In this paper, we study the 2D stochastic quasi-geostrophic equation on $\mathbb{T}^2$ for general parameter $\alpha \in (0, 1)$ and multiplicative noise. We prove the existence of weak solutions and Markov selections for multiplicative noise for all $\alpha \in (0, 1)$. In the subcritical case $\alpha > 1/2$, we prove existence and uniqueness of (probabilistically) strong solutions. Moreover, we prove ergodicity for the solution of the stochastic quasi-geostrophic equations in the subcritical case driven by possibly degenerate noise. The law of large numbers for the solution of the stochastic quasi-geostrophic equations in the subcritical case is also established. In the case of nondegenerate noise and $\alpha > 2/3$ in addition exponential ergodicity is proved.

1. Introduction. Consider the following two-dimensional (2D) stochastic quasi-geostrophic equation in the periodic domain $\mathbb{T}^2 = \mathbb{R}^2/(2\pi\mathbb{Z})^2$:

$$\frac{\partial \theta(t,\xi)}{\partial t} = -u(t,\xi)\cdot\nabla\theta(t,\xi) - \kappa(-\triangle)^\alpha\theta(t,\xi) + \bigl(G(\theta)\eta\bigr)(t,\xi), \tag{1.1}$$

with initial condition

$$\theta(0,\xi) = \theta_0(\xi), \tag{1.2}$$

where $\theta(t,\xi)$ is a real-valued function of $\xi \in \mathbb{T}^2$ and $t \geq 0$, $0 < \alpha < 1, \kappa > 0$ are real numbers. u is determined by θ via the following relation:

$$u = (u_1, u_2) = (-R_2\theta, R_1\theta) = R^\perp\theta. \tag{1.3}$$

Here, R_j is the jth periodic Riesz transform and $\eta(t,\xi)$ is a Gaussian random field, white noise in time, subject to the restrictions imposed below. The case $\alpha = \frac{1}{2}$ is called the critical case, the case $\alpha > \frac{1}{2}$ subcritical and the case $\alpha < \frac{1}{2}$ supercritical.

Received February 2013.

[1]Research supported in part by NSFC (11301026) and China Postdoctoral Science Foundation funded project (2012M520153) and DFG through IRTG 1132 and CRC 701.

[2]Corresponding author.

MSC2010 subject classifications. 60H15, 60H30, 35R60.

Key words and phrases. Stochastic quasi-geostrophic equation, well posedness, martingale problem, Markov property, strong Feller property, Markov selections, ergodicity for the subcritical case, degenerate noise.

In the deterministic case ($G \equiv 0$), such equations are important models in geophysical fluid dynamics. Indeed, they are special cases of general quasi-geostrophic approximations for atmospheric and oceanic fluid flows with small Rossby and Ekman numbers. These models arise under the assumptions of fast rotation, uniform stratification and uniform potential vorticity. The case $\alpha = 1/2$ exhibits similar features (singularities) as the 3D Navier–Stokes equations and can therefore serve as a model case for the latter. For more details about the geophysical background, see, for instance, [6, 42]. In the deterministic case, this equation has been intensively investigated because of both its mathematical importance and its background in geophysical fluid dynamics (see, e.g., [5, 7, 8, 23–26, 44] and the references therein). In the deterministic case, the global existence of weak solutions has been obtained in [44] and one most remarkable result in [5] gives the existence of a classical solution for $\alpha = 1/2$. In [26], another very important result is proved, namely that solutions for $\alpha = 1/2$ with periodic C^∞ data remain C^∞ for all times.

There is another model considering a simplified geophysical fluid model at asymptotically high rotation rate or with small Rossby number. This geophysical model with random perturbation has been studied in [2, 22] and the references therein. The equation is of a different type compared with our equation.

In this paper, we study the 2D stochastic quasi-geostrophic equation on the torus $\mathbb{T}^2$ for general parameter $\alpha \in (0, 1)$ and for both additive as well as multiplicative noise. Here, since the dissipation term is not strong enough to control the nonlinear term, we have to work in L^p and to prove appropriate L^p-norm estimates. This leads to considerable complications in comparison to the stochastic Navier–Stokes equation, for example, when one wants to prove L^p-norm estimates for the weak solutions (see Theorem 3.3), which are essential to obtain pathwise uniqueness, and the improved positivity lemma to obtain uniform L^p-norm estimates (see Lemma 5.5 and Proposition 5.6) which will be used to prove ergodicity.

Main results for general $\alpha \in (0, 1)$: We prove the existence of weak solutions for multiplicative noise (Theorem 3.3). In order to prove the existence of (probabilistically strong) solutions and ergodicity in subsequent sections, we need L^p norm estimates for the solutions, which are obtained using the L^p-Itô formula proved in [29]. But these L^p-norm estimates we cannot prove by Galerkin approximation; instead, we use another approximation which can be seen as a piecewise linear equation on small subintervals [see (3.4)]. To piece together martingale solutions on each subinterval and to get the existence of a martingale solution for the approximation, we first use the measurable selection theorem to find a martingale solution measurable with respect to the initial condition and apply a classical theorem from [48] (see Theorem 3.2). Using an abstract result for obtaining Markov selections from [20], we prove the existence of an a.s. Markov family in Appendix C (Theorem C.5).

Main results for the subcritical case $\alpha > 1/2$: We obtain pathwise uniqueness in a larger space by using L^p-norm estimates (Theorem 4.2) and, therefore, get a (probabilistically strong) solution (Theorem 4.3) by the Yamada–Watanabe theorem. In particular, it follows that the laws of the solutions form a Markov process. Subsequently, in Section 5 we use a coupling method to study the long time behavior of the solution for the 2D stochastic quasi-geostrophic equation and we obtain ergodicity, that is, the existence (Theorem 5.12) and uniqueness (Theorem 5.9) of an invariant measure, for the solution to the 2D stochastic quasi-geostrophic equation (in case $\alpha > 1/2$) driven by possibly degenerate noise. Furthermore, the Markov semigroup P_t converges to the unique invariant measure polynomially fast (Theorem 5.13). Finally, we prove that a law of large numbers holds in our case, that is, the times averages $\frac{1}{T}\int_0^T \psi(\theta_t)\,dt$ converge to a constant in probability if $\psi : H^1 \mapsto \mathbb{R}$ is smooth (Theorem 5.14).

We add a detailed discussion on our approach to ergodicity via coupling, in particular, on its justification and on its relation to other approaches in Remark 5.10 below. In this paper, we are inspired by [40] to construct an intermediate process $\tilde{\theta}$ such that $\theta - \tilde{\theta}$ has a strong dissipation term and $\|\theta(t) - \tilde{\theta}(t)\|_{H^{-1/2}} \to 0$ as $t \to \infty$. Using this intermediate process, we can prove $E\|\theta_1(t, \theta_0^1, \theta_0^2) - \theta_2(t, \theta_0^1, \theta_0^2)\|_{H^{-1/2}}$ converges to zero polynomially fast when time goes to infinity, where $(\theta_1(t, \theta_0^1, \theta_0^2), \theta_2(t, \theta_0^1, \theta_0^2))$ denotes a coupling of two solutions to (3.1) starting from two different initial values $\theta_0^i \in H^1, i = 1, 2$. Then we can deduce the uniqueness of invariant measures (Theorem 5.9). Also by a suitable choice of the metrics the asymptotically strong Feller property of the semigroup associated with the solution to the 2D stochastic quasi-geostrophic equation is also established (Remark 5.10). Here, we want to emphasize that although we consider the semigroup in H^1, the convergence is in $H^{-1/2}$ norm. Moreover, we obtain the existence of the invariant measure, which lives on H^1, by using the uniform L^p-estimates (Theorem 5.12), which require the improved positivity lemma (Lemma 5.5). Thus, we obtain ergodicity for the solution of the quasi-geostrophic equation in the subcritical case (Theorem 5.13).

Additional results in the subcritical case $\alpha > 2/3$: In Section 6, we prove the exponential convergence of the solution under a stronger condition on the noise and on α. In order to prove the exponential convergence (Theorem 6.13), we first show the strong Feller property of the associated semigroup (Theorem 6.3), which follows from employing the weak-strong uniqueness principle in [18] (Theorem 6.4) and the Bismut–Elworthy–Li formula. As the dynamics only exist in the (analytically) weak sense and standard tools of stochastic analysis are not available, the computations are made for an approximating cutoff dynamics, which are equal to the original dynamics on a small random time interval. Since in our case $\alpha < 1$, it is more difficult to use the H^α-norm to control the nonlinear term even though the equation is on $\mathbb{T}^2$. To prove the weak-strong uniqueness principle, we need some regularity for the trajectories of the noise. Therefore, we need conditions on G so that it is enough regularizing. However, in order to apply the Bismut–Elworthy–Li formula, we also need G^{-1} to be regularizing enough. As a result, $\alpha > 2/3$ is required (see Remark 6.2 below for details). It seems difficult to use the Kolmogorov equation method as in [10, 14] (see Remark 6.2 below).

This paper is organized as follows. In Section 2, we introduce some notation as preparation. In Section 3, we prove the existence of weak solutions for general parameter $\alpha \in (0, 1)$ and multiplicative noise. In Section 4, we prove pathwise uniqueness for all $\alpha \in (\frac{1}{2}, 1)$. Furthermore, we get the existence and uniqueness of (probabilistically strong) solutions for multiplicative noise in the subcritical case. Moreover, we prove the Markov property for this unique solution. In Section 5, we use the coupling method to prove the uniqueness of an invariant measure in the subcritical case. Moreover, we obtain that the semigroup P_t converges to the invariant measure polynomially fast. The law of large numbers for the solution to the 2D stochastic quasi-geostrophic equation is also established in this section. In Section 6, for $\alpha > 2/3$, and provided the noise is nondegenerate, we prove the exponential convergence to the (unique) invariant measure. Appendix A is devoted to a measurability problem (see Theorem A.4) which arises in implementing the coupling method in Section 5. In Appendix B, we prove existence of measurable selections for the solutions to the martingale problem in Section 3, and finally Appendix C is devoted to the existence of the corresponding Markov selection.

2. Notations and preliminaries. In the following, we will restrict ourselves to flows which have zero average on the torus $\mathbb{T}^2$, that is,

$$\int_{\mathbb{T}^2} \theta \, d\xi = 0,$$

where $d\xi$ denotes the volume measure on $\mathbb{T}^2$. Thus, (1.3) can be restated as

$$u = \left(-\frac{\partial \psi}{\partial \xi_2}, \frac{\partial \psi}{\partial \xi_1}\right) \quad \text{and} \quad (-\Delta)^{1/2}\psi = -\theta.$$

Set $H = \{f \in L^2(\mathbb{T}^2) : \int_{\mathbb{T}^2} f \, d\xi = 0\}$ and let $|\cdot|$ and $\langle \cdot, \cdot \rangle$ denote the norm and inner product in H, respectively. $L^p(\mathbb{T}^2)$, $p \in (0, \infty]$ denote the standard L^p spaces on $\mathbb{T}^2$ with norm $\|\cdot\|_{L^p}$. On the periodic domain $\mathbb{T}^2$, $\{\sin\langle k, \cdot\rangle_{\mathbb{R}^2} | k \in \mathbb{Z}^2_+\} \cup \{\cos\langle k, \cdot\rangle_{\mathbb{R}^2} | k \in \mathbb{Z}^2_-\}$ form an eigenbasis of $-\Delta$ (we denote it by $\{e_k\}$). Here, $\mathbb{Z}^2_+ = \{(k_1, k_2) \in \mathbb{Z}^2 | k_2 > 0\} \cup \{(k_1, 0) \in \mathbb{Z}^2 | k_1 > 0\}$, $\mathbb{Z}^2_- = \{(k_1, k_2) \in \mathbb{Z}^2 | -k \in \mathbb{Z}^2_+\}$, $\xi \in \mathbb{T}^2$, and the corresponding eigenvalues are $|k|^2$. For $s > 0$, define

$$\|f\|^2_{H^s} = \sum_k |k|^{2s} \langle f, e_k\rangle^2$$

and let H^s denote the Sobolev space of all $f \in H$ for which $\|f\|_{H^s}$ is finite. For $s < 0$, define H^s to be the dual of H^{-s}. Set $\Lambda = (-\Delta)^{1/2}$. Then

$$\|f\|_{H^s} = |\Lambda^s f|.$$

For $s \geq 0$, $p \in [1, +\infty]$ we use $H^{s,p}$ to denote a subspace of $L^p(\mathbb{T}^2)$, consisting of all f which can be written in the form $f = \Lambda^{-s} g$, $g \in L^p(\mathbb{T}^2)$ and the $H^{s,p}$ norm of f is defined to be the L^p norm of g, that is, $\|f\|_{H^{s,p}} := \|\Lambda^s f\|_{L^p}$.

By the singular integral theory of Calderón and Zygmund (cf. [47], Chapter 3), for any $s \geq 0$, $p \in (1, \infty)$, there is a constant $C_R = C_R(s, p)$, such that

$$\|\Lambda^s u\|_{L^p} \leq C_R(s, p)\|\Lambda^s \theta\|_{L^p}. \tag{2.1}$$

Fix $\alpha \in (0, 1)$ and define the linear operator $A_\alpha : D(A_\alpha) = H^{2\alpha}(\mathbb{T}^2) \subset H \to H$ as $A_\alpha u := \kappa(-\Delta)^\alpha u$. The operator A_α is positive definite and self-adjoint with the same eigenbasis as that of $-\Delta$ mentioned above. Denote the eigenvalues of A_α by $0 < \lambda_1 \leq \lambda_2 \leq \cdots$, and renumber the above eigenbasis correspondingly as $e_1, e_2, \ldots$.

First, we recall the following important product estimates (cf. [44], Lemma A.4):

LEMMA 2.1. *Suppose that $s > 0$ and $p \in (1, \infty)$. If $f, g \in C^\infty(\mathbb{T}^2)$ then*

$$\|\Lambda^s(fg)\|_{L^p} \leq C\big(\|f\|_{L^{p_1}}\|\Lambda^s g\|_{L^{p_2}} + \|g\|_{L^{p_3}}\|\Lambda^s f\|_{L^{p_4}}\big), \tag{2.2}$$

with $p_i \in (1, \infty]$, $i = 1, \ldots, 4$ such that

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p_3} + \frac{1}{p_4}.$$

We shall use as well the following standard Sobolev inequality (cf. [47], Chapter V):

LEMMA 2.2. *Suppose that $q > 1$, $p \in [q, \infty)$ and*

$$\frac{1}{p} + \frac{\sigma}{2} = \frac{1}{q}.$$

Suppose that $\Lambda^\sigma f \in L^q$, then $f \in L^p$ and there is a constant $C_S \geq 0$ independent of f such that

$$\|f\|_{L^p} \leq C_S\|\Lambda^\sigma f\|_{L^q}.$$

The following commutator estimate from [23], Lemma 3.1, is very important for later use.

LEMMA 2.3 (Commutator estimates). *Suppose that $s > 0$ and $p \in (1, \infty)$. If $f, g \in C^\infty(\mathbb{T}^2)$, then*

$$\|\Lambda^s(fg) - f\Lambda^s g\|_{L^p} \leq C\big(\|\nabla f\|_{L^{p_1}}\|\Lambda^{s-1} g\|_{L^{p_2}} + \|g\|_{L^{p_3}}\|\Lambda^s f\|_{L^{p_4}}\big),$$

with $p_i \in (1, \infty)$, $i = 1, \ldots, 4$ such that

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p_3} + \frac{1}{p_4}.$$

We will also use the following classical interpolation inequality (see, e.g., [9], (5.5)).

LEMMA 2.4. *For $f \in C^\infty(\mathbb{T}^2)$, we have*

$$\|f\|_{H^s} \le C\|f\|_{H^{s_1}}^{(s_2-s)/(s_2-s_1)}\|f\|_{H^{s_2}}^{(s-s_1)/(s_2-s_1)}, \qquad s_1 < s < s_2. \tag{2.3}$$

3. Weak solutions in the general case.

In this section, we consider the following abstract stochastic evolution equation in place of equations (1.1)–(1.3):

$$\begin{cases} d\theta(t) + A_\alpha\theta(t)\,dt + u(t)\cdot\nabla\theta(t)\,dt = G(\theta(t))\,dW(t), \\ \theta(0) = \theta_0 \in H, \end{cases} \tag{3.1}$$

where u satisfies (1.3) and $W(t)$, $t \in [0,T]$, is a cylindrical Wiener process in a separable Hilbert space U defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, P)$. Here, G is a measurable mapping from H^α to $L_2(U,H)$ (= all Hilbert–Schimit operators from U to H). Let f_n, $n \in \mathbb{N}$, be an ONB of U.

In the following, we assume the following conditions on G:

HYPOTHESIS G.1. (i) $\|G(\theta)\|^2_{L_2(U,H)} \le \lambda_0|\theta|^2 + \rho_1|\Lambda^\alpha\theta|^2 + \rho_2$, $\theta \in H^\alpha$, *for some positive real numbers* λ_0, ρ_2 *and* $\rho_1 < 2\kappa$. *Moreover, for some* $\beta > 3$, $\|G(\theta)\|^2_{L_2(U,H^{-\beta})} \le \rho_3(|\theta|^2+1)$, $\theta \in H^\alpha$, *for some positive real numbers* ρ_3.

(ii) *If* $\theta, \theta_n \in H^\alpha$ *such that* $\theta_n \to \theta$ *in* H, *then* $\lim_{n\to\infty}\|G(\theta_n)^*(v) - G(\theta)^*(v)\|_U = 0$ *for all* $v \in C^\infty(\mathbb{T}^2)$, *where the asterisk denotes the adjoint operator of* $G(\theta)$.

First, we introduce the following definition of a weak solution.

DEFINITION 3.1. We say that there exists a weak solution of equation (3.1) if there exists a stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, P)$, a cylindrical Wiener process W on the space U and a progressively measurable process $\theta : [0,T]\times\Omega \to H$, such that for P-a.e. $\omega \in \Omega$,

$$\theta(\cdot,\omega) \in L^\infty([0,T];H) \cap L^2([0,T];H^\alpha) \cap C([0,T];H^{-\beta}),$$

where β in Hypothesis G.1, and such that P-a.s.

$$\begin{aligned} \langle\theta(t),\phi\rangle &+ \int_0^t \langle A_\alpha^{1/2}\theta(s), A_\alpha^{1/2}\phi\rangle\,ds - \int_0^t \langle u(s)\cdot\nabla\phi, \theta(s)\rangle\,ds \\ &= \langle\theta_0,\phi\rangle + \left\langle \int_0^t G(\theta(s))\,dW(s), \phi\right\rangle \end{aligned}$$

for $t \in [0,T]$ and all $\phi \in C^1(\mathbb{T}^2)$.

REMARK. (i) Note that, because $\operatorname{div} u = 0$ for smooth functions θ and ψ, we have

$$\langle u(s) \cdot \nabla\theta(s), \psi\rangle = -\langle u(s) \cdot \nabla\psi, \theta(s)\rangle.$$

Thus, the integral equation in Definition 3.1 corresponds to equation (3.1).

(ii) Note that since the solution $\theta \in L^2(0, T; H^\alpha)$ we only need $\theta, \theta_n \in H^\alpha$ instead of $\theta, \theta_n \in H$ in Hypothesis G.1(ii).

(iii) A typical example satisfying Hypothesis G.1 is the following: For $y \in U$,

$$G(\theta)y = \sum_{k=1}^{\infty} \big(c_k \Lambda^\alpha \theta + b_k g(\theta)\big)\langle y, f_k\rangle_U, \qquad \theta \in H^\alpha,$$

where g is continuous function on $\mathbb{R}$ of at most linear growth and $b_k, c_k \in C^\infty(\mathbb{T}^2)$ satisfy $\sum_k c_k^2(\xi) < 2\kappa$, $\sum_k b_k^2(\xi) \leq M$, $\xi \in \mathbb{T}^2$, and $\sum_k |\Lambda^\alpha c_k|^2 \leq M$.

It is standard to show that under Hypothesis G.1 there exists a weak solution to (3.1) by using the Galerkin approximation. However, as mentioned in the Introduction, we also need L^p norm estimates for the solutions, more precise that they belong to $L^p(\Omega; L^\infty([0, T]); L^p(\mathbb{T}^2))$, provided so do their initial values. This will be essential to the proof of pathwise uniqueness. For this, we have to use another approximation instead of the Galerkin approximation and the following theorem from [48], Theorem 6.1.2.

Let $\Omega_0 := C([0, \infty), H^1)$, $\Omega_0^t := C([t, \infty), H^1)$ for $t > 0$ and $\mathcal{P}(\Omega_0)$ denote the set of all probability measures on $(\Omega_0, \mathcal{B})$ with $\mathcal{B}$ being the Borel σ-algebra coming from the topology of locally uniform convergence on Ω_0. Define the canonical process $x : \Omega_0 \to H^1$ as

$$x_t(\omega) = \omega(t).$$

Also define the σ-algebra $\mathcal{B}_t := \sigma\{x(s), s \leq t\}$ and $\mathcal{B}^t := \sigma\{x(s), s \geq t\}$.

THEOREM 3.2. *Fix $t > 0$. Let $x \mapsto Q_x$ be a mapping from Ω_0 to $\mathcal{P}(\Omega_0^t)$ such that for any $A \in \mathcal{B}^t$, $x \mapsto Q_x(A)$ is $\mathcal{B}_t$-measurable, and for any $x \in \Omega_0$*

$$Q_x\big(y \in \Omega_0^t : y(t) = x(t)\big) = 1.$$

Then for any $P \in \mathcal{P}(\Omega_0)$, there exists a unique $P \otimes_t Q \in \mathcal{P}(\Omega_0)$ such that

$$(P \otimes_t Q)(A) = P(A), \qquad \forall A \in \mathcal{B}_t,$$

and for $P \otimes_t Q$-almost all $x \in \Omega_0$

$$Q_x = (P \otimes_t Q)(\cdot|\mathcal{B}_t)(x).$$

Now we will prove the existence of a martingale solution under Hypothesis G.1.

THEOREM 3.3. *Let* $\alpha \in (0,1)$. *If* G *satisfies Hypothesis* G.1, *then there exists a weak solution* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P, W, \theta)$ *to* (3.1). *Moreover, assume that* G *satisfies the following condition*:

(Gp.1) *There exists some* $p \in (2,\infty)$ *such that for all* $\theta \in H^\alpha \cap L^p(\mathbb{T}^2)$,

$$\int \Big(\sum_j |G(\theta)(f_j)|^2\Big)^{p/2} d\xi \le C\Big(\int |\theta|^p \, d\xi + 1\Big), \qquad \forall t > 0 \tag{3.2}$$

for some constant $C := C(p) > 0$ *and* $\theta_0 \in L^p(\mathbb{T}^2)$. *Then*

$$E \sup_{t\in[0,T]} \|\theta(t)\|_{L^p}^p < \infty.$$

REMARK 3.4. Typical examples for G satisfying (Gp.1) have the following form: for $\theta \in H^\alpha$

$$G(\theta)y = \sum_{k=1}^\infty b_k \langle y, f_k\rangle_U g(\theta), \qquad y \in U,$$

where g is a continuous function on $\mathbb{R}$ of at most linear growth and b_k are C^∞ functions on $\mathbb{T}^2$ satisfying $\sum_{k=1}^\infty b_k^2(\xi) \le M$.

PROOF OF THEOREM 3.3. *Step* 1: We first establish the existence of martingale solutions of the following equation:

$$\begin{aligned} &d\theta(t) + A_\alpha\theta(t)\,dt + w(t)\cdot\nabla\theta(t)\,dt = k_\delta * G(\theta)\,dW(t),\\ &\theta(0) = \theta_0 \in H^3, \end{aligned} \tag{3.3}$$

with a given smooth function $w(t)$ which satisfies $\operatorname{div} w(t) = 0$ for all $t \in [0,T]$ and

$$\sup_{t\in[0,T]} \|w(t)\|_{C^3(\mathbb{T}^2)} \le C.$$

Here, $k_\delta * G(\theta)$ means for $y \in U$, $k_\delta * G(\theta)(y) := k_\delta * (G(\theta)(y))$, where k_δ is the periodic Poisson kernel in $\mathbb{T}^2$ given by $\widehat{k_\delta}(\zeta) = e^{-\delta|\zeta|}$, $\zeta \in \mathbb{Z}^2$. By [20], Theorem 4.7, this equation has a martingale solution $P \in \mathcal{P}(C([0,\infty); H^1))$ with initial value θ_0 in the following sense:

(M1) $P(x(0) = \theta_0) = 1$ and for any $n \in \mathbb{N}$

$$P\Big\{x \in C([0,\infty); H^1) : \int_0^n \|\Lambda^{2\alpha}x(s) + w(s)\cdot\nabla x(s)\|_{H^1}\,ds + \int_0^n \|k_\delta * G(x(s))\|_{L_2(U;H^3)}^2\,ds < +\infty\Big\} = 1.$$

(M2) For every e_i, the process

$$\langle x(t), e_i\rangle - \int_0^t \langle -w(s)\cdot\nabla x(s) - A_\alpha x(s), e_i\rangle\, ds$$

is a continuous square-integrable $\mathcal{B}_t$-martingale under P, whose quadratic variation process is given by

$$\int_0^t \|(k_\delta * G)^*(x(s))(e_i)\|_U^2\, ds,$$

where the asterisk denotes the adjoint operator of $k_\delta * G(x(s))$.

(M3) For any $q \in \mathbb{N}$ there exists a continuous positive real function $t \to C_{t,q}$ such that

$$E^P\Big(\sup_{r\in[0,t]} |\Lambda^3 x(r)|^{2q} + \int_0^t |\Lambda^3 x(r)|^{2q-2} |\Lambda^{\alpha+3} x(r)|^2\, dr\Big) \le C_{t,q}(|\Lambda^3\theta_0|^{2q} + 1),$$

where E^P denotes the expectation under P.

Indeed, we only need to check conditions (C1)–(C3) in [20]. The demicontinuity condition (C1) is obvious by Hypothesis G.1(ii) and the linearity of the equation. For (C2), we have that for $x \in H^4$

$$\langle -w\cdot\nabla x - A_\alpha x, x\rangle_{H^3} \le -\kappa |\Lambda^{3+\alpha} x|^2 + |\langle \Lambda^3(w\cdot\nabla x), \Lambda^3 x\rangle|.$$

By Lemma 2.3 and because $\langle w\cdot\nabla\Lambda^3 x, \Lambda^3 x\rangle = 0$ for $x \in H^4$ we have that for $x \in H^4$

$$\begin{aligned}|\langle \Lambda^3(w\cdot\nabla x), \Lambda^3 x\rangle| &= |\langle \Lambda^3(w\cdot\nabla x) - w\cdot\nabla\Lambda^3 x, \Lambda^3 x\rangle| \\ &\le \|w\|_{C^3(\mathbb{T}^2)} |\Lambda^3 x||\Lambda^{3+\alpha} x|.\end{aligned}$$

Thus, the coercivity condition (C2) follows from the above two inequalities and Young's inequality. Also by Hypothesis G.1, we have for $x \in H^4$

$$\|k_\delta * G(x)\|_{L_2(U,H^3)}^2 \le C(\delta)\|G(x)\|_{L_2(U,H)}^2 \le C(|\Lambda^\alpha x|^2 + 1),$$

and by Lemma 2.1

$$\|w\cdot\nabla x + A_\alpha x\|_{H^1}^2 \le 2\|A_\alpha x\|_{H^1}^2 + C\|w\|_{C^3(\mathbb{T}^2)}^2 |\Lambda^3 x|^2 \le C|\Lambda^3 x|^2,$$

which implies the growth condition (C3).

Step 2: Now we construct an approximation of (3.1).

We pick a smooth $\phi \ge 0$, with $\operatorname{supp}\phi \subset [1,2]$, $\int_0^\infty \phi = 1$, and for $\delta > 0$ let

$$U_\delta[\theta](t) := \int_0^\infty \phi(\tau)(k_\delta * R^\perp\theta)(t - \delta\tau)\, d\tau,$$

where k_δ is the periodic Poisson kernel in $\mathbb{T}^2$ given by $\widehat{k_\delta}(\zeta) = e^{-\delta|\zeta|}$, $\zeta \in \mathbb{Z}^2$, and we set $\theta(t) = 0$, $t < 0$. We take a sequence $\delta_n \to 0$ and consider the equation

$$d\theta_n(t) + A_\alpha \theta_n(t)\,dt + u_n(t) \cdot \nabla \theta_n(t)\,dt = k_{\delta_n} * G(\theta_n)\,dW(t), \tag{3.4}$$

with initial data $\theta_n(0) = k_{\delta_n} * \theta_0$ and $u_n = U_{\delta_n}[\theta_n]$. For a fixed n, this is a linear equation in θ_n on each subinterval $[t_k^n, t_{k+1}^n]$ with $t_k^n = k\delta_n$, since u_n is determined by the values of θ_n on the two previous subintervals. By Step 1, we obtain the existence of a martingale solution to (3.4) for fixed n. Indeed, we obtain the martingale solution $P_n^1 \in \mathcal{P}(C([0,\infty), H^1))$ with initial condition $k_{\delta_n} * \theta_0$ on the subinterval $[0, t_1^n]$ by Step 1. Also, by Step 1, we get that for $x_0 \in B_0$ with $B_0 := \{x \in \Omega_0 : \sup_{0 \le t \le t_1^n} \|x(t)\|_{H^3} < \infty\}$, there exists a $Q_{x_0} \in \mathcal{P}(C([t_1^n, t_2^n], H^1))$ satisfying the following:

(M1) $Q_{x_0}(x(t_1^n) = x_0(t_1^n)) = 1$

$$Q_{x_0}\Big\{x \in C([t_1^n, t_2^n]; H^1) : \int_{t_1^n}^{t_2^n} \|\Lambda^{2\alpha} x(s) + U_{\delta_n}[x_0](s) \cdot \nabla x(s)\|_{H^1}\,ds + \int_{t_1^n}^{t_2^n} \|k_\delta * G(x(s))\|_{L_2(U;H^3)}^2\,ds < +\infty\Big\} = 1.$$

(M2) For every e_i, $i \in \mathbb{N}$, the process

$$M_i(t \wedge t_2^n, x) := \langle x(t \wedge t_2^n), e_i\rangle - \langle x_0(t_1^n), e_i\rangle - \int_{t_1^n}^{t \wedge t_2^n} \langle -U_{\delta_n}[x_0](s) \cdot \nabla x(s) - A_\alpha x, e_i\rangle\,ds, \qquad t \ge t_1^n$$

is a continuous square-integrable $\mathcal{B}_t$-martingale under Q_{x_0}, whose quadratic variation process is given by

$$\langle M_i\rangle(t \wedge t_2^n, x) := \int_{t_1^n}^{t \wedge t_2^n} \|(k_{\delta_n} * G)^*(x(s))(e_i)\|_U^2\,ds,$$

where the asterisk denotes the adjoint operator of $k_{\delta_n} * G(x(s))$.

(M3) For any $q \in \mathbb{N}$, there exists a constant C_q depending on $\sup_{t \in [0,t_1^n]} \|x_0(t)\|_{H^1}$ such that

$$E^{Q_{x_0}}\Big(\sup_{r \in [t_1^n, t_2^n]} |\Lambda^3 x(r)|^{2q} + \int_{t_1^n}^{t_2^n} |\Lambda^3 x(r)|^{2q-2} |\Lambda^{\alpha+3} x(r)|^2\,dr\Big) \le C_q(|\Lambda^3 x_0(t_1^n)|^{2q} + 1).$$

Now we extend Q_{x_0} to a probability measure on $C([t_1^n, +\infty), H^1)$ by $Q_{x_0} \circ \psi^{-1}$ with $\psi : C([t_1^n, t_2^n], H^1) \to C([t_1^n, +\infty), H^1)$ by $\psi x(s) := x(s \wedge t_2^n)$, $s \in [t_1^n, +\infty)$. The set of all such martingale solutions is denoted by $\mathcal{Q}_{x_0}$. Now we can find $Q_{x_0} \in \mathcal{Q}_{x_0}$ satisfying (M1)–(M3) such that the map $x_0 \mapsto Q_{x_0}$ from B_0

to $\mathcal{P}(\Omega_0^{t_1^n})$ is measurable with respect to $\mathcal{B}_{t_1^n}$. This will be proved in Lemma B.1 in Appendix B.

For $x_0 \in B_0^c$, define $Q_{x_0} := \delta_{x_0|_{[t_1^n,\infty)}}$. Thus, by Theorem 3.2 we get that there exists $P_n^1 \otimes_{t_1^n} Q \in \mathcal{P}(C([0,\infty), H^1))$ such that

$$\big(P_n^1 \otimes_{t_1^n} Q\big)(A) = P_n^1(A), \qquad \forall A \in \mathcal{B}_{t_1^n},$$

and for $P_n^1 \otimes_{t_1^n} Q$-almost all $x \in \Omega_0$

$$Q_x = \big(P_n^1 \otimes_{t_1^n} Q\big)(\cdot|\mathcal{B}_{t_1^n})(x).$$

Here, Q_{x_0} extends to a probability measure on $C([0,\infty), H^1)$ by the following: Let δ_{x_0} be the point-mass on $C([0,t_1^n], H^1)$ at $x_0|_{[0,t_1^n]}$, that is,

$$\delta_{x_0}\big(x \in C\big([0,t_1^n], H^1\big) : x(t) = x_0(t), 0 \le t \le t_1^n\big) = 1.$$

Define $\tilde{Q} = \delta_{x_0} \times Q_{x_0}$ on $\tilde{X} := C([0,t_1^n], H^1) \times C([t_1^n,\infty), H^1)$ and set $X := \{(x_1,x_2) \in C([0,t_1^n], H^1) \times C([t_1^n,\infty), H^1) : x_1(t_1^n) = x_2(t_1^n)\}$. Then X is a measurable subset of $\tilde{X}$ and $\tilde{Q}(X) = 1$. Then $\tilde{Q}$ can be restricted to X. Finally, $\Psi : X \to C([0,\infty), H^1)$ defined by $\Psi((x_1,x_2))(t) := x_1(t)$, if $0 \le t \le t_1^n$, $\Psi((x_1,x_2))(t) := x_2(t)$, if $t > t_1^n$, is a measurable map form X onto $C([0,\infty), H^1)$. Then $\tilde{Q}|_X \circ \Psi^{-1}$ is the desired measure, which still be denoted Q_{x_0}.

By (M2), we have for every e_i, $i \in \mathbb{N}$, that the process

$$\begin{aligned} M_i\big(t \wedge t_2^n, x\big) &= \big\langle x\big(t \wedge t_2^n\big), e_i\big\rangle - \big\langle x_0\big(t_1^n\big), e_i\big\rangle \\ &\quad - \int_{t_1^n}^{t \wedge t_2^n} \big\langle -U_{\delta_n}[x_0](s) \cdot \nabla x(s) - A_\alpha x, e_i\big\rangle\, ds \\ &= \big\langle x\big(t \wedge t_2^n\big), e_i\big\rangle - \big\langle x_0\big(t_1^n\big), e_i\big\rangle \\ &\quad - \int_{t_1^n}^{t \wedge t_2^n} \big\langle -U_{\delta_n}[x](s) \cdot \nabla x(s) - A_\alpha x, e_i\big\rangle\, ds \end{aligned}$$

is a continuous square-integrable $\mathcal{B}_t$-martingale under Q_{x_0}. Thus, by [48], Theorem 1.2.10, we obtain for every e_i, $i \in \mathbb{N}$, that the process

$$\big\langle x\big(t \wedge t_2^n\big), e_i\big\rangle - \int_0^{t \wedge t_2^n} \big\langle -U_{\delta_n}[x](s) \cdot \nabla x(s) - A_\alpha x, e_i\big\rangle\, ds$$

is a continuous square-integrable $\mathcal{B}_t$-martingale under $P_n^1 \otimes_{t_1^n} Q$, whose quadratic variation process is given by

$$\int_0^{t \wedge t_2^n} \big\|(k_{\delta_n} * G)^*\big(x(s)\big)(e_i)\big\|_U^2\, ds.$$

Thus, we construct a martingale solution $P_n^1 \otimes_{t_1^n} Q \in \mathcal{P}(C([0,\infty), H^1))$ of (3.4) on $[0, t_2^n]$. Then step by step we can construct a martingale solution $P_n \in \mathcal{P}(C([0,\infty), H^1))$ of (3.4) on $[0, T]$ for any given T in the following sense:

(M1′) $P_n(x(0) = k_{\delta_n} * \theta_0) = 1$ and

$$P_n\bigg\{x \in C([0,+\infty); H^1): \int_0^T \|\Lambda^{2\alpha} x(s) + U_{\delta_n}[x](s) \cdot \nabla x(s)\|_{H^1} \, ds + \int_0^T \|k_{\delta_n} * G(x(s))\|^2_{L_2(U;H^3)} \, ds < +\infty\bigg\} = 1.$$

(M2′) For every e_i, the process

$$\langle x(t \wedge T), e_i\rangle - \int_0^{t\wedge T} \langle -U_{\delta_n}[x](s) \cdot \nabla x(s) - A_\alpha x, e_i\rangle \, ds$$

is a continuous square-integrable $\mathcal{B}_t$-martingale under P_n, whose quadratic variation process is given by

$$\int_0^{t\wedge T} \|(k_{\delta_n} * G)^*(x(s))(e_i)\|_U^2 \, ds,$$

where the asterisk denotes the adjoint operator of $k_{\delta_n} * G(x(s))$.

(M3′) $P_n(L^\infty_{\mathrm{loc}}([0,+\infty), H^3) \cap \Omega_0) = 1$.

Then by the martingale representation theorem (cf. [41], Theorem 2, [11], Theorem 8.2) we can find a new probability space (Ω^n, P^n, W_n) and θ_n such that (θ_n, W_n) is a weak solution of (3.4) and θ_n has the same law as P_n.

Step 3: Now we show that θ_n converge to the solution of (3.1). Since we have

$$\langle u_n(t) \cdot \nabla\theta_n(t), \theta_n(t)\rangle = 0,$$

by Itô's formula we have

$$d|\theta_n|^p + p\kappa|\theta_n|^{p-2}|\Lambda^\alpha\theta_n|^2 \, dt \le p|\theta_n|^{p-2}\langle k_{\delta_n} * G(\theta_n)\, dW_n, \theta_n\rangle + \frac{p(p-1)}{2}|\theta_n|^{p-2}\|k_{\delta_n} * G(\theta_n)\|^2_{L_2(U,H)} \, dt.$$

By classical arguments, we easily show that there exist positive constants C_1, C_2 independent of n, such that (cf. [16], Appendix 1) for $2 \le p < 1 + \frac{2\kappa}{\rho_1}$ if $\rho_1 > 0$ and for $2 \le p < \infty$ if $\rho_1 = 0$, the following are satisfied:

$$E^{P^n}\Big(\sup_{0\le s\le T} |\theta_n(s)|^p\Big) \le C_1 \tag{3.5}$$

and

$$E^{P^n} \int_0^T \|\theta_n(s)\|^2_{H^\alpha} \, ds \le C_2. \tag{3.6}$$

Now we prove that the family $\mathcal{D}(\theta_n), n \in \mathbb{N}$, is tight in $C([0,T]; H^{-\beta})$, for all $\beta > 3$. Here, $\mathcal{D}(\theta_n)$ means the law of θ_n. By (3.5) for each $t \in [0,T]$, $\mathcal{D}(\theta_n(t))$ is tight on $H^{-\beta}$. Then by Aldous' criterion in [1], it suffices to check that for all stopping times $\tau_n \le T$ and $\eta_n \to 0$,

$$\lim_n E^{P^n} \|\theta_n(\tau_n + \eta_n) - \theta_n(\tau_n)\|_{H^{-\beta}} = 0. \tag{3.7}$$

We have P^n-a.s.

$$\theta_n(\tau_n + \eta_n) - \theta_n(\tau_n) = -\int_{\tau_n}^{\tau_n+\eta_n} A_\alpha \theta_n(s)\,ds - \int_{\tau_n}^{\tau_n+\eta_n} u_n(s)\cdot\nabla\theta_n(s)\,ds + \int_{\tau_n}^{\tau_n+\eta_n} k_{\delta_n} * G(\theta_n(s))\,dW_n(s).$$

It is easy to obtain the following:

$$E^{P^n}\left\|\int_{\tau_n}^{\tau_n+\eta_n} A_\alpha\theta_n(s)\,ds\right\|_{H^{-\beta}} \le C\eta_n E^{P^n} \sup_{t\in[0,T]}|\theta_n(t)|. \tag{3.8}$$

And since $H^2 \subset L^\infty$, we obtain that for $v \in H^3$,

$$|\langle u_n\cdot\nabla\theta_n, v\rangle| = |\langle u_n\cdot\nabla v, \theta_n\rangle| \le |\theta_n||u_n|\|\nabla v\|_{L^\infty} \le |\theta_n||u_n|\|v\|_{H^3}.$$

Since $\sup_{[0,t]}|u_n| \le C\sup_{[0,t]}|\theta_n|$, we get that

$$E^{P^n}\left\|\int_{\tau_n}^{\tau_n+\eta_n} u_n(s)\cdot\nabla\theta_n(s)\,ds\right\|_{H^{-\beta}} \le C\eta_n E^{P^n}\sup_{t\in[0,T]}|\theta_n(t)|^2. \tag{3.9}$$

In addition by Hypothesis G.1, we have

$$\begin{aligned} &E^{P^n}\left\|\int_{\tau_n}^{\tau_n+\eta_n} k_{\delta_n} * G(\theta_n(s))\,dW(s)\right\|_{H^{-\beta}}^2 \\ &\quad\le CE^{P^n}\int_{\tau_n}^{\tau_n+\eta_n}\|G(\theta_n(s))\|_{L_2(U,H^{-\beta})}^2\,ds \\ &\quad\le C\eta_n\Big(E^{P^n}\sup_{t\in[0,T]}|\theta_n(t)|^2 + 1\Big) \to 0 \qquad \text{as } \eta_n \to 0. \end{aligned} \tag{3.10}$$

Thus, (3.7) follows by (3.8), (3.9) and (3.10), which implies the tightness of $\mathcal{D}(\theta_n)$ in $C([0,T], H^{-\beta})$. This yields that for each $\eta > 0$

$$\lim_{\delta\to 0}\sup_n P^n\Big(\sup_{|s-t|\le\delta, s,t\le T}|\theta_n(t) - \theta_n(s)|_{H^{-\beta}} > \eta\Big) = 0.$$

By this and (3.5), (3.6), it is easy to get that $\mathcal{D}(\theta_n)$ is tight in $L^2([0,T]; H) \cap C([0,T], H^{-\beta})$ (cf. [37], Lemma 2.7). Therefore, we find a subsequence, still denoted by θ_n, such that $\mathcal{D}(\theta_n)$ converges weakly in

$$L^2([0,T]; H) \cap C([0,T], H^{-\beta}).$$

By Skorohod's representation theorem, there exist a stochastic basis $(\tilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}_t\}_{t\in[0,T]}, \tilde{P})$ and, on this basis, $L^2([0,T];H)\cap C([0,T],H^{-\beta})$-valued random variables $\tilde{\theta}, \tilde{\theta}_n, n\geq 1$, such that $\tilde{\theta}_n$ has the same law as θ_n on $L^2([0,T];H)\cap C([0,T],H^{-\beta})$, and $\tilde{\theta}_n \to \tilde{\theta}$ in $L^2([0,T];H)\cap C([0,T],H^{-\beta})$ $\tilde{P}$-a.s. For $\tilde{\theta}_n$ we also have (3.5) and (3.6). Hence, it follows that

$$\tilde{\theta}(\cdot,\omega) \in L^2\big([0,T];H^\alpha\big)\cap L^\infty\big([0,T];H\big) \qquad \text{for } \tilde{P}\text{-a.e. } \omega\in\Omega.$$

For each $\tilde{\theta}_n$ we define $\tilde{u}_n := U_{\delta_n}[\tilde{\theta}_n]$ and for each $n \geq 1$ we define the process

$$\tilde{M}_n(t) := \tilde{\theta}_n(t) - k_{\delta_n} * \theta_0 + \int_0^t A_\alpha \tilde{\theta}_n(s)\,ds + \int_0^t \tilde{u}_n(s)\cdot\nabla\tilde{\theta}_n(s)\,ds.$$

In fact $\tilde{M}_n$ is a square integrable martingale with respect to the filtration

$$\{\mathcal{G}_n\}_t = \sigma\big\{\tilde{\theta}_n(s), s\leq t\big\}.$$

For all $r \leq t \in [0,T]$, all bounded continuous functions ϕ on $C([0,r];H^{-\beta})\cap L^2([0,r];H)$, and all $v\in C^\infty(\mathbb{T}^2)$, we have

$$\tilde{E}\big(\langle \tilde{M}_n(t) - \tilde{M}_n(r), v\rangle \phi(\tilde{\theta}_n|_{[0,r]})\big) = 0$$

and

$$\tilde{E}\bigg(\bigg(\langle \tilde{M}_n(t), v\rangle^2 - \langle \tilde{M}_n(r), v\rangle^2 - \int_r^t \big\|(k_{\delta_n} * G)^*(\tilde{\theta}_n)v\big\|_U^2\,ds\bigg)\phi(\tilde{\theta}_n|_{[0,r]})\bigg) = 0.$$

By the B–D–G inequality, we have for $1 < p < \frac{1}{2} + \frac{\kappa}{\rho_1}$ if $\rho_1 > 0$ and $1 < p < \infty$ if $\rho_1 = 0$, that

$$\sup_n \tilde{E}\big|\langle \tilde{M}_n(t), v\rangle\big|^{2p} \leq C \sup_n \tilde{E}\bigg(\int_0^t \big\|(k_{\delta_n} * G)^*(\tilde{\theta}_n)v\big\|_U^2\,ds\bigg)^p < \infty.$$

Since $\tilde{\theta}_n \to \tilde{\theta}$ in $L^2(0,T;H)\cap C(0,T,H^{-\beta})$, we also have

$$\lim_{n\to\infty} \tilde{E}\big|\langle \tilde{M}_n(t) - M(t), v\rangle\big| = 0$$

and

$$\lim_{n\to\infty} \tilde{E}\big|\langle \tilde{M}_n(t) - M(t), v\rangle\big|^2 = 0,$$

where

$$M(t) := \tilde{\theta}(t) - \theta_0 + \int_0^t \tilde{u}\cdot\nabla\tilde{\theta} + A_\alpha\tilde{\theta}\,ds.$$

Here, $\tilde{u}$ is defined by (1.3) with θ replaced by $\tilde{\theta}$. Taking the limit, we obtain that for all $r\leq t\in[0,T]$, all bounded continuous functions ϕ on $C([0,r];H^{-\beta})\cap L^2([0,r];H)$, and $v\in C^\infty(\mathbb{T}^2)$,

$$\tilde{E}\big(\langle M(t) - M(r), v\rangle\phi(\tilde{\theta}|_{[0,r]})\big) = 0$$

and

$$\tilde{E}\Big(\Big(\langle M(t), v\rangle^2 - \langle M(r), v\rangle^2 - \int_r^t \|G(\theta)^* v\|_U^2 \, ds\Big)\phi(\tilde{\theta}|_{[0,r]})\Big) = 0.$$

Thus, the existence of a weak solution for (3.1) follows by the martingale representation theorem (cf. [11], Theorem 8.2, [41], Theorem 2).

Step 4: Now we prove the last statement. It is sufficient to prove that

$$E^{P^n} \sup_{t\in[0,T]} \|\theta_n(t)\|_{L^p}^p \le C,$$

where C is a constant independent of n. We write for simplicity $\theta(t) = \theta_n(t)$, $u(t) = u_n(t)$, $W(t) = W_n(t)$, $P = P^n$. By [29], Lemma 5.1, or [4], Theorem 2.4, we have

$$\begin{aligned}
\|\theta(t)\|_{L^p}^p &= \|k_{\delta_n} * \theta_0\|_{L^p}^p \\
&\quad + \int_0^t \Big[-p \int_{\mathbb{T}^2} |\theta(s)|^{p-2}\theta(s)\big(\Lambda^{2\alpha}\theta(s) + u(s)\cdot\nabla\theta(s)\big)\, d\xi \\
&\qquad + \frac{1}{2}p(p-1)\int_{\mathbb{T}^2} |\theta(s)|^{p-2}\Big(\sum_j \big|k_{\delta_n} * G\big(\theta(s)\big)(f_j)\big|^2\Big)\, d\xi\Big]\, ds \\
&\quad + p\int_0^t \int_{\mathbb{T}^2} |\theta(s)|^{p-2}\theta(s) k_{\delta_n} * G\big(\theta(s)\big)\, d\xi\, dW(s) \\
&\le \|k_{\delta_n} * \theta_0\|_{L^p}^p \\
&\quad + \int_0^t \frac{1}{2}p(p-1)\int_{\mathbb{T}^2} |\theta(s)|^{p-2}\Big(\sum_j \big|k_{\delta_n} * G\big(\theta(s)\big)(f_j)\big|^2\Big)\, d\xi\, ds \\
&\quad + p\int_0^t \int_{\mathbb{T}^2} |\theta(s)|^{p-2}\theta(s) k_{\delta_n} * G\big(\theta(s)\big)\, d\xi\, dW(s) \\
&\le \|k_{\delta_n} * \theta_0\|_{L^p}^p \\
&\quad + \int_0^t \Big(\varepsilon\int_{\mathbb{T}^2} |\theta(s)|^p\, d\xi \\
&\qquad + C(\varepsilon)\int \Big(\sum_j \big|k_{\delta_n} * G\big(\theta(s)\big)(f_j)\big|^2\Big)^{p/2} d\xi\Big)\, ds \\
&\quad + p\int_0^t \int_{\mathbb{T}^2} |\theta(s)|^{p-2}\theta(s) k_{\delta_n} * G\big(\theta(s)\big)\, d\xi\, dW(s),
\end{aligned}$$

where in the first inequality we used $\operatorname{div} u = 0$ and $\int |\theta|^{p-2}\theta\Lambda^{2\alpha}\theta \ge 0$ (cf. [44], Lemma 3.2) as well as Young's inequality in the second inequality. Then by the Burkholder–Davis–Gundy inequality and Minkowski's inequality, we obtain

$$
\begin{aligned}
&E \sup_{s\in[0,t]} \|\theta(s)\|_{L^p}^p \\
&\quad\leq E\|\theta_0\|_{L^p}^p \\
&\qquad + E\int_0^t \biggl(\varepsilon \int_{\mathbb{T}^2} |\theta(s)|^p \, d\xi + C\int \biggl(\sum_j |k_{\delta_n} * G(\theta(s))(f_j)|^2\biggr)^{p/2} d\xi\biggr) ds \\
&\qquad + pE\biggl(\int_0^t \biggl(\int_{\mathbb{T}^2} |\theta(s)|^{p-1} \biggl(\sum_j |k_{\delta_n} * G(\theta(s))(f_j)|^2\biggr)^{1/2} d\xi\biggr)^2 ds\biggr)^{1/2} \\
&\quad\leq E\|\theta_0\|_{L^p}^p \\
&\qquad + E\int_0^t \biggl(\varepsilon \int_{\mathbb{T}^2} |\theta(s)|^p \, d\xi + C\int \biggl(\sum_j |k_{\delta_n} * G(\theta(s))(f_j)|^2\biggr)^{p/2} d\xi\biggr) ds \\
&\qquad + pE \sup_{s\in[0,t]} \|\theta(s)\|_{L^p}^{p-1} \\
&\qquad\qquad \times \biggl(\int_0^t \biggl(\int_{\mathbb{T}^2} \biggl(\sum_j |k_{\delta_n} * G(\theta(s))(f_j)|^2\biggr)^{p/2} d\xi\biggr)^{2/p} ds\biggr)^{1/2} \qquad (3.11) \\
&\quad\leq E\|\theta_0\|_{L^p}^p \\
&\qquad + E\int_0^t \biggl(\varepsilon \int_{\mathbb{T}^2} |\theta(s)|^p \, d\xi + C\int \biggl(\sum_j |G(\theta(s))(f_j)|^2\biggr)^{p/2} d\xi\biggr) ds \\
&\qquad + C(T)E \sup_{s\in[0,t]} \|\theta(s)\|_{L^p}^{p-1} \biggl(\int_0^t \biggl(\int_{\mathbb{T}^2} \biggl(\sum_j |G(\theta(s))(f_j)|^2\biggr)^{p/2} d\xi\biggr) ds\biggr)^{1/p} \\
&\quad\leq E\|\theta_0\|_{L^p}^p + \varepsilon E \sup_{s\in[0,t]} \|\theta(s)\|_{L^p}^p + C_1 E\int_0^t \|\theta(s)\|_{L^p}^p \, ds + C_2 \\
&\quad\leq E\|\theta_0\|_{L^p}^p + \varepsilon E \sup_{s\in[0,t]} \|\theta(s)\|_{L^p}^p + C_1 \int_0^t E \sup_{s\in[0,\sigma]} \|\theta(s)\|_{L^p}^p \, d\sigma + C_2.
\end{aligned}
$$

Here, in the fourth inequality, we used (Gp.1) and Young's inequality. By Gronwall's lemma, the assertion follows. □

4. Existence and uniqueness of probabilistically (strong) solutions in the subcritical case. In this section, we assume $\alpha > 1/2$ and prove pathwise uniqueness for equation (3.1), and hence by the Yamada–Watanabe theorem the existence of a unique (probabilistically) strong solution to (3.1) in the subcritical case. Let us first give the definition of a (probabilistically) strong solution to (3.1).

DEFINITION 4.1. We say that there exists a (probabilistically) strong solution to (3.1) over the time interval $[0,T]$ if for every probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, P)$ with an $\mathcal{F}_t$-Wiener process W, there exists an $\mathcal{F}_t$-adapted process $\theta : [0,T]\times\Omega \to H$ such that for P-a.e. $\omega \in \Omega$

$$\theta(\cdot,\omega) \in L^\infty(0,T;H) \cap L^2(0,T;H^\alpha) \cap C([0,T];H^{-\beta})$$

and P-a.e.

$$\begin{aligned}\langle\theta(t),\varphi\rangle + \int_0^t \langle A_\alpha^{1/2}\theta(s), A_\alpha^{1/2}\varphi\rangle\,ds - \int_0^t \langle u(s)\cdot\nabla\varphi, \theta(s)\rangle\,ds \\ = \langle\theta_0,\varphi\rangle + \left\langle \int_0^t G(\theta(s))\,dW(s), \varphi\right\rangle\end{aligned} \tag{4.1}$$

for all $t \in [0,T]$ and all $\varphi \in C^1(\mathbb{T}^2)$ (assuming also that all integrals in the equation are defined).

THEOREM 4.2. *Assume $\alpha > \frac{1}{2}$. If G satisfies the following condition*:

$$\begin{aligned}\|\Lambda^{-1/2}(G(u)-G(v))\|^2_{L_2(U,H)} &\le \beta|\Lambda^{-1/2}(u-v)|^2 \\ &\quad + \beta_1|\Lambda^{\alpha-1/2}(u-v)|^2\end{aligned} \tag{GL.1}$$

for all $u, v \in H^\alpha$, for some $\beta \in \mathbb{R}$ independent of u, v, and $\beta_1 < 2\kappa$, then (3.1) *admits at most one probabilistically strong solution in the sense of Definition* 4.1 *such that*

$$\sup_{t\in[0,T]} \|\theta(t)\|_{L^p} < \infty, \quad P\text{-a.s.}$$

for some $p \in ((\alpha - \frac{1}{2})^{-1}, \infty)$, and

$$E \sup_{t\in[0,T]} |\Lambda^{-1/2}\theta(t)|^2 < \infty.$$

REMARK. The examples in Remark 3.4 with g being a Lipschitz function on $\mathbb{R}$ satisfy (GL.1) since

$$\begin{aligned}\|\Lambda^{-1/2}(G(u)-G(v))\|^2_{L_2(K,H)} &= \sum_k |\Lambda^{-1/2}(b_k(g(u)-g(v)))|^2 \\ &\le \int_{\mathbb{T}^2}\sum_k b_k^2(g(u)-g(v))^2\,d\xi \\ &\le C|u-v|^2 \\ &\le C|\Lambda^{-1/2}(u-v)|^2 + \varepsilon|\Lambda^{\alpha-1/2}(u-v)|^2.\end{aligned}$$

PROOF OF THEOREM 4.2. Let θ_1, θ_2 be two solutions of (3.1), and let $\{e_k\}_{k\in\mathbb{N}}$ be the eigenbasis of A_α from above. Then their difference $\theta = \theta_1 - \theta_2$ satisfies for $\psi \in C^1(\mathbb{T}^2)$

$$\begin{aligned}\langle\psi, \theta(t)\rangle - \int_0^t \langle u \cdot \nabla\psi, \theta_1\rangle\, ds - \int_0^t \langle u_2 \cdot \nabla\psi, \theta\rangle\, ds + \kappa \int_0^t \langle\theta, \Lambda^{2\alpha}\psi\rangle\, ds \\ = \int_0^t \langle\psi, \bigl(G(\theta_1) - G(\theta_2)\bigr)\, dW\rangle.\end{aligned} \tag{4.2}$$

Here, u_1, u_2, u satisfy (1.3) with θ replaced by $\theta_1, \theta_2, \theta$, respectively. Now set $\phi_k = \langle e_k, \theta(t)\rangle$, $\varphi_k = \langle\Lambda^{-1}e_k, \theta(t)\rangle$. Itô's formula and (4.2) yield

$$\begin{aligned}\phi_k\varphi_k &= \int_0^t \phi_k\, d\varphi_k + \int_0^t \varphi_k\, d\phi_k + \langle\varphi_k, \phi_k\rangle(t) \\ &= 2\int_0^t \langle u \cdot \nabla e_k, \theta_1\rangle\langle\Lambda^{-1}\theta, e_k\rangle + \langle u_2 \cdot \nabla e_k, \theta\rangle\langle\Lambda^{-1}\theta, e_k\rangle \\ &\qquad - \kappa\langle\Lambda^{2\alpha}e_k, \theta\rangle\langle\Lambda^{-1}\theta, e_k\rangle\, ds \\ &\quad + 2\int_0^t \langle\Lambda^{-1}\theta, e_k\rangle\langle e_k, \bigl(G(\theta_1) - G(\theta_2)\bigr)\, dW(s)\rangle \\ &\quad + \int_0^t \langle\bigl(G(\theta_1) - G(\theta_2)\bigr)^* e_k, \bigl(G(\theta_1) - G(\theta_2)\bigr)^* \Lambda^{-1}e_k\rangle_U\, ds.\end{aligned} \tag{4.3}$$

Here, $\langle\varphi_k, \phi_k\rangle(t)$ denotes the covariation process of φ_k, ϕ_k. The dominated convergence theorem implies

$$\sum_{k\le N} \int_0^t \langle u \cdot \nabla e_k, \theta_1\rangle\langle\Lambda^{-1}\theta, e_k\rangle\, ds \to \int_0^t {}_{H^{-1}}\langle u \cdot \nabla\theta_1, \Lambda^{-1}\theta\rangle_{H^1}\, ds, \qquad N \to \infty,$$

$$\sum_{k\le N} \int_0^t \langle u_2 \cdot \nabla e_k, \theta\rangle\langle\Lambda^{-1}\theta, e_k\rangle\, ds \to \int_0^t {}_{H^{-1}}\langle u_2 \cdot \nabla\theta, \Lambda^{-1}\theta\rangle_{H^1}\, ds, \qquad N \to \infty$$

and

$$\sum_{k\le N} \int_0^t \langle\Lambda^{2\alpha}e_k, \theta\rangle\langle\Lambda^{-1}\theta, e_k\rangle\, ds \to \int_0^t \langle\theta, \Lambda^{2\alpha-1}\theta\rangle\, ds, \qquad N \to \infty.$$

Furthermore, since

$$\begin{aligned}&\int_0^t |\Lambda^{-1/2}\theta|^2 \|\Lambda^{-1/2}(G(\theta_1) - G(\theta_2))\|^2_{L_2(U,H)}\, ds \\ &\qquad \le C \sup_{s\le t}|\theta(s)|^2 \int_0^t \|\Lambda^{-1/2}(G(\theta_1) - G(\theta_2))\|^2_{L_2(U,H)}\, ds < \infty,\end{aligned}$$

we obtain

$$\sum_{k\le N}\int_0^t \langle \Lambda^{-1}\theta, e_k\rangle\langle e_k, (G(\theta_1)-G(\theta_2))\,dW(s)\rangle$$

$$\to M_t := \int_0^t \langle \Lambda^{-1/2}\theta, \Lambda^{-1/2}(G(\theta_1)-G(\theta_2))\,dW(s)\rangle, \qquad N\to\infty,$$

in probability. Finally, the following inequality holds:

$$\sum_{k\le N}\int_0^t \langle (G(\theta_1)-G(\theta_2))^* e_k, (G(\theta_1)-G(\theta_2))^*\Lambda^{-1}e_k\rangle_U\,ds$$

$$\le \int_0^t \|\Lambda^{-1/2}(G(\theta_1)-G(\theta_2))\|^2_{L_2(U,H)}\,ds.$$

Thus, summing up over $k\le N$ in (4.3) and letting $N\to\infty$, we obtain

$$\begin{aligned}&|\Lambda^{-1/2}\theta|^2 + 2\kappa\int_0^t |\Lambda^{\alpha-1/2}\theta|^2\,ds\\ &\quad\le 2M(t) + 2\int_0^t {}_{H_{-1}}\langle u\cdot\nabla\theta_1, \Lambda^{-1}\theta\rangle_{H_1} + {}_{H_{-1}}\langle u_2\cdot\nabla\theta, \Lambda^{-1}\theta\rangle_{H_1}\,ds\\ &\qquad + \int_0^t \|\Lambda^{-1/2}(G(\theta_1)-G(\theta_2))\|^2_{L_2(U,H)}\,ds.\end{aligned}$$

By [44], we have

$${}_{H^{-1}}\langle u\cdot\nabla\theta_1, \Lambda^{-1}\theta\rangle_{H^1} = 0$$

and

$$\begin{aligned}&|{}_{H_{-1}}\langle u_2\cdot\nabla\theta, \Lambda^{-1}\theta\rangle_{H_1}|\\ &\quad\le \|u_2\|_{L^p}\|\theta\|_{L^{p_1}}\|\nabla\Lambda^{-1}\theta\|_{L^{p_1}} \le C\|u_2\|_{L^p}\|\theta\|_{H^{1/p}}\|\nabla\Lambda^{-1}\theta\|_{H^{1/p}}\\ &\quad\le C\|\theta_2\|_{L^p}\|\Lambda^{-1}\theta\|^2_{H^{1+1/p}} \le C\|\theta_2\|_{L^p}\|\Lambda^{-1}\theta\|^{2/r}_{H^{1/2}}\|\Lambda^{-1}\theta\|^{2(1-1/r)}_{H^{1/2+\alpha}}\\ &\quad\le \varepsilon|\Lambda^{\alpha-1/2}\theta|^2 + C\|\theta_2\|^r_{L^p}|\Lambda^{-1/2}\theta|^2,\end{aligned}$$

where $\frac{1}{p}+\frac{2}{p_1}=1$ for $p\in((\alpha-\frac{1}{2})^{-1},+\infty)$, $r=\frac{\alpha}{\alpha-1/2-1/p}$. Here we use $\operatorname{div} u_2 = 0$ in the first inequality, that $H^{1/p}\hookrightarrow L^{p_1}$ continuously in the second inequality, the interpolation inequality (2.3) in the fourth inequality and Young's inequality in the last equality.

Now by (GL.1) we have

$$|\Lambda^{-1/2}\theta|^2 \le 2M(t) + \int_0^t C\|\theta_2\|^r_{L^p}|\Lambda^{-1/2}\theta|^2\,ds + \beta\int_0^t |\Lambda^{-1/2}(\theta_1-\theta_2)|^2\,ds.$$

Let

$$\tau_n^1 := \inf\{t>0, \|\theta_2(t)\|_{L^p} > n\}.$$

Then by the weak continuity of θ_2, τ_n^1 are stopping times with respect to $\mathcal{F}_{t+}$, ($\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$) and $\|\theta_2(t \wedge \tau_n^1)\|_{L^p} \leq n$ for large n. Furthermore, let τ_n^2 be a localizing sequence of stopping times for M and $\tau_n := \tau_n^1 \wedge \tau_n^2$. Then, since $M(t \wedge \tau_n)$ is a martingale with respect to $\mathcal{F}_{t+}$, we get

$$
\begin{aligned}
E\left|\Lambda^{-1/2}\theta(t \wedge \tau_n)\right|^2 &\leq Cn^r E \int_0^{t \wedge \tau_n} \left|\Lambda^{-1/2}\theta\right|^2 ds + \beta E \int_0^{t \wedge \tau_n} \left|\Lambda^{-1/2}\theta\right|^2 ds \\
&= C(n) \int_0^t E\left|\Lambda^{-1/2}\theta(s \wedge \tau_n)\right|^2 ds \\
&\quad + \beta \int_0^t E\left|\Lambda^{-1/2}\theta(s \wedge \tau_n)\right|^2 ds.
\end{aligned}
$$

By Gronwall's inequality, we get $|\Lambda^{-1/2}\theta(t \wedge \tau_n)|^2 = 0$ P-a.s., and recalling that $\tau_n \to T$ P-a.s. as $n \to \infty$, we obtain that $\theta(t) = 0$ P-a.s. for $t \leq T$. By the weak continuity of θ, we obtain the zero set does not depend on t, thus completing the proof. □

REMARK. From the proof of Theorem 4.2, we immediately obtain that if there exists a probabilistically strong solution θ in the sense of Definition 3.1 satisfying

$$
\sup_{t \in [0,T]} \|\theta(t)\|_{L^p} < \infty, \qquad P\text{-a.s.}
$$

for some $p \in ((\alpha - \frac{1}{2})^{-1}, +\infty)$ and G satisfies (GL.1), then for any other solution $\tilde{\theta}$ such that

$$
E \sup_{t \in [0,T]} \left|\Lambda^{-1/2}\tilde{\theta}(t)\right|^2 < \infty,
$$

it follows that $\tilde{\theta} = \theta$, which implies that

$$
\sup_{t \in [0,T]} \|\tilde{\theta}(t)\|_{L^p} < \infty.
$$

THEOREM 4.3. *Assume $\alpha > \frac{1}{2}$ and that G satisfies Hypothesis* G.1, (GL.1) *and* (Gp.1) *for some $p \in ((\alpha - \frac{1}{2})^{-1}, +\infty)$. Then for each initial condition $\theta_0 \in L^p$, there exists a pathwise unique probabilistically strong solution θ of equation* (3.1) *over $[0, T]$ with initial condition $\theta(0) = \theta_0$ such that*

$$
E \sup_{t \in [0,T]} \left|\Lambda^{-1/2}\theta(t)\right|^2 < \infty.
$$

Moreover, the solution satisfies

$$
E \sup_{t \in [0,T]} \|\theta(t)\|_{L^p}^p + E \int_0^T \left|\Lambda^\alpha \theta(t)\right|^2 dt < \infty.
$$

PROOF. By Theorem 4.2, Theorem 3.3 and the Yamada–Watanabe theorem (cf. [45] or [33, 43]), we get that for each initial condition $\theta_0 \in L^p$, there exists a pathwise unique probabilistically strong solution θ of equation (3.1) over $[0, T]$ with initial condition $\theta(0) = \theta_0$ such that

$$\sup_{t\in[0,T]} \|\theta(t)\|_{L^p} < \infty, \qquad P\text{-a.s.},$$

and

$$E \sup_{t\in[0,T]} \left|\Lambda^{-1/2}\theta(t)\right|^2 < \infty.$$

By the remark before Theorem 4.3, the first result follows. By Theorem 3.3 and (3.6), the last part of the assertion follows. □

THEOREM 4.4 (Markov property). *Assume $\alpha > \frac{1}{2}$ and that G satisfies Hypothesis* G.1, (GL.1) *and* (Gp.1) *for some $p \in ((\alpha - \frac{1}{2})^{-1}, +\infty)$. If $\theta_0 \in L^p$, then for every bounded, $\mathcal{B}(H)$-measurable $F : H \to \mathbb{R}$, and all $s, t \in [0, T]$, $s \le t$*

$$E\big(F\big(\theta(t)\big)|\mathcal{F}_s\big)(\omega) = E\big(F\big(\theta\big(t, s, \theta(s)(\omega)\big)\big)\big) \qquad \text{for } P\text{-a.s. } \omega \in \Omega.$$

Here, $\theta(t, s, \theta(s)(\omega))$ denotes the solution to (3.1) *starting from $\theta(s)$ at time s satisfying*

$$E \sup_{t\in[s,T]} \left|\Lambda^{-1/2}\theta(t)\right|^2 < \infty.$$

PROOF. By Theorem 4.3, we have $\theta(t) = \theta(t, s, \theta(s))$ P-a.s. Then by the Yamada–Watanabe theorem in [45], we have P-a.s.

$$\begin{aligned} E\big(F\big(\theta(t)\big)|\mathcal{F}_s\big)(\omega) &= E\big(F\big(\theta\big(t, s, \theta(s)\big)\big)|\mathcal{F}_s\big)(\omega) \\ &= E\big(F\big(\mathbf{H}\big(\theta(s), W(\cdot + s) - W(s)\big)\big)|\mathcal{F}_s\big)(\omega) \\ &= E\big(F\big(\mathbf{H}\big(\theta(s)(\omega), W(\cdot + s) - W(s)\big)\big)\big) \\ &= E\big(F\big(\theta\big(t, s, \theta(s)(\omega)\big)\big)\big), \end{aligned}$$

where $\mathbf{H}$ is the functional obtained by the Yamada–Watanabe theorem such that $\mathbf{H}(\theta(0), W)$ is a strong solution to (3.1). □

We set for $\mathcal{B}(H)$-measurable $F : H \to \mathbb{R}$, and $t \in [0, T]$, $x \in L^p$

$$P_t F(x) := E F\big(\theta(t, x)\big).$$

Here, and in the following, we use $\theta(t, x)$ to denote a solution with initial value x. Then by Theorem 4.4, we have for $F : H \to \mathbb{R}$, bounded and $\mathcal{B}(H)$-measurable, $s, t \ge 0$,

$$P_s(P_t F)(x) = P_{s+t} F(x), \qquad x \in L^p, p \in \big((\alpha - \tfrac{1}{2})^{-1}, +\infty\big).$$

5. Ergodicity in the subcritical case. Now fix $\alpha > \frac{1}{2}$ and we assume $U = H$, $W(t)$ is a cylindrical Wiener process in H defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. We make the following assumptions on G.

HYPOTHESIS E.1. *G does not depend on θ and there exists $\sigma > 0$ such that $G \in L_2(H; H^{2-\alpha+\sigma})$ that is,*

$$\mathcal{E}_0 := \mathrm{Tr}\big(\Lambda^{4-2\alpha+2\sigma} GG^*\big) < \infty.$$

HYPOTHESIS E.2. *There exist $N \in \mathbb{N}$ and $g \in L(H)$ such that $Gg = P_N$.*

For $\varepsilon_0 > 0$ and any $\overline{W} \in C(\mathbb{R}^+, H^{-1-\varepsilon_0})$, we define

$$z(\overline{W})(t) := \sum_{i,j=1}^{\infty} \left(g_{ij}\beta_i(t) - \lambda_j \int_0^t e^{-\lambda_j(t-s)} g_{ij}\beta_i(s)\,ds \right) e_j,$$

if the convergence of the sum is uniformly with respect to t in every bounded time interval, otherwise set $z(\overline{W}) := +\infty$. Here, $\beta_i(t) := {}_{H^{1+\varepsilon_0}}\langle e_i, \overline{W}(t)\rangle_{H^{-1-\varepsilon_0}}$, $g_{ij} = \langle Ge_i, e_j\rangle$. Under Hypothesis E.1, there exists $\Omega' \subset \Omega$ such that $P(\Omega') = 1$ and for $\omega \in \Omega'$, $z(W(\omega)) \in C([0,\infty), H^{2+\varepsilon})$ for some $0 < \varepsilon < \sigma$, and on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, $z(W)$ is the mild solution of the equation: $dz + A_\alpha z = G\,dW$ with initial condition $z(0) = 0$.

Now for $v_0 \in H^1$, $\overline{W} \in C(\mathbb{R}^+, H^{-1-\varepsilon_0})$ we define

$$v(t, \overline{W}, v_0) := \begin{cases} v\big(t, v_0, z(\overline{W})\big), & \text{if } z(\overline{W}) \in C\big(\mathbb{R}^+, H^m\big) \text{ for } m < 2+\sigma, \\ 0, & \text{otherwise,} \end{cases}$$

where $v(t, v_0, z(\overline{W}))$ is the solution to (A.1) we obtained in Theorem A.1. Then by Theorem A.4 in Appendix A, v is a measurable mapping from $\mathbb{R}^+ \times C(\mathbb{R}^+, H^{-1-\varepsilon_0}) \times H^1$ into H^1, $(t, \overline{W}, \theta_0) \mapsto v(t, \overline{W}, \theta_0)$. We can now define

$$\theta(t, \overline{W}, \theta_0) := v(t, \overline{W}, \theta_0) + z(t, \overline{W}),$$

which is a measurable map from $\mathbb{R}^+ \times C(\mathbb{R}^+, H^{-1-\varepsilon_0}) \times H^1$ into H^1. Then for the cylindrical Wiener process W, $\theta(t, W, \theta_0)$ is a solution to (3.1), whose laws $P_{\theta_0}, \theta_0 \in H^1$ form a Markov process on H^1, since H^1 is an invariant space for (3.1) under assumption Hypothesis E.1. Let $(P_t)_{t\geq 0}$ be the associated transition semigroup on $\mathcal{B}_b(H^1)$. Now we want to study the long time behavior of the semigroup P_t.

REMARK 5.1. (i) Hypothesis E.1 obviously implies Hypothesis G.1, (Gp.1) for all $p \in ((\alpha - \frac{1}{2})^{-1}, \infty)$ and (GL.1). For $x := \theta_0 \in L^p$, let P_x denote the law of the corresponding solution θ to (3.1). Then by Theorems 4.3 and 4.4, the measures P_x, $x \in L^p$ form a Markov process.

(ii) The existence of a map g such that $Gg = P_N$ is equivalent to the following property:

$$P_N H \subset \mathrm{Im}(G).$$

(iii) Hypothesis E.1 is to make sure that the associated O–U process has a version $z \in C([0,\infty); H^{1,\infty}(\mathbb{T}^2))$ (see, e.g., [11], the proof of Theorem 5.16, and use Sobolev embedding). If we consider the stochastic integral taking values in a Banach space [e.g., $L^p(\mathbb{T}^2)$, $p > 1$] and use the theory developed in [3], we can change Hypothesis E.1 to the following condition: $G \in L_2(H; H^{1-\alpha+\varepsilon_1/2})$ and for some ε_1, q satisfying $\varepsilon_1 q > 2$

$$\left\| \left[\sum_k (\Lambda^{1-\alpha+\varepsilon_1} G e_k)^2 \right]^{1/2} \right\|_{L^q} + \left\| \left[\sum_k (G e_k)^2 \right]^{1/2} \right\|_{L^{(\alpha+1)/(\alpha-1/2)}} < \infty.$$

By this and similar arguments as in [3], we obtain for $\varepsilon < \varepsilon_1$ and $\varepsilon q > 2$ that the O–U process has a version $z \in C([0,\infty); H^{1+\varepsilon,q}) \subset C([0,\infty); H^{1,\infty}(\mathbb{T}^2))$, but in this paper we stay in the Hilbert space framework for simplicity.

(iv) For more general noise, we do not know how to obtain Proposition 5.7 since we cannot control $E \exp \|\theta\|_{L^p}^p$. Therefore, we restrict ourselves to additive noise.

5.1. *Preliminaries and some useful estimates.* First, we want to collect some useful and fundamental results about coupling from [34] and [36] which we will use later. Let (Λ_1, Λ_2) be two probability measures on a Polish space E. Let (Z_1, Z_2) be a couple of random variables $(\Omega, \mathcal{F}) \to E \times E$. We say that (Z_1, Z_2) is a coupling of (Λ_1, Λ_2) if $\Lambda_i = \mathcal{D}(Z_i)$ for $i = 1, 2$, where we use $\mathcal{D}(Z_i)$ to denote the distribution of Z_i.

LEMMA 5.2. *Let (Λ_1, Λ_2) be two probability measures on a Polish space $(E, \mathcal{B}(E))$. Then*

$$\|\Lambda_1 - \Lambda_2\|_{\mathrm{var}} = \min P(Z_1 \neq Z_2),$$

where the minimum is taken over all couplings (Z_1, Z_2) of (Λ_1, Λ_2). There exists a coupling for which the minimum value is attained and it is called a maximal coupling. Moreover, the maximal coupling has the following property:

$$P(Z_1 = Z_2, Z_1 \in \Gamma) = (\Lambda_1 \wedge \Lambda_2)(\Gamma), \qquad \Gamma \in \mathcal{B}(E).$$

LEMMA 5.3 (cf. [36], Lemma C.1). *Let Λ_1 and Λ_2 be two equivalent probability measures on E. Then for any $p > 1$ and any measurable subset $A \subset E$*

$$I_p(A) := \int_A \left(\frac{d\Lambda_1}{d\Lambda_2} \right)^p d\Lambda_1 < \infty$$

implies

$$(\Lambda_1 \wedge \Lambda_2)(A) \geq \left(1 - \frac{1}{p}\right) \left(\frac{\Lambda_1(A)^p}{p I_p(A)} \right)^{1/(p-1)}.$$

PROPOSITION 5.4 (cf. [40], Proposition 1.4). *Let E and F be two Polish spaces, $f_0 : E \to F$ be a measurable map and (Λ_1, Λ_2) be two probability measures on E. Set $\lambda_i = f_0^* \Lambda_i$, $i = 1, 2$. Then there exists a coupling (V_1, V_2) of (Λ_1, Λ_2) such that $(f_0(V_1), f_0(V_2))$ is a maximal coupling of (λ_1, λ_2).*

Now we give some useful estimates which will be used in the next two subsections. Let θ_n denote the approximation in the proof of Theorem 3.3. As will be seen below, we shall need uniform L^p-estimates, and a crucial ingredient to prove them is the following improved version of the "positivity lemma," that is, Lemma 3.2 in [44].

LEMMA 5.5 (Improved positivity lemma). *For $\alpha \in (0, 1)$, and $\theta \in L^p$ with $\Lambda^{2\alpha}\theta \in L^p$, for some $2 < p < \infty$,*

$$\int |\theta|^{p-2}\theta\left(\kappa\Lambda^{2\alpha} - \frac{2\lambda_1}{p}\right)\theta \geq 0.$$

PROOF. Denote the semigroup with respect to $-\kappa\Lambda^{2\alpha} + \frac{2\lambda_1}{p}$ and $-\kappa\Lambda^{2\alpha}$ in L^2 by P_t^0 and P_t^1, respectively. Then we have $P_t^0 f = e^{2t\lambda_1/p} P_t^1 f$. Since

$$\|P_t^1 f\|_{L^2} \leq e^{-\lambda_1 t}\|f\|_{L^2}$$

and

$$\|P_t^1 f\|_{L^\infty} \leq \|f\|_{L^\infty},$$

by the interpolation theorem, we have

$$\|P_t^1 f\|_{L^p} \leq e^{-2\lambda_1 t/p}\|f\|_{L^p},$$

which implies that

$$\|P_t^0 f\|_{L^p} \leq \|f\|_{L^p}.$$

Then we get that

$$\frac{d}{dt}\|P_t^0\theta\|_{L^p}^p = \int |P_t^0\theta|^{p-2}(P_t^0\theta)\left(P_t^0\left(-\kappa\Lambda^{2\alpha} + \frac{2\lambda_1}{p}\right)\theta\right)dx \leq 0.$$

Letting $t \to 0$, we obtain the result. □

PROPOSITION 5.6. *Let $\alpha > \frac{1}{2}$. Suppose Hypothesis E.1 holds. For $x \in L^p$, let θ denote the solution of equation (3.1) with the initial value x. Then for $2 < p < \infty$*

$$E\|\theta(t)\|_{L^p}^p \leq \|x\|_{L^p}^p e^{-\lambda_1 t} + C_S^p\left[\tfrac{1}{2}p(p-1)\right]^{p/2}\lambda_1^{-p/2}\mathcal{E}_0^{p/2}(1 - e^{-\lambda_1 t}),$$

where C_S is the constant for the Sobolev embedding.

PROOF. Using [29], Lemma 5.1, or [4], Theorem 2.4, for θ_n, we obtain

$$
\begin{aligned}
\|\theta(t)\|_{L^p}^p &= \|\theta(s)\|_{L^p}^p \\
&\quad + \int_s^t \biggl[-p\int_{\mathbb{T}^2}|\theta(l)|^{p-2}\theta(l)\bigl(\kappa\Lambda^{2\alpha}\theta(l)+u(l)\cdot\nabla\theta(l)\bigr)\,d\xi \\
&\qquad + \frac{1}{2}p(p-1)\int_{\mathbb{T}^2}|\theta(l)|^{p-2}\Bigl(\sum_j|k_{\delta_n}*G(e_j)|^2\Bigr)\,d\xi\biggr]\,dl \\
&\quad + p\int_s^t\int_{\mathbb{T}^2}|\theta(l)|^{p-2}\theta(l)k_{\delta_n}*G\,d\xi\,dW(l) \\
&\le \|\theta(s)\|_{L^p}^p - 2\lambda_1\int_s^t\int_{\mathbb{T}^2}|\theta(l)|^p\,d\xi\,dl \\
&\quad + \int_s^t \frac{1}{2}p(p-1)\int_{\mathbb{T}^2}|\theta(l)|^{p-2}\Bigl(\sum_j|k_{\delta_n}*G(e_j)|^2\Bigr)\,d\xi\,dl \qquad (5.1)\\
&\quad + p\int_s^t\int_{\mathbb{T}^2}|\theta(l)|^{p-2}\theta(l)k_{\delta_n}*G\,d\xi\,dW(l) \\
&\le \|\theta(s)\|_{L^p}^p - 2\lambda_1\int_s^t\int_{\mathbb{T}^2}|\theta(l)|^p\,d\xi\,dl \\
&\quad + \int_s^t\biggl(\lambda_1\int_{\mathbb{T}^2}|\theta(l)|^p\,d\xi \\
&\qquad + \Bigl[\frac{1}{2}p(p-1)\Bigr]^{p/2}\lambda_1^{-(p-2)/2}\int\Bigl(\sum_j|k_{\delta_n}*G(e_j)|^2\Bigr)^{p/2}\,d\xi\biggr)\,dl \\
&\quad + p\int_s^t\int_{\mathbb{T}^2}|\theta(l)|^{p-2}\theta(l)k_{\delta_n}*G\,d\xi\,dW(l),
\end{aligned}
$$

where we used Lemma 5.5 to get the first inequality and Young's inequality to get the last inequality. Here, for simplicity, we write $\theta(t)=\theta_n(t,x)$. Taking expectation, we obtain

$$
\begin{aligned}
E\|\theta_n(t)\|_{L^p}^p &\le E\|\theta_n(s)\|_{L^p}^p - E\lambda_1\int_s^t\int_{\mathbb{T}^2}|\theta_n(l)|^p\,d\xi\,dl \\
&\quad + C_S^p\Bigl[\frac{1}{2}p(p-1)\Bigr]^{p/2}\lambda_1^{-(p-2)/2}\mathcal{E}_0^{p/2}(t-s).
\end{aligned}
$$

Here, we use $\int_{\mathbb{T}^2}(\sum_j|G(e_j)|^2)^{p/2}\,d\xi \le (\sum_j(\int_{\mathbb{T}^2}|G(e_j)|^p\,d\xi)^{2/p})^{p/2} \le C_S^p\mathcal{E}_0^{p/2}$. Then Gronwall's lemma yields that

$$
E\|\theta_n(t)\|_{L^p}^p \le \|\theta_n(0)\|_{L^p}^p e^{-\lambda_1 t} + C_S^p[\tfrac{1}{2}p(p-1)]^{p/2}\lambda_1^{-p/2}\mathcal{E}_0^{p/2}(1-e^{-\lambda_1 t}).
$$

Letting $n\to\infty$ in the above inequality, we deduce

$$
E\|\theta(t)\|_{L^p}^p \le \|x\|_{L^p}^p e^{-\lambda_1 t} + C_S^p[\tfrac{1}{2}p(p-1)]^{p/2}\lambda_1^{-p/2}\mathcal{E}_0^{p/2}(1-e^{-\lambda_1 t}). \qquad \square
$$

5.2. *Uniqueness of the invariant measure.* In this subsection, we assume conditions Hypotheses E.1 and E.2 to hold. To prove uniqueness of invariant measure is much harder and in this section we first concrete on proving this. Existence will be shown in the next subsection. In addition, we shall prove polynomial convergence of the semigroup to the invariant measure in Section 5.3 below. If the dissipation term is strong enough (i.e., $\alpha > \frac{2}{3}$) we actually obtain exponential convergence (see Section 6).

Now we build an auxiliary process $\tilde{\theta}$. The aim is to find a shift h belonging to Cameron–Martin space of the driving process such that $E\|\theta(t) - \tilde{\theta}(t)\|_{H^{-1/2}} \to 0$ as $t \to \infty$. Fix θ, and consider

$$\begin{cases} d\tilde{\theta}(t) + A_\alpha\tilde{\theta}(t)\,dt + \tilde{u}(t)\cdot\nabla\tilde{\theta}(t)\,dt + K_0 P_N\big(\tilde{\theta} - \theta(t, W, \theta_0)\big)\,dt \\ \qquad = G\,dW(t), \\ \tilde{\theta}(0) = \tilde{\theta}_0 \in H^1, \end{cases} \tag{5.2}$$

where $\tilde{u}$ satisfies (1.3) with θ replaced by $\tilde{\theta}$ and K_0 is a constant to be determined later. Since $\|P_N\tilde{\theta}\|_{L^p} \le C_N\|\tilde{\theta}\|_{L^p}$ for $p \ge 2$, by a similar argument as in the proof of Theorems A.4 in Appendix A we obtain that there exists a measurable mapping from $\mathbb{R}^+ \times C(\mathbb{R}^+, H^{-1-\varepsilon}) \times H^1 \times H^1$ into H^1, $(t, \overline{W}, \theta_0, \tilde{\theta}_0) \mapsto \tilde{\theta}(t, \overline{W}, \theta_0, \tilde{\theta}_0)$, such that $\tilde{\theta}(t, W, \theta_0, \tilde{\theta}_0)$ is the solution of (5.2). Moreover, by the ω-wise uniqueness of (3.1) and (5.2) (which can be easily checked by a similar argument as the proof of Theorem 4.2), we have

$$\big(\theta(t, \theta_0), \tilde{\theta}(t, \theta_0, \tilde{\theta}_0)\big) = \big(\theta\big(t, s, \theta(s)\big), \tilde{\theta}\big(t, s, \theta(s, \theta_0), \tilde{\theta}(s, \theta_0, \tilde{\theta}_0)\big)\big) \qquad P\text{-a.s.},$$

which implies that $(\theta(t), \tilde{\theta}(t)) = (\theta(t, W, \theta_0), \tilde{\theta}(t, W, \theta_0, \tilde{\theta}_0))$ defines a Markov process. Here, for simplicity, we omit W and $\theta(t, s, \theta(s))$, $\tilde{\theta}(t, s, \theta(s, \theta_0), \tilde{\theta}(s, \theta_0, \tilde{\theta}_0))$ denote the solutions to (3.1), (5.2) starting from $\theta(s), \tilde{\theta}(s)$ at time s, respectively.

Now we derive a uniform $|\cdot|^4$ estimate for $\tilde{\theta}$. Here, we give formal calculations which can be made rigorous by using Galerkin approximations:

$$\begin{aligned} d\big|\tilde{\theta}(t)\big|^4 &+ 4\kappa\big|\tilde{\theta}(t)\big|^2\|\tilde{\theta}\|_{H^\alpha}^2\,dt + 4K_0\big|\tilde{\theta}(t)\big|^2|P_N\tilde{\theta}|^2\,dt \\ &\le 4\big|\tilde{\theta}(t)\big|^2\big\langle G\,dW(t), \tilde{\theta}\big\rangle + 4K_0\big|\tilde{\theta}(t)\big|^2|P_N\tilde{\theta}||\theta|\,dt + 6|\tilde{\theta}|^2\|G\|_{L_2(H,H)}^2\,dt \\ &\le 4\big|\tilde{\theta}(t)\big|^2\big\langle G\,dW(t), \tilde{\theta}\big\rangle + \varepsilon\big|\tilde{\theta}(t)\big|^4\,dt + C(\varepsilon)\big(|\theta|^4 + 1\big)\,dt. \end{aligned}$$

Taking expectation and by Proposition 5.6, we obtain

$$E\big|\tilde{\theta}(t)\big|^4 \le C, \qquad \forall t \ge 0, \tag{5.3}$$

where C is a constant independent of t.

Define $h(\theta, \tilde{\theta}) := -gK_0P_N(\tilde{\theta} - \theta)$ for g in Hypothesis E.2. Then for any $(t, \theta_0, \tilde{\theta}_0) \in \mathbb{R}^+ \times H^1 \times H^1$ and the cylindrical Wiener process W we have for

$\omega \in \Omega'$ that $z(W(\omega)), z(W(\omega) + \int_0^\cdot h(\theta(s, W(\omega), \theta_0), \tilde{\theta}(s, W(\omega), \theta_0, \tilde{\theta}_0))\,ds) \in C([0, \infty), H^{2+\varepsilon})$, $\varepsilon < \sigma$. Then for $\omega \in \Omega'$,

$$\theta\Big(t, W(\omega) + \int_0^\cdot h\big(\theta\big(s, W(\omega), \theta_0\big), \tilde{\theta}\big(s, W(\omega), \theta_0, \tilde{\theta}_0\big)\big)\,ds, \tilde{\theta}_0\Big) - z\big(W(\omega)\big)$$

is a solution to the following equation:

$$d\tilde{v}(t) + A_\alpha \tilde{v}(t)\,dt + u_{\tilde{v}+z}(t) \cdot \nabla(\tilde{v} + z)(t)\,dt + K_0 P_N\big(\tilde{v} - v(t, W, \theta_0)\big)\,dt = 0,$$

where $u_{\tilde{v}+z}$ satisfies (1.3) with θ replaced by $\tilde{v} + z$. Since for every $\omega \in \Omega'$ the above equation admits at most one solution, for $\omega \in \Omega'$ we have

$$\begin{aligned} &\tilde{\theta}\big(t, W(\omega), \theta_0, \tilde{\theta}_0\big) \\ &\quad = \theta\Big(t, W(\omega) + \int_0^\cdot h\big(\theta\big(s, W(\omega), \theta_0\big), \tilde{\theta}\big(s, W(\omega), \theta_0, \tilde{\theta}_0\big)\big)\,ds, \tilde{\theta}_0\Big). \end{aligned} \tag{5.4}$$

Now for $\rho = \tilde{\theta}(t, W, \theta_0, \tilde{\theta}_0) - \theta(t, W, \theta_0)$, we have the following results. Here, we want to emphasize that although the initial value $\theta_0 \in H^1$, we can only obtain that ρ converges to 0 in $H^{-1/2}$ norm.

PROPOSITION 5.7. *Fix $\alpha > 1/2$. Let $\delta_0 := \lambda_{N+1} - 2^{p/2} C_R^p C_S^{2p} \kappa^{1-p} [p(p-1)]^{p/2} \lambda_1^{-p/2} \mathcal{E}_0^{p/2} > 0$ for $p = \frac{\alpha+1}{\alpha-1/2}$, where N is as in Hypothesis E.2, and C_S, C_R are the constants for the Sobolev embedding and Riesz transform, respectively. Then for $\|\theta_0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\tilde{\theta}_0\|_{L^{2m(p-1)}}^{2m(p-1)} \le 2C_0$ for some $m > 5$, $K_0 > \lambda_{N+1}$ and $1 < q < \frac{m-1}{4}$, there exists a positive constant $\overline{C}$ such that for any $t > 0$*

$$E\big|\Lambda^{-1/2}\rho(t)\big|^2 \le \frac{\overline{C}}{(t+1)^{2q}}$$

(*where we can choose C_0 large enough such that $C_0 > 4C_S^p[\frac{1}{2}p(p-1)]^{p/2}\lambda_1^{-p/2} \times \mathcal{E}_0^{p/2}$*).

REMARK 5.8. From the condition $\lambda_{N+1} - 2^{p/2} C_R^p C_S^{2p} \kappa^{1-p} [p(p-1)]^{p/2} \times \lambda_1^{-p/2} \mathcal{E}_0^{p/2} > 0$, which also appears in the main theorem, we know that if the viscosity constant κ is large enough or $\mathcal{E}_0$ is small enough we could even take $N = 0$.

PROOF OF PROPOSITION 5.7. In the proof, we omit W for simplicity. From (3.1) and (5.2), we obtain that ρ satisfies the following equation in the weak sense:

$$\begin{aligned} \frac{d\rho(t)}{dt} &= -A_\alpha \rho - K_0 P_N \rho - \tilde{u} \cdot \nabla\tilde{\theta} + u \cdot \nabla\theta \\ &= -A_\alpha \rho - K_0 P_N \rho - u \cdot \nabla\rho - u_\rho \cdot \nabla\tilde{\theta}, \end{aligned}$$

where u_ρ satisfies (1.3) with θ replaced by ρ. Taking the inner product with $\Lambda^{-1}\rho$ in H, and using that

$$_{H^{-1}}\langle u_\rho \cdot \nabla\tilde{\theta}, \Lambda^{-1}\rho\rangle_{H^1} = 0$$

(cf. [44]), we obtain

$$\frac{1}{2}\frac{d}{dt}|\Lambda^{-1/2}\rho|^2 = -\kappa|\Lambda^{\alpha-1/2}\rho|^2 - K_0|P_N\Lambda^{-1/2}\rho|^2 - {}_{H^{-1}}\langle u \cdot \nabla\rho, \Lambda^{-1}\rho\rangle_{H^1}.$$

We have

$$\begin{aligned}
&\big|{}_{H^{-1}}\langle u \cdot \nabla\rho, \Lambda^{-1}\rho\rangle_{H^1}\big| \\
&\quad \le \|u\|_{L^p}\|\rho\|_{L^{p_1}}\|\nabla\Lambda^{-1}\rho\|_{L^{p_1}} \le C_S\|u\|_{L^p}\|\rho\|_{H^{1/p}}\|\nabla\Lambda^{-1}\rho\|_{H^{1/p}} \\
&\quad \le C_S C_R\|\theta\|_{L^p}\|\Lambda^{-1}\rho\|^2_{H^{1+1/p}} \le C_S C_R\|\theta\|_{L^p}\|\Lambda^{-1}\rho\|^{2/r}_{H^{1/2}}\|\Lambda^{-1}\rho\|^{2(1-1/r)}_{H^{1/2+\alpha}} \\
&\quad \le \frac{\kappa}{2}|\Lambda^{\alpha-1/2}\rho|^2 + C_1^r\left(\frac{\kappa}{2}\right)^{1-r}\|\theta\|^r_{L^p}|\Lambda^{-1/2}\rho|^2,
\end{aligned}$$

where C_S, C_R are the constants for Sobolev embedding and Riesz transform, respectively, and $C_1 = C_S C_R$. Here, $\frac{1}{p} + \frac{2}{p_1} = 1$ for $p > \frac{1}{\alpha-1/2}$, $r = \frac{\alpha}{\alpha-1/2-1/p}$ and we use Hölder's inequality and that $\operatorname{div} u = 0$ in the first inequality and $H^{1/p} \hookrightarrow L^{p_1}$ continuously in the second inequality, the interpolation inequality (2.3) in the fourth inequality and Young's inequality in the last equality. Then we obtain

$$\begin{aligned}
\frac{d}{dt}|\Lambda^{-1/2}\rho|^2 &\le -\kappa|\Lambda^{\alpha-1/2}\rho|^2 - K_0|P_N\Lambda^{-1/2}\rho|^2 \\
&\qquad + 2C_1^r\left(\frac{\kappa}{2}\right)^{1-r}\|\theta\|^r_{L^p}|\Lambda^{-1/2}\rho|^2.
\end{aligned}$$

Since, because $K_0 > \lambda_{N+1}$, we have

$$\begin{aligned}
\lambda_{N+1}|\Lambda^{-1/2}\rho|^2 &\le \kappa|Q_N\Lambda^{\alpha-1/2}\rho|^2 + K_0|P_N\Lambda^{-1/2}\rho|^2 \\
&\le \kappa|\Lambda^{\alpha-1/2}\rho|^2 + K_0|P_N\Lambda^{-1/2}\rho|^2,
\end{aligned}$$

it follows that

$$\frac{d}{dt}|\Lambda^{-1/2}\rho|^2 + \left(\lambda_{N+1} - 2C_1^r\left(\frac{\kappa}{2}\right)^{1-r}\|\theta\|^r_{L^p}\right)|\Lambda^{-1/2}\rho|^2 \le 0.$$

Thus, by Gronwall's lemma, we obtain

$$|\Lambda^{-1/2}\rho(t)|^2 \le e^{t\Gamma(t,\theta_0)}|\Lambda^{-1/2}\rho(0)|^2,$$

where

$$\Gamma(t,\theta_0) = -\lambda_{N+1} + 2C_1^r\left(\frac{\kappa}{2}\right)^{1-r}\frac{1}{t}\int_0^t\|\theta(s)\|^r_{L^p}\,ds.$$

By the same arguments as in the proof of Theorem A.1 in Appendix A, we have $\theta_n \to \theta$ in $L^2([0,T], H^1)$ a.s. Letting $n \to \infty$ in (5.1), by (3.11) we obtain

$$\begin{aligned}\|\theta(t)\|_{L^p}^p &+ \lambda_1 \int_0^t \int_{\mathbb{T}^2} |\theta(l)|^p \, d\xi \, dl \\ &\le \|\theta_0\|_{L^p}^p + C_S^p \left[\frac{1}{2} p(p-1)\right]^{p/2} \lambda_1^{-(p-2)/2} \mathcal{E}_0^{p/2} t \\ &\quad + p \int_0^t \int_{\mathbb{T}^2} |\theta(l)|^{p-2} \theta(l) G \, d\xi \, dW(l).\end{aligned}$$

Here, we use that $\int_{\mathbb{T}^2} (\sum_j |G(e_j)|^2)^{p/2} \, d\xi \le (\sum_j (\int_{\mathbb{T}^2} |G(e_j)|^p \, d\xi)^{2/p})^{p/2} \le C_S^p \mathcal{E}_0^{p/2}$.

Since $p = \frac{\alpha+1}{\alpha-1/2}$ implies $p = r$, we get

$$\begin{aligned}\Gamma(t, \theta_0) &\le -\lambda_{N+1} + 2C_1^p \left(\frac{\kappa}{2}\right)^{1-p} \frac{1}{t} \int_0^t \|\theta(s)\|_{L^p}^p \, ds \\ &\le -\lambda_{N+1} + 2C_1^p \left(\frac{\kappa}{2}\right)^{1-p} \frac{1}{t\lambda_1} \|\theta_0\|_{L^p}^p \\ &\quad + 2^{p/2} C_1^p C_S^p \kappa^{1-p} [p(p-1)]^{p/2} \lambda_1^{-p/2} \mathcal{E}_0^{p/2} \\ &\quad + 2C_1^p \left(\frac{\kappa}{2}\right)^{1-p} \frac{p}{t\lambda_1} \int_0^t \int_{\mathbb{T}^2} |\theta(l)|^{p-2} \theta(l) G \, d\xi \, dW(l).\end{aligned}$$

For $M(t) := p \int_0^t \int_{\mathbb{T}^2} |\theta(l)|^{p-2} \theta(l) G \, d\xi \, dW(l)$, we have

$$\langle M \rangle_t \le p^2 \mathcal{E}_0 C_S^2 \int_0^t \left(\int_{\mathbb{T}^2} |\theta(s)|^{p-1} \, d\xi \right)^2 ds,$$

where we use that $\sum_j |G(e_j)|^2(\xi) \le \sum_j \|G(e_j)\|_{L^\infty}^2 \le C_S^2 \mathcal{E}_0$. Then for any $m > 1$

$$\begin{aligned}\langle M \rangle_t^m &\le C_S^{2m} p^{2m} \mathcal{E}_0^m \left(\int_0^t \left(\int_{\mathbb{T}^2} |\theta(s)|^{p-1} \, d\xi \right)^2 ds \right)^m \\ &\le C_S^{2m} p^{2m} \mathcal{E}_0^m t^{m-1} \int_0^t \left(\int_{\mathbb{T}^2} |\theta(s)|^{2m(p-1)} \, d\xi \right) ds.\end{aligned}$$

Since $\|\theta_0\|_{L^{2m(p-1)}}^{2m(p-1)} \le 2C_0$ by Proposition 5.6 there exists a constant $C_{p,m}(C_0)$ independent of t such that $E\|\theta(t)\|_{L^{2m(p-1)}}^{2m(p-1)} \le C_{p,m}$ for $t \ge 0$. Thus, for $M_n = \sup_{n-1 \le t < n} M(t)$, we have

$$P\left(|M_n| > \frac{\varepsilon \lambda_1}{4C_1^p (\kappa/2)^{1-p}} n \right) \le \frac{p^{2m} \mathcal{E}_0^m C_{p,m} n^m C_S^{2m}}{(\varepsilon \kappa^{p-1} \lambda_1 / (2^{p+1} C_1^p))^{2m} n^{2m}}.$$

Now define the following random times:

$$T_{\text{bound}} := \sup\left\{n : |M_n| > \frac{\varepsilon\lambda_1}{4C_1^p(\kappa/2)^{1-p}}n\right\}.$$

By [35], Lemma 5, we have that if $m > 1$, then T_{bound} is finite almost surely. Set

$$\tau := \max\left(T_{\text{bound}}, \frac{2^{p+1}C_0^{p/(2m(p-1))}C_1^p}{\kappa^{p-1}\lambda_1\varepsilon}\right),$$

then we have

$$t > \tau \quad \Rightarrow \quad \Gamma(t, \theta_0) - (-\delta_0) < \varepsilon,$$

where $\delta_0 = \lambda_{N+1} - 2^{p/2}C_1^p C_S^p \kappa^{1-p}[p(p-1)]^{p/2}\lambda_1^{-p/2}\mathcal{E}_0^{p/2}$, which implies that for $\delta \in (0, \delta_0)$ and $t > \tau$,

$$\left|\Lambda^{-1/2}\rho(t)\right|^2 \le \left|\Lambda^{-1/2}(\theta_0 - \tilde{\theta}_0)\right|^2 e^{-\delta t}.$$

For $p_0 \in (0, m-1)$, by [35], Lemma 5, $E\tau^{p_0}$ is finite. Moreover, we obtain that for $1 < q < \frac{m-1}{4}$, there exists $\overline{C} > 0$ such that for any $t > 0$

$$\begin{aligned} &E\left|\Lambda^{-1/2}\rho(t)\right|^2 \\ &\quad \le Ce^{-\delta t} + \left(E\left|\Lambda^{-1/2}\rho(t)\right|^4\right)^{1/2} P(\tau > t)^{1/2} \\ &\quad \le \overline{C}\frac{1}{(t+1)^{2q}}, \end{aligned} \tag{5.5}$$

where we used (5.3) in the last inequality. $\square$

Now we fix $m > 35$ and $8 < q < \frac{m-3}{4}$. Proposition 5.7 still holds for such m, q. Moreover, we also have for any $t_0 \ge 0$

$$\begin{aligned} &P\left(\int_{t_0}^{\infty} \left|h(t)\right|^2 dt \ge \overline{C}\frac{1}{(t_0+1)^q}\right) \\ &\quad \le C\frac{(t_0+1)^q}{\overline{C}} \int_{t_0}^{\infty} E\left|\Lambda^{-1/2}\rho(t)\right|^2 dt \\ &\quad \le \overline{C}\frac{1}{(t_0+1)^q}, \end{aligned} \tag{5.6}$$

where $h(t) = h(\theta(t, W, \theta_0), \tilde{\theta}(t, W, \theta_0, \tilde{\theta}_0))$ and we used Proposition 5.7 in the last inequality. Moreover, by Theorem 5.9, we obtain that there exists $p_2 > 0$ such that

$$\begin{aligned} P\left(\int_0^{\infty} \left|h(t)\right|^2 \ge \overline{C}\right) &\le \frac{C_1}{\overline{C}} E\int_0^{\infty} \left|\Lambda^{-1/2}\rho(t)\right|^2 dt \\ &\le 1 - p_2, \end{aligned} \tag{5.7}$$

where $\overline{C}$ can be chosen large enough such that (5.5), (5.6) and (5.7) are satisfied.

Now we use a similar coupling method as in [40] to deduce the uniqueness of the invariant measure. More precisely, we have the following result.

THEOREM 5.9. *Fix $\alpha > 1/2$. Assume Hypotheses* E.1 *and* E.2 *hold. Let* $\delta_0 := \lambda_{N+1} - 2^{p/2} C_R^p C_S^{2p} \kappa^{1-p} [p(p-1)]^{p/2} \lambda_1^{-p/2} \mathcal{E}_0^{p/2} > 0$ *for* $p = \frac{\alpha+1}{\alpha-1/2}$, *where N is as in Hypothesis* E.2, *and C_S, C_R are the constants for Sobolev embedding and Riesz transform, respectively. Then there exists at most one invariant measure for the Markov semigroup P_t on H^1.*

PROOF. *Step* 1. Construction of a coupling of the solutions.

For $\theta_0^1, \theta_0^2 \in H^1$ and $T > 0$, we apply [40], Corollary 1.5, to $(\theta(\cdot, W, \theta_0^1), \theta(\cdot, W, \theta_0^2), \tilde{\theta}(\cdot, W, \theta_0^1, \theta_0^2))$ on $[0, T]$ and obtain $(\theta_1^0(\cdot, \theta_0^1, \theta_0^2), \theta_2^0(\cdot, \theta_0^1, \theta_0^2), \tilde{\theta}^0(\cdot, \theta_0^1, \theta_0^2))$ on $[0, T]$ such that the law of $(\theta_1^0(\cdot, \theta_0^1, \theta_0^2), \tilde{\theta}^0(\cdot, \theta_0^1, \theta_0^2))$ is the same as $(\theta(\cdot, W, \theta_0^1), \tilde{\theta}(\cdot, W, \theta_0^1, \theta_0^2))$ and $(\theta_2^0(\cdot, \theta_0^1, \theta_0^2), \tilde{\theta}^0(\cdot, \theta_0^1, \theta_0^2))$ is a maximal coupling of $(\mathcal{D}(\theta(\cdot, W, \theta_0^2)), \mathcal{D}(\tilde{\theta}(\cdot, W, \theta_0^1, \theta_0^2)))$ on $[0, T]$.

Then we obtain a sequence of independent versions of the mapping

$$(\theta_0^1, \theta_0^2) \to (\theta_1^0(\cdot, \theta_0^1, \theta_0^2), \theta_2^0(\cdot, \theta_0^1, \theta_0^2), \tilde{\theta}^0(\cdot, \theta_0^1, \theta_0^2)).$$

We denote this sequence by $(\theta_1^n, \theta_2^n, \tilde{\theta}^n)_n$ and define recursively

$$\begin{cases} \theta_1(nT + \cdot, \theta_0^1, \theta_0^2) = \theta_1^n(\cdot, \theta_1(nT), \theta_2(nT)), \\ \theta_2(nT + \cdot, \theta_0^1, \theta_0^2) = \theta_2^n(\cdot, \theta_1(nT), \theta_2(nT)), \\ \tilde{\theta}(nT + \cdot, \theta_0^1, \theta_0^2) = \tilde{\theta}^n(\cdot, \theta_1(nT), \theta_2(nT)). \end{cases}$$

Then $\theta_1(t, \theta_0^1, \theta_0^2), \theta_2(t, \theta_0^1, \theta_0^2), \tilde{\theta}(t, \theta_0^1, \theta_0^2)$ is defined for all $t \in [0, \infty)$ such that $(\theta_1(\cdot, \theta_0^1, \theta_0^2), \theta_2(\cdot, \theta_0^1, \theta_0^2))$ is a coupling of $(\mathcal{D}(\theta(\cdot, W, \theta_0^1)), \mathcal{D}(\theta(\cdot, W, \theta_0^2)))$. We denote the associated probability space by $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Moreover, $(\theta_1(nT, \theta_0^1, \theta_0^2), \theta_2(nT, \theta_0^1, \theta_0^2), \tilde{\theta}(nT, \theta_0^1, \theta_0^2))_n$ is a Markov chain and $\theta_1(\cdot, \theta_0^1, \theta_0^2), \theta_2(\cdot, \theta_0^1, \theta_0^2), \tilde{\theta}(\cdot, \theta_0^1, \theta_0^2)$ satisfy the following property:

$$E^{(\theta_0^1, \theta_0^2)}[f(\theta_1, \theta_2, \tilde{\theta}) \circ \Phi_{kT} | \mathcal{F}_{kT}] = E^{(\theta_1(kT), \theta_2(kT))} f(\theta_1, \theta_2, \tilde{\theta}),$$

where Φ_t is the shift operator.

Step 2. Introduction of l_0.

We set

$$l_0(k) = \min\{l \le k | P_{l,k}\},$$

where $\min \varnothing = \infty$ and

$$(P_{l,k}) \begin{cases} \tilde{\theta}(\cdot, \theta_0^1, \theta_0^2) = \theta_2(\cdot, \theta_0^1, \theta_0^2) \qquad \text{on } (lT, kT), \\ \|\theta_1(lT, \theta_0^1, \theta_0^2)\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2(lT, \theta_0^1, \theta_0^2)\|_{L^{2m(p-1)}}^{2m(p-1)} \le 2C_0. \end{cases}$$

Then by (5.5) and the Markov property of $\theta_1(\cdot,\theta_0^1,\theta_0^2),\theta_2(\cdot,\theta_0^1,\theta_0^2),\tilde{\theta}(\cdot,\theta_0^1,\theta_0^2)$ we have for $t>lT$

$$
\begin{aligned}
&E\big(\big|\Lambda^{-1/2}\big(\theta_2\big(t,\theta_0^1,\theta_0^2\big)-\theta_1\big(t,\theta_0^1,\theta_0^2\big)\big)\big|1_{l_0(\infty)\le l}\big)\\
&\quad=\sum_{k=0}^{l}E\big(\big|\Lambda^{-1/2}\big(\theta_2\big(t,\theta_0^1,\theta_0^2\big)-\theta_1\big(t,\theta_0^1,\theta_0^2\big)\big)\big|1_{l_0(\infty)=k}\big)\\
&\quad=\sum_{k=0}^{l}E\big[E\big(\big|\Lambda^{-1/2}\big(\theta_2\big(t-kT+kT,\theta_0^1,\theta_0^2\big)-\theta_1\big(t-kT+kT,\theta_0^1,\theta_0^2\big)\big)\big|\\
&\qquad\qquad\cdot 1_{l_0(\infty)=k}|\mathcal{F}_{kT}\big)\big]\\
&\quad=\sum_{k=0}^{l}E\big[E^{(\theta_1(kT),\theta_2(kT))}\big[\big|\Lambda^{-1/2}\big(\theta_2\big(t-kT,\theta_1(kT),\theta_2(kT)\big)\\
&\qquad\qquad-\theta_1\big(t-kT,\theta_1(kT),\theta_2(kT)\big)\big)\big|\\
&\qquad\qquad\cdot 1_{\{\theta_2(\cdot-kT,\theta_1(kT),\theta_2(kT))=\tilde{\theta}(\cdot-kT,\theta_1(kT),\theta_2(kT))\}}\big]\\
&\qquad\qquad\cdot 1_{\{\|\theta_1(kT)\|_{L^{2m(p-1)}}^{2m(p-1)}+\|\theta_2(kT)\|_{L^{2m(p-1)}}^{2m(p-1)}\le 2C_0\}}\big]\\
&\quad\le\sum_{k=0}^{l}E\big[E^{(\theta_1(kT),\theta_2(kT))}\big[\big|\Lambda^{-1/2}\big(\tilde{\theta}\big(t-kT,W,\theta_1(kT),\theta_2(kT)\big)\\
&\qquad\qquad-\theta\big(t-kT,W,\theta_1(kT)\big)\big)\big|\big]\\
&\qquad\qquad\cdot 1_{\{\|\theta_1(kT)\|_{L^{2m(p-1)}}^{2m(p-1)}+\|\theta_2(kT)\|_{L^{2m(p-1)}}^{2m(p-1)}\le 2C_0\}}\big]\\
&\quad\le\overline{C}\sum_{k=0}^{l}(t-kT+1)^{-q}\le C(t-lT+1)^{-q+1},
\end{aligned}
$$

where we used $\theta_i(kT)$ to denote $\theta_i(kT,\theta_0^1,\theta_0^2)$ for simplicity.

Step 3. Construction of Wiener processes.

Now we want to estimate $P(l_0(k+1)=0|l_0(k)=0)$. As in most papers using coupling methods for SPDEs, our tool is the Girsanov transform. Set

$$
\begin{cases}
h(t,W)=h\big(\theta\big(t-kT,W,\theta_1\big(kT,\theta_0^1,\theta_0^2\big)\big),\\
\qquad\qquad\tilde{\theta}\big(t-kT,W,\theta_1\big(kT,\theta_0^1,\theta_0^2\big),\theta_2\big(kT,\theta_0^1,\theta_0^2\big)\big)\big),\\
\tau_1(W)=\inf\Big\{t\in\big(kT,(k+1)T\big]\Big|\displaystyle\int_{kT}^{t}\big|h(t,W)\big|^2\,dt>\overline{C}(kT+1)^{-q}\Big\}.
\end{cases}
$$

Then by Proposition 5.4, we obtain cylindrical Wiener processes W_1,W_2 on $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ such that

$$
\Big(W_2,W_1+\int_{kT}^{\tau_1(W_1)\wedge\cdot}h(t,W_1)\,dt\Big),
$$

is a maximal coupling of $(\mathcal{D}(W), \mathcal{D}(W+\int_{kT}^{\tau_1(W)\wedge\cdot} h(t,W)\,dt))$ on $[kT,(k+1)T]$.

If $l_0(k)=0$, by construction in Step 1, we have

$$
\begin{aligned}
&P\big(l_0(k+1)=0|\mathcal{F}_{kT}\big)\\
&\quad = P\big(\tilde\theta\big(t,\theta_0^1,\theta_0^2\big)=\theta_2\big(t,\theta_0^1,\theta_0^2\big) \text{ for } t\in\big[kT,(k+1)T\big]|\mathcal{F}_{kT}\big)\\
&\quad \ge \tilde P\big(\tilde\theta\big(\cdot-kT,W_1,\theta_1\big(kT,\theta_0^1,\theta_0^2\big),\theta_2\big(kT,\theta_0^1,\theta_0^2\big)\big)\\
&\quad = \theta\big(\cdot-kT,W_2,\theta_2\big(kT,\theta_0^1,\theta_0^2\big)\big) \text{ for } t\in\big[kT,(k+1)T\big]\big)\\
&\quad \ge \tilde P\bigg(W_2=W_1+\int_{kT}^{\tau_1(W_1)\wedge\cdot} h(t,W_1)\,dt \text{ and } \tau_1(W_1)=(k+1)T\bigg),
\end{aligned}
\tag{5.8}
$$

where we used that $(\tilde\theta(\cdot,\theta_0^1,\theta_0^2),\theta_2(\cdot,\theta_0^1,\theta_0^2))$ is a maximal coupling of $(\tilde\theta(\cdot-kT,W_1,\theta_1(kT,\theta_0^1,\theta_0^2),\theta_2(kT,\theta_0^1,\theta_0^2)),\theta(\cdot-kT,W_2,\theta_2(kT,\theta_0^1,\theta_0^2)))$ in the first inequality and (5.4) in the last inequality.

Now set $A:=\{W|\tau_1(W)=(k+1)T\}$, $\Lambda_1:=\mathcal{D}(W)$, $\Lambda_2:=\mathcal{D}(W+\int_{kT}^{\tau_1(W)\wedge\cdot}\times h(t,W)\,dt)$. Then the Novikov condition is satisfied for Λ_1 and Λ_2, which by the Girsanov transform implies that

$$
\bigg(\frac{d\Lambda_1}{d\Lambda_2}\bigg)(W)=\exp\bigg(-\int_{kT}^{\tau_1(W)} h(t,W)\,dW(t)-\frac{1}{2}\int_{kT}^{\tau_1(W)}\big|h(t,W)\big|^2\,dt\bigg).
$$

Thus, we have

$$
\int\bigg(\frac{d\Lambda_1}{d\Lambda_2}\bigg)^2 d\Lambda_1\le E\big(M_2 e^{\int_{kT}^{\tau_1(W)}|h(t,W)|^2\,dt}\big)\le e^{\overline{C}(kT+1)^{-q}},
$$

where $M_2=\exp(-2\int_{kT}^{\tau_1(W)} h(t,W)\,dW(t)-2\int_{kT}^{\tau_1(W)}|h(t,W)|^2\,dt)$ and $EM_2\le 1$. By this, (5.7), (5.8) and Lemmas 5.2 and 5.3, we obtain

$$
\begin{aligned}
P\big(l_0(1)=0\big)&\ge(\Lambda_1\wedge\Lambda_2)(A)\\
&\ge\frac{1}{4}\bigg(\int\bigg(\frac{d\Lambda_1}{d\Lambda_2}\bigg)^2 d\Lambda_1\bigg)^{-1}\Lambda_1(A)^2\ge\frac{p_2^2}{4}e^{-\overline{C}}.
\end{aligned}
\tag{5.9}
$$

Step 4. Estimate for $P(l_0(k+1)\ne 0, l_0(k)=0)$.

By (5.8), we obtain

$$
\begin{aligned}
&P\big(l_0(k+1)\ne 0|\mathcal{F}_{kT}\big)\\
&\quad\le\tilde P\bigg(W_2=W_1+\int_{kT}^{\tau_1(W_1)\wedge\cdot} h(t,W_1)\,dt \text{ and } \tau_1(W_1)<(k+1)T\bigg)\\
&\qquad+\tilde P\bigg(W_2\ne W_1+\int_{kT}^{\tau_1(W_1)\wedge\cdot} h(t,W_1)\,dt\bigg).
\end{aligned}
$$

Since $(W_2, W_1 + \int_0^{\tau_1(W_1)\wedge\cdot} h(t, W_1)\,dt)$ is a maximal coupling, it follows from Lemma 5.2 and the construction of τ_1 that

$$\begin{aligned}
&\tilde{P}\biggl(W_2 \neq W_1 + \int_{kT}^{\tau_1(W_1)\wedge\cdot} h(t, W_1)\,dt\biggr)\\
&\quad = \|\Lambda_1 - \Lambda_2\|_{\mathrm{var}}\\
&\quad \le \frac{1}{2}\sqrt{\int\biggl(\frac{d\Lambda_1}{d\Lambda_2}\biggr)^2 d\Lambda_2 - 1} \le \frac{1}{2}\sqrt{\int\biggl(\biggl(\frac{d\Lambda_1}{d\Lambda_2}\biggr)^2 d\Lambda_1\biggr)^{1/2} - 1}\\
&\quad \le e^{\overline{C}/4}(kT+1)^{-q/2}.
\end{aligned} \tag{5.10}$$

Since by the Markov property of $(\theta_1(\cdot, \theta_0^1, \theta_0^2), \tilde{\theta}(\cdot, \theta_0^1, \theta_0^2))$, we have

$$\begin{aligned}
&\tilde{P}\biggl(W_2 = W_1 + \int_{kT}^{\tau_1(W_1)\wedge\cdot} h(t, W_1)\,dt \text{ and } \tau_1(W_1) < (k+1)T\biggr)\\
&\quad \le \tilde{P}\Bigl(\theta\bigl(\cdot - kT, W_2, \theta_2(kT, \theta_0^1, \theta_0^2)\bigr)\\
&\qquad\quad = \tilde{\theta}\bigl(\cdot - kT, W_1, \theta_1(kT, \theta_0^1, \theta_0^2), \theta_2(kT, \theta_0^1, \theta_0^2)\bigr)\\
&\qquad\qquad\qquad\qquad \text{and } \tau_1(W_1) < (k+1)T\Bigr)\\
&\quad \le \tilde{P}\biggl(\int_{kT}^{(k+1)T} \bigl|h(t, W_1)\bigr|^2 dt > \overline{C}(kT+1)^{-q}\biggr)\\
&\quad \le P\biggl(\int_{kT}^{(k+1)T} \bigl|h\bigl(\theta(t, W, \theta_0^1), \tilde{\theta}(t, W, \theta_0^1, \theta_0^2)\bigr)\bigr|^2 dt > \overline{C}(kT+1)^{-q}\biggr),
\end{aligned}$$

by (5.6) and (5.10), we obtain

$$P\bigl(l_0(k+1) \neq 0 \text{ and } l_0(k) = 0\bigr) \le C(kT+1)^{-q/2}, \tag{5.11}$$

where C depends on $\overline{C}$.

Step 5. Estimate for $El_0(\infty)^q$.

Since $l_0(k) = 0$ implies $l_0(l) = 0$ for any $0 \le l \le k \le \infty$,

$$P\bigl(l_0(\infty) \neq 0\bigr) \le \sum_{k=0}^{\infty} P\bigl(l_0(k+1) \neq 0 \text{ and } l_0(k) = 0\bigr).$$

By (5.9) and (5.11), we obtain

$$P\bigl(l_0(\infty) \neq 0\bigr) \le 1 - \frac{p_2^2}{4}e^{-\overline{C}} + C\sum_{k=1}^{\infty}(kT+1)^{-q/2}.$$

Then there exists T_0 such that for $T \ge T_0$ we have

$$P\bigl(l_0(\infty) = 0\bigr) \ge p_0 = \frac{p_2^2}{8}e^{-\overline{C}}. \tag{5.12}$$

Now fix $T = T_0$. Define

$$\sigma := \inf\{n \in \mathbb{N} | l_0(n) > 0\}.$$

It follows from (5.11) that

$$P(\sigma = k+1) \le C(kT+1)^{-q/2}.$$

Now for $1 < q_1 < \frac{q}{2} - 1$,

$$E\sigma^{q_1} 1_{\sigma<\infty} \le K_1, \tag{5.13}$$

where K_1 is a constant. For $\delta := \min\{n \in \mathbb{N} | \|\theta_1(nT)\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2(nT)\|_{L^{2m(p-1)}}^{2m(p-1)} \le 2C_0\}$, by Proposition 5.6 we obtain that there exist $\gamma > 0$ and $c > 0$ such that

$$E(e^{\gamma\delta}) \le c(1 + \|\theta_1^0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2^0\|_{L^{2m(p-1)}}^{2m(p-1)}) \tag{5.14}$$

(cf. [36], [38], (1.56)), where we used $C_0 > 4C_S^p[\frac{1}{2}p(p-1)]^{p/2}\lambda_1^{-p/2}\mathcal{E}_0^{p/2}$. Set

$$\begin{cases} \delta_0 := \delta, & & \\ \sigma_{k+1} := \infty & \text{if } \delta_k = \infty; & \sigma_{k+1} := \sigma \circ \Phi_{\delta_k T} + \delta_k & \text{else}, \\ \delta_k := \infty & \text{if } \sigma_k = \infty; & \delta_k := \delta \circ \Phi_{\sigma_k T} + \sigma_k & \text{else}, \end{cases}$$

where Φ_t is the shift operator. Set $\eta := \sigma + \delta \circ \Phi_{\sigma T}$. If $l_0(0) = 0$, by the Markov property, (5.13) and (5.14)

$$\begin{aligned} E(\eta^{q_1} 1_{\eta<\infty}) &\le C(E(\sigma^{q_1} 1_{\sigma<\infty}) + E((\delta \circ \Phi_{\sigma T})^{q_1} 1_{\delta\circ\Phi_\sigma<\infty} 1_{\sigma<\infty})) \\ &\le C(E(\sigma^{q_1} 1_{\sigma<\infty}) \\ &\qquad + cE(1 + \|\theta_1(\sigma T)\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2(\sigma T)\|_{L^{2m(p-1)}}^{2m(p-1)}) 1_{\sigma<\infty}) \\ &\le C(1 + \|\theta_1^0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2^0\|_{L^{2m(p-1)}}^{2m(p-1)}), \end{aligned}$$

where we used Proposition 5.6 in the last inequality. Since $\delta_k = \delta_{k-1} + \eta \circ \Phi_{\delta_{k-1}T}$, we obtain for $1 < q_1 < \frac{q}{2} - 1$,

$$\begin{aligned} E(\delta_k^{q_1} 1_{\delta_k<\infty}) &\le (k+1)^{q_1-1}\left(E\delta^{q_1} + \sum_{n=0}^{k-1} E(\eta \circ \Phi_{\delta_n T})^{q_1} 1_{\eta\circ\Phi_{\delta_n T}<\infty}\right) \\ &\le C(k+1)^{q_1}(1 + \|\theta_1^0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2^0\|_{L^{2m(p-1)}}^{2m(p-1)}). \end{aligned} \tag{5.15}$$

Moreover, if $\delta_k < \infty$, then $\sigma_{k+1} = \infty$ deduces that $l_0(\infty) = \delta_k$. Define

$$k_0 := \inf\{k \in \mathbb{Z}^+ | \sigma_{k+1} = \infty\}.$$

Then (5.12) implies that

$$P(k_0 \ge n) \le (1 - p_0)^n. \tag{5.16}$$

By (5.16), we obtain $k_0 < \infty$ a.s., which implies $l_0(\infty) < \infty$ a.s. Moreover, we have for $1 < q_2 < \frac{q}{2} - 1$,

$$E\big(l_0(\infty)^{q_2}\big) \le \sum_{n=0}^{\infty} E\big(\delta_n^{q_2} 1_{\delta_n<\infty} 1_{k_0=n}\big).$$

Then by Hölder's inequality, we have for $\frac{1}{p_1} + \frac{1}{p_1'} = 1$, $p_1, p_1' > 1$, satisfying $p_1 q_2 < \frac{q}{2} - 1$

$$E\big(l_0(\infty)^{q_2}\big) \le \sum_{n=0}^{\infty} \big(E\delta_n^{p_1 q_2} 1_{\delta_n<\infty}\big)^{1/p_1} P(k_0 = n)^{1/p_1'}.$$

By (5.15) and (5.16), we obtain

$$E\big(l_0(\infty)^{q_2}\big) \le C\left(\sum_{n=0}^{\infty} (n+1)^{q_2}(1-p_0)^{n/p_1'}\right)\big(1 + \|\theta_1^0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2^0\|_{L^{2m(p-1)}}^{2m(p-1)}\big)$$
$$< \infty.$$

Step 6. Conclusion.

By Step 2 and Step 5, we have for $t > 0$ and $1 < q_2 < \frac{q}{2} - 1$

$$\begin{aligned}
&E\big|\Lambda^{-1/2}\big(\theta_2(t, \theta_0^1, \theta_0^2) - \theta_1(t, \theta_0^1, \theta_0^2)\big)\big| \\
&\quad \le E\big(\big|\Lambda^{-1/2}\big(\theta_2(t, \theta_0^1, \theta_0^2) - \theta_1(t, \theta_0^1, \theta_0^2)\big)\big| 1_{l_0(\infty)\le l}\big) \\
&\qquad + CP\big(l_0(\infty) \ge l + 1\big)^{1/2} \\
&\quad \le C\big(1 + \|\theta_1^0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2^0\|_{L^{2m(p-1)}}^{2m(p-1)}\big)\big[(t + 1 - lT)^{-q+1} + (l+1)^{-q_2/2}\big],
\end{aligned}$$

where we used Proposition 5.6 in the first inequality. Choosing $l = [\frac{t+1}{2T}]$, we obtain for $1 < q_3 < \frac{q}{4} - 1$

$$\begin{aligned}
&E\big|\Lambda^{-1/2}\big(\theta_2(t, \theta_0^1, \theta_0^2) - \theta_1(t, \theta_0^1, \theta_0^2)\big)\big| \\
&\quad \le C\big(1 + \|\theta_1^0\|_{L^{2m(p-1)}}^{2m(p-1)} + \|\theta_2^0\|_{L^{2m(p-1)}}^{2m(p-1)}\big)(t+1)^{-q_3}.
\end{aligned} \tag{5.17}$$

Thus, for $\psi \in C(H^1)$ with $C_\psi := \sup_{x,y\in H^1} \frac{|\psi(x)-\psi(y)|}{|\Lambda^{-1/2}(x-y)|} < \infty$, we have

$$\begin{aligned}
&\big|P_t\psi(x) - P_t\psi(y)\big| \\
&\quad \le C_\psi E\big|\Lambda^{-1/2}\big(\theta_2(t, x, y) - \theta_1(t, x, y)\big)\big| \\
&\quad \le CC_\psi\big(1 + \|x\|_{L^{2m(p-1)}}^{2m(p-1)} + \|y\|_{L^{2m(p-1)}}^{2m(p-1)}\big)(t+1)^{-q_3}.
\end{aligned} \tag{5.18}$$

By Proposition 5.6, we obtain that for $2 < p_2 < \infty$

$$\begin{aligned}
E\|\theta(t)\|_{L^{p_2}}^{p_2} &\le \|x\|_{L^{p_2}}^{p_2} e^{-\lambda_1 t} \\
&\quad + C_S^{p_2}\left[\frac{1}{2}p_2(p_2-1)\right]^{p_2/2} \lambda_1^{-p_2/2} \mathcal{E}_0^{p_2/2}\big(1 - e^{-\lambda_1 t}\big).
\end{aligned}$$

Since for any invariant measure μ on H^1 and any $\varepsilon > 0$, there exists $b_\varepsilon > 0$ such that $\mu(x \in H^1 : \|x\|_{L^{p_2}}^{p_2} > b_\varepsilon) \le \varepsilon$, we obtain that for any $L > 0$

$$\begin{aligned}\int (\|x\|_{L^{p_2}}^{p_2} \wedge L)\, d\mu &\le \int_{\{x:\|x\|_{L^{p_2}}^{p_2} \le b_\varepsilon\}} (E^x \|\theta(t)\|_{L^{p_2}}^{p_2} \wedge L)\, d\mu + L\varepsilon \\ &\le b_\varepsilon e^{-\lambda_1 t} + C_S^{p_2}\left[\frac{1}{2}p_2(p_2-1)\right]^{p_2/2} \lambda_1^{-p_2/2} \mathcal{E}_0^{p_2/2}(1-e^{-\lambda_1 t}) \\ &\quad + L\varepsilon.\end{aligned}$$

Letting $t \to \infty$, $\varepsilon \to 0$ and $L \to \infty$, we obtain that for any invariant measure μ

$$\int \|x\|_{L^{p_2}}^{p_2}\, d\mu(x) \le C_S^{p_2}\left[\frac{1}{2}p_2(p_2-1)\right]^{p_2/2} \lambda_1^{-p_2/2} \mathcal{E}_0^{p_2/2}. \tag{5.19}$$

Then by (5.18), (5.19) for any invariant measures μ_1, μ_2 we obtain for $\psi \in C(H^1)$ with $C_\psi < +\infty$ and $1 < q_3 < \frac{q}{4} - 1$,

$$\begin{aligned}&\left|\int \psi(x)\mu_1(dx) - \int \psi(x)\mu_2(dx)\right| \\ &\quad \le C C_\psi \left(1 + \int \|x\|_{L^{2m(p-1)}}^{2m(p-1)} \mu_1(dx) + \int \|x\|_{L^{2m(p-1)}}^{2m(p-1)} \mu_2(dx)\right)(t+1)^{-q_3}.\end{aligned}$$

Letting $t \to \infty$, we get that $\mu_1 = \mu_2$. □

REMARK 5.10. (i) The coupling method has been introduced, for example, in [13, 30–32, 36] to study ergodicity for stochastic partial differential equations. In these papers, they decompose the process into the sum of a strongly dissipative process h and another finite dimensional dynamics l driven by a nondegenerate noise. The process is uniquely determined by the nondegenerate part l which can be treated by probabilistic arguments. However, in our case, we cannot decompose the process into the two desired parts since the uniqueness of the process h depends on the L^p-norm estimate, which cannot be obtained for h.

(ii) It is not clear how to directly use the results in [40] for the following two reasons: Although we consider the semigroup in H^1, the convergence we used in Theorem 5.9 is in $H^{-1/2}$. In [40], only one state space has been considered. If we choose the general Hilbert space in [40] as H^1, we cannot get the estimate (5.5) for the H^1-norm. If we choose the general Hilbert space in [40] as $H^{-1/2}$, the estimate (5.5) does also not hold for rough initial values in $H^{-1/2}$. The second reason is that, since Theorem 5.9 depends on the L^p-norm estimate, we can only prove $E\|\theta_1(t,\theta_0^1,\theta_0^2) - \theta_2(t,\theta_0^1,\theta_0^2)\|_{H^{-1/2}}$ converges to zero polynomially fast instead of exponentially fast, when time goes to infinity, where $(\theta_1(t,\theta_0^1,\theta_0^2), \theta_2(t,\theta_0^1,\theta_0^2))$ denotes a coupling of two solutions to (3.1) with different initial values $\theta_0^i \in H^1$, $i = 1, 2$.

(iii) In the situation of Theorem 5.9, we also obtain that P_t on H^1 is asymptotically strong Feller. In fact, for $x, y \in H^1$, define $d_n(x, y) := 1 \wedge n|\Lambda^{-1/2}(x - y)|$. For any two probabilities on H^1 μ_1, μ_2, we denote the set of positive measures on $H^1 \times H^1$ with marginals μ_1 and μ_2 by $\mathcal{C}(\mu_1, \mu_2)$. Define the Wasserstein distance

$$\|\mu_1 - \mu_2\|_d := \inf_{\mu \in \mathcal{C}(\mu_1,\mu_2)} \int_{H^1 \times H^1} d(x, y)\mu(dx, dy).$$

By definition and (5.17), we obtain

$$\begin{aligned}\|P_n(x, \cdot) - P_n(y, \cdot)\|_{d_n} &\le nE\big|\Lambda^{-1/2}\big(\theta_2(n, x, y) - \theta_1(n, x, y)\big)\big| \\ &\le C\big(\|x\|_{L^{2m(p-1)}}, \|y\|_{L^{2m(p-1)}}\big)nn^{-q_3}.\end{aligned}$$

Then we have

$$\lim_{\gamma\to 0} \limsup_{n\to\infty} \sup_{y\in B(x,\gamma)} \|P_n(x, \cdot) - P_n(y, \cdot)\|_{d_n} = 0,$$

where $B(x, \gamma)$ denotes the ball in H^1 with center x and radius γ, which implies that P_t on H^1 is asymptotically strong Feller.

(iv) It seems difficult to directly verify the gradient estimate for the semigroup as [21] did for the 2D Navier–Stokes equation. By their method, we need to consider an infinitesimal perturbation to the initial condition and to estimate the derivative of the solution $D\theta$ with respect to the initial value, which requires a good estimate for $E \exp \|\theta\|_{L^p}^p$. However, this cannot be obtained for $\alpha > \frac{1}{2}$. Even if the noise is nondegenerate and we use the Bismut–Elworthy–Li formula to compute the gradient of the semigroup, the ergodicity results only holds for $\alpha > \frac{2}{3}$ by delicate estimates (see Section 6). We cannot directly use the criterion in [27], since it is not clear how to verify the e-property in [27] for the semigroup associated with the 2D stochastic quasi-geostrophic equation.

5.3. *Existence of invariant measures for* $\alpha > \frac{1}{2}$. Assume that G satisfies condition Hypothesis E.1.

LEMMA 5.11. *Let* $\alpha > \frac{1}{2}$. *If* $\theta_0 \in H^1$, $t > 0$, *then*:

(i) $E(|\theta(t)|^2) + E\int_0^t |\Lambda^\alpha\theta(r)|^2\,dr \le |\theta_0|^2 + t\,\mathrm{Tr}[GG^*]$,

(ii) *for* $\delta \le 1$ *and* $q \ge \frac{2\alpha+2}{2\alpha-1}$, $p \ge 1$, *we have*

$$E\int_0^t \frac{|\Lambda^{\delta+\alpha}\theta(r)|^2}{(1 + |\Lambda^\delta\theta(r)|^2)^{p+1}}\,dr \le C\bigg(\int_0^t E\|\theta(r)\|_{L^q}^q\,dr + 1\bigg) \le Ct\big(\|\theta_0\|_{L^q}^q + 1\big),$$

(iii) *for* $q \ge \frac{2\alpha+2}{2\alpha-1}$, *there exist* $0 < \delta_1 < 1 - \alpha$ *and* $0 < \gamma_0 < 1$ *such that*

$$E\bigg[\int_0^t \big|A_\alpha^{\delta_1}\theta(r)\big|_{H^1}^{2\gamma_0}\,dr\bigg] \le C(1 + t)\big(\|\theta_0\|_{L^q}^q + 1\big).$$

PROOF. (i) is well known and follows from Itô's formula applied to $|\theta(t)|^2$. By Theorems A.1, A.2 in Appendix A, we obtain $\theta \in C([0,\infty), H^1) \cap L^2_{\mathrm{loc}}([0,\infty), H^{1+\alpha})$ P-a.s. By a similar argument as in the proof of Theorem 4.2, we obtain for $\delta \le 1$

$$\tfrac{1}{2}d|\Lambda^\delta\theta|^2 + \kappa|\Lambda^{\delta+\alpha}\theta|^2\,dt + \langle\Lambda^{\delta-\alpha}(u\cdot\nabla\theta), \Lambda^{\delta+\alpha}\theta\rangle\,dt$$
$$= \langle\Lambda^\delta\theta, \Lambda^\delta G\,dW_t\rangle + \tfrac{1}{2}\operatorname{Tr}[GG^*\Lambda^{2\delta}]\,dt.$$

Then we apply Itô's formula to the function $(1+|\Lambda^\delta\theta|^2)^{-p}$ and get

$$\frac{1}{(1+|\Lambda^\delta\theta(t)|^2)^p} - \frac{1}{(1+|\Lambda^\delta\theta_0|^2)^p}$$
$$= 2p\kappa\int_0^t \frac{|\Lambda^{\delta+\alpha}\theta|^2}{(1+|\Lambda^\delta\theta|^2)^{p+1}}\,dr + 2p\int_0^t \frac{\langle\Lambda^{\delta-\alpha}(u\cdot\nabla\theta), \Lambda^{\delta+\alpha}\theta\rangle}{(1+|\Lambda^\delta\theta|^2)^{p+1}}\,dr$$
$$- 2p\int_0^t \frac{\langle\Lambda^\delta\theta, \Lambda^\delta G\,dW_r\rangle}{(1+|\Lambda^\delta\theta|^2)^{p+1}} - p\int_0^t \frac{\operatorname{Tr}[GG^*\Lambda^{2\delta}]}{(1+|\Lambda^\delta\theta|^2)^{p+1}}\,dr$$
$$+ 2p(p+1)\int_0^t \frac{|G^*\Lambda^{2\delta}\theta|^2}{(1+|\Lambda^\delta\theta|^2)^{p+2}}\,dr,$$

where the last term is meaningful since $|G^*\Lambda^{2\delta}\theta|^2 \le |\Lambda^\delta\theta|^2\|\Lambda^\delta G\|^2_{L_2(H,H)}$. For $q \ge \frac{2\alpha+2}{2\alpha-1}$ and $\sigma := \frac{2}{q} < 2\alpha - 1$, we have

$$|\langle\Lambda^{\delta-\alpha}(u\cdot\nabla\theta), \Lambda^{\delta+\alpha}\theta\rangle| = |\langle\Lambda^{\delta-\alpha}\nabla\cdot(u\theta), \Lambda^{\delta+\alpha}\theta\rangle|$$
$$\le C|\Lambda^{\delta-\alpha+1+\sigma}\theta|\cdot\|\theta\|_{L^q}|\Lambda^{\delta+\alpha}\theta|$$
$$\le C\|\theta\|_{L^q}^{2\alpha/(2\alpha-1-\sigma)}|\Lambda^\delta\theta|^2 + \kappa|\Lambda^{\delta+\alpha}\theta|^2,$$

where we used $\operatorname{div} u = 0$ in the first equality and Lemmas 2.1 and 2.2 in the first inequality and Young's together with the interpolation inequality (2.3) in the last inequality.

Hence, we obtain

$$E\int_0^t \frac{|\Lambda^{\delta+\alpha}\theta|^2}{(1+|\Lambda^\delta\theta|^2)^{p+1}}\,dr \le C\left(\int_0^t E\|\theta\|^q_{L^q}\,dr + t\right) \le Ct(\|\theta_0\|^q_{L^q} + 1),$$

where we used Proposition 5.6 in the last step.

(iii) Since by Young's inequality for some $\gamma_0 > 0$, we have

$$|\Lambda^{\delta+\alpha}\theta|^{2\gamma_0} \le c\left[\frac{|\Lambda^{\delta+\alpha}\theta|^2}{(1+|\Lambda^\delta\theta|^2)^{p+1}} + 1 + |\Lambda^\delta\theta|^2\right],$$

we obtain for $\delta + \alpha > 1$

$$E\left[\int_0^t |\Lambda^{\delta+\alpha}\theta|^{2\gamma_0}\,dr\right] \le C(1+t)(\|\theta_0\|^q_{L^q} + 1).$$

□

THEOREM 5.12. *Let $\alpha > \frac{1}{2}$ and suppose Hypothesis E.1 holds. Then $(P_t)_{t\geq 0}$ is H^1-Feller, that is, for every $t > 0$ and $\psi \in \mathcal{C}_b(H^1)$, $P_t\psi \in C_b(H^1)$. Furthermore, there exists an invariant measure ν on H^1 of the transition semigroup $(P_t)_{t\geq 0}$. Moreover, there are $0 < \delta_1 < 1-\alpha$ and $0 < \gamma_0 < 1$ such that*

$$\int |A_\alpha^{\delta_1} x|_{H^1}^{2\gamma_0}\, d\nu < \infty.$$

PROOF. Choose $x_0 \in H^1$ and define for $t > 0$

$$\mu_t := \frac{1}{t}\int_0^t P_r^* \delta_{x_0}\, dr.$$

By Lemma 5.11(iii), we have for $t > 1$ that

$$\int |A_\alpha^{\delta_1} x|_{H^1}^{2\gamma_0} \mu_t(dx) \leq C.$$

This implies that $\{\mu_t | t > 0\}$ is tight on H^1. By Theorem A.3 in Appendix A, we obtain that $(P_t)_{t\geq 0}$ is H^1-Feller. Hence, any limit point of μ_t is an invariant measure for $(P_t)_{t\geq 0}$. □

Combining Theorem 5.9 and Theorem 5.12, we obtain the following results.

THEOREM 5.13. *Fix $\alpha > 1/2$. Assume Hypotheses E.1 and E.2 hold. Let $\delta_0 = \lambda_{N+1} - 2^{p/2} C_R^p C_S^{p+1} \kappa^{1-p} [p(p-1)]^{p/2} \lambda_1^{-p/2} \mathcal{E}_0^{p/2} > 0$ for $p = \frac{\alpha+1}{\alpha-1/2}$, where N is as in Hypothesis E.2, C_S, C_R are the constants for Sobolev embedding and Riesz transform, respectively. Then there exists exactly one invariant probability measure ν for P_t.*

Moreover, for $\psi \in C(H^1)$ with $C_\psi := \sup_{x,y\in H^1} \frac{|\psi(x)-\psi(y)|}{|\Lambda^{-1/2}(x-y)|} < \infty$ and any initial distribution μ_0 on H^1 with $\int \|x\|_{L^{2m(p-1)}}^{2m(p-1)}\, d\mu_0 < \infty$ for some $m > 35$, the following polynomial bound is satisfied for $1 < q_3 < \frac{m-19}{16}$:

$$\begin{aligned} &\left| \int P_t\psi(x)\mu_0(dx) - \int \psi(x)\nu(dx) \right| \\ &\quad \leq CC_\psi \left(1 + \int \|x\|_{L^{2m(p-1)}}^{2m(p-1)} \mu_0(dx) \right) (t+1)^{-q_3}. \end{aligned} \tag{5.20}$$

PROOF. (5.20) can be easily deduced from (5.18) and (5.19). □

5.4. *Law of large numbers.* In this section, we establish the law of large numbers for the solution of the stochastic quasi-geostrophic equation. The proof is mainly inspired by the approach used in [28].

THEOREM 5.14. *Fix $\alpha > 1/2$. Assume Hypotheses E.1 and E.2 hold. Set $\delta_0 := \lambda_{N+1} - 2^{p/2} C_R^p C_S^{2p} \kappa^{1-p} [p(p-1)]^{p/2} \lambda_1^{-p/2} \mathcal{E}_0^{p/2} > 0$ for $p = \frac{\alpha+1}{\alpha-1/2}$, where N is as in Hypothesis E.2, C_S, C_R are the constants for the Sobolev embedding and Riesz transform, respectively. Then for $\psi \in C(H^1)$ with $C_\psi := \sup_{x,y\in H^1} \frac{|\psi(x)-\psi(y)|}{|\Lambda^{-1/2}(x-y)|} < \infty$ and any initial distribution μ_0 on H^1 with $\int \|x\|_{L^{2m(p-1)}}^{2m(p-1)} d\mu_0 < \infty$ for some $m > 35$,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \psi(\theta(s))\, ds = \int \psi\, d\nu \qquad \textit{in probability}.$$

PROOF. (5.20) implies that for $\psi \in C(H^1)$ with $C_\psi < \infty$

$$\lim_{T\to\infty} \left| \frac{1}{T} \int_0^T E\psi(\theta(t))\, dt - \int \psi(x)\nu(dx) \right| = 0. \tag{5.21}$$

Now we want to prove that for bounded $\psi \in C(H^1)$ with $C_\psi < \infty$

$$\lim_{T\to\infty} \left| \frac{1}{T^2} E\left(\int_0^T \psi(\theta(t))\, dt \right)^2 - \left(\int \psi(x)\nu(dx) \right)^2 \right| = 0. \tag{5.22}$$

We have

$$\begin{aligned} \frac{1}{T^2} E\left(\int_0^T \psi(\theta(t))\, dt \right)^2 &= \frac{1}{T^2} E\left(\int_0^T \psi(\theta(t))\, dt \int_0^T \psi(\theta(s))\, ds \right) \\ &= \frac{2}{T^2} \int_0^T \int_0^t E[\psi(\theta(t))\psi(\theta(s))]\, dt\, ds \\ &= \frac{2}{T^2} \int_0^T \int_0^t \langle \mu_0 P_s, \psi P_{t-s}\psi \rangle\, dt\, ds. \end{aligned}$$

Moreover, we have that for $B := \{\|x\|_{L^{2m(p-1)}} \le R\}$,

$$\begin{aligned} &\left| \frac{2}{T^2} \int_0^T \int_0^t \left\langle \mu_0 P_s, \psi\left(P_{t-s}\psi - \int \psi(x)\nu(dx) \right) \right\rangle dt\, ds \right| \\ &\quad \le \left| \frac{2}{T^2} \int_0^T \int_0^t \left\langle \mu_0 P_s, 1_B \psi\left(P_{t-s}\psi - \int \psi(x)\nu(dx) \right) \right\rangle dt\, ds \right| \\ &\qquad + \left| \frac{2}{T^2} \int_0^T \int_0^t \left\langle \mu_0 P_s, 1_{B^c} \psi\left(P_{t-s}\psi - \int \psi(x)\nu(dx) \right) \right\rangle dt\, ds \right| \\ &\quad := I_T + II_T. \end{aligned}$$

By (5.20), we obtain that there exists $T_1 > 0$ such that for any $T > T_1$

$$\sup_{x\in B} \left| \frac{1}{T} \int_0^T P_t\psi(x) - \int \psi\, d\nu \right| < \varepsilon. \tag{5.23}$$

Thus, for the first term we have the following:

$$\begin{aligned}
I_T &= \left|\frac{2}{T^2}\int_0^T (T-s)\left\langle \mu_0 P_s, 1_B\psi\left[\frac{1}{T-s}\int_0^{T-s}\left(P_t\psi - \int \psi(x)\nu(dx)\right)dt\right]\right\rangle ds\right| \\
&\le \left|\frac{2}{T^2}\int_0^{T_1} s\left\langle \mu_0 P_{T-s}, 1_B\psi\left[\frac{1}{s}\int_0^{s}\left(P_t\psi - \int \psi(x)\nu(dx)\right)dt\right]\right\rangle ds\right| \\
&\quad + \left|\frac{2}{T^2}\int_{T_1}^T s\left\langle \mu_0 P_{T-s}, 1_B\psi\left[\frac{1}{s}\int_0^{s}\left(P_t\psi - \int \psi(x)\nu(dx)\right)dt\right]\right\rangle ds\right| \\
&\le 4\|\psi\|_{L^\infty}^2\left(\frac{T_1}{T}\right)^2 + \varepsilon\|\psi\|_{L^\infty},
\end{aligned}$$

where we used (5.23) in the last step. For the second term by Proposition 5.6, we have

$$\begin{aligned}
II_T &\le \frac{4\|\psi\|_{L^\infty}^2}{T^2}\int_0^T\int_0^t \mu_0 P_s(B^c)\,ds\,dt \\
&\le \|\psi\|_{L^\infty}^2\frac{C}{R}.
\end{aligned}$$

Choosing R large enough, we obtain for any $\varepsilon > 0$ that there exists T_0 such that for $T \ge T_0$

$$\left|\frac{2}{T^2}\int_0^T\int_0^t\left\langle \mu_0 P_s, \psi\left(P_{t-s}\psi - \int\psi(x)\nu(dx)\right)\right\rangle dt\,ds\right| \le \varepsilon.$$

The latter implies

$$\begin{aligned}
&\lim_{T\to\infty}\left|\frac{1}{T^2}E\left(\int_0^T \psi(\theta(t))\,dt\right)^2 - \left(\int\psi(x)\nu(dx)\right)^2\right| \\
&\quad \le \lim_{T\to\infty}\left|\frac{2}{T^2}\int\psi(x)\nu(dx)\int_0^T\int_0^t\langle\mu_0 P_s,\psi\rangle\,dt\,ds - \left(\int\psi(x)\nu(dx)\right)^2\right| \\
&\quad = \left|\int\psi(x)\nu(dx)\right|\lim_{T\to\infty}\left|\frac{2}{T^2}\int_0^T t\,dt\left[\frac{1}{t}\int_0^t\langle\mu_0 P_s,\psi\rangle\,ds - \int\psi(x)\nu(dx)\right]\right| \\
&\quad = 0.
\end{aligned}$$

Now by (5.21) and (5.22) we obtain for bounded ψ with $C_\psi < \infty$,

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\psi(\theta(s))\,ds = \int\psi\,d\nu \qquad \text{in probability.}$$

In general, we can remove the restriction of the boundedness of ψ by defining $\psi_L = \psi \wedge L \vee (-L)$ for $L \in \mathbb{R}^+$. Since for $x, y \in H^1$

$$|\psi_L(x) - \psi_L(y)| \le |\psi(x) - \psi(y)| \le C_\psi|\Lambda^{-1/2}(x-y)|,$$

we have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \psi_L\big(\theta(s)\big)\,ds=\int \psi_L\,d\nu \qquad \text{in probability.} \tag{5.24}$$

Since $\int|\psi|\,d\nu<\infty$, it is clear that

$$\lim_{L\to\infty}\int \psi_L\,d\nu=\int \psi\,d\nu.$$

Applying (5.21) for $|\psi-\psi_L|$, we have

$$\lim_{L\to\infty}\lim_{T\to\infty}E\frac{1}{T}\int_0^T \big|\psi_L\big(\theta(s)\big)-\psi\big(\theta(s)\big)\big|\,ds=\lim_{L\to\infty}\int|\psi_L-\psi|\,d\nu=0.$$

Now the result follows by taking the limit on both sides of (5.24). □

6. Exponential convergence for $\alpha>\frac{2}{3}$. Under the conditions (Hypotheses E.1, E.2) on G we only obtain the semigroup converges to the invariant measure polynomially fast [see (5.20)]. In this section, we prove that the convergence is exponentially fast, however, under stronger conditions for α and G. We assume that $\alpha>\frac{2}{3}$, and that G satisfies:

HYPOTHESIS E.3. *There are an isomophism Q_0 of H and a number $s\geq 1$ such that $G=A_\alpha^{-(s+\alpha)/(2\alpha)}Q_0^{1/2}$, and furthermore, G satisfies* (Gp.1) *for some fixed $p\in((\alpha-\frac{1}{2})^{-1},\infty)$ (which is, e.g., always the case if $Q_0=I$).*

For $x:=\theta_0\in L^p$, let P_x denote the law of the corresponding solution $\theta(\cdot,x)$ to (3.1). Since Hypothesis G.1, (Gp.1) and (GL.1) are satisfied under Hypothesis E.3, by Theorems 4.3 and 4.4 the measures P_x, $x\in L^p$, form a Markov process. Let $(P_t)_{t\geq 0}$ be the associated transition semigroup on $\mathcal{B}_b(H)$, defined as

$$P_t(\varphi)(x):=E\big[\varphi\big(\theta(t,x)\big)\big],\qquad x\in L^p,\varphi\in\mathcal{B}_b(H). \tag{6.1}$$

REMARK 6.1. If Hypothesis E.3 is satisfied with $s>3-2\alpha$, then Hypotheses E.1, E.2 hold for G and (Gp.1) holds for any $p\in(0,\infty)$.

6.1. *The strong Feller property for $\alpha>\frac{2}{3}$.* In this subsection, we prove that its transition semigroup has the strong Feller property under Hypothesis E.3.

REMARK 6.2. (i) Since in our case $\alpha<1$, the linear part $(-\Delta)^\alpha$ in (1.1) is less regularizing. As $G=A_\alpha^{-(s+\alpha)/(2\alpha)}Q_0^{1/2}$, we get the trajectories z of the associated O–U process to be in $C([0,\infty),H^{s+2\alpha-1-\varepsilon_0})$ for every $\varepsilon_0>0$ (cf. [11], Theorem 5.16, [14], Proposition 3.1). However, in order to prove the weak-strong uniqueness principle (see Theorem 6.4 below) and the strong Feller property of the semigroup associated with the solution of the cutoff equation (see Proposition 6.5

below), we need $z \in C([0,\infty), H^{s+1-\alpha+\sigma_1})$ for some $\sigma_1 > 0$. Therefore, we need $s + 2\alpha - 1 > s + 1 - \alpha$, that is, $\alpha > \frac{2}{3}$. The situation of the 3D Navier–Stokes equation is different. While in our case the needed regularity of z is higher than the regularity of our solution space $C((0,\infty), H^s)$ for the cutoff equation (6.2), for the 3D Navier–Stokes equation the needed regularity of z is the same as for the solution of the cutoff equation.

(ii) Since $\alpha < 1$, we cannot apply the same type of estimate as in [18] (cf. [18], Lemma D.2). Instead, we use Lemma 2.1 and choose suitable parameters (s, σ_1, σ_2) such that the approach in [18] can be modified to apply here [see (6.6)–(6.10) and so on].

(iii) It seems difficult to use the Kolmogorov equation method as in [10, 14] or a coupling approach as in [39] in our situation. In fact, to get a uniform H^s-norm estimate for the solutions of the Galerkin approximations of equation (1.1) for some $s > 0$, the regularity, needed for the trajectories of the associated Ornstein–Uhlenbeck (O–U) process z is higher than H^s, which is entirely different from the situation of the 3D Navier–Stokes equation. According to the method in [10, 14] and [39], we should use the solutions' $H^{s+\alpha}$-norm to control the $H^{s+\alpha}$-norm of the derivative of the solutions as required for the Bismut–Elworthy–Li formula. In particular, the associated O–U process z should be also in $H^{s+\alpha}$. However, under Hypothesis E.3 for the noise, the O–U process z is only in $L^2([0,T], H^{s+2\alpha-1})$. As a consequence, for their method to apply here, we need even $\alpha \geq 1$.

Fix $s > 1$ as in Hypothesis E.3 and set $\mathcal{W} := H^s$ and $|x|_{\mathcal{W}} := \|x\|_{H^s}$. In this subsection, we choose

$$\Omega := C\big([0,\infty); H^{-\beta}\big)$$

for some $\beta > 3$ and let $\mathcal{B}$ denote the Borel σ-algebra on Ω.

Now we state the main result of this section.

THEOREM 6.3. *Fix $\alpha > \frac{2}{3}$. Under Hypothesis* E.3, $(P_t)_{t\geq 0}$ *is $\mathcal{W}$-strong Feller, that is, for every $t > 0$ and $\psi \in \mathcal{B}_b(H)$, $P_t\psi \in C_b(\mathcal{W})$.*

We shall use [18], Theorem 5.4, which is an abstract result to prove the strong Feller property. In order to use [18], Theorem 5.4, we follow the idea of [18], Theorem 5.11, to construct $P_x^{(R)}$. We introduce an equation which differs from the original one by a cut-off only, so that with large probability they have the same trajectories on a small random time interval [see (6.3) below]. We consider the equation

$$d\theta(t) + A_\alpha\theta(t)\,dt + \chi_R\big(|\theta|_{\mathcal{W}}^2\big)u(t)\cdot\nabla\theta(t)\,dt = G\,dW(t), \tag{6.2}$$

where $\chi_R : \mathbb{R} \to [0,1]$ is of class C^∞ such that $\chi_R(|\theta|) = 1$ if $|\theta| \leq R$, $\chi_R(|\theta|) = 0$ if $|\theta| > R + 1$ and with its first derivative bounded by 1. Then, if we can prove the following Theorem 6.4 and Proposition 6.5, Theorem 6.3 follows.

THEOREM 6.4 (Weak–strong uniqueness). *Fix $\alpha > \frac{2}{3}$. Suppose Hypothesis E.3 holds. Then for every $x \in \mathcal{W}$, equation (6.2) has a unique martingale solution $P_x^{(R)}$, with*

$$P_x^{(R)}\big[C\big([0,\infty);\mathcal{W}\big)\big] = 1.$$

Let $\tau_R : \Omega \to [0,\infty]$ be defined by

$$\tau_R(\omega) := \inf\big\{t \geq 0 : |\omega(t)|_{\mathcal{W}}^2 \geq R\big\},$$

and $\tau_R(\omega) := \infty$ if this set is empty. If $x \in \mathcal{W}$ and $|x|_{\mathcal{W}}^2 < R$, then

$$\lim_{\varepsilon\to 0} P_{x+h}^{(R)}[\tau_R \geq \varepsilon] = 1, \qquad \text{uniformly in } h \in \mathcal{W}, |h|_{\mathcal{W}} < 1. \tag{6.3}$$

Moreover,

$$E^{P_x^{(R)}}\big[\varphi(\omega_t)1_{[\tau_R \geq t]}\big] = E^{P_x}\big[\varphi(\omega_t)1_{[\tau_R \geq t]}\big] \tag{6.4}$$

for every $t \geq 0$ and $\varphi \in \mathcal{B}_b(H)$.

PROOF. Let z denote the solution to

$$dz(t) + A_\alpha z(t)\,dt = G\,dW(t),$$

with initial data $z(0) = 0$ and let $v_x^{(R)}$ be the solution to the auxiliary problem

$$\frac{dv^{(R)}(t)}{dt} + A_\alpha v^{(R)}(t) + u^{(R)}(t) \cdot \nabla\big(v^{(R)}(t) + z(t)\big)\chi_R\big(|v^{(R)} + z|_{\mathcal{W}}^2\big) = 0, \tag{6.5}$$

with $v^{(R)}(0) = x$. Here, $u^{(R)}(t) = u_{v^{(R)}}(t) + u_z(t)$, $u_{v^{(R)}}$ and u_z satisfy (1.3) with θ replaced by $v^{(R)}$ and z, respectively. Moreover, define $\theta^{(R)} := v^{(R)} + z$, which is a weak solution to equation (6.2). We denote its law on Ω by $P_x^{(R)}$. By Hypothesis E.3, the trajectories of the noise belong to

$$\Omega^* := \bigcap_{\beta\in(0,1/2),\eta\in[0,(s+\alpha)/(2\alpha)-1/(2\alpha))} C^\beta\big([0,\infty); D\big(A_\alpha^\eta\big)\big),$$

with probability one. Hence, the analyticity of the semigroup generated by A_α implies that for each $\omega \in \Omega^*$, $z(\omega) \in C([0,\infty), H^{s+2\alpha-1-\varepsilon_0})$ for every $\varepsilon_0 > 0$.

Now, for $\omega \in \Omega^*$ we prove that equation (6.5) with $z(\omega)$ replacing z has a unique global weak solution in the space $C([0,\infty);\mathcal{W})$. First, we obtain the following a priori estimate for suitable $\sigma_1, \sigma_2 > 0$ with $\sigma_2 \leq s$, $\sigma_2 + \sigma_1 = 1$, $s + \sigma_1 - \alpha + 1 < s + 2\alpha - 1 < s + \alpha$, where we used that $\alpha > \frac{2}{3}$ since $0 < \sigma_1 < 3\alpha - 2$:

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\big|\Lambda^s v^{(R)}\big|^2 + \kappa\big|\Lambda^{s+\alpha}v^{(R)}\big|^2 \\
&\quad = \chi_R\big(\big|\theta^{(R)}\big|_{\mathcal{W}}^2\big)\big\langle\Lambda^{s-\alpha}\nabla\cdot\big(u^{(R)}\theta^{(R)}\big), \Lambda^{s+\alpha}v^{(R)}\big\rangle
\end{aligned}$$

$$
\begin{aligned}
&\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big|\Lambda^{s-\alpha+1}\big(u^{(R)}\theta^{(R)}\big)\big|\cdot\big|\Lambda^{s+\alpha}v^{(R)}\big| \\
&\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big|\Lambda^{s-\alpha+1+\sigma_1}\theta^{(R)}\big|\big|\Lambda^{\sigma_2}\theta^{(R)}\big|\cdot\big|\Lambda^{s+\alpha}v^{(R)}\big| \\
&\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big(\big|\Lambda^{s-\alpha+1+\sigma_1}v^{(R)}\big|+\big|\Lambda^{s-\alpha+1+\sigma_1}z\big|\big)\cdot\big|\Lambda^{s+\alpha}v^{(R)}\big| \\
&\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big(C\big|\Lambda^{s}v^{(R)}\big|^{1-r_1}\big|\Lambda^{s+\alpha}v^{(R)}\big|^{r_1}+\big|\Lambda^{s-\alpha+1+\sigma_1}z\big|\big) \\
&\qquad\cdot\big|\Lambda^{s+\alpha}v^{(R)}\big| \\
&\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big(\big|\Lambda^{s}v^{(R)}\big|^2+\big|\Lambda^{s-\alpha+1+\sigma_1}z\big|^2\big)+\frac{\kappa}{2}\big|\Lambda^{s+\alpha}v^{(R)}\big|^2 \\
&\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big(C(R)+\big|\Lambda^{s-\alpha+1+\sigma_1}z\big|^2\big)+\frac{\kappa}{2}\big|\Lambda^{s+\alpha}v^{(R)}\big|^2,
\end{aligned}
\tag{6.6}
$$

where $r_1 := \frac{1-\alpha+\sigma_1}{\alpha}$. Here, in the first equality, we used $\operatorname{div} u = 0$, and in the second inequality we used Lemmas 2.1 and 2.2, and in the fourth inequality we used the interpolation inequality (2.3) and that $s-\alpha+1+\sigma_1 < s+2\alpha-1$, and in the fifth inequality we used Young's inequality and in the last inequality we used $|\Lambda^s v^{(R)}| \leq |\Lambda^s\theta^{(R)}| + |\Lambda^{s-\alpha+1+\sigma_1}z|$. Then as in the proof of Theorem A.1 in Appendix A, we prove (6.5) has a weak solution in $L^\infty([0,T],\mathcal{W}) \cap L^2([0,T],H^{s+\alpha})$.

Continuity. For each $\omega \in \Omega^*$, σ_1 and σ_2 as in (6.6), since $s-\alpha+1+\sigma_1 < s+2\alpha-1$, we have $z \in C([0,\infty); H^{s-\alpha+1+\sigma_1})$. Since $s > 3-3\alpha$, multiplying the equations (6.5) by $\frac{d}{dt}\Lambda^{2(s-\alpha)}v^{(R)}$, we obtain

$$
\begin{aligned}
&\frac{\kappa}{2}\frac{d}{dt}\big|\Lambda^s v^{(R)}\big|^2+\big|\Lambda^{s-\alpha}\dot{v}^{(R)}\big|^2 \\
&\quad= C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big\langle\Lambda^{s-\alpha}\nabla\cdot\big(u^{(R)}\theta^{(R)}\big),\Lambda^{s-\alpha}\dot{v}^{(R)}\big\rangle \\
&\quad\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big|\Lambda^{s-\alpha+1}\big(u^{(R)}\theta^{(R)}\big)\big|\cdot\big|\Lambda^{s-\alpha}\dot{v}^{(R)}\big| \\
&\quad\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big|\Lambda^{s-\alpha+1+\sigma_1}\theta^{(R)}\big|\big|\Lambda^{\sigma_2}\theta^{(R)}\big|\cdot\big|\Lambda^{s-\alpha}\dot{v}^{(R)}\big| \\
&\quad\leq C\chi_R\big(|\theta^{(R)}|^2_{\mathcal{W}}\big)\big(\big|\Lambda^{s+\alpha}v^{(R)}\big|^2+\big|\Lambda^{s}v^{(R)}\big|^2+\big|\Lambda^{s-\alpha+1+\sigma_1}z\big|^2\big) \\
&\qquad+\frac{1}{2}\big|\Lambda^{s-\alpha}\dot{v}^{(R)}\big|^2.
\end{aligned}
\tag{6.7}
$$

Here, $\dot{v}^{(R)} = \frac{dv^{(R)}}{dt}$ and in the first equality we used $\operatorname{div} u = 0$, in the second inequality we used Lemmas 2.1 and 2.2, and in the third inequality we used the interpolation inequality (2.3), that $s-\alpha+1+\sigma_1 \leq s+\alpha$ and Young's inequality.

As $\int_0^T |\Lambda^{s+\alpha}v^{(R)}(t)|^2\,dt$ can be dominated by (6.6), we get an a priori estimate for the time derivative $\frac{d}{dt}v^{(R)}$ in $L^2(0,T;H^{s-\alpha})$. Then by [49], we obtain $v^{(R)} \in C([0,T],\mathcal{W})$.

Uniqueness. Let v_1, v_2 be two solutions of equation (6.5) in $C([0,\infty);\mathcal{W})$ and set $w := v_1 - v_2$ and $u_w := u_1 - u_2$, where u_1, u_2 satisfy (1.3) with θ replaced by $\theta_1 = v_1 + z, \theta_2 = v_2 + z$. Then by a similar argument as in the proof of Theorem 4.2, we have for small $0 < \varepsilon_1 < (2\alpha - 1 - \sigma_1) \wedge \sigma_1$ with σ_1 as in (6.6)

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}|\Lambda^{s-\alpha}w|^2 + \kappa|\Lambda^s w|^2 \\
&\quad = -\big(\chi_R(|\theta_1|_{\mathcal{W}}^2) - \chi_R(|\theta_2|_{\mathcal{W}}^2)\big)\langle\Lambda^{s+\varepsilon_1-2\alpha}(u_1\cdot\nabla\theta_1), \Lambda^{s-\varepsilon_1}w\rangle \\
&\qquad - \chi_R(|\theta_2|_{\mathcal{W}}^2)\langle\Lambda^{s-2\alpha}(u_1\cdot\nabla w), \Lambda^s w\rangle \\
&\qquad - \chi_R(|\theta_2|_{\mathcal{W}}^2)\langle\Lambda^{s-2\alpha}(u_w\cdot\nabla\theta_2), \Lambda^s w\rangle \\
&\quad = I + II + III.
\end{aligned}$$

As

$$\big|\chi_R(|\theta_1|_{\mathcal{W}}^2) - \chi_R(|\theta_2|_{\mathcal{W}}^2)\big| \le C(R)|w|_{\mathcal{W}}\big[1_{[0,R+1]}(|\theta_1|_{\mathcal{W}}^2) + 1_{[0,R+1]}(|\theta_2|_{\mathcal{W}}^2)\big],$$

we get for σ_1, σ_2 as in (6.6),

$$\begin{aligned}
I &= -\big(\chi_R(|\theta_1|_{\mathcal{W}}^2) - \chi_R(|\theta_2|_{\mathcal{W}}^2)\big)\langle\Lambda^{s+\varepsilon_1-2\alpha}\nabla\cdot(u_1\theta_1), \Lambda^{s-\varepsilon_1}w\rangle \\
&\le C\big[1_{[0,R+1]}(|\theta_1|_{\mathcal{W}}^2) + 1_{[0,R+1]}(|\theta_2|_{\mathcal{W}}^2)\big] \\
&\qquad \times |w|_{\mathcal{W}}|\Lambda^{s-2\alpha+\varepsilon_1+1+\sigma_1}\theta_1||\Lambda^{\sigma_2}\theta_1||\Lambda^{s-\varepsilon_1}w| \\
&\le C(R, |\theta_1|_{\mathcal{W}}, |\theta_2|_{\mathcal{W}})|w|_{\mathcal{W}}|\Lambda^{s-\varepsilon_1}w| \\
&\le C(R, |\theta_1|_{\mathcal{W}}, |\theta_2|_{\mathcal{W}})|\Lambda^{s-\alpha}w|^2 + \frac{\kappa}{4}|w|_{\mathcal{W}}^2,
\end{aligned} \tag{6.8}$$

where in the first equality we used $\operatorname{div} u_1 = 0$ and in the first inequality we used Lemmas 2.1 and 2.2, in the second inequality we used that $s - 2\alpha + \varepsilon_1 + 1 + \sigma_1 < s$, that is, $\varepsilon_1 < 2\alpha - 1 - \sigma_1$ and in the third inequality we used the interpolation inequality (2.3) and Young's inequality. In a similar way, we obtain

$$\begin{aligned}
II &\le |\Lambda^s w||\Lambda^{s-2\alpha+1}(u_1 w)| \\
&\le C|\Lambda^s w|\big[|\Lambda^{s-2\alpha+1+\sigma_1}\theta_1||\Lambda^{s-\varepsilon_1}w| + |\Lambda^{s-2\alpha+1+\sigma_1}w||\Lambda^s\theta_1|\big] \\
&\le C(R, |\theta_1|_{\mathcal{W}})|\Lambda^{s-\alpha}w|^2 + \frac{\kappa}{4}|w|_{\mathcal{W}}^2,
\end{aligned}$$

where in the first inequality we used $\operatorname{div} u_1 = 0$ and in the second inequality we used Lemmas 2.1 and 2.2 and $s - \varepsilon_1 \ge 1 - \sigma_1$, and in the third inequality we used the interpolation inequality (2.3) and Young's inequality. Similarly,

$$III \le C(R, |\theta_2|_{\mathcal{W}})|\Lambda^{s-\alpha}w|^2 + \frac{\kappa}{4}|w|_{\mathcal{W}}^2.$$

Then we obtain

$$\frac{1}{2}\frac{d}{dt}|\Lambda^{s-\alpha}w|^2+\kappa|\Lambda^s w|^2 \\ \le C\Big(R,\sup_{t\in[0,T]}|\theta_1(t)|_{\mathcal{W}},\sup_{t\in[0,T]}|\theta_2(t)|_{\mathcal{W}}\Big)|\Lambda^{s-\alpha}w|^2+\frac{3\kappa}{4}|w|^2_{\mathcal{W}}.$$

Gronwall's lemma now yields that $|\Lambda^{s-\alpha}w|=0$, which implies $w=0$.

So, equation (6.5) has a unique global weak solution in the space $C([0,\infty);\mathcal{W})$.

Next, we prove (6.3). In order to do so, it is sufficient to show that $P_x^{(R)}[\tau_R<\varepsilon]\le C(\varepsilon,R)$ with $C(\varepsilon,R)\downarrow 0$ as $\varepsilon\downarrow 0$, for all $x\in\mathcal{W}$, with $|x|^2_{\mathcal{W}}\le\frac{R}{8}$. So, fix $\varepsilon>0$ small enough, let $\Theta_{\varepsilon,R}:=\sup_{t\in[0,\varepsilon]}|\Lambda^{s-\alpha+1+\sigma_1}z(t)|$ and assume that $\Theta^2_{\varepsilon,R}\le\frac{R}{8}$. Setting $\varphi(t):=|v^{(R)}|^2_{\mathcal{W}}+\Theta^2_{\varepsilon,R}$, by (6.6) we get $\dot\varphi\le C(R)$. This implies, together with the bounds on x and $\Theta_{\varepsilon,R}$, that

$$\sup_{t\in[0,\varepsilon]}|\theta^{(R)}(t)|^2_{\mathcal{W}}\le R$$

for ε small enough. It follows that $\tau_R\ge\varepsilon$. Hence,

$$P_x^{(R)}[\tau_R<\varepsilon]\le P_x^{(R)}\Big[\sup_{t\in[0,\varepsilon]}|\Lambda^{s+1+\sigma_1-\alpha}z(t)|^2>\frac{R}{8}\Big].$$

Letting $\varepsilon\downarrow 0$, we have $P_x^{(R)}[\tau_R<\varepsilon]\to 0$, and the claim is proved, since the probability above is independent of x.

Finally, the same arguments as in the proof of Theorem 4.2 imply that

$$\theta\big(t\wedge\tau_R(\theta^{(R)})\big)=\theta^{(R)}\big(t\wedge\tau_R(\theta^{(R)})\big)\qquad\forall t,\ P\text{-a.s.}$$

Moreover, since θ is H-valued weakly continuous, we obtain $\tau_R(\theta^{(R)})=\tau_R(\theta)$. □

In order to apply [18], Theorem 5.4, we now only need the following result.

PROPOSITION 6.5. *Fix $\alpha>\frac{2}{3}$. Suppose Hypothesis E.3 holds. For every $R>0$, the transition semigroup $(P_t^{(R)})_{t\ge 0}$ associated to equation (6.2) is $\mathcal{W}$-strong Feller.*

PROOF. We shall provide formal estimates, that can, however, be made rigorous through Galerkin approximations. Let $(\Sigma,\mathcal{F},(\mathcal{F}_t)_{t\ge 0},\mathbb{P})$ be a filtered probability space, $(W_t)_{t\ge 0}$ a cylindrical Wiener process on H and, for every $x\in\mathcal{W}$, let $\theta_x^{(R)}$ be the solution to equation (6.2) with initial value $x\in\mathcal{W}$. By the Bismut–Elworthy–Li formula,

$$D_y\big(P_t^{(R)}\psi\big)(x)=\frac{1}{t}E^{\mathbb{P}}\Big[\psi\big(\theta_x^{(R)}(t)\big)\int_0^t\langle G^{-1}D_y\theta_x^{(R)}(l),dW(l)\rangle\Big],$$

where $D_y(P_t^{(R)}\psi)$ denotes $\langle D(P_t^{(R)}\psi), y\rangle$ for $y \in H$, $D_y\theta_x^{(R)} = D\theta_x^{(R)} \cdot y$ and $D\theta_x^{(R)}$ denotes the derivative of $\theta_x^{(R)}$ with respect to the initial value. Then for $\|\psi\|_\infty \le 1$, by the B–D–G inequality

$$\begin{aligned}
&\big|(P_t^{(R)}\psi)(x_0+h) - (P_t^{(R)}\psi)(x_0)\big| \\
&\quad \le \frac{C}{t} \sup_{\eta\in[0,1]} E^{\mathbb{P}}\bigg[\bigg(\int_0^t \big|G^{-1}D_h\theta_{x_0+\eta h}^{(R)}(l)\big|^2\,dl\bigg)^{1/2}\bigg].
\end{aligned}$$

The proposition is proved once we prove that the right-hand side of the above inequality converges to 0 as $|h|_{\mathcal{W}} \to 0$.

Fix $x \in \mathcal{W}$, $h \in H$ and write $\theta = \theta_x^{(R)}$, $v = v^{(R)}$, $u = u^{(R)}$, $D\theta = D_h\theta$ for simplicity. The term $D\theta$ solves the following equation:

$$\begin{aligned}
&\frac{d}{dt}D\theta + \kappa\Lambda^{2\alpha}(D\theta) \\
&\quad = -\big[\chi_R\big(|\theta|_{\mathcal{W}}^2\big)[Du\cdot\nabla\theta + u\cdot\nabla D\theta] + 2\chi_R'\big(|\theta|_{\mathcal{W}}^2\big)\langle\theta, D\theta\rangle_{\mathcal{W}} u\cdot\nabla\theta\big],
\end{aligned}$$

with initial value $D\theta(0) = h$ and Du satisfying (1.3) with θ replaced by $D\theta$. Multiplying the above equation with $\Lambda^{2s}D\theta$ and taking the inner product in L^2, we have

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}\big|\Lambda^s D\theta\big|^2 + \kappa\big|\Lambda^{s+\alpha}(D\theta)\big|^2 \\
&\quad = -\big\langle\big[\chi_R\big(|\theta|_{\mathcal{W}}^2\big)[Du\cdot\nabla\theta + u\cdot\nabla D\theta] + 2\chi_R'\big(|\theta|_{\mathcal{W}}^2\big)\langle\theta, D\theta\rangle_{\mathcal{W}} u\cdot\nabla\theta\big], \\
&\qquad \Lambda^{2s}D\theta\big\rangle.
\end{aligned}$$

For the first term on the right-hand side, we have for $|\theta|_{\mathcal{W}}^2 \le R$

$$\begin{aligned}
\big|\big\langle Du\cdot\nabla\theta, \Lambda^{2s}D\theta\big\rangle\big| &= \big|\big\langle\Lambda^{s-\alpha}\nabla\cdot(Du\theta), \Lambda^{s+\alpha}D\theta\big\rangle\big| \\
&\le C\big|\Lambda^{s-\alpha+1+\sigma_1}\theta\big|\cdot\big|\Lambda^{\sigma_2}D\theta\big|\cdot\big|\Lambda^{s+\alpha}D\theta\big| \\
&\quad + C\big|\Lambda^{s-\alpha+1+\sigma_1}D\theta\big|\cdot\big|\Lambda^{\sigma_2}\theta\big|\cdot\big|\Lambda^{s+\alpha}D\theta\big| \\
&\le \varepsilon\big|\Lambda^{s+\alpha}D\theta\big|^2 \\
&\quad + C\big(C(R) + \big|\Lambda^{s+\alpha}v\big|^2 + \big|\Lambda^{s-\alpha+1+\sigma_1}z\big|^2\big)\big|\Lambda^s D\theta\big|^2
\end{aligned} \tag{6.9}$$

for σ_1, σ_2 as (6.6), where we used $\operatorname{div} Du = 0$ in the first equality and Lemmas 2.1 and 2.2 in the first inequality as well as the interpolation inequality (2.3) and Young's inequality in the second inequality.

The second term can be estimated similarly. For the third term, by Lemmas 2.1 and 2.2, we have

$$\begin{aligned}
\big|\big\langle u\cdot\nabla\theta, \Lambda^{2s}D\theta\big\rangle\big| &= \big|\big\langle\Lambda^{s-\alpha}\nabla\cdot(u\theta), \Lambda^{s+\alpha}D\theta\big\rangle\big| \\
&\le C\big|\Lambda^{s-\alpha+1+\sigma_1}\theta\big|\big|\Lambda^{\sigma_2}\theta\big|\cdot\big|\Lambda^{s+\alpha}D\theta\big| \\
&\le C\big(\big|\Lambda^{s+\alpha}v\big| + \big|\Lambda^{s-\alpha+1+\sigma_1}z\big|\big)\big|\Lambda^s\theta\big|\big|\Lambda^{s+\alpha}D\theta\big|,
\end{aligned} \tag{6.10}$$

where in the first equality we used $\operatorname{div} u = 0$. Then we obtain

$$\frac{1}{2}\frac{d}{dt}\left|\Lambda^s D\theta\right|^2 + \kappa\left|\Lambda^{s+\alpha}(D\theta)\right|^2 \\ \le \frac{\kappa}{2}\left|\Lambda^{s+\alpha}(D\theta)\right|^2 + C\left(C(R) + \left|\Lambda^{s+\alpha}v\right|^2 + \left|\Lambda^{s-\alpha+1+\sigma_1}z\right|^2\right)\left|\Lambda^s D\theta\right|^2.$$

From Gronwall's inequality and (6.6), we finally get

$$\int_0^t \left|\Lambda^{s+\alpha}(D\theta(l))\right|^2 dl \\ \le C\left|\Lambda^s h\right|^2 + \exp\left(C\int_0^t \left(C(R) + \left|\Lambda^{s+\alpha}v\right|^2 + \left|\Lambda^{s-\alpha+1+\sigma_1}z\right|^2 dl\right)\right)\left|\Lambda^s h\right|^2 \\ \le C\left|\Lambda^s h\right|^2 + \exp\left(C\left(\left|\Lambda^s x\right|^2 + \int_0^t \left(C(R) + \left|\Lambda^{s-\alpha+1+\sigma_1}z\right|^2 dl\right)\right)\right)\left|\Lambda^s h\right|^2.$$

Since by $s - \alpha + 1 + \sigma_1 < s + 2\alpha - 1$, z is a Gaussian random variable in $C([0,\infty); H^{s-\alpha+1+\sigma_1})$ (cf. [9], Proposition 2.15), by Fernique's theorem we could choose t_0 small enough and obtain

$$E\int_0^{t_0} \left|\Lambda^{s+\alpha}(D\theta(l))\right|^2 dl \le c(t_0, R)\left|\Lambda^s h\right|^2,$$

which, as $G^{-1} = Q_0^{-1/2}\Lambda^{s+\alpha}$, implies the assertion for t_0. For general t, by the semigroup property the assertion follows easily. □

6.2. *A support theorem for $\alpha > 2/3$.* A Borel probability measure μ on H is fully supported on $\mathcal{W}$ if $\mu(U) > 0$ for every nonempty open set $U \subset \mathcal{W}$. Set $\mathcal{W}_1 := H^{s-\alpha+1+\sigma_1}$, where σ_1 is the same as (6.6) and we will use it below.

LEMMA 6.6 (Approximate controllability). *Let $R > 0$, $T > 0$. Let $x \in \mathcal{W}$ and $y \in \mathcal{W}$, with $A_\alpha y \in \mathcal{W}_1$, such that*

$$|x|_{\mathcal{W}}^2 \le \frac{R}{2}, \qquad |y|_{\mathcal{W}}^2 \le \frac{R}{2}.$$

Then there exist (a control function) $\omega \in \operatorname{Lip}([0,T]; \mathcal{W}_1)$ and

$$\theta \in C([0,T]; \mathcal{W}) \cap L^2([0,T]; H^{s+\alpha}),$$

such that θ solves the equation

$$\theta(t) - x + \int_0^t A_\alpha\theta(r) + \chi_R\left(|\theta|_{\mathcal{W}}^2\right)u(r)\cdot\nabla\theta(r)\,dr \\ = \omega(t) \qquad dt\text{-a.e. } t \in [0,T], \tag{6.11}$$

with $\theta(0) = x$ and $\theta(T) = y$, and

$$\sup_{t\in[0,T]} \left|\theta(t)\right|_{\mathcal{W}}^2 \le R. \tag{6.12}$$

PROOF. First consider $\omega = 0$. By similar arguments as in Theorems A.1 and A.2, there exist a unique solution $\theta \in C([0, T], \mathcal{W})$. Then by a similar calculation as (6.6), we get

$$\frac{d}{dt}|\theta|_{\mathcal{W}}^2 + \kappa|\Lambda^\alpha \theta|_{\mathcal{W}}^2 \le C(R).$$

Hence, $\theta(t) \in H^{s+\alpha}$ for almost every $t \in [0, T]$ and, by solving again the equation with one of these regular points as initial condition and using Lemmas 2.1 and 2.2 we have

$$\begin{aligned}\frac{d}{dt}|\Lambda^{\alpha+s}\theta|^2 + \kappa|\Lambda^{2\alpha+s}\theta|_{\mathcal{W}}^2 &= \chi_R(|\theta|_{\mathcal{W}}^2)\langle \Lambda^s \nabla \cdot (u\theta), \Lambda^{2\alpha+s}\theta\rangle \\ &\le C\chi_R(|\theta|_{\mathcal{W}}^2)|\Lambda^{2\alpha+s}\theta||\Lambda^{s+1+\sigma_3}\theta|\|\theta\|_{L^p} \\ &\le C(R)|\Lambda^{s+\alpha}\theta|^2 + \frac{\kappa}{2}|\Lambda^{s+2\alpha}\theta|^2,\end{aligned}$$

where $\sigma_3 = \frac{2}{p} < 2\alpha - 1$ and we used $\operatorname{div} u = 0$ in the first equality and $H^s \subset L^p$ and the interpolation inequality (2.3), Young's inequality in the last step. Then by a boot strapping argument, we find a small $T_* \in (0, \frac{T}{2})$ such that $|\theta(t)|_{\mathcal{W}}^2 \le R$ and $A_\alpha \theta(T_*) \in \mathcal{W}_1$ for all $t \le T_*$. Define θ to be the solution above for $t \in [0, T_*]$ and extended by linear interpolation between y and $\theta(T_*)$ in $[T_*, T]$. Then obviously (6.12) follows.

Next, if we set

$$\eta := \partial_t \theta + A_\alpha \theta + \chi_R(|\theta|_{\mathcal{W}}^2) u \cdot \nabla\theta, \qquad T_* \le t \le T,$$

$\omega := 0$ for $t \le T_*$ and $\omega(t) = \int_{T_*}^t \eta_s\, ds$ for $t \in [T_*, T]$, we also have (6.11). It remains to prove that $\eta \in L^\infty(0, T; \mathcal{W}_1)$. For the first two terms of η, this is obvious. For the nonlinear term, we have that

$$|u \cdot \nabla\theta|_{\mathcal{W}_1} = |\nabla \cdot (u\theta)|_{\mathcal{W}_1} \le C|\Lambda^{2\alpha}\theta|_{\mathcal{W}_1}^2$$

for any $\theta \in \mathcal{W}_1$, where in the first equality we used $\operatorname{div} u = 0$ and in the last step we used Lemma 2.1. □

Let $l \in (0, \frac{1}{2})$ and $p > 1$ such that $l - \frac{1}{p} > 0$. Under Hypothesis E.3, we see that for every $\alpha_1 < \frac{s+\alpha-1}{2\alpha}$ the map

$$\omega \mapsto z(\cdot, \omega) : W^{l,p}([0, T]; D(A_\alpha^{\alpha_1})) \to C([0, T]; D(A_\alpha^{\alpha_1+l-1/p-\varepsilon}))$$

is continuous, for all $\varepsilon > 0$ (cf. [12]), where z is the solution to the following equation:

$$z(t) + \int_0^t A_\alpha z(s)\, ds = \omega(t). \tag{6.13}$$

In particular, it is possible to find $\alpha_1 \in (0, \frac{s+\alpha-1}{2\alpha})$ and p such that the above map is continuous from $W^{l,p}([0, T]; D(A_\alpha^{\alpha_1}))$ to $C([0, T]; H^{s-\alpha+1+\sigma_1})$ since $\alpha > \frac{2}{3}$ and $\sigma_1 < 3\alpha - 2$.

LEMMA 6.7 (Continuity with respect to the control functions). *Let l, p and α_1 be chosen as above, and let $\omega_n \to \omega$ in $W^{l,p}([0,T]; D(A_\alpha^{\alpha_1}))$. Let θ be the solution to equation (6.11) corresponding to ω and some initial condition $x \in \mathcal{W}$ (the solution exists by the same arguments as the proof of Theorem A.1), and let*

$$\tau = \inf\{t \geq 0 : |\theta(t)|_{\mathcal{W}}^2 \geq R\},$$

where as usual we set $\inf \varnothing = \infty$. For each $n \in \mathbb{N}$, define similarly θ_n and τ_n corresponding to ω_n with the same initial condition x. If $\tau > T$, then $\tau_n > T$ for n large enough and

$$\theta_n \to \theta \qquad \text{in } C([0,T]; \mathcal{W}).$$

PROOF. Set $v_n := \theta_n - z_n$ for each $n \in \mathbb{N}$, and $v := \theta - z$, where z_n, z are the solutions to (6.13) corresponding to ω_n, ω, respectively. Since $\omega_n \to \omega$ in $W^{l,p}([0,T]; D(A_\alpha^{\alpha_1}))$, we can find a common lower bound for $(\tau_n)_{n\in\mathbb{N}}$ and τ by (6.6). For every time smaller than this lower bound t_0, by (6.6), we have

$$\sup_{(0,t_0)} |\Lambda^s \theta_n|^2 \leq R, \qquad \sup_{(0,t_0)} |\Lambda^s \theta|^2 \leq R, \qquad \sup_{(0,t_0)} |\Lambda^{s-\alpha+1+\sigma_1} z_n| \leq C,$$

and

$$\sup_{(0,t_0)} |\Lambda^{s-\alpha+1+\sigma_1} z| \leq C,$$

$$\int_0^{t_0} |\Lambda^{s+\alpha} v_n(l)|^2 \, dl \leq C(R), \qquad \int_0^{t_0} |\Lambda^{s+\alpha} v(l)|^2 \, dl \leq C(R),$$

where $C(R)$ is a constant depending only on R. Moreover, we obtain for $t \leq t_0$

$$\begin{aligned}
&\frac{d}{dt}|v - v_n|_{\mathcal{W}}^2 + 2\kappa |\Lambda^\alpha (v_n - v)|_{\mathcal{W}}^2 \\
&\quad = \langle u_n \cdot \nabla\theta_n, \Lambda^{2s}(v - v_n)\rangle - \langle u \cdot \nabla\theta, \Lambda^{2s}(v - v_n)\rangle \\
&\quad = \big[\langle (u_{v_n} - u_v) \cdot \nabla\theta_n, \Lambda^{2s}(v - v_n)\rangle + \langle u \cdot \nabla(v_n - v), \Lambda^{2s}(v - v_n)\rangle \\
&\qquad + \langle (u_{z_n} - u_z) \cdot \nabla\theta_n, \Lambda^{2s}(v - v_n)\rangle + \langle u \cdot \nabla(z_n - z), \Lambda^{2s}(v - v_n)\rangle\big],
\end{aligned}$$

where u_{v_n}, u_{z_n} satisfy (1.3) with θ replaced by v_n, z_n, respectively. For the first term on the right-hand side, we have

$$\begin{aligned}
&|\langle (u_{v_n} - u_v) \cdot \nabla\theta_n, \Lambda^{2s}(v - v_n)\rangle| \\
&\quad = |\langle \Lambda^{s-\alpha}\nabla \cdot ((u_{v_n} - u_v)\theta_n), \Lambda^{s+\alpha}(v - v_n)\rangle| \\
&\quad \leq C|\Lambda^{s+\alpha}(v - v_n)||\Lambda^{s-\alpha+1+\sigma_1}(v - v_n)||\Lambda^{\sigma_2}\theta_n| \\
&\qquad + C|\Lambda^{s+\alpha}(v - v_n)||\Lambda^{s-\alpha+1+\sigma_1}\theta_n||\Lambda^{\sigma_2}(v - v_n)| \\
&\quad \leq \frac{\kappa}{4}|\Lambda^{s+\alpha}(v - v_n)|^2 + C(C(R) + |\Lambda^{s+\alpha} v_n|^2)|\Lambda^s(v - v_n)|^2 \\
&\qquad + c|\Lambda^{s-\alpha+1+\sigma_1} z_n|^2 |\Lambda^s(v - v_n)|^2.
\end{aligned}$$

Here, σ_1, σ_2 are as (6.6) and we used $\operatorname{div}(u_{v_n} - u_v) = 0$ in the first equality and Lemmas 2.1 and 2.2 in the first inequality and the interpolation inequality (2.3) and Young's inequality in the last step. The other term can be estimated similarly. Then we obtain

$$\begin{aligned}
&\frac{d}{dt}|v - v_n|_{\mathcal{W}}^2 + 2\kappa\big|\Lambda^\alpha(v_n - v)\big|_{\mathcal{W}}^2 \\
&\quad \le \kappa\big|\Lambda^\alpha(v_n - v)\big|_{\mathcal{W}}^2 \\
&\qquad + C\big(C(R) + \big|\Lambda^\alpha v_n\big|_{\mathcal{W}}^2 + \big|\Lambda^\alpha v\big|_{\mathcal{W}}^2\big)\big(|v - v_n|_{\mathcal{W}}^2 + \big|\Lambda^{s-\alpha+1+\sigma_1}(z - z_n)\big|^2\big).
\end{aligned}$$

Then Gronwall's lemma yields that

$$\begin{aligned}
|v - v_n|_{\mathcal{W}}^2 \le \Theta_n \exp\bigg(C\int_0^t \big(C(R) + \big|\Lambda^\alpha v_n\big|_{\mathcal{W}}^2 + \big|\Lambda^\alpha v\big|_{\mathcal{W}}^2\big)\,dl\bigg) \\
\times \int_0^t \big(C(R) + \big|\Lambda^\alpha v_n\big|_{\mathcal{W}}^2 + \big|\Lambda^\alpha v\big|_{\mathcal{W}}^2\big)\,dl,
\end{aligned}$$

where $\Theta_n = \sup_{[0,T]} |\Lambda^{s-\alpha+1+\sigma_1}(z - z_n)|$. We conclude $\theta_n \to \theta$ in $C([0, T]; \mathcal{W})$. Now, since $\tau > T$, if $S = \sup_{t\in[0,T]} |\Lambda^s\theta(t)|^2$, then $S < R$ and we find $\delta > 0$ (depending only on R and S) and $n_0 \in \mathbb{N}$ such that $\Theta_n^2 < \delta$ and $|v_n - v|_{\mathcal{W}}^2 < \delta$ for all $n \ge n_0$, and so

$$\big|\theta_n(t)\big|_{\mathcal{W}} \le \big|v_n(t) - v(t)\big|_{\mathcal{W}} + \Theta_n + \big|\theta(t)\big|_{\mathcal{W}} \le 2\sqrt{\delta} + \sqrt{S} \le \sqrt{R - \delta}.$$

Then $\tau_n > T$ for all $n \ge n_0$. $\square$

THEOREM 6.8. *Fix $\alpha > \frac{2}{3}$. Suppose Hypothesis E.3 holds and for $x \in \mathcal{W}$ let P_x be the distribution of the solution of (3.1) with initial value $\theta(0) = x$. Then for every $x \in \mathcal{W}$ and every $T > 0$, the image measure of P_x at time T is fully supported on $\mathcal{W}$.*

PROOF. Fix $x \in \mathcal{W}$ and $T > 0$. We need to show that for every $y \in \mathcal{W}$ and $\varepsilon > 0$, $P_x[|\theta_T - y|_{\mathcal{W}} < \varepsilon] > 0$. Let $\bar{y} \in \mathcal{W} \cap D(A_\alpha)$ such that $A_\alpha\bar{y} \in \mathcal{W}_1$ and $|y - \bar{y}|_{\mathcal{W}} < \frac{\varepsilon}{2}$. Choose $R > 0$ such that $3|x|_{\mathcal{W}}^2 < R$ and $3|y|_{\mathcal{W}}^2 < R$. Then by Theorem 6.4,

$$\begin{aligned}
P_x\big[|\theta_T - y|_{\mathcal{W}} < \varepsilon\big] \ge P_x\bigg[|\theta_T - \bar{y}|_{\mathcal{W}} < \frac{\varepsilon}{2}\bigg] \ge P_x\bigg[|\theta_T - \bar{y}|_{\mathcal{W}} < \frac{\varepsilon}{2}, \tau_R > T\bigg] \\
= P_x^{(R)}\bigg[|\theta_T - \bar{y}|_{\mathcal{W}} < \frac{\varepsilon}{2}, \tau_R > T\bigg].
\end{aligned}$$

By Lemma 6.6, there is a control $\bar{\omega} \in W^{l,p}([0, T]; D(A_\alpha^{\alpha_1}))$, with l, p and α_1 chosen as in Lemma 6.7, such that the solution $\bar{\theta}$ to the control problem (6.11) corresponding to $\bar{\omega}$ satisfies $\bar{\theta}(0) = x, \bar{\theta}(T) = \bar{y}$ and $|\bar{\theta}(t)|_{\mathcal{W}}^2 \le \frac{2}{3}R$. By

Lemma 6.7, there exists $\delta > 0$ such that for all $\omega \in W^{l,p}([0,T]; D(A_\alpha^{\alpha_1}))$ with $|\omega - \bar{\omega}|_{W^{l,p}([0,T];D(A_\alpha^{\alpha_1}))} < \delta$, we have

$$\left|\theta(T,\omega) - \bar{y}\right|_{\mathcal{W}} < \frac{\varepsilon}{2} \quad \text{and} \quad \sup_{t\in[0,T]} \left|\theta(t,\omega)\right|_{\mathcal{W}}^2 < R,$$

where $\theta(\cdot,\omega)$ is the solution to the control problem (6.11) corresponding to ω and starting at x. Hence,

$$P_x^{(R)}\left[|\theta_T - \bar{y}|_{\mathcal{W}} < \frac{\varepsilon}{2}, \tau_R > T\right] \geq P_x^{(R)}\left[|\eta - \bar{\omega}|_{W^{l,p}([0,T];D(A_\alpha^{\alpha_1}))} < \delta\right],$$

where $\eta_t = \theta_t - x + \int_0^t (A_\alpha \theta_s + \chi_R(|\theta_s|_{\mathcal{W}}^2) u \cdot \nabla\theta_s)\, ds$, hence $\theta_T = \theta(T,\eta)$, and the right-hand side of the inequality above is strictly positive since by Hypothesis E.3 η is a Gaussian process in $D(A_\alpha^{\alpha_1})$. □

THEOREM 6.9. *Let $\alpha > \frac{2}{3}$ and suppose Hypothesis* E.3 *holds. Then there exists a unique invariant measure ν on $\mathcal{W}$ for the transition semigroup $(P_t)_{t\geq 0}$. Moreover*:

(i) *The invariant measure ν is ergodic.*
(ii) *The transition semigroup $(P_t)_{t\geq 0}$ is $\mathcal{W}$-strong Feller, irreducible and, therefore, strongly mixing. Furthermore, $P_t(x,dy), t > 0, x \in \mathcal{W}$, are mutually equivalent.*
(iii) *There exist $0 < \delta_1 < \frac{s+\alpha-1}{2\alpha}$ and $0 < \gamma_0 < 1$ such that*

$$\int |A_\alpha^{\delta_1} x|_{\mathcal{W}}^{2\gamma_0}\, d\nu < \infty.$$

PROOF. By similar methods as the proof of Theorem 5.12, we obtain the existence of the invariant measures. In fact, under Hypothesis E.3, we could choose the following approximation:

$$d\theta_n(t) + A_\alpha\theta_n(t)\, dt + u_n(t)\cdot\nabla\theta_n(t)\, dt = k_{\delta_n} * G\, dW(t),$$

with initial data $\theta_n(0) = x \in H^s$, u_n satisfying (1.3) with θ replaced by θ_n and k_{δ_n} is the periodic Poisson kernel as in the proof of Theorem 3.3. By the same arguments as Theorems A.1 and A.2, we obtain that there exist a unique solution to the above equation with $\theta_n \in C([0,\infty), H^s) \cap L^2_{\mathrm{loc}}([0,\infty), H^{s+\alpha})$ P-a.s. Then do the same calculations for θ_n as in Lemma 5.11, we obtain that there exists $0 < \gamma_0 < 1$, $0 < \tilde{\delta}_1 < s + \alpha - 1$ such that

$$E\left[\int_0^t \left|\Lambda^{\tilde{\delta}_1+s}\theta_n\right|^{2\gamma_0} dr\right] \leq C(1+t)\left(\|x\|_{L^q}^q + 1\right).$$

Choose $x_0 \in H^1$ and define

$$\mu_t = \frac{1}{t}\int_0^t P_r^* \delta_{x_0}\, dr.$$

Since by similar arguments as in the proof of Theorem A.1, we have P-a.s. $\theta_n \to \theta$ in $L^2([0,T],H)$ and for $2\alpha\delta_1 \le \tilde{\delta}_1$, $0<\gamma_0<1$

$$\int |A_\alpha^{\delta_1} x|_{H^s}^{2\gamma_0} \mu_t(dx) = \frac{1}{t} E_{x_0}\left[\int_0^t |A_\alpha^{\delta_1}\theta|_{H^s}^{2\gamma_0}\,dr\right],$$

by the above estimates we have for $t>1$

$$\int |A_\alpha^{\delta_1} x|_{H^s}^{2\gamma_0} \mu_t(dx) \le C.$$

This implies that μ_t is tight on H^s. Hence, any limit point of μ_t is an invariant measure for $(P_t)_{t\ge 0}$. Therefore, by Doob's theorem, the strongly mixing property is a consequence of Theorem 6.3 and Theorem 6.8. □

REMARK 6.10 (Mildly degenerate noise). We can also consider the ergodicity of the equation driven by a mildly degenerate noise as in [15]. For this, we have to use an extension of the Bismut–Elworthy–Li formula. We have the same problem as explained in Remark 6.2. So, we can just get the result for $\alpha > 2/3$.

6.3. *Exponential convergence for* $\alpha > \frac{2}{3}$. In this subsection, we assume that $\alpha > \frac{2}{3}$ and $s > 3 - 2\alpha$. Then under Hypothesis E.3 the associated O–U process $z \in C([0,\infty), H^{2+\delta_0})$ for some $0 < \delta_0 < s + 2\alpha - 3$.

LEMMA 6.11. *Fix* $\alpha > 2/3$. *Let* θ *denote the solution of* (3.1) *and take* $p > \frac{2}{3\alpha-2}$, *then for every* $R_0 \ge 1$, *there exist values* $T_1 = T_1(R_0)$ *and* $\tilde{C}_1 = \tilde{C}_1(R_0)$ *such that if* $\sup_{t\in[0,T_1]} \|\theta(t)\|_{L^p}^p \le R_0$, *and* $\sup_{t\in[0,T_1]} |\Lambda^{s+2\alpha-1-\varepsilon} z(t)|^2 \le R_0$ *for some* $0 < \varepsilon < 3\alpha - 2 - \frac{2}{p}$, *then* $|\Lambda^{s+\delta}\theta(T_1)|^2 \le \tilde{C}_1$ *for some* $\delta > 0$.

PROOF. For $v = \theta - z$, we have the following estimate:

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}|v|^2 + \kappa|\Lambda^\alpha v|^2 &= \langle -u\cdot\nabla(v+z), v\rangle = \langle -u\cdot\nabla z, v\rangle \\ &\le C\|\nabla z\|_{L^\infty}\big[|v|^2 + |v|\cdot|z|\big],\end{aligned}$$

which implies that there exist $\tilde{C}_0 = \tilde{C}_0(R_0) > 0$ and for P-a.s. ω, $\exists 0 < t_0(\omega) < 1$ such that

$$|\Lambda^\alpha \theta(t_0)|^2 \le \tilde{C}_0.$$

For any $\tilde{r} > 0$ with $\tilde{r} - \alpha + 1 + \sigma_3 < s + 2\alpha - 1 - \varepsilon$ for $\sigma_3 = \frac{2}{p}$, we have the following a priori estimate for v, $r = \frac{\alpha}{\alpha - 1/2 - 1/p}$:

$$\begin{aligned}&\frac{d}{dt}|\Lambda^{\tilde{r}} v|^2 + 2\kappa|\Lambda^{\tilde{r}+\alpha} v|^2 \\ &\quad\le 2|\langle \Lambda^{\tilde{r}-\alpha}\nabla\cdot(u\theta), \Lambda^{\tilde{r}+\alpha} v\rangle| \qquad (6.14)\end{aligned}$$

$$\leq C\big|\Lambda^{\tilde r+\alpha}v\big|\cdot\big|\Lambda^{\tilde r-\alpha+1+\sigma_3}\theta\big|\cdot\|\theta\|_{L^p}$$
$$\leq \frac{\kappa}{4}\big|\Lambda^{\tilde r+\alpha}v\big|^2+C\|\theta\|_{L^p}^r\big|\Lambda^{\tilde r}v\big|^2+C\big|\Lambda^{\tilde r-\alpha+1+\sigma_3}z\big|^2\cdot\|\theta\|_{L^p}^2,$$

where we used $\operatorname{div} u=0$ in the first inequality and Lemmas 2.1, 2.2 in the second inequality and the interpolation inequality (2.3) and Young's inequality in the last inequality. We choose the approximation v_n as in the proof of Theorem A.1 with initial time $t=0$ replaced by initial time $t=t_0(\omega)$. Then by a similar argument as in the proof of Theorem A.1 we have the following L^p-norm estimate of v_n,

$$\frac{d}{dt}\|v_n\|_{L^p}^p\leq Cp\|\nabla z\|_\infty\big(\|v_n\|_{L^p}^p+\|z\|_{L^p}\|v_n\|_{L^p}^{p-1}\big).$$

Thus, we have

$$\frac{d}{dt}\|v_n\|_{L^p}\leq C\|\nabla z\|_\infty\big(\|v_n\|_{L^p}+\|z\|_{L^p}\big).$$

Then by Gronwall's lemma and $s>3-2\alpha$, we obtain the uniform L^p-norm estimates as (A.6) for v_n. Moreover, by (6.14) and Gronwall's lemma, we obtain the uniform H^α-norm estimates as (A.7) for v_n. By a similar argument as in the proof of Theorem A.1, we have v_n converges to some process $\tilde v$ in $L^2([t_0,T],H)$ such that $\tilde v+z$ is the solution of (3.1) in $[t_0,T]$. Then by the uniqueness proof in Theorem 4.2, we have $\tilde v=v$, which implies for P-a.s. ω, $v\in L^\infty_{\mathrm{loc}}([t_0,\infty),H^\alpha)\cap L^2_{\mathrm{loc}}([t_0,\infty),H^{2\alpha})$. Therefore, (6.14) also holds for v with $\tilde r=\alpha$, which implies that

$$\big|\Lambda^\alpha v(t)\big|^2+\kappa\int_{t_0}^t\big|\Lambda^{2\alpha}v(l)\big|^2\,dl\leq\big(\big|\Lambda^\alpha v(t_0)\big|^2+C(R_0)\big)\big(\exp\big[C(R_0)t\big]+1\big),$$

which implies that there exist $\tilde C_1=\tilde C_1(R_0)>0$ and $\tilde T_0(R_0)$ such that $|\Lambda^\alpha v(\tilde T_0)|\leq\tilde C_1(R_0)$. Moreover, there exists $t_1=t_1(\omega)>t_0(\omega)$ such that $|\Lambda^{2\alpha}v(t_1)|\leq\tilde C_1$. Using (6.14) for $\tilde r=2\alpha$ and by similar arguments as above, we obtain that there exists $T_0=T_0(R_0)$ independent of ω such that $|\Lambda^{2\alpha}v(T_0)|\leq\tilde C_1$. Then we proceed analogously and obtain that there exists $T_1=T_1(R_0)>T_0(R_0)$ such that $|\Lambda^{s+\delta}v(T_1)|\leq\tilde C_1$ for some $0<\delta<3\alpha-2-\sigma_3-\varepsilon$. □

LEMMA 6.12. *Let $\alpha>2/3$. Suppose Hypothesis E.3 holds with $s>3-2\alpha$. Then for each $R\geq 1$ there exist $T_1>0$ and a compact subset $K\subset\mathcal{W}$ such that*

$$\inf_{\|x\|_{L^p}\leq R}P_{T_1}(x,K)>0$$

for p in Lemma 6.11.

PROOF. Define $K:=\{x:|\Lambda^{s+\delta}x|^2\leq\tilde C_1(R_0)\}$, where $\tilde C_1(R_0),\delta$ comes from the previous lemma. By Lemma 6.11, for $R\leq R_0$, we have

$$\inf_{\|x\|_{L^p}\leq R}P_{T_1}(x,K)\geq\inf_{\|x\|_{L^p}\leq R}\Big(1-P_x\Big[\sup_{t\in[0,T_1]}\big|\Lambda^{s+2\alpha-1-\varepsilon}z(t)\big|^2>R_0\Big]$$
$$-P_x\Big[\sup_{t\in[0,T_1]}\|\theta(t)\|_{L^p}^p>R_0\Big]\Big).$$

Under Hypothesis E.3, since z is a Gaussian process, one deduces that there exist $\eta, C > 0$ such that

$$P_x\Big[\sup_{t\in[0,T_1]} \big|\Lambda^{s+2\alpha-1-\varepsilon} z(t)\big|^2 > R_0\Big] \le Ce^{-\eta(R_0^2/T_1)}$$

(see, e.g., [17], Proposition 15). Also by Theorem 3.3, we obtain

$$\sup_{\|x\|_{L^p}\le R} P_x\Big[\sup_{t\in[0,T_1]} \|\theta(t)\|_{L^p}^p > R_0\Big] \le \sup_{\|x\|_{L^p}\le R} \frac{E_x[\sup_{t\in[0,T_1]} \|\theta(t)\|_{L^p}^p]}{R_0} \le \frac{C(R)}{R_0}.$$

Choosing R_0 big enough, we prove the assertion. □

The exponential convergence now follows from Lemma 6.12 and an abstract result of [19], Theorem 3.1. For $p > \frac{2}{3\alpha-2}$ let $V : L^p \to \mathbb{R}$ be a measurable function and define $\|\phi\|_{\mathrm{V}} := \sup_{x\in L^p} \frac{|\phi(x)|}{V(x)}$ and $\|\nu\|_{\mathrm{V}} := \sup_{\|\phi\|_{\mathrm{V}}\le 1}\langle \nu, \phi\rangle$ for a signed measure ν.

THEOREM 6.13. *Let $\alpha > 2/3$. Assume that Hypothesis E.3 holds with $s > 3 - 2\alpha$ and let $V(x) := 1 + \|x\|_{L^p}^p$ for $p > \frac{2}{3\alpha-2}$. Then there exist $C_{\exp} > 0$ and $a > 0$ such that*

$$\|P_t^*\delta_{x_0} - \mu\|_{\mathrm{var}} \le \|P_t^*\delta_{x_0} - \mu\|_{\mathrm{V}} \le C_{\exp}(1 + \|x_0\|_{L^p}^p)e^{-at}$$

for all $t > 0$ and $x_0 \in L^p$, where $\|\cdot\|_{\mathrm{var}}$ is the total variation distance on measures.

PROOF. By [19], Theorem 3.1, we need to verify the following four conditions:

1. the measures $(P_t(x, \cdot))_{t>0, x\in L^p}$ are equivalent,
2. $x \to P_t(x, \Gamma)$ is continuous in $\mathcal{W}$ for all $t > 0$ and all Borel sets $\Gamma \subset H$,
3. for each $R \ge 1$ there exist $T_1 > 0$ and a compact subset $K \subset \mathcal{W}$ such that
$$\inf_{\|x\|_{L^p}\le R} P_{T_1}(x, K) > 0,$$
4. there exist $k, b, c > 0$ such that for all $t \ge 0$,
$$E^{P_x}\big[\|\theta(t)\|_{L^p}^p\big] \le k\|x\|_{L^p}^p e^{-bt} + c.$$

Condition 1 can be verified by [19], Lemma 3.2, and $P_t(x, \mathcal{W}) = 1$ for $x \in L^p$ since for fixed $t > 0$ the solution θ will go into H^s space if the initial value $x \in L^p$. Other conditions can be verified by Theorem 6.9, Lemma 6.12 and Proposition 5.6. □

REMARK 6.14. For $\alpha > \frac{3}{4}$, we could get a better result following a similar argument as in [46]. Namely, there exist $C_{\exp} > 0$ and $a > 0$ such that

$$\|P_t^*\delta_{x_0} - \mu\|_{\mathrm{TV}} \le \|P_t^*\delta_{x_0} - \mu\|_{\mathrm{V}} \le C_{\exp}(1 + |x_0|^2)e^{-at}$$

for all $t > 0$ and $x_0 \in H$. Here, P_t could be every Markov selection obtained in Theorem C.5 associated to the solution of equation (3.1). The reason why $\alpha > \frac{3}{4}$ is needed is as follows.

As in Theorem 6.3, we can prove P_t is H^s-strong Feller with $s > 3 - 3\alpha$. And for a solution θ of equation (3.1) starting from $x \in H$, we can only prove that it will enter H^α under Hypothesis E.3. If the process θ enters H^s, we can prove that it satisfies the above four conditions. Hence, to obtain exponential convergence for every $x \in H$, we need the process starting from $x \in H$ to enter H^s. Hence, we need $3 - 3\alpha < s \le \alpha$, that is, $\alpha > \frac{3}{4}$.

APPENDIX A

In this appendix, we construct a measurable map associated with the stochastic quasi-geostrophic equation, which will be used in the proof of Section 6. This proof is similar as done in [50], Section 3. Here, we give it for the reader's convenience.

Assume that for any $m < 2 + \sigma$, $z \in C((0, \infty), H^m)$ with σ in Hypothesis E.1. Then consider the following equation:

$$\frac{dv}{dt} + A_\alpha v + (u_v + u_z) \cdot \nabla(v + z) = 0. \tag{A.1}$$

For (A.1), we obtain the following existence and uniqueness result if the initial value starts from H^1.

THEOREM A.1. *Fix $\alpha > 1/2$. Suppose that for any $m < 2 + \sigma$, $z \in C((0, \infty), H^m)$. For any $v_0 \in H^1$, there exists a unique solution $v \in L^\infty_{\mathrm{loc}}([0, \infty); H^1) \cap L^2_{\mathrm{loc}}([0, \infty); H^{1+\alpha})$ of equation (A.1) with $v(0) = v_0$, that is, for any $\varphi \in C^1(\mathbb{T}^2)$*

$$\langle v(t), \varphi\rangle - \langle v_0, \varphi\rangle + \int_0^t \langle A_\alpha^{1/2} v(r), A_\alpha^{1/2}\varphi\rangle\, dr - \int_0^t \langle (u_v + u_z)(r) \cdot \nabla\varphi, (v + z)(r)\rangle\, dr = 0,$$

where u_v, u_z satisfy (1.3) with θ replaced by v, z, respectively.

PROOF. We construct an approximation of (A.1) by a similar construction as in the proof of Theorem 3.3.

We pick a smooth $\phi \ge 0$, with $\operatorname{supp}\phi \subset [1, 2]$, $\int_0^\infty \phi = 1$, and for $\delta > 0$ let

$$U_\delta[\theta](t) := \int_0^\infty \phi(\tau)\big(k_\delta * R^\perp\theta\big)(t - \delta\tau)\, d\tau,$$

where k_δ is the periodic Poisson kernel in $\mathbb{T}^2$ given by $\widehat{k_\delta}(\zeta) = e^{-\delta|\zeta|}$, $\zeta \in \mathbb{Z}^2$, and we set $\theta(t) = 0, t < 0$. We take a zero sequence $\delta_n, n \in \mathbb{N}$, and consider the equation

$$dv_n(t) + A_\alpha v_n(t)\, dt + u_n(t) \cdot \nabla\big(v_n(t) + z\big)\, dt = 0, \tag{A.2}$$

with initial data $v_n(0) = v_0$ and $u_n = U_{\delta_n}[v_n + z]$. For a fixed n, this is a linear equation in v_n on each subinterval $[t_k, t_{k+1}]$ with $t_k = k\delta_n$, since u_n is determined by the values of v_n on the two previous subintervals. By a similar argument as in the proof of Theorem 3.3, we obtain the existence and uniqueness of a solution $v_n \in C([0, T], H^1) \cap L^2([0, T], H^{1+\alpha})$ to (A.2) (for more details, we refer to [50]). Now we take any p satisfying $\frac{2}{2\alpha-1} < p < \infty$. From now on, we fix such p and we have $H^1 \subset L^p$ by Lemma 2.2. Since the periodic Riesz transform is bounded on L^p, we have for $t > 0$

$$\sup_{[0,t]} \|U_\delta[\theta]\|_{L^p} \le C \sup_{[0,t]} \|\theta\|_{L^p}, \tag{A.3}$$

and also

$$\int_0^t \|U_\delta[\theta]\|_{L^p}^p \, d\tau \le C \int_0^t \|\theta\|_{L^p}^p \, d\tau. \tag{A.4}$$

By Lemma 5.5, we obtain for v_n the following inequality by taking inner product with $|v_n|^{p-2}v_n$ in L^2:

$$\begin{aligned} \frac{d}{dt}\|v_n\|_{L^p}^p + 2\lambda_1\|v_n\|_{L^p}^p &\le p\left|\langle u_n \cdot \nabla(v_n + z), |v_n|^{p-2}v_n\rangle\right| \\ &\le p\|\nabla z\|_\infty \|u_n\|_{L^p}\|v_n\|_{L^p}^{p-1}, \end{aligned} \tag{A.5}$$

where we used $\operatorname{div} u_n = 0$ and Hölder's inequality in the last inequality. Therefore,

$$\begin{aligned} &\|v_n(t)\|_{L^p}^p - \|v_n(0)\|_{L^p}^p + \int_0^t 2\lambda_1 \|v_n(\tau)\|_{L^p}^p \, d\tau \\ &\quad\le \varepsilon \int_0^t \left(\|u_n\|_{L^p}^p + \|v_n\|_{L^p}^p\right) d\tau + pC(\varepsilon) \int_0^t \|\nabla z\|_\infty^{p/(p-1)} \|v_n\|_{L^p}^p \, d\tau \\ &\quad\le \varepsilon \int_0^t \|v_n\|_{L^p}^p \, d\tau + pC(\varepsilon) \int_0^t \|\nabla z\|_\infty^{p/(p-1)} \|v_n\|_{L^p}^p \, d\tau + C \int_0^t \|z\|_{L^p}^p \, d\tau, \end{aligned}$$

where we used (A.4) in the last inequality. Then Gronwall's lemma and $H^1 \subset L^p$ yield that for any $T \ge 0$

$$\sup_{t \in [0,T]} \|v_n(t)\|_{L^p} \le C, \tag{A.6}$$

where C is a constant independent of n.

Moreover, we get the following estimate by taking the inner product in L^2 with Λe_k for (A.2), multiplying both sides by $\langle v, \Lambda e_k\rangle$ and summing up over k:

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}|\Lambda v_n|^2 + \kappa\left|\Lambda^{1+\alpha} v_n\right|^2 \\ &\quad\le \left|\Lambda^{1-\alpha}\left(u_n \cdot \nabla(v_n + z)\right)\right|\left|\Lambda^{1+\alpha} v_n\right| \\ &\quad\le C\left|\Lambda^{1+\alpha} v_n\right|\left[\left|\Lambda^{2-\alpha+\sigma_1}(v_n + z)\right|\|u_n\|_{L^p} + \left|\Lambda^{2-\alpha+\sigma_1} u_n\right|\|v_n + z\|_{L^p}\right], \end{aligned}$$

where $\sigma_1 = 2/p < (2\alpha - 1)$ and we used Lemma 2.1 in the last inequality. Hence, we obtain that for $r = \frac{2\alpha}{2\alpha-1-\sigma_1}$,

$$\begin{aligned}
&\frac{1}{2}\big(|\Lambda v_n(t)|^2 - |\Lambda v_n(0)|^2\big) + \kappa \int_0^t |\Lambda^{1+\alpha} v_n|^2 \, d\tau \\
&\quad \le C \int_0^t |\Lambda^{1+\alpha} v_n| \big[|\Lambda^{2-\alpha+\sigma_1}(v_n + z)| \|u_n\|_{L^p} + |\Lambda^{2-\alpha+\sigma_1} u_n| \|v_n + z\|_{L^p} \big] \, d\tau \\
&\quad \le \frac{\kappa}{2} \int_0^t |\Lambda^{1+\alpha} v_n|^2 d\tau \\
&\qquad + C\Big[\sup_{t\in[0,T]} \big(\|v_n(t) + z(t)\|_{L^p}^r + \|v_n(t) + z(t)\|_{L^p}^2 \big) + 1 \Big] \\
&\qquad \times \int_0^t |\Lambda v_n|^2 + |\Lambda^{2-\alpha+\sigma_1} z|^2 \, d\tau,
\end{aligned}$$

where we used (A.3), (A.4), the interpolation inequality (2.3) and Young's inequality in the last inequality. By Gronwall's lemma and (A.6), we get that for $v_0 \in H^1$

$$\sup_{0\le t\le T} |\Lambda v_n(t)|^2 + \kappa \int_0^T |\Lambda^{1+\alpha} v_n|^2 \, d\tau \le C, \tag{A.7}$$

where C is a constant independent of n. Now decompose v_n as

$$v_n(t) = v_0 - \int_0^t A_\alpha v_n(s) \, ds - \int_0^t \big(u_n(s) \cdot \nabla\big(v_n(s) + z(s)\big)\big) \, ds.$$

By (A.7), we obtain

$$\left\| \int_0^\cdot A_\alpha v_n(s) \, ds \right\|_{W^{1,2}(0,T,H^{-\alpha})} \le C$$

and

$$\left\| \int_0^\cdot \big(u_n(s) \cdot \nabla\big(v_n(s) + z(s)\big)\big) \, ds \right\|_{W^{1,2}(0,T,H^{-3})} \le C.$$

So, we have proved

$$\|v_n\|_{W^{1,2}([0,T],H^{-3})} \le C,$$

where C is a constant independent of n. By the compactness embedding $W^{1,2}([0,T], H^{-3}) \cap L^2([0,T], H^{1+\alpha}) \subset L^2([0,T], H^1)$ we have that there exists a subsequence of v_n converging in $L^2([0,T], H^1)$ to a solution $v \in L^\infty_{\mathrm{loc}}([0,\infty); H^1) \cap L^2_{\mathrm{loc}}([0,\infty); H^{1+\alpha})$ of equation (A.1). Thus, (A.7) is also satisfied for v. Uniqueness can be deduced from a similar argument as in the proof of Theorem 4.2. □

THEOREM A.2. *Fix $\alpha > 1/2$. Suppose that for any $m < 2 + \sigma$, $z \in C([0,\infty), H^m)$. The solution v obtained in Theorem A.1 is in $C([0,\infty); H^1)$.*

PROOF. It is sufficient to show that

$$\Lambda \frac{dv}{dt} \in L^2_{\mathrm{loc}}\big([0,\infty); H^{-\alpha}\big).$$

For φ smooth enough, we have

$$\begin{aligned}\left|\left\langle \frac{dv}{dt}, \Lambda\varphi\right\rangle\right| &= |\kappa\langle -\Lambda^\alpha v, \Lambda^{1+\alpha}\varphi\rangle - \langle (u\cdot\nabla(\Lambda\varphi)), v+z\rangle\\ &\le \big[\kappa|\Lambda^{1+\alpha}v| + C|\Lambda^{2-\alpha}(u\cdot(v+z))|\big]|\Lambda^\alpha\varphi|\\ &\le C\big[|\Lambda^{1+\alpha}v| + |\Lambda^{2-\alpha+\sigma_1}(v+z)|\|v+z\|_{L^p}\big]|\Lambda^\alpha\varphi|,\end{aligned}$$

where $0<\sigma_1<2\alpha-1$, $p=\frac{2}{\sigma_1}$ and we used Lemma 2.1 in the last inequality. Then

$$\left\|\Lambda\frac{dv}{dt}\right\|_{H^{-\alpha}} \le C\big(\|v+z\|_{L^p}+1\big)|\Lambda^{1+\alpha}v| + C\|v+z\|_{L^p}|\Lambda^{2-\alpha+\sigma_1}z|.$$

By (A.6) and (A.7), we obtain for $0<T<\infty$

$$\int_0^T \left\|\Lambda\frac{dv}{dt}(\tau)\right\|^2_{H^{-\alpha}}\,d\tau<\infty,$$

which implies that $v\in C([0,\infty);H^1)$. $\square$

THEOREM A.3. *Fix $\alpha>1/2$. Suppose that for any $m<2+\sigma$, $z\in C([0,\infty), H^m)$. For any fixed $t>0$, the map $v_0\mapsto v(t,v_0)$ is a continuous map from H^1 into itself, where $v(t,v_0)$ is the solution of equation (A.1) with $v(0)=v_0$.*

PROOF. Let v_1, v_2 be two solutions of (A.1) and $\zeta=v_1-v_2$, $\theta_1=v_1+z$, $\theta_2=v_2+z$. Then ζ satisfies the following equation:

$$\left(\frac{d}{dt}\zeta,\varphi\right) + \kappa\big(\Lambda^\alpha\zeta,\Lambda^\alpha\varphi\big) = -(u_1\cdot\nabla\zeta,\varphi) - (u_\zeta\cdot\nabla\theta_2,\varphi),$$

where u_1, u_ζ satisfy (1.3) with θ replaced by θ_1,ζ, respectively.

Taking $\varphi=\Lambda e_k$, multiplying both sides by $\langle\zeta,\Lambda e_k\rangle$ and summing up over k we have the following estimate since $v_i\in C([0,\infty);H^1)\cap L^2_{\mathrm{loc}}([0,\infty);H^{1+\alpha})$, $i=1,2$, by Theorems A.1 and A.2:

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}|\Lambda\zeta|^2 + \kappa|\Lambda^{1+\alpha}\zeta|^2\\ &\quad= -\langle\Lambda(u_1\cdot\nabla\zeta),\Lambda\zeta\rangle - \langle u_\zeta\cdot\nabla\theta_2,\Lambda^2\zeta\rangle\\ &\quad\le C|\Lambda^{1+\alpha}\zeta|\big[|\Lambda^{2-\alpha}(u_\zeta\theta_2)| + |\Lambda^{2-\alpha}(u_1\zeta)|\big]\\ &\quad\le C|\Lambda^{1+\alpha}\zeta|\big[|\Lambda^{2-\alpha+\sigma_1}\zeta||\Lambda^{\sigma_2}\theta_2| + |\Lambda^{2-\alpha+\sigma_1}\theta_2||\Lambda^{\sigma_2}\zeta|\\ &\qquad\qquad + |\Lambda^{2-\alpha+\sigma_1}\theta_1||\Lambda^{\sigma_2}\zeta| + |\Lambda^{2-\alpha+\sigma_1}\zeta||\Lambda^{\sigma_2}\theta_1|\big]\end{aligned}$$

$$\le \frac{\kappa}{2}|\Lambda^{1+\alpha}\zeta|^2$$
$$+ C\big[|\Lambda\theta_2|^r + |\Lambda\theta_1|^r + \big|\Lambda^{1+\alpha}v_2\big|^2 + \big|\Lambda^{2-\alpha+\sigma_1}z\big|^2 + \big|\Lambda^{s+\alpha}v_1\big|^2\big]|\Lambda\zeta|^2,$$

where $r = \frac{2\alpha}{2\alpha-1-\sigma_1}$, $\sigma_2 = 1-\sigma_1$ for some $0 < \sigma_1 < (2\alpha-1)$ and we used Lemma 2.1 in the second inequality and Lemma 2.2, the interpolation inequality (2.3), $H^1 \subset H^{\sigma_2}$ and Young's inequality in the last inequality. Then Gronwall's lemma yields that

$$|\Lambda\zeta|^2 \le C\big|\Lambda\zeta(0)\big|^2 \exp\biggl\{\int_0^T \big|\Lambda\theta_2(\tau)\big|^r + \big|\Lambda\theta_1(\tau)\big|^r + \big|\Lambda^{2-\alpha+\sigma}z\big|^2 + \big|\Lambda^{1+\alpha}v_1(\tau)\big|^2 + \big|\Lambda^{1+\alpha}v_2(\tau)\big|^2\,d\tau\biggr\}.$$

Thus, the result follows. □

Now for $v_0 \in H^1$, $\overline{W} \in C(\mathbb{R}^+, H^{-1-\varepsilon_0})$ we define

$$v(t, \overline{W}, v_0) := \begin{cases} v\big(t, v_0, z(\overline{W})\big), & \text{if } z(\overline{W}) \in C\big(\mathbb{R}^+, H^m\big) \text{ for } m < 2+\sigma, \\ 0, & \text{otherwise,}\end{cases}$$

where $v(t, v_0, z(\overline{W}))$ is the solution to (A.1) we obtained in Theorem A.1.

Combining Theorems A.1–A.3 we obtain the following results.

THEOREM A.4. *Fix $\alpha > 1/2$. $v : \mathbb{R}^+ \times C(\mathbb{R}^+, H^{-1-\varepsilon_0}) \times H^1 \mapsto H^1$, $(t, \overline{W}, v_0) \mapsto v(t, \overline{W}, v_0)$ is a measurable map.*

PROOF. By Theorems A.1–A.3 $t \mapsto v(t, \overline{W}, v_0)$ and $v_0 \mapsto v(t, \overline{W}, v_0)$ is continuous. Then it is sufficient to prove that if $z_n \to z$ in $C(\mathbb{R}^+, H^m)$, $m < 2+\sigma$, $v_n \to v$ in $C([0,T], H^1)$, where $v_n = v(\cdot, v_0, z_n)$, $v = v(\cdot, v_0, z)$. By the same arguments as in the proof of Theorem A.1, we have the following estimate:

$$\sup_{[0,T]} |\Lambda v_n|^2 \le C(T), \qquad \sup_{[0,T]} |\Lambda v|^2 \le C(T),$$

and

$$\int_0^T \big|\Lambda^{1+\alpha}v_n(l)\big|^2\,dl \le C(T), \qquad \int_0^T \big|\Lambda^{1+\alpha}v(l)\big|^2\,dl \le C(T).$$

Since $v, v_n \in C([0,+\infty), H^1) \cap L^2_{\text{loc}}((0,+\infty), H^{1+\alpha})$, we obtain

$$\begin{aligned}\frac{d}{dt}\big|\Lambda(v-v_n)\big|^2 &+ 2\kappa\big|\Lambda^{1+\alpha}(v_n-v)\big|^2 \\ &= \big\langle (u_{v_n}+u_{z_n})\cdot\nabla(v_n+z_n), \Lambda^2(v-v_n)\big\rangle \\ &\quad - \big\langle (u_v+u_z)\cdot\nabla(v+z), \Lambda^2(v-v_n)\big\rangle \\ &= \big[\big\langle (u_{v_n}-u_v)\cdot\nabla(v_n+z_n), \Lambda^2(v-v_n)\big\rangle\end{aligned}$$

$$+\langle (u_v+u_z)\cdot\nabla(v_n-v),\Lambda^2(v-v_n)\rangle$$
$$+\langle (u_{z_n}-u_z)\cdot\nabla(v_n+z_n),\Lambda^2(v-v_n)\rangle$$
$$+\langle (u_v+u_z)\cdot\nabla(z_n-z),\Lambda^2(v-v_n)\rangle\big],$$

where u_{v_n}, u_{z_n} satisfy (1.3) with θ replaced by v_n, z_n respectively. For the first term on the right-hand side, we have

$$\begin{aligned}
&\big|\langle (u_{v_n}-u_v)\cdot\nabla(v_n+z_n),\Lambda^2(v-v_n)\rangle\big|\\
&\quad=\big|\langle \Lambda^{1-\alpha}\nabla\cdot\big((u_{v_n}-u_v)(v_n+z_n)\big),\Lambda^{1+\alpha}(v-v_n)\rangle\big|\\
&\quad\le C\big|\Lambda^{1+\alpha}(v-v_n)\big|\big|\Lambda^{2-\alpha+\sigma_1}(v-v_n)\big|\big|\Lambda^{\sigma_2}(v_n+z_n)\big|\\
&\qquad+C\big|\Lambda^{1+\alpha}(v-v_n)\big|\big|\Lambda^{2-\alpha+\sigma_1}(v_n+z_n)\big|\big|\Lambda^{\sigma_2}(v-v_n)\big|\\
&\quad\le\frac{\kappa}{4}|\Lambda^{1+\alpha}(v-v_n)|^2+C\big(C(T)+\big|\Lambda^{1+\alpha}v_n\big|^2\big)\big|\Lambda(v-v_n)\big|^2\\
&\qquad+c\big|\Lambda^{2-\alpha+\sigma_1}z_n\big|^2\big|\Lambda(v-v_n)\big|^2.
\end{aligned}$$

Here, σ_1,σ_2 are as (6.6) and we used $\operatorname{div}(u_{v_n}-u_v)=0$ in the first equality and Lemmas 2.1 and 2.2 in the first inequality and the interpolation inequality (2.3) and Young's inequality in the last step. The other term can be estimated similarly. Then we obtain

$$\begin{aligned}
&\frac{d}{dt}\big|\Lambda(v-v_n)\big|^2+2\kappa\big|\Lambda^{1+\alpha}(v_n-v)\big|^2\\
&\quad\le\kappa\big|\Lambda^{1+\alpha}(v_n-v)\big|^2+C\big(C(T)+\big|\Lambda^{1+\alpha}v_n\big|^2+\big|\Lambda^{1+\alpha}v\big|^2\big)\\
&\qquad\times\big(\big|\Lambda(v-v_n)\big|^2+\big|\Lambda^{2-\alpha+\sigma_1}(z-z_n)\big|^2\big).
\end{aligned}$$

Gronwall's lemma yields that

$$\begin{aligned}
\big|\Lambda(v-v_n)(t)\big|^2&\le\Theta_n\exp\biggl(C\int_0^t\big(C(T)+\big|\Lambda^{1+\alpha}v_n\big|^2+\big|\Lambda^{1+\alpha}v\big|^2\big)\,dl\biggr)\\
&\quad\times\int_0^t\big(C(T)+\big|\Lambda^{1+\alpha}v_n\big|^2+\big|\Lambda^{1+\alpha}v\big|^2\big)\,dl,
\end{aligned}$$

where $\Theta_n=\sup_{[0,T]}|\Lambda^{2-\alpha+\sigma_1}(z-z_n)|$. Then the results follow. □

APPENDIX B

In this appendix, we prove the following lemma to complete the proof of Theorem 3.3.

LEMMA B.1. *For any $x_0\in B_0$ defined in the proof of Theorem 3.3, there exists $Q_{x_0}\in\mathcal{Q}_{x_0}$ such that the map $x_0\mapsto Q_{x_0}$ from B_0 to $\mathcal{P}(\Omega_0^{t_1^n})$ is measurable with respect to $\mathcal{B}_{t_1^n}$.*

PROOF. Let $\mathcal{B}^1_{t_1^n}$ be the Borel σ-algebra on $\tilde{B}_0 := \{x(\cdot)1_{[0,t_1^n]}(\cdot) + x(t_1^n) \times 1_{[t_1^n,\infty)}(\cdot) : x \in B_0\}$ with the topology induced by $\sup_{0\le t\le t_1^n} \|x(t)\|_{H^3}$. Since $\{\sup_{0\le t\le t_1^n} \|x(t)\|_{H^3} < a\} \in \mathcal{B}_{t_1^n}$, we know $\mathcal{B}^1_{t_1^n} \subset \mathcal{B}_{t_1^n}$. It suffices to prove that if for $\{x_m, m \in \mathbb{N}\cup\{0\}\} \subset \tilde{B}_0$, $\sup_{0\le t\le t_1^n} \|x_m(t) - x_0(t)\|_{H^3} \to 0$ and $Q_m \in \mathcal{Q}_{x_m}$, then for some subsequence m_k, Q_{m_k} weakly converges to some $Q \in \mathcal{Q}_{x_0}$, because then [48], Lemma 12.1.8, Theorem 12.1.10, implies the existence of a $Q_{x_0} \in \mathcal{Q}_{x_0(\cdot)1_{[0,t_1^n]}(\cdot)+x_0(t_1^n)1_{[t_1^n,\infty)}(\cdot)}$ such that the map $x_0 \mapsto x_0(\cdot)1_{[0,t_1^n]}(\cdot) + x(t_1^n)1_{[t_1^n,\infty)}(\cdot) \mapsto Q_{x_0}$ from B_0 to $\tilde{B}_0$ to$\mathcal{P}(\Omega_0^{t_1^n})$ is measurable with respect to $\mathcal{B}_{t_1^n}$. Moreover, by $Q_{x_0} \in \mathcal{Q}_{x_0}$, the result follows.

Step 1: We prove that $(Q_m)_{m\in\mathbb{N}}$ is tight in $\mathbb{S} := C([t_1^n, +\infty), H^1) \cap L^q_{\mathrm{loc}}([t_1^n, +\infty), H^3)$ for some $q \in \mathbb{N}$. Define for each $m \in \mathbb{N}$,

$$M^m(t,x) := \sum_{i=1}^{\infty} M_i^m(t,x)e_i,$$

where M_i^m is given in the proof of Theorem 3.3 (Step 2) with x_0 replaced by x_m. Then $(M^m(t,x))_{t\ge t_1^n}$ is a continuous H^3-valued $\mathcal{B}_t$-martingale with respect to Q_m and the following equality holds in H^1:

$$\begin{aligned} x(t) = x_m(t_1^n) - \int_{t_1^n}^{t\wedge t_2^n} \big(A_\alpha x(s) + U_{\delta_n}[x_m](s)\cdot\nabla x(s)\big)\,ds \\ + M^m(t), \qquad Q_m\text{-a.s.} \end{aligned} \tag{B.1}$$

By Hölder's inequality and (M3), (M1) for Q_m, we have

$$\begin{aligned} &E^{Q_m}\Big[\sup_{s\ne t\in[t_1^n,t_2^n]}\Big(\Big\|\int_s^t A_\alpha x(r) + U_{\delta_n}[x_m](r)\cdot\nabla x(r)\,dr\Big\|_{H^1}^{\gamma}\Big/|t-s|^{\gamma-1}\Big)\Big] \\ &\le C E^{Q_m}\Big[\int_{t_1^n}^{t_2^n} \|A_\alpha x(r) + U_{\delta_n}[x_m](r)\cdot\nabla x(r)\|_{H^1}^\gamma\,dr\Big] \\ &\le C E^{Q_m}\Big[\sup_{t_1^n\le r\le t_2^n}\|x(r)\|_{H^3}^\gamma\Big(1+\sup_{0\le r\le t_1^n}\|x_m(r)\|_{H^1}^\gamma\Big)\Big] \\ &\le C\big(\|x_m(t_1^n)\|_{H^3}^\gamma + 1\big)\Big(1+\sup_{0\le r\le t_1^n}\|x_m(r)\|_{H^1}^\gamma\Big), \end{aligned} \tag{B.2}$$

where C is independent of m. For $t_1^n \le s < t \le t_2^n$ and $q \in \mathbb{N}$, we have

$$\begin{aligned} E^{Q_m}\|M^m(t,x) - M^m(s,x)\|_{H^1}^{2q} &\le C_q E^{Q_m}\Big(\int_s^t \|\Lambda(k_{\delta_n} * G(x(r)))\|_{L_2(U;H)}^2\,dr\Big)^q \\ &\le C_q|t-s|^{q-1}\int_s^t E^{Q_m}\|G(x(r))\|_{L_2(U;H)}^{2q}\,dr \end{aligned}$$

$$\leq C_q |t-s|^{q-1} \int_s^t E^{Q_m}\big(|x(r)|^{2q}+1\big)\,dr$$
$$\leq C_q |t-s|^{q}\big(|\Lambda^3 x_m(t_1^n)|^{2q}+1\big),$$

where we used Hypothesis G.1 in the third inequality and (M3) in the last inequality. By Kolmogorov's criterion for any $\beta \in (0, \frac{q-1}{2q})$, we get

$$E^{Q_m}\bigg(\sup_{s\neq t\in[t_1^n,t_2^n]} \frac{\|M^m(t,x)-M^m(s,x)\|_{H^1}^{2q}}{|t-s|^{q\beta}}\bigg) \leq C\big(|\Lambda^3 x_m(t_1^n)|^{2q}+1\big). \tag{B.3}$$

Combining (B.1)–(B.3) and $Q_m(\{x : x(s) = x(t_2^n), s \in [t_2^n, +\infty)\}) = 1$, we obtain for $\beta_1 = 1 - \frac{1}{\gamma}$ and any $T > 0$

$$\sup_{m\in\mathbb{N}} E^{Q_m}\bigg(\sup_{s\neq t\in[t_1^n,T]} \frac{\|x(t)-x(s)\|_{H^1}}{|t-s|^{\beta_1}}\bigg) < \infty.$$

Thus, by (M3) for Q_m and [20], Lemma 4.3, $(Q_m)_{m\in\mathbb{N}}$ is tight in $\mathbb{S}$.

Without loss of generality, we assume that Q_m weakly converges to some probability measure Q in $\mathbb{S}$. We need to prove $Q \in \mathcal{Q}_{x_0}$.

Step 2: By Skorohod's representation theorem, there exist a probability space $(\tilde{\Omega}, \tilde{\mathcal{B}}, \tilde{P})$ and $\mathbb{S}$-valued random variable $\tilde{x}_m$ and $\tilde{x}$ such that:

(i) $\tilde{x}_m$ has the law Q_m for each $m \in \mathbb{N}$;
(ii) $\tilde{x}_m \to \tilde{x}$ in $\mathbb{S}$, $\tilde{P}$-a.e., and $\tilde{x}$ has the law Q.

First, we easily deduce that

$$\begin{aligned} Q\big(x(t_1^n) = x_0(t_1^n)\big) &= \tilde{P}\big(\tilde{x}(t_1^n) = x_0(t_1^n)\big) \\ &= \lim_{m\to\infty} Q_m\big(x(t_1^n) = x_m(t_1^n)\big) = 1, \\ Q\big(x(t) = x(t_2^n), t \geq t_2^n\big) &= \tilde{P}\big(\tilde{x}(t) = \tilde{x}(t_2^n), t \geq t_2^n\big) \\ &= \lim_{m\to\infty} Q_m\big(x(t) = x(t_2^n), t \geq t_2^n\big) = 1. \end{aligned}$$

For $q \in \mathbb{N}$, set

$$\xi_q(x) := \sup_{r\in[t_1^n,t_2^n]} \|x(r)\|_{H^3}^{2q} + \int_{t_1^n}^{t_2^n} \|x(r)\|_{H^3}^{2(q-1)} \|x(r)\|_{H^{3+\alpha}}^2\,dr.$$

Then

$$\begin{aligned} E^Q\big(\xi_q(x)\big) = E^{\tilde{P}}\big(\xi_q(\tilde{x})\big) &\leq \liminf_{m\to\infty} E^{Q_m}\big(\xi_q(x)\big) \leq \liminf_{m\to\infty} C\big(\|x_m(t_1^n)\|_{H^3}^{2q}+1\big) \\ &\leq C\big(\|x_0(t_1^n)\|_{H^3}^{2q}+1\big). \end{aligned}$$

Thus, (M1) and (M3) follow.

Now we want to show that $(M_i(t,x))_{t\ge t_1^n}$ in the proof of Theorem 3.2 (Step 2) is a continuous $\mathcal{B}_t$-martingale with respect to Q, whose square variation process is given by

$$\langle M_i\rangle(t,x)=\int_{t_1^n}^{t\wedge t_2^n}\big\|(k_{\delta_n}*G)^*\big(x(s)\big)(e_i)\big\|_U^2\,ds.$$

Since $\sup_{0\le t\le t_1^n}\|x_m(t)-x_0(t)\|_{H^3}\to 0$ and $\tilde{x}_m\to\tilde{x}$ in $\mathbb{S}$, we have

$$\begin{aligned}
\lim_{m\to\infty} E^{\tilde{P}}\int_{t_1^n}^{t_2^n}\big|\big\langle U_{\delta_n}[x_m](s)\cdot\nabla\tilde{x}_m(s)+A_\alpha\tilde{x}_m(s)&\\
-U_{\delta_n}[x_0](s)\cdot\nabla\tilde{x}(s)-A_\alpha\tilde{x}(s),e_i\big\rangle\big|\,ds&\\
\le \lim_{m\to\infty} E^{\tilde{P}}\int_{t_1^n}^{t_2^n}\big|\big\langle\big(U_{\delta_n}[x_m](s)-U_{\delta_n}[x_0](s)\big)\cdot\nabla\tilde{x}_m(s)&\\
+U_{\delta_n}[x_0](s)\cdot\nabla\big(\tilde{x}_m(s)-\tilde{x}(s)\big)&\\
+A_\alpha\big(\tilde{x}_m(s)-\tilde{x}(s)\big),e_i\big\rangle\big|\,ds&\\
=0,&
\end{aligned}$$

which implies that for $t\ge t_1^n$

$$\lim_{m\to\infty}E^{\tilde{P}}\big|M_i^m(t,\tilde{x}_m)-M_i(t,\tilde{x})\big|=0. \tag{B.4}$$

Then we obtain for $t_1^n\le s<t$,

$$E^Q\big(M_i(t,x)|\mathcal{B}_s\big)=M_i(s,x).$$

On the other hand, by the B–D–G inequality, we have

$$\sup_m E^{\tilde{P}}\big|M_i^m(t,\tilde{x}_m)\big|^{2q}\le C\sup_m\int_{t_1^n}^{t_2^n}E^{\tilde{P}}\big(\big\|(k_{\delta_n}*G)^*\big(\tilde{x}_m(s)\big)(e_i)\big\|_U^{2q}\big)\,ds<+\infty.$$

By (B.4), we have

$$\lim_{m\to\infty}E^{\tilde{P}}\big|M_i(t,\tilde{x}_m)-M_i(t,\tilde{x})\big|^2=0.$$

Then we obtain

$$\begin{aligned}
E^Q\bigg(M_i^2(t,x)-\int_{t_1^n}^{t}\big\|(k_{\delta_n}*G)^*\big(x(r)\big)(e_i)\big\|_U^2\,dr|\mathcal{B}_s\bigg)&\\
=M_i^2(s,x)-\int_{t_1^n}^{s}\big\|(k_{\delta_n}*G)^*\big(x(r)\big)(e_i)\big\|_U^2\,dr.&
\end{aligned}$$

Now the results follow. □

APPENDIX C: MARKOV SELECTIONS IN THE GENERAL CASE

In this appendix, we will use [20], Theorem 4.7, to get an almost sure Markov family $(P_x)_{x\in L^2}$ for equation (3.1). Here, we will use the same notation as in [20]. Below we choose

$$H = \mathbb{Y} = L^2(\mathbb{T}^2)$$

and

$$\mathbb{X} = (H^{2+2\alpha})^*, \qquad \mathbb{X}^* = H^{2+2\alpha}.$$

Then $\mathbb{X}$ is a Hilbert space and $\mathbb{X}^* \subset \mathbb{Y}$ compactly. Let $\mathcal{E} = \{e_i, i \in \mathbb{N}\}$ be the orthonormal basis of H introduced in Section 2. We define the operator $\mathcal{A}$ as follows: for $\theta \in C^\infty(\mathbb{T}^2)$

$$\mathcal{A}(\theta) := -\kappa(-\Delta)^\alpha \theta - u \cdot \nabla\theta,$$

where u satisfies (1.3). Then by Lemma C.3 below, $\mathcal{A}$ can be extended to an operator $\mathcal{A}: H \to \mathbb{X}$. For θ not in H define $\mathcal{A}(\theta) := \infty$.

Set

$$\Omega := C([0,\infty); \mathbb{X}),$$

and let $\mathcal{B}$ denote the σ-field of Borel sets of Ω and let $\mathcal{P}(\Omega)$ denote the set of all probability measures on $(\Omega, \mathcal{B})$. Define the canonical process $x : \Omega \to \mathbb{X}$ as

$$x_t(\omega) = \omega(t).$$

For each t, $\mathcal{B}_t = \sigma(x_s : 0 \le s \le t)$. Given $P \in \mathcal{P}(\Omega)$ and $t > 0$, let $P(\cdot|\mathcal{B}_t)(\omega)$ denote a regular conditional probability distribution of P given $\mathcal{B}_t$. In particular, $P(\cdot|\mathcal{B}_t)(\omega) \in \mathcal{P}(\Omega)$ for every $\omega \in \Omega$ and for any bounded $\mathcal{B}$-measurable function f on Ω

$$E^P[f|\mathcal{B}_t] = \int_\Omega f(y)P(dy|\mathcal{B}_t), \qquad P\text{-a.s.},$$

and there exists a P-null set $N \in \mathcal{B}_t$ such that for every ω not in N

$$P(\cdot|\mathcal{B}_t)(\omega)_{|\mathcal{B}_t} = \delta_\omega \qquad (= \text{Dirac measure at } \omega),$$

hence,

$$P(\{y : y(s) = \omega(s), s \in [0,t]\}|\mathcal{B}_t)(\omega) = 1.$$

In particular, we can consider $P(\cdot|\mathcal{B}_t)(\omega)$ as a measure on $(\Omega^t, \mathcal{B}^t)$, that is,

$$P(\cdot|\mathcal{B}_t)(\omega) \in \mathcal{P}(\Omega^t),$$

where $\Omega^t := C([t,\infty); \mathbb{X})$ and $\mathcal{B}^t := \sigma(x_s : s \ge t)$.

We say $P \in \mathcal{P}(\Omega)$ is concentrated on the paths with values in H, if there exists $A \in \mathcal{B}$ with $P(A) = 1$ such that $A \subset \{\omega \in \Omega : x_t(\omega) \in H, \forall t \ge 0\}$. The set of such measures is denoted by $\mathcal{P}_H(\Omega)$. The shift operator $\Phi_t : \Omega \to \Omega^t$ is defined by

$$\Phi_t(\omega)(s) = \omega(s-t), \qquad s \ge t.$$

Following [20], Definitions 2.5, we introduce the following notions.

DEFINITION C.1. A family $(P_x)_{x\in H}$ of probability measures in $\mathcal{P}_H(\Omega)$, is called an *almost sure Markov family* if for any $A \in \mathcal{B}$, $x \mapsto P_x(A)$ is $\mathcal{B}(H)/\mathcal{B}([0,1])$-measurable, and for each $x \in H$ there exists a Lebesgue null set $T_{P_x} \subset (0,\infty)$ such that for all t not in T_{P_x} and P_x-almost all $\omega \in \Omega$

$$P_x(\cdot|\mathcal{B}_t)(\omega) = P_{\omega(t)} \circ \Phi_t^{-1}.$$

We now introduce the following notion of a martingale solution to equation (3.1) and write $x(t)$ instead of x_t.

DEFINITION C.2. Let $x_0 \in H$. A probability measure $P \in \mathcal{P}(\Omega)$ is called a martingale solution of equation (3.1) with initial value x_0, if:

(M1) $P(x(0) = x_0) = 1$ and for any $n \in \mathbb{N}$

$$P\left\{x \in \Omega : \int_0^n \|\mathcal{A}(x(s))\|_{\mathbb{X}}\, ds + \int_0^n \|G(x(s))\|_{L_2(U;H)}^2\, ds < +\infty\right\} = 1;$$

(M2) for every $l \in \mathcal{E}$, the process

$$M_l(t,x) := {}_{\mathbb{X}}\langle x(t), l\rangle_{\mathbb{X}^*} - \int_0^t {}_{\mathbb{X}}\langle \mathcal{A}(x(s)), l\rangle_{\mathbb{X}^*}\, ds$$

is a continuous square-integrable $\mathcal{B}_t$-martingale under P, whose quadratic variation process is given by

$$\langle M_l\rangle(t,x) := \int_0^t \|G^*(x(s))(l)\|_U^2\, ds,$$

where the asterisk denotes the adjoint operator of $G(x(s))$;

(M3) for any $p \in \mathbb{N}$, there exist a continuous positive real function $t \mapsto C_{t,p}$ (only depending on p and $\mathcal{A}$, G), a lower semicontinuous positive real functional $\mathcal{N}_p : \mathbb{Y} \to [0,\infty]$, and a Lebesgue null set $T_P \subset (0,\infty)$ such that for all $0 \le s \in [0,\infty)\setminus T_P$ and for all $t \ge s$

$$E^P\left[\sup_{r\in[s,t]} |x(r)|^{2p} + \int_s^t \mathcal{N}_p(x(r))\, dr \Big| \mathcal{B}_s\right] \le C_{t-s}(|x(s)|^{2p} + 1).$$

First, we prove the following lemma.

LEMMA C.3. *For any* $\theta_1, \theta_2 \in C^\infty(\mathbb{T}^2)$,

$$\|(-\Delta)^\alpha\theta_1 - (-\Delta)^\alpha\theta_2\|_{\mathbb{X}} \le C_1|\theta_1 - \theta_2|,$$
$$\|u_1 \cdot \nabla\theta_1 - u_2 \cdot \nabla\theta_2\|_{\mathbb{X}} \le C_2(|\theta_1| + |\theta_2|)|\theta_1 - \theta_2|$$

for constants C_1, C_2. *In particular, the operator* $\mathcal{A} : C^\infty(\mathbb{T}^2) \to \mathbb{X}$ *extends to an operator* $\mathcal{A} : H \to \mathbb{X}$ *by continuity.*

PROOF. We only prove the second assertion, the first can be proved analogously. By the Sobolev embedding theorem, we have

$$
\begin{aligned}
&\|u_1 \cdot \nabla\theta_1 - u_2 \cdot \nabla\theta_2\|_{\mathbb{X}} \\
&\quad = \sup_{w\in C^\infty(\mathbb{T}^2):\|w\|_{H^{2+2\alpha}}\le 1} \big|\langle u_1 \cdot \nabla\theta_1 - u_2 \cdot \nabla\theta_2, w\rangle\big| \\
&\quad = \sup_{w\in C^\infty(\mathbb{T}^2):\|w\|_{H^{2+2\alpha}}\le 1} \big|\langle u_1 \cdot \nabla w, \theta_1\rangle - \langle u_2 \cdot \nabla w, \theta_2\rangle\big| \\
&\quad = \sup_{w\in C^\infty(\mathbb{T}^2):\|w\|_{H^{2+2\alpha}}\le 1} \big|\big\langle (u_1 - u_2) \cdot \nabla w, \theta_1\big\rangle + \langle u_2 \cdot \nabla w, \theta_1 - \theta_2\rangle\big| \\
&\quad \le C\Big[\sup_{w\in C^\infty(\mathbb{T}^2):\|w\|_{H^{2+2\alpha}}\le 1} \|\nabla w\|_{C(\mathbb{T}^2)}\Big]\big(|u_1 - u_2| \cdot |\theta_1| + |\theta_1 - \theta_2| \cdot |u_2|\big) \\
&\quad \le C\big(|\theta_1| + |\theta_2|\big)|\theta_1 - \theta_2|.
\end{aligned}
$$

In the last inequality, we use (2.1) and the constant C changes from line to line. □

In order to use [20], Theorem 4.7, we define the functional $\mathcal{N}_1$ on $\mathbb{Y}$ as follows:

$$
\mathcal{N}_1(\theta) := \begin{cases} |\Lambda^\alpha\theta|^2, & \text{if } \theta \in H^\alpha, \\ +\infty, & \text{otherwise.} \end{cases}
$$

It is obvious that $\mathcal{N}_1 \in \mathfrak{U}^2$, defined in [20], Section 4. We recall that a lower semicontinuous function $\mathcal{N}: \mathbb{Y} \to [0, \infty]$ belongs to $\mathfrak{U}^2$ if $\mathcal{N}(x) = 0$ implies $x = 0$, $\mathcal{N}(cy) \le c^2\mathcal{N}(y), \forall c \ge 0, y \in \mathbb{Y}$ and $\{y \in \mathbb{Y}: \mathcal{N}(y) \le 1\}$ is relatively compact in $\mathbb{Y}$.

THEOREM C.4. *Let $\alpha \in (0, 1)$ and assume G satisfies Hypothesis* G.1 *with $\rho_1 = 0$. Then for each $x_0 \in H$, there exists a martingale solution $P \in \mathcal{P}(\Omega)$ starting from x_0 to equation* (3.1) *in the sense of Definition* C.2.

PROOF. We only need to check (C1)–(C3) in [20], Section 4, for the above $\mathcal{A}$ and G.

The demi-continuity condition (C1) holds since Lemma C.3 and Hypothesis G.1 imply demi-continuity of $\mathcal{A}$ and G.

The coercivity condition (C2) follows, because noting that for $\theta \in \mathbb{X}^*$

$$
\langle u \cdot \nabla\theta, \theta\rangle = 0,
$$

we have

$$
\big\langle \mathcal{A}(\theta), \theta\big\rangle = -\mathcal{N}_1(\theta).
$$

Also the growth condition (C3) is clear since by Lemma C.3

$$\|\mathcal{A}(\theta)\|_{\mathbb{X}} \le C|\theta|^2$$

and

$$\|G(\theta)\|_{L_2(K;H)} \le C(|\theta|+1). \qquad \square$$

The set of all such martingale solutions with initial value x_0 is denoted by $\mathcal{C}(x_0)$. Using [20], Theorem 4.7, we now obtain the following.

THEOREM C.5. *Let $\alpha \in (0,1)$. Assume G satisfies Hypothesis G.1 with $\rho_1 = 0$. Then there exists an almost sure Markov family $(P_{x_0})_{x_0 \in H}$ for equation (3.1) and $P_{x_0} \in \mathcal{C}(x_0)$ for each $x_0 \in H$.*

REFERENCES

[1] ALDOUS, D. (1978). Stopping times and tightness. *Ann. Probab.* **6** 335–340. MR0474446

[2] BRANNAN, J. R., DUAN, J. and WANNER, T. (1998). Dissipative quasi-geostrophic dynamics under random forcing. *J. Math. Anal. Appl.* **228** 221–233. MR1659913

[3] BRZEŹNIAK, Z. (1997). On stochastic convolution in Banach spaces and applications. *Stochastics Stochastics Rep.* **61** 245–295. MR1488138

[4] BRZEŹNIAK, Z., VAN NEERVEN, J. M. A. M., VERAAR, M. C. and WEIS, L. (2008). Itô's formula in UMD Banach spaces and regularity of solutions of the Zakai equation. *J. Differential Equations* **245** 30–58. MR2422709

[5] CAFFARELLI, L. A. and VASSEUR, A. (2010). Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation. *Ann. of Math.* (2) **171** 1903–1930. MR2680400

[6] CONSTANTIN, P., MAJDA, A. J. and TABAK, E. (1994). Formation of strong fronts in the 2-D quasigeostrophic thermal active scalar. *Nonlinearity* **7** 1495–1533. MR1304437

[7] CONSTANTIN, P. and WU, J. (1999). Behavior of solutions of 2D quasi-geostrophic equations. *SIAM J. Math. Anal.* **30** 937–948. MR1709781

[8] CÓRDOBA, A. and CÓRDOBA, D. (2004). A maximum principle applied to quasi-geostrophic equations. *Comm. Math. Phys.* **249** 511–528. MR2084005

[9] DA PRATO, G. (2004). *Kolmogorov Equations for Stochastic PDEs*. Birkhäuser, Basel. MR2111320

[10] DA PRATO, G. and DEBUSSCHE, A. (2003). Ergodicity for the 3D stochastic Navier–Stokes equations. *J. Math. Pures Appl.* (9) **82** 877–947. MR2005200

[11] DA PRATO, G. and ZABCZYK, J. (1992). *Stochastic Equations in Infinite-Dimensions*. *Encyclopedia of Mathematics and Its Applications* **44**. Cambridge Univ. Press, Cambridge. MR1207136

[12] DA PRATO, G. and ZABCZYK, J. (1996). *Ergodicity for Infinite-Dimensional Systems*. *London Mathematical Society Lecture Note Series* **229**. Cambridge Univ. Press, Cambridge. MR1417491

[13] DEBUSSCHE, A. and ODASSO, C. (2005). Ergodicity for a weakly damped stochastic nonlinear Schrödinger equation. *J. Evol. Equ.* **5** 317–356. MR2174876

[14] DEBUSSCHE, A. and ODASSO, C. (2006). Markov solutions for the 3D stochastic Navier–Stokes equations with state dependent noise. *J. Evol. Equ.* **6** 305–324. MR2227699

[15] ECKMANN, J.-P. and HAIRER, M. (2001). Uniqueness of the invariant measure for a stochastic PDE driven by degenerate noise. *Comm. Math. Phys.* **219** 523–565. MR1838749

[16] FLANDOLI, F. and GATAREK, D. (1995). Martingale and stationary solutions for stochastic Navier–Stokes equations. *Probab. Theory Related Fields* **102** 367–391. MR1339739

[17] FLANDOLI, F. and ROMITO, M. (2007). Regularity of transition semigroups associated to a 3D stochastic Navier–Stokes equation. In *Stochastic Differential Equations*: *Theory and Applications*. *Interdiscip*. *Math*. *Sci*. **2** 263–280. World Sci. Publ., Hackensack, NJ. MR2393580

[18] FLANDOLI, F. and ROMITO, M. (2008). Markov selections for the 3D stochastic Navier–Stokes equations. *Probab. Theory Related Fields* **140** 407–458. MR2365480

[19] GOLDYS, B. and MASLOWSKI, B. (2005). Exponential ergodicity for stochastic Burgers and 2D Navier–Stokes equations. *J. Funct*. *Anal*. **226** 230–255. MR2158741

[20] GOLDYS, B., RÖCKNER, M. and ZHANG, X. (2009). Martingale solutions and Markov selections for stochastic partial differential equations. *Stochastic Process*. *Appl*. **119** 1725–1764. MR2513126

[21] HAIRER, M. and MATTINGLY, J. C. (2006). Ergodicity of the 2D Navier–Stokes equations with degenerate stochastic forcing. *Ann*. *of Math*. (2) **164** 993–1032. MR2259251

[22] HUANG, D., GUO, B. and HAN, Y. (2008). Random attractors for a quasi-geostrophic dynamical system under stochastic forcing. *Int*. *J*. *Dyn*. *Syst*. *Differ. Equ*. **1** 147–154. MR2446950

[23] JU, N. (2004). Existence and uniqueness of the solution to the dissipative 2D quasi-geostrophic equations in the Sobolev space. *Comm*. *Math*. *Phys*. **251** 365–376. MR2100059

[24] JU, N. (2005). On the two dimensional quasi-geostrophic equations. *Indiana Univ. Math*. *J*. **54** 897–926. MR2151238

[25] KISELEV, A. and NAZAROV, F. (2010). A variation on a theme of Caffarelli and Vasseur. *J. Math*. *Sci*. **166** 31–39. MR2749211

[26] KISELEV, A., NAZAROV, F. and VOLBERG, A. (2007). Global well-posedness for the critical 2D dissipative quasi-geostrophic equation. *Invent*. *Math*. **167** 445–453. MR2276260

[27] KOMOROWSKI, T., PESZAT, S. and SZAREK, T. (2010). On ergodicity of some Markov processes. *Ann*. *Probab*. **38** 1401–1443. MR2663632

[28] KOMOROWSKI, T. and WALCZUK, A. (2012). Central limit theorem for Markov processes with spectral gap in the Wasserstein metric. *Stochastic Process*. *Appl*. **122** 2155–2184. MR2921976

[29] KRYLOV, N. V. (2010). Itô's formula for the L_p-norm of stochastic W_p^1-valued processes. *Probab. Theory Related Fields* **147** 583–605. MR2639716

[30] KUKSIN, S., PIATNITSKI, A. and SHIRIKYAN, A. (2002). A coupling approach to randomly forced nonlinear PDEs. II. *Comm*. *Math*. *Phys*. **230** 81–85. MR1927233

[31] KUKSIN, S. and SHIRIKYAN, A. (2001). A coupling approach to randomly forced nonlinear PDE's. I. *Comm*. *Math*. *Phys*. **221** 351–366. MR1845328

[32] KUKSIN, S. and SHIRIKYAN, A. (2002). Coupling approach to white-forced nonlinear PDEs. *J. Math*. *Pures Appl*. (9) **81** 567–602. MR1912412

[33] KURTZ, T. G. (2007). The Yamada–Watanabe–Engelbert theorem for general stochastic equations and inequalities. *Electron*. *J*. *Probab*. **12** 951–965. MR2336594

[34] LINDVALL, T. (1992). *Lectures on the Coupling Method*. Wiley, New York. MR1180522

[35] MATTINGLY, J. C. (1999). Ergodicity of 2D Navier–Stokes equations with random forcing and large viscosity. *Comm*. *Math*. *Phys*. **206** 273–288. MR1722141

[36] MATTINGLY, J. C. (2002). Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics. *Comm*. *Math*. *Phys*. **230** 421–462. MR1937652

[37] MIKULEVICIUS, R. and ROZOVSKII, B. L. (2005). Global L_2-solutions of stochastic Navier–Stokes equations. *Ann*. *Probab*. **33** 137–176. MR2118862

[38] ODASSO, C. (2006). Ergodicity for the stochastic complex Ginzburg–Landau equations. *Ann*. *Inst*. *Henri Poincaré Probab*. *Stat*. **42** 417–454. MR2242955

[39] ODASSO, C. (2007). Exponential mixing for the 3D stochastic Navier–Stokes equations. *Comm. Math. Phys.* **270** 109–139. MR2276442

[40] ODASSO, C. (2008). Exponential mixing for stochastic PDEs: The non-additive case. *Probab. Theory Related Fields* **140** 41–82. MR2357670

[41] ONDREJÁT, M. (2005). Brownian representations of cylindrical local martingales, martingale problem and strong Markov property of weak solutions of SPDEs in Banach spaces. *Czechoslovak Math. J.* **55** 1003–1039. MR2184381

[42] PEDLOSKY, J. (1987). *Geophysical Fluid Dynamics*. Springer, New York.

[43] PRÉVÔT, C. and RÖCKNER, M. (2007). *A Concise Course on Stochastic Partial Differential Equations*. *Lecture Notes in Math.* **1905**. Springer, Berlin. MR2329435

[44] RESNICK, S. G. (1995). Dynamical problems in non-linear advective partial differential equations. Ph.D. thesis, Univ. of Chicago. MR2716577

[45] RÖCKNER, M., SCHMULAND, B. and ZHANG, X. (2008). Yamada–Watanabe theorem for stochastic evolution equations in infinite dimensions. *Condensed Matter Physics* **54** 247–259.

[46] ROMITO, M. (2008). Analysis of equilibrium states of Markov solutions to the 3D Navier–Stokes equations driven by additive noise. *J. Stat. Phys.* **131** 415–444. MR2386571

[47] STEIN, E. M. (1970). *Singular Integrals and Differentiability Properties of Functions*. *Princeton Mathematical Series* **30**. Princeton Univ. Press, Princeton, N.J. MR0290095

[48] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. *Grundlehren der Mathematischen Wissenschaften* [*Fundamental Principles of Mathematical Sciences*] **233**. Springer, Berlin. MR0532498

[49] TEMAM, R. (1984). *Navier–Stokes Equations*: *Theory and Numerical Analysis*, 3rd ed. *Studies in Mathematics and Its Applications* **2**. North-Holland, Amsterdam. MR0769654

[50] ZHU, R. C. and ZHU, X. C. (2012). Random attractor associated with the quasi-geostrophic equation. Preprint. Available at arXiv:1303.5970.

M. RÖCKNER
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BIELEFELD
D-33615 BIELEFELD
GERMANY
E-MAIL: roeckner@math.uni-bielefeld.de

R. ZHU
DEPARTMENT OF MATHEMATICS
BEIJING INSTITUTE OF TECHNOLOGY
BEIJING 100081
CHINA
E-MAIL: zhurongchan@126.com

X. ZHU
SCHOOL OF SCIENCE
BEIJING JIAOTONG UNIVERSITY
BEIJING 100044
CHINA
E-MAIL: zhuxiangchan@126.com