Journal of Theoretical Probability (2022) 35:1736–1781
https://doi.org/10.1007/s10959-021-01125-1

Large and Moderate Deviations Principles and Central Limit Theorem for the Stochastic 3D Primitive Equations with Gradient-Dependent Noise

Jakub Slavík[1]

Received: 28 August 2020 / Revised: 23 June 2021 / Accepted: 8 August 2021 /
Published online: 25 August 2021

Abstract

We establish the large deviations principle (LDP), the moderate deviations principle (MDP), and an almost sure version of the central limit theorem (CLT) for the stochastic 3D viscous primitive equations driven by multiplicative white noise allowing dependence on the spatial gradient of velocity with initial data in H^2. We establish the LDP using the weak convergence approach by Budjihara and Dupuis and a uniform version of the stochastic Gronwall lemma. The result corrects a minor technical issue in Dong et al. (J Differ Equ 263(5):3110–3146, 2017) and establishes the result for a more general noise. The MDP is established by a similar argument.

Keywords Large deviations principle · Moderate deviations principle · Primitive equations · Weak convergence approach

Mathematics Subject Classification (2020) 60H15 · 60F10 · 35Q86

1 Introduction

The primitive equations are one of the fundamental models in geophysical fluid dynamics, see, e.g. [39,45] and the references therein. They can be derived from the Navier–Stokes equations using the Boussinesq approximation and hydrostatic balance. The aim of this paper is to study the behaviour of solutions of the stochastic 3D primitive equations driven by multiplicative white noise w.r.t. small noise limit. In particular, we establish large and moderate deviations principles and an almost sure version of the central limit theorem.

✉ Jakub Slavík
slavik@utia.cas.cz

1 Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodárenskou věží 4, 182 00 Prague 8, Czech Republic

A systematic study of the deterministic primitive equations began in the 1990s [32–34]. The local and global existence and the uniqueness of strong solutions were established in [19,22] and [6,28,29], respectively. More recent results include well-posedness in L^p spaces [21] and well-posedness of the primitive equations with partial viscosity and/or diffusivity, see [23,30] and the references therein.

The stochastic 3D primitive equations driven by additive noise were studied in [20]. Local and global existence of strong solutions (both in the stochastic and PDE senses) were established in [9,10]. The existence and regularity of invariant measures are shown in [16]. The time discretization of weak solutions (both in the stochastic and PDE senses) was treated in [17]. In this paper, we use the existence theorem from Brzeźniak and Slavík [2], which allows noise terms σ depending on the spatial gradient of the velocity. Such noise terms are physically reasonable, see, e.g. [37,38] and the references therein.

The stochastic 3D primitive equations driven by scaled multiplicative noise can be written in an abstract form as

$$dU^\varepsilon + \left[AU^\varepsilon + B\left(U^\varepsilon\right) + A_{\mathrm{pr}}U^\varepsilon + EU^\varepsilon + F_U\right] dt = \sqrt{\varepsilon}\sigma\left(U^\varepsilon\right) dW, \quad U^\varepsilon(0) = u_0, \tag{1.1}$$

with $U^\varepsilon = (v^\varepsilon, T^\varepsilon)$, where v^ε is the horizontal velocity and T^ε is the temperature. The full form of Eq. (1.1) can be found in Sect. 2 together with the definitions of the operators and function spaces used here. We will study the convergence of U^ε to the solution of the deterministic equation

$$dU^0 + \left[AU^0 + B\left(U^0\right) + A_{\mathrm{pr}}U^0 + EU^0 + F_U\right] dt = 0, \quad U^0(0) = u_0, \tag{1.2}$$

in various scales. Let $\lambda(\varepsilon)$ be a certain deviations scale and let

$$R^\varepsilon = \frac{U^\varepsilon - U^0}{\sqrt{\varepsilon}\lambda(\varepsilon)}.$$

If $\lambda(\varepsilon) = 1/\sqrt{\varepsilon}$, the asymptotic behaviour of R^ε as $\varepsilon \to 0+$ is known as the large deviations principle (LDP), see [4,13,35,47] and the references therein. For the stochastic primitive equations, the LDP was established in [12,15] in 2D and 3D, respectively. In the last mentioned reference, the LDP result is obtained in the setting of Debussche et al. [10], in particular with noise term that does not allow dependence on gradients. Also, the proof of the main result of Dong et al. [12] contains certain minor technical issues which we address in Sect. 3.3. Other related results include the LDP for the stochastic 2D Navier–Stokes equations [7] and the stochastic 2D quasi-geostrophic equations [36].

Let us formulate the main results of this paper. Let $T > 0$ be fixed. Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be a fixed stochastic basis satisfying the usual conditions and let W be a cylindrical $\mathbb{F}$-Wiener process. The following theorem will be proved in Sect. 3.3.

Theorem 1.1 (LDP) *Let σ satisfy assumptions [(2.14)–(2.21)] and let F_U satisfy (2.10). Moreover, let σ be such that the solution of the skeleton equation*

$$dU_h + [AU_h + B(U_h) + F(U_h)]\, dt = \sigma(U_h)h\, dt, \quad U_h(0) = u_0, \tag{1.3}$$

satisfy $U_h \in L^\infty(0, T; H^2)$ for all $T > 0$, $u_0 \in V \cap H^2$ and $h \in L^2(0, T; \mathcal{U})$. Then, there exists $\varepsilon_0 > 0$ such that for all $u_0 \in V$ the solutions $\{U^\varepsilon\}_{\varepsilon\in(0,\varepsilon_0]}$ of (1.1) satisfy the LDP with a good rate function I given by (3.22).

The additional regularity of solutions of the skeleton Eq. (1.3) will be used to get compactness, see the proof of Proposition 3.3. An example of additional assumptions on σ guaranteeing the desired regularity of solutions of (1.3) can be found in Proposition 3.2.

If $\lambda(\varepsilon) \to \infty$ and $\sqrt{\varepsilon}\lambda(\varepsilon) \to 0$, the asymptotic behaviour of R^ε as $\varepsilon \to 0+$ is known as the moderate deviations principle (MDP). The MDP for the stochastic 2D Navier–Stokes equations was established in [48]. Recently, the MDP for weak solutions (in the PDE sense) of the stochastic 2D primitive equations is shown in [49]. The difficulty in obtaining analogous result in 3D and for strong solutions lies in more delicate estimates of the nonlinear term. For these, we require higher regularity of the solution of deterministic Eq. (1.2). We will prove the following theorem in Sect. 4.3.

Theorem 1.2 (MDP) *Let σ satisfy assumptions [(2.14)–(2.21)] and let F_U satisfy (2.10). Then for all $u_0 \in V\cap H^2$, there exists $\varepsilon_0 > 0$ such that the solutions $\{U^\varepsilon\}_{\varepsilon\in(0,\varepsilon_0]}$ of the stochastic 3D periodic Eq. (1.1) satisfy the MDP with a good rate function given by (4.60).*

For the MDP, additional regularity of solutions of (1.3) is not needed, see Sect. 4.1.

If $\lambda(\varepsilon) \equiv 1$, the limiting process $\varepsilon \to 0+$ corresponds to the central limit theorem (CLT). The CLT for weak solutions of the stochastic 2D primitive equations was recently established in [49]. We prove a weaker version of the CLT with convergence only $\mathbb{P}$-a.s. instead of convergence in $L^2(\Omega)$. The weaker form of the CLT results from a different definition of a solution, which does not guarantee that our solutions considered over a time interval $[0, t]$ form an L^2-integrable random variable for any deterministic $t > 0$. We will establish the following theorem in Sect. 5.

Theorem 1.3 (Almost sure CLT) *Let σ, F and u_0 be as in Theorem 1.2. Let $\hat{U}$ be the solution of*

$$d\hat{U} + \left[A\hat{U} + B(U^0, \hat{U}) + B(\hat{U}, U^0) + F_d(\hat{U})\right] ds = \sigma(\hat{U})\, dW, \quad \hat{U}(0) = u_0.$$

Then,

$$\frac{U^\varepsilon - U^0}{\sqrt{\varepsilon}} \to \hat{U} \text{ as } \varepsilon \to 0+ \text{ in } C([0, T], V) \cap L^2(0, T; D(A)) \ \mathbb{P}\text{-a.s.}$$

We assume the Neumann boundary condition on the top and bottom parts of the boundary. On the lateral part, we assume periodic boundary condition similarly as in, e.g. [5]. The lateral periodic boundary condition allows us to use previous result to establish additional regularity of the solution U^0, i.e. the solution of a deterministic Eq. (1.2), see Sect. 4.2. In particular, the existence result from Cao et al. [5] provides sufficient regularity for the estimates in Sect. 4.2. We emphasize that we do not require any additional regularity of the solutions of the stochastic 3D primitive Eq. (1.1).

The paper is organized as follows: In Sect. 2, we present the functional setting and recall the sufficient condition for the LDP by Budhiraja and Dupuis [3]. The LDP, MDP, and CLT are then established in Sects. 3, 4, and 5, respectively. Since the proofs of Sects. 3 and 4 are quite similar, we choose to present the crucial stochastic argument in full detail only for the LDP in Sect. 3. The MDP counterpart in Sect. 4 is discussed only briefly. Similarly, since the estimates in Sect. 4 are more involved than the estimates in Sect. 3, we include full proofs only for the estimates in Sect. 4.

2 Preliminaries

Let $L, h > 0$ and let $\mathcal{M}_0 = (0, L) \times (0, L) \subseteq \mathbb{R}^2$, $\mathcal{M} = \mathcal{M}_0 \times (-h, 0) \subseteq \mathbb{R}^3$. We decompose the boundary of $\mathcal{M}$ into

$$\Gamma_i = \overline{\mathcal{M}_0} \times \{0\}, \quad \Gamma_l = \partial\mathcal{M}_0 \times (-h, 0), \quad \Gamma_b = \overline{\mathcal{M}_0} \times \{-h\}.$$

2.1 Functional Setting

The reformulated[1] stochastic 3D primitive equations[2] are given by

$$\begin{aligned} &\partial_t v + (v \cdot \nabla)\, v + w(v)\partial_z v + \tfrac{1}{\rho_0}\nabla p_S - \beta_T g \nabla \int_z^0 T \,\mathrm{d}z' + f\vec{k} \times v \\ &\quad - \mu_v \Delta v - \nu_v \partial_{zz} v = F_v + \sigma_1 \left(v, \nabla_3 v, T, \nabla_3 T\right) \dot{W}_1, && (2.1)\\ &\mathrm{div} \int_{-h}^0 v(x, y, z')\,\mathrm{d}z' = 0, && (2.2)\\ &\partial_t T + (v \cdot \nabla)\, T + w\partial_z T - \mu_T \Delta T - \nu_T \partial_{zz} T = F_T + \sigma_2 \left(v, \nabla_3 v, T, \nabla_3 T\right) \dot{W}_2, && (2.3) \end{aligned}$$

where

$$w(v)(x, y, z) = -\int_{-h}^{z} \mathrm{div}\, v(x, y, z')\,\mathrm{d}z' \tag{2.4}$$

is the vertical velocity, $v = (v_1, v_2)$ denotes the horizontal velocity, p_S is the surface pressure, f is the Coriolis parameter, μ_v, ν_v and μ_T and ν_T are the horizontal and vertical viscosity and diffusivity coefficients, respectively. The equations are being

[1] For the original system and the reformulation procedure, see, e.g. [40, Section 2.1].

[2] Usually, the primitive equations of ocean also contain salinity. However, since it does not introduce any additional mathematical difficulties, it is omitted here.

driven by deterministic non-autonomous forces F_v, F_T and stochastic terms σ_1 and σ_2 with multiplicative white noise in time. In the whole manuscript, the symbols div, ∇ and Δ denote the two-dimensional (that is w.r.t. the horizontal coordinates x and y) divergence, gradient and Laplacian, respectively. Their three-dimensional variants will be denoted by div_3, ∇_3 and Δ_3. Equations [(2.1)–(2.3)] are supplied with the initial data

$$v(0) = v_0, \quad T(0) = T_0,$$

and the boundary conditions

$$\begin{aligned} \text{on } \Gamma_i : &\quad \partial_z v = 0, \quad \partial_z T + \alpha T = 0, \\ \text{on } \Gamma_l : &\quad v \text{ and } T \text{ are periodic}, \\ \text{on } \Gamma_b : &\quad \partial_z v = 0, \quad \partial_z T = 0. \end{aligned}$$

Unless specified otherwise, all the function spaces are tacitly considered to contain functions with domain $\mathcal{M}$. The Lebesgue space $L^p(\mathcal{M})$, $p \in [1, \infty]$ will be often denoted by L^p. The norms on $L^p(\mathcal{M})$ and $L^p(\mathcal{M}_0)$ may both be denoted by $|\cdot|_{L^p}$. The norm and the inner product on L^2 will be denoted by $|\cdot|$ and $(\cdot, \cdot)$, respectively. We will also often omit the range of the function spaces as it should be clear from the context. Therefore, assuming $k \in \mathbb{N}$ and $p \in [1, \infty]$, both the Sobolev spaces $W^{k,p}(\mathcal{M})$ and $W^{k,p}(\mathcal{M}; \mathbb{R}^2)$ might be denoted by $W^{k,p}$. We will also use the notation $W^{k,2} = H^k$, $k \in \mathbb{N}$. The norm on H^1 will be denoted by $\|\cdot\|$.

Let H_1 and H_2 be the spaces defined by

$$\begin{aligned} H_1 &= \left\{ v \in L^2\left(\mathcal{M}; \mathbb{R}^2\right) \mid \mathrm{div} \int_{-h}^{0} v \, \mathrm{d}z = 0, \int_{-h}^{0} v \, \mathrm{d}z \text{ is periodic in } \mathcal{M}_0 \right\}, \\ H_2 &= L^2(\mathcal{M}), \end{aligned}$$

and let $H = H_1 \times H_2$. Equipped with the inner product of L^2, the spaces H_1, H_2 and H are Hilbert spaces. Let $P_{H_1} : L^2 \to H_1$ be the hydrostatic Helmholtz–Leray projection, see [34, Lemma 2.2] and [21, Proposition 4.3], and let

$$P_H = \begin{pmatrix} P_{H_1} \\ I \end{pmatrix},$$

where I is the identity on L^2. Let V_1 be the space defined by

$$\begin{aligned} V_1 &= \left\{ v \in H^1\left(\mathcal{M}; \mathbb{R}^2\right) \mid \mathrm{div} \int_{-h}^{0} v \, \mathrm{d}z = 0, v \text{ is periodic w.r.t. } x \text{ and } y \right\}, \\ H_2 &= H^1(\mathcal{M}), \end{aligned}$$

and let $V = V_1 \times V_2$. The spaces V_1, V_2 and V, equipped with the inner product of H^1, are Hilbert spaces. Clearly, $V \hookrightarrow H$.

Let $A_1 : V_1 \to V_1'$, $A_2 : V_2 \to V_2'$ be the symmetric linear operators given by the bilinear forms

$$(A_1 v, v^\sharp) = a_1(v, v^\sharp) = \int_{\mathcal{M}} \mu_v \nabla v \cdot \nabla v^\sharp + \nu_v \partial_z v \partial_z v^\sharp \,\mathrm{d}\mathcal{M},$$
$$(A_2 T, T^\sharp) = a_2(T, T^\sharp) = \int_{\mathcal{M}} \mu_T \nabla T \cdot \nabla T^\sharp + \nu_T \partial_z T \partial_z T^\sharp \,\mathrm{d}\mathcal{M} + \alpha \int_{\Gamma_i} T T^\sharp \,\mathrm{d}\Gamma_i.$$

By [25, Theorem 3.4], $A_1 = P_H \Delta$ in $L(V, V')$. Let

$$AU = \begin{pmatrix} A_1 v \\ A_2 T \end{pmatrix}.$$

The operators A_1, A_2 and A can be extended to self-adjoint unbounded operators on H_1, H_2 and H, respectively, see [34, Lemma 2.4]. Then, we have

$$D(A) = \{U \in V \mid AU \in H\}.$$

Let b be the trilinear form defined by

$$b(U, U^\sharp, U^\flat) = \int_{\mathcal{M}} \left[(v \cdot \nabla) v^\sharp + w(v)\partial_z v^\sharp\right] v^\flat + \left[v \nabla T^\sharp + w(v)\partial_z T^\sharp\right] T^\flat \,\mathrm{d}\mathcal{M}.$$

From [40, Lemma 2.1], we have

$$\left|b(U, U^\sharp, U^\flat)\right| \le C\|U\|\|U^\sharp\|_{H^2}\|U^\flat\|, \qquad U, U^\flat \in H^1, U^\sharp \in H^2. \tag{2.5}$$

By [40, Lemma 3.1], for $U, U^\sharp \in H^2$ and $U^\flat \in H$ we have

$$\left|b\left(U, U^\sharp, U^\flat\right)\right| \le C\left(|v|_{L^6}\|U^\sharp\|^{1/2}\|U^\sharp\|_{H^2}^{1/2} + \|v\|^{1/2}\|v\|_{H^2}^{1/2}|\partial_z U^\sharp|^{1/2}\|\partial_z U^\sharp\|^{1/2}\right)|U^\flat|. \tag{2.6}$$

Using a similar argument as in [40, Lemma 3.1], one can establish

$$\left|b(U, U^\sharp, U^\flat)\right| \le C\|U\|^{1/2}\|U\|_{H^2}^{1/2}\|U^\sharp\|^{1/2}\|U^\sharp\|_{H^2}^{1/2}|U^\flat|, \qquad U, U^\sharp \in H^2, U^\flat \in L^2. \tag{2.7}$$

Similarly, we can improve (2.5) to

$$\left|b\left(U, U^\sharp, U^\flat\right)\right| \le C\|v\|\|U^\sharp\|^{1/2}\|U^\sharp\|_{H^2}^{1/2}|U^\flat|^{1/2}\|U^\flat\|^{1/2}, \qquad U, U^\flat \in H^1, U^\sharp \in H^2. \tag{2.8}$$

A similar estimate for $U^\flat \in H^3$ has been established in [43, Lemma 2.3]. The form b has the anti-symmetry property

$$b(U, U^\sharp, U^\flat) = -b(U, U^\flat, U^\sharp), \quad U \in V, U^\sharp, U^\flat \in V \cap H^2,$$

in particular

$$b(U, U^\sharp, U^\sharp) = 0, \quad U \in V, U^\sharp \in V \cap H^2. \tag{2.9}$$

We define the bilinear operator B by

$$B(U, U^\sharp) = P_H \begin{pmatrix} (v \cdot \nabla)v^\sharp + w(v)\partial_z v^\sharp \\ v\nabla T^\sharp + w(v)\partial_z T^\sharp \end{pmatrix}$$

and write $B(U) = B(U, U)$.

Let $A_{\mathrm{pr}} : V \to H$ be the linear operator

$$A_{\mathrm{pr}} U = P_H \begin{pmatrix} -\beta_T g \nabla \int_z^0 T z, \mathrm{d}z' \\ 0 \end{pmatrix}.$$

Clearly, A_{pr} is continuous. We define the linear operator $E : H \to H$ by

$$EU = P_H \begin{pmatrix} f\vec{k} \times v \\ 0 \end{pmatrix}$$

The operator E is continuous and $(EU, U) = 0$. Let

$$F_U = P_H \begin{pmatrix} F_v \\ F_T \end{pmatrix} \in L^2_{\mathrm{loc}}(0, \infty; H). \tag{2.10}$$

We denote

$$F(U) = A_{\mathrm{pr}} U + EU + F_U.$$

To summarize the above, we assume that $F : V \to H$ satisfies

$$\int_s^t |F(U)|^2 \, \mathrm{d}r \leq C \left(\int_s^t |F_U|^2 + \|U\|^2 \, \mathrm{d}r \right), \quad U \in V, 0 \leq s \leq t < \infty, \tag{2.11}$$

$$|F(U) - F(U^\sharp)| \leq C\|U - U^\sharp\|, \quad U \in V, \tag{2.12}$$

with the constant C in (2.11) independent of s, t.

Let $\mathcal{A}_2$, $\mathcal{A}_3$ and $\mathcal{R}$ be the averaging operators and the remainder defined for $v : \mathcal{M} \to \mathbb{R}^2$ by

$$(\mathcal{A}_2 v)(x, y) = \frac{1}{h} \int_{-h}^0 v(x, y, z') \, \mathrm{d}z', \quad (\mathcal{A}_3 v)(x, y, z) = (\mathcal{A}_2 v)(x, y), \quad \mathcal{R} = I - \mathcal{A}_3. \tag{2.13}$$

It is straightforward to check that $\|\mathcal{A}_3\|_{L(H_1)} \leq 1$ and $\|\mathcal{A}_2\|_{L(H_1, \overline{H})} \leq h^{-1/2}$, $\mathcal{R} : H_1 \to H_1$ and $\|\mathcal{R}\|_{L(H_1)} \leq 2$. Since the spaces H and $\overline{H}$ have the norm of $L^2\left(\mathcal{M}; \mathbb{R}^2\right)$

and $L^2\left(\mathcal{M}_0;\mathbb{R}^2\right)$, respectively, we observe that the operators $\mathcal{A}_2$, $\mathcal{A}_3$ and $\mathcal{R}$ remain bounded also if considered with $L^2(\mathcal{M})$ and $L^2(\mathcal{M}_0)$ in place of H_1 and $\overline{H}$.

Let $\sigma \in \mathrm{Lip}\left(V, L_2(\mathcal{U}, H)\right) \cap \mathrm{Lip}\left(D(A), L_2(\mathcal{U}, V)\right)$, in particular

$$\|\sigma(U)\|^2_{L_2(\mathcal{U},H)} \le C\left(1+|U|^2\right)+\eta_0\|U\|^2, \quad U \in H, \tag{2.14}$$

$$\|\sigma(U)\|^2_{L_2(\mathcal{U},V)} \le C\left(1+\|U\|^2\right)+\eta_1|AU|^2, \quad U \in V, \tag{2.15}$$

$$\|\sigma(U)-\sigma(U^\sharp)\|^2_{L_2(\mathcal{U},H)} \le C\|U-U^\sharp\|^2, \quad U, U^\sharp \in V, \tag{2.16}$$

$$\|\sigma(U)-\sigma(U^\sharp)\|^2_{L^2(\mathcal{U},V)} \le C\|U-U^\sharp\|^2+\gamma|AU-AU^\sharp|^2, \quad U, U^\sharp \in D(A), \tag{2.17}$$

for some $\eta_0, \eta_1, \gamma \ge 0$. Due to the nature of the estimates in Sects. 3 and 4 , we also assume that σ satisfies the following structural assumptions

$$\sum_{k=1}^{\infty} |\mathcal{R}\sigma_1(U)e_k|^2_{L^6} \le C\left(1+|\mathcal{R}v|^2_{L^6}\right), \tag{2.18}$$

$$\sum_{k=1}^{\infty} |\sigma_2(U)e_k|^2_{L^6} \le C\left(1+|T|^2_{L^6}\right), \tag{2.19}$$

$$\|\mathcal{A}_2\sigma_1(U)\|^2_{L_2(\mathcal{U},\overline{V})} \le C\left(1+\|U\|^2\right)+\eta_2|A_S\mathcal{A}_2 v|^2_{\overline{H}}, \tag{2.20}$$

$$\|\partial_z\sigma_i(U)\|^2_{L_2(\mathcal{U},L^2)} \le C\left(1+\|U\|^2\right)+\eta_3|\nabla_3\partial_z U_i|^2. \tag{2.21}$$

with $\eta_2, \eta_3 \ge 0$. A non-trivial example of a noise term σ depending on the horizontal gradient of the vertically averaged velocity satisfying the assumptions [(2.14)–(2.21)] can be found in [2, Section 2.5].

Using notation from above, we are able to rewrite equations [(2.1)–(2.3)] in the abstract form

$$dU+\left[AU+B(U)+F(U)\right]dt=\sigma(U)\,dW, \quad U(0)=u_0. \tag{2.22}$$

In [19], anisotropic spaces were used to establish estimates on the nonlinear term B. For $1 \le p, q < \infty$, we denote

$$|v|_{L^q_x L^p_z}=\left(\int_{\mathcal{M}_0}\left(\int_{-h}^{0}\left(|v_1|^p+|v_2|^p\right)\,\mathrm{d}z\right)^{q/p}\mathrm{d}\mathcal{M}_0\right)^{1/q}.$$

Lemma 2.1 *The following anisotropic estimates hold with constant C depending only on $\mathcal{M}$:*

$$|v|_{L^q_x L^2_z} \le C|v|^{2/q}\|v\|^{1-2/q}, \quad v \in H^1, \quad q \ge 2, \tag{2.23}$$

$$|v|_{L^q_x L^2_z} \le C|v|^{6/q}_{L^6}\|v^3\|^{1/3-2/q}, \quad v^3 \in H^1, \quad q \ge 6, \tag{2.24}$$

$$\mathrm{esssup}_{z\in(-h,0)}|v(\cdot,z)|_{L^2(\mathcal{M}_0)} \equiv |v|_{L^\infty_z L^2_x} \le C|v|^{1/2}\|v\|^{1/2}, \quad v\in H^1, \tag{2.25}$$

$$|v^5|_{L^3_x L^2_z} \le C|v|_{L^6}\|v^3\|^{4/3}, \quad v^3\in H^1, \tag{2.26}$$

$$|v^2|_{L^4_x L^3_z} \le C|v|_{L^6}^{3/2}\|v^3\|^{1/6}, \quad v^3\in H^1. \tag{2.27}$$

Proof The estimate (2.23) can be shown by repeating the argument from the proof of [40, Lemma 3.1]. The inequality (2.25) has been established in [19, Lemma 3.3(a)]. The estimate (2.24) follows from (2.23). Denoting $s = 1 - 6/q$, we use the Hölder inequality and (2.23) and obtain

$$|v|_{L^q_x L^2_z} \le |v^3|^{1/3}_{L^{q/3}_x L^2_z} \le C|v^3|^{(1-s)/3}\|v^3\|^{s/3} = C|v|^{1-s}_{L^6}\|v^3\|^{s/3}.$$

Regarding (2.26), we employ the Minkowski inequality, the Gagliardo–Nirenberg inequality in 2D and the estimate (2.25) to deduce

$$\begin{aligned}
|v^5|^2_{L^3_x L^2_z} &= \left(\int_{\mathcal{M}_0}\left(\int_{-h}^0 |v|^{10}\,\mathrm{d}z\right)^{3/2}\mathrm{d}\mathcal{M}_0\right)^{2/3} \le \int_{-h}^0\left(\int_{\mathcal{M}_0}|v|^{15}\,\mathrm{d}\mathcal{M}_0\right)^{2/3} \\
&= \int_{-h}^0 |v^3|^{10/3}_{L^5_x}\,\mathrm{d}z \\
&\le C\int_{-h}^0 |v^3|^{4/3}_{L^2_x}|\nabla(v^3)|^2_{L^2_x} + |v^3|^{10/3}_{L^2_x}\,\mathrm{d}z \le C|v^3|^{4/3}_{L^\infty_z L^2_x}\|v^3\|^2 \\
&\le C|v^3|^{2/3}\|v^3\|^{8/3} \le C|v|^2_{L^6}\|v^3\|^{8/3}.
\end{aligned}$$

The final inquality (2.27) can be established similarly as (2.26). ◻

For the sake of completeness, let us recall the Burkholder–Davis–Gundy inequality. Let X be a separable Hilbert space and $r \ge 2$. Let W be a cylindrical Wiener process with reproducing kernel Hilbert space $\mathcal{U}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for all $\Phi \in L^2\left(\Omega, L^2\left(0, T; L_2\left(\mathcal{U}, X\right)\right)\right)$

$$\mathbb{E}\sup_{t\in[0,T]}\left|\int_0^t \Phi\,\mathrm{d}W\right|_X^r \le C_{BDG}\,\mathbb{E}\left(\int_0^T \|\Phi\|^2_{L_2(\mathcal{U},X)}\,\mathrm{d}t\right)^{r/2}, \tag{2.28}$$

where the constant C_{BDG} depends only on r. For proof, see, e.g. [27, Theorem 3.28, p. 166].

2.2 Definition of Solution

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, $\mathbb{F} = (\mathcal{F}_t)_{t\ge0}$, be a stochastic basis satisfying the usual conditions. We will consider only solutions strong both in the stochastic and PDE senses.

Definition 2.2 A progressively measurable stochastic process $U : \Omega\times[0,\infty) \to V$ is a *global solution* of Eq. (2.22) if there exists a nondecreasing sequence of $\mathbb{F}$-stopping times τ_N such that

1. for all $N \in \mathbb{N}$ and $T \geq 0$,

$$U(\cdot \wedge \tau_N) \in L^2(\Omega; C([0,T], V)), \quad \mathbb{1}_{[0,\tau_N]}(\cdot)U \in L^2\left(\Omega; L^2(0,T; D(A))\right), \tag{2.29}$$

2. for all $N \in \mathbb{N}$ and $t \geq 0$, the stopped process satisfies the following equation $\mathbb{P}$-a.s. in H

$$U(t \wedge \tau_N) + \int_0^{t\wedge\tau_N} AU + B(U) + A_{\text{pr}}U + EU - F_U \,\mathrm{d}s = U(0) + \int_0^{t\wedge\tau_N} \sigma(U)\,\mathrm{d}W, \tag{2.30}$$

3. $\tau_N \to \infty$ $\mathbb{P}$-almost surely.

In the following Sections, equations similar to (2.22) are studied and the above definition needs to be adjusted accordingly. The following well-posedness result was established in [2, Theorem 2.6].

Theorem 2.3 *Let $u_0 \in V$. Let F_U satisfy (2.10) and $F_T \in L^2\left(\Omega; L^2\left(0,T; L^2(\Gamma_i)\right)\right)$. Let σ satisfy [(2.14)–(2.21)]. If the constants γ, η_i, $i = 1,2,3,4$, in (2.14), (2.15), (2.17), (2.20), and (2.21) are sufficiently small, then there exists a unique global solution of Eq. (2.22). Moreover, if τ_N is the sequence of stopping times from Definition 2.2, the solution U satisfies*

$$\mathbb{1}_{[0,\tau_N]}U \in L^p\left(\Omega; L^\infty(0,T;V)\right), \quad \mathbb{1}_{[0,\tau_N]}\|U\|^{p-2}|AU|^2 \in L^1\left(\Omega; L^1(0,T)\right),$$

for all $N \in \mathbb{N}$, $p \geq 2$ and $T > 0$.

2.3 Large Deviations Principle

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, let X be a separable Banach space, and let W be a cylindrical Wiener process with reproducing kernel Hilbert space $\mathcal{U}$.

Definition 2.4 A function $I : X \to [0,\infty]$ is called a *rate function* if it is lower semicontinuous. If for all $M > 0$ the set $\{U \in X \mid I(U) \leq M\}$ is compact, a rate function I is called a *good rate function*.

Definition 2.5 Let $\varepsilon_0 > 0$ and let $\{U^\varepsilon\}_{\varepsilon\in(0,\varepsilon_0]}$ be X-valued stochastic processes, $U^0 \in X$ and let I be a rate function. We say that the processes $\{U^\varepsilon\}_{\varepsilon\in(0,\varepsilon_0]}$ satisfy

1. the *large deviations principle* (LDP) with the rate function I if for each $A \in \text{Borel}(X)$

$$-\inf_{x\in A^o} I(x) \leq \liminf_{\varepsilon\to 0+} \varepsilon \log \mathbb{P}(\{U^\varepsilon \in A\}) \leq \limsup_{\varepsilon\to 0+} \varepsilon \log \mathbb{P}(\{U^\varepsilon \in A\}) \leq -\inf_{x\in \bar{A}} I(x),$$

2. the *moderate deviations principle* (MDP) the with rate function I if the processes

$$R^\varepsilon = \{(U^\varepsilon - U^0)/(\sqrt{\varepsilon}\lambda(\varepsilon)\}_{\varepsilon\in(0,\varepsilon_0]}$$

satisfy the LDP with the rate function I.

It is well known that if $\mathcal{U}_0$ is a separable Hilbert space such that the embedding $\mathcal{U} \hookrightarrow \mathcal{U}_0$ is Hilbert-Schmidt, the cylindrical Wiener process W has continuous trajectories in $\mathcal{U}_0$. Let $\mathcal{G}^\varepsilon : C([0,t],\mathcal{U}_0) \to X$ be a measurable map such that $U^\varepsilon = \mathcal{G}^\varepsilon(W(\cdot))$. For $M \geq 0$, let

$$\begin{aligned}\mathcal{T}_M &= \left\{h \in L^2(0,T;\mathcal{U}) \mid \int_0^T |h(s)|^2_{\mathcal{U}}\,\mathrm{d}s \leq M\right\},\\ \mathcal{A}_M &= \{\phi : \Omega\times[0,T] \to \mathcal{U} \mid \phi \text{ is an } \mathbb{F}\text{-predictable process s.t. } \phi(\omega,\cdot)\\ &\qquad \in \mathcal{T}_M \text{ for } \mathbb{P}\text{-a.s. } \omega\in\Omega\}.\end{aligned}$$

The space $L^2(0,T;\mathcal{U})$ is equipped with the weak topology which makes the set $\mathcal{T}_M \subseteq L^2(0,T;\mathcal{U})$ compact for arbitrary $M \geq 0$.

The following theorem from Budhiraja and Dupuis [3, Theorem 4.4] establishes a sufficient condition for the LDP.

Theorem 2.6 *Let $\mathcal{G}^0 : C([0,T],\mathcal{U}_0) \to X$ be a measurable map satisfying the following conditions:*

1. *Let $M > 0$ and let $\{h_\varepsilon \mid \varepsilon \in (0,\varepsilon_0]\} \subseteq \mathcal{A}_M$ be such that $h_\varepsilon \to h$ in distribution for some $\mathcal{T}_M$-valued random variable h. Then,*

$$\mathcal{G}^\varepsilon\left(W(\cdot) + \frac{1}{\sqrt{\varepsilon}}\int_0^\cdot h_\varepsilon(s)\,\mathrm{d}s\right) \to \mathcal{G}^0\left(\int_0^\cdot h(s)\,\mathrm{d}s\right) \quad \text{in distribution.} \tag{2.31}$$

2. *The set*

$$K_M = \left\{\mathcal{G}^0\left(\int_0^\cdot h(s)\,\mathrm{d}s\right) \mid h\in\mathcal{T}_M\right\} \tag{2.32}$$

is compact in X for all $M > 0$.

Then, Y^ε satisfies the large deviations principle with a good rate function

$$I(U) = \inf\left\{\frac{1}{2}\int_0^T |h|^2_{\mathcal{U}}\,\mathrm{d}s \mid h\in L^2(0,T;\mathcal{U}) \text{ s.t. } U = \mathcal{G}^0\left(\int_0^\cdot h\,\mathrm{d}s\right)\right\}. \tag{2.33}$$

3 Large Deviations Principle

The proofs in this section are modifications of the ones in [12], the modifications consisting in considering less regular noise term σ and taking into account the definition of solution of the stochastic primitive equations. We will concentrate mainly on the differences between our setting and the one of Dong et al. [12].

In this Section, we assume $u_0 \in V$. Let U^ε be the solution of the equation

$$dU^\varepsilon + \left[AU^\varepsilon + B\left(U^\varepsilon\right) + F\left(U^\varepsilon\right)\right] dt = \sqrt{\varepsilon}\sigma\left(U^\varepsilon\right) dW, \quad U^\varepsilon(0) = u_0. \quad (3.1)$$

Equation (3.1) is well-posed by Theorem 2.3 for arbitrary γ, η_i, $i = 0, 1, 2, 3$, and sufficiently small $\varepsilon > 0$. Indeed, the analogues of the constants γ, η_i, $i = 0, 1, 2, 3$, for $\tilde{\sigma}(U^\varepsilon) = \sqrt{\varepsilon}\sigma(U^\varepsilon)$ can be made sufficiently small by choosing $\varepsilon_0 > 0$ small enough and considering only $\varepsilon \in (0, \varepsilon_0]$.

By the result of Röckner et al. [42], there exists a measurable map $\mathcal{G}^\varepsilon : C([0, T], \mathcal{U}_0) \to C([0, T], V) \cap L^2(0, T; D(A))$ such that $U^\varepsilon = \mathcal{G}^\varepsilon(W(\cdot))$.

3.1 Skeleton Equation

Let $h \in L^2(0, T; \mathcal{U})$. By the *skeleton equation*, we understand the equation

$$dU_h + [AU_h + B(U_h) + F(U_h)]\, dt = \sigma(U_h)\, h\, dt, \quad U_h(0) = u_0. \quad (3.2)$$

Proposition 3.1 *The skeleton equation (3.2) is well-posed. Moreover, the solution U_h satisfies*

$$\sup_{s\in[0,T]} \|U_h\|^2 + \int_0^T |AU_h|^2\, \mathrm{d}s \le C_{T,M,u_0}. \quad (3.3)$$

Proof The proof runs along the same lines as in [12, Section 5], that is by using the approaches of Petcu et al. [40] and Cao and Titi [6] for the local and global existence results, respectively, with only minor additional complications in regard to the presence of the control term $\sigma(U_h)h$. Similar estimates will be seen in Sect. 4.2 for a more complicated stochastic equation. □

Before we proceed to the compactness argument, let us return to the regularity assumption $U_h \in L^\infty(0, T; H^2)$ in Theorem 1.1. Existence of local regular solutions of (3.2) has been established in [5] for $\sigma \equiv 0$ by a fixed point argument. In the following Proposition, we introduce additional assumptions on σ to ensure $U_h \in L^\infty(0, T; H^2)$. These assumptions are more or less a technical issue in the sense that the LDP result from Theorem 1.1 depends only on the regularity of U_h and not on the particular additional assumptions formulated below.

Proposition 3.2 *Let $\sigma : [0, \infty) \times V \to H$ be such that the bounds* [(2.14)–(2.15)] *and* [(2.18)–(2.21)] *are satisfied with $\sigma(t, U)$ instead of $\sigma(U)$ with constants C, η_i, $i = 0, 1, 2, 3$, independent of t. Let $\gamma_H, \gamma_V : [0, \infty) \to [0, \infty)$ be continuous functions such that $\gamma_H(0) = \gamma_V(0) = 0$ and*

$$\|\sigma(t, U_1) - \sigma(t, U_2)\|_{L_2(\mathcal{U},H)} \le \gamma_H(t)\|U_1 - U_2\|, \quad U_1, U_2 \in V, \quad (3.4)$$

$$\|\sigma(t, U_1) - \sigma(t, U_2)\|_{L_2(\mathcal{U},V)} \le \gamma_V(t)\|U_1 - U_2\|_{H^2}, \quad U_1, U_2 \in H^2. \quad (3.5)$$

Then the solution U_h of (3.2) *satisfies*

$$U_h \in C\left([0,T], H^2\right) \cap L^2\left(0,T; H^3\right)$$

for all $T > 0$, $h \in L^2(0,T;\mathcal{U})$ and $u_0 \in V \cap H^2$.

Proof We briefly comment on the extension of the fixed point argument from Cao et al. [5, Section 2]. For simplicity, we will not include all the details, such as the reflection argument. Let for $t > 0$

$$X_T = \left\{U = (v,T) \in C\left([0,T], H^2\right) \cap L^2\left(0,T; H^3\right) \mid \operatorname{div} \mathcal{A}_3 v = 0\right\},$$
$$\|U\|^2_{X_T} = \sup_{s\in[0,T]} \|U(s)\|^2_{H^2} + \int_0^T \|U(s)\|^2_{H^3}\,\mathrm{d}s.$$

Let

$$D(U) = -B(U) - A_{\mathrm{pr}}U - EU - F_U, \qquad U \in X_T.$$

Using the decomposition into barotropic and baroclinic modes as in [6] and standard regularity theory for the Stokes and parabolic equations, one can show that the equation

$$d\mathcal{V} + A\mathcal{V}\,dt = (D(U) + \sigma(U)h)\,dt, \qquad \mathcal{V}(0) = U(0), \tag{3.6}$$

has a solution in X_T for all $U \in X_T$. Let $h \in L^2(0,T;\mathcal{U})$ with $\|h\|_{L^2(0,T;\mathcal{U})} \le M$ for some $M \ge 0$. Let $U_1, U_2 \in X_T$ such that $\|U_1\|_{X_T}, \|U_2\|_{X_T} \le K$ for some $K \ge 0$ and let $\mathcal{V}_1, \mathcal{V}_2 \in X_T$ be the respective solutions of (3.6) with $U = U_1$ and U_2, respectively. Subtracting the equations for $\mathcal{V}_1$ and $\mathcal{V}_2$, testing by $\mathcal{V} = \mathcal{V}_1 - \mathcal{V}_2$ and $-\nabla\Delta\mathcal{V}$ and using the bounds on $D(U)$ from Cao et al. [5, Proposition 2.2] and [(3.4)–(3.5)], we obtain the estimate

$$\sup_{s\in[0,T]} \|\mathcal{V}\|^2_{H^2} + \int_0^T \|\nabla\mathcal{V}\|^2_{H^2}\,\mathrm{d}s \le C\|U_1 - U_2\|^2_{X_T}$$
$$\left(K^2T^{1/2} + M^2 \sup_{s\in[0,T]} \max\{\gamma_H(s), \gamma_V(s)\}\right).$$

It is now easy to see that, thanks to the continuity of γ_H, γ_V and $\gamma_H(0) = \gamma_V(0) = 0$, the solution operator is a contraction if we choose $T > 0$ sufficiently small. The rest of the fixed point argument follows similarly as in [5, Propositions 2.1 and 2.3].

The global existence follows from apriori estimates analogous to those in [26, Section 3]. □

Proposition 3.3 *Let $\{h_n\}_{n=1}^\infty \subseteq \mathcal{T}_M$ and let $U^n = U_{h_n}$ be the respective solutions of* (3.2)*. Then, there exist $h \in L^2(0,T;\mathcal{U})$ and a subsequence (for simplicity* *not relabelled) such that* $h_n \rightharpoonup h$ *in* $L^2(0, T; \mathcal{U})$ *and* $U^n \to U_h$ *in* $L^\infty(0, T; V) \cap L^2(0, T; D(A))$.

Moreover, there exists a measurable map $\mathcal{G}^0 : C([0, T], \mathcal{U}_0) \to C([0, T], V) \cap L^2(0, T; D(A))$ *such that* $U_h = \mathcal{G}^0\Big(\int_0^\cdot h_s \, \mathrm{d}s\Big)$.

Proof Recalling the growth estimate of σ (2.14) in $L_2(\mathcal{U}; H)$, we may repeat the argument from Dong et al. [12, Proposition 5.3] to show that there exists a (not relabelled) subsequence U^n and $\tilde{U} \in L^\infty(0, T; V) \cap L^2(0, T; D(A))$ such that

$$U^n \to \tilde{U} \text{ in } L^2(0, T; V), \quad U^n \overset{*}{\rightharpoonup} \tilde{U} \text{ in } L^\infty(0, T; V), \quad U^n \overset{*}{\rightharpoonup} \tilde{U} \text{ in } L^2(0, T; D(A)). \tag{3.7}$$

The continuity of $\tilde{U}$ can be shown using a maximal regularity type argument based on the Lions–Magenes lemma, see, e.g. [44, Lemma 1.2, Chapter 3], similarly as in [24, Section 4] for the deterministic or [9, Section 7.3] for the stochastic case. The case of σ satisfying assumptions [(2.14)–(2.15)] is established in [2, Proposition 3.3].

Next, we show $\tilde{U} = U_h$. We proceed similarly as in, e.g. [9, Section 7]. Let $U^\sharp \in D(A)$ and $t \in [0, T]$ be fixed. Since A is self-adjoint, we may use the Cauchy–Schwartz inequality and the first convergence in (3.7) to get

$$\left| \int_0^t \Big(AU^n - A\tilde{U}, U^\sharp \Big) \, \mathrm{d}s \right| \le C_t \|U^\sharp\| \left(\int_0^t \|U^n - \tilde{U}\|^2 \, \mathrm{d}s \right)^{1/2} \to 0.$$

By the Lipschitz continuity of F in (2.12) and the first convergence in (3.7), we have

$$\left| \int_0^t \Big(F(U^n) - F(\tilde{U}), U^\sharp \Big) \, \mathrm{d}s \right| \le C_t |U^\sharp| \left(\int_0^t \|U^n - \tilde{U}\|^2 \, \mathrm{d}s \right) \to 0.$$

Using the estimate (2.5) and the first two convergences (3.7), we deduce

$$\begin{aligned} &\left| \int_0^t \Big(B(U^n) - B(\tilde{U}), U^\sharp \Big) \, \mathrm{d}s \right| \\ &\quad \le C \int_0^t \|U^n\| \|U^n - \tilde{U}\| |AU^\sharp| + \|U^n - \tilde{U}\| \|\tilde{U}\| |AU^\sharp| \, \mathrm{d}s \\ &\quad \le C_t \left(\sup_{s \in [0,t]} \|U^n\| + \sup_{s \in [0,t]} \|\tilde{U}\| \right) \left(\int_0^t \|U^n - \tilde{U}\|^2 \, \mathrm{d}s \right)^{1/2} \to 0. \end{aligned}$$

By the triangle inequality, we have

$$\begin{aligned} &\left| \int_0^t \Big(\sigma(U^n) h_n - \sigma(\tilde{U}) h, U^\sharp \Big) \, \mathrm{d}s \right| \\ &\quad \le \int_0^t \left| \Big(\Big[\sigma(U^n) - \sigma(\tilde{U}) \Big] h_n, U^\sharp \Big) \right| \, \mathrm{d}s \\ &\qquad + \left| \int_0^t \Big(\sigma(\tilde{U}) \left[h^n - h \right], U^\sharp \Big) \, \mathrm{d}s \right| = \left| I_1^n \right| + \left| I_2^n \right|. \end{aligned}$$

The first integral can be estimated directly using the Lipschitz continuity (2.16) of σ in $L_2(\mathcal{U}, H)$, the boundedness of h_n and the first convergence in (3.7) by

$$\left|I_1^n\right| \leq C|U^\sharp| \left(\int_0^t |h_n|^2_{\mathcal{U}}\, \mathrm{d}s\right)^{1/2} \left(\int_0^t \|U^n - \tilde{U}\|^2\, \mathrm{d}s\right)^{1/2} \to 0.$$

Since $\sigma(\tilde{U}) : L^2(0, T; \mathcal{U}) \to L^2(0, T; H)$ is a linear bounded operator, the integral I_2^n converges to zero using the weak convergence $\sigma(\tilde{U})[h_n - h] \rightharpoonup 0$ in $L^2(0, T; H)$. Collecting the above, we get

$$\begin{aligned} &\left(U^n(t), U^\sharp\right) + \int_0^t \left(AU^n, U^\sharp\right) + \left(B(U^n), U^\sharp\right) + \left(F(U^n), U^\sharp\right) - \left(\sigma(U^n), U^\sharp\right) \\ &\quad \to \left(\tilde{U}(t), U^\sharp\right) + \int_0^t \left(A\tilde{U}, U^\sharp\right) + \left(B(\tilde{U}), U^\sharp\right) + \left(F(\tilde{U}), U^\sharp\right) - \left(\sigma(\tilde{U}), U^\sharp\right) \end{aligned}$$

and thus $\tilde{U}$ satisfies the skeleton Eq. (3.2) in $D(A)'$. Using the density of $D(A)$ in H, we get that $\tilde{U}$ satisfies (3.2) in H and thus by uniqueness $\tilde{U} = U_h$.

It remains to establish the strong convergence in $L^\infty(0, T; V) \cap L^2(0, T; D(A))$. To that end, we adapt the proof of [12, Theorem 5.8]. Let $w^n = U^n - U_h$. Clearly, w_n satisfies

$$\begin{aligned} &dw^n + \left[Aw^n + B(U^n) - B(U_h) + F(U^n) - F(U_h)\right] dt \\ &\quad = \left[\sigma(U^n)h_n - \sigma(U_h)\right] dt, \quad w^n(0) = 0. \end{aligned}$$

By the Itô lemma, see, e.g. [2, Theorem A.1], we have

$$\begin{aligned} d\|w^n\|^2 + 2|Aw^n|^2\, dt &= -2\left(A^{1/2}\left[B(U^n) - B(U_h)\right], A^{1/2}w^n\right) dt \\ &\quad -2\Big[\left(A^{1/2}\left[F(U^n) - F(U_h)\right], A^{1/2}w^n\right) \\ &\quad + \left(A^{1/2}\left[\sigma(U^n)h_n - \sigma(U_h)h\right], A^{1/2}w^n\right)\Big]\, dt. \end{aligned}$$

Fix $\eta > 0$ and let $t \in [0, T]$. Since A is self-adjoint, we recall the Lipschitz continuity of F in (2.12) and use the Cauchy–Schwartz and the Young inequalities to obtain

$$\left|2\int_0^t \left(A^{1/2}\left[F(U^n) - F(U_h)\right], A^{1/2}w^n\right) \mathrm{d}s\right| \leq \frac{\eta}{3}\int_0^t |Aw^n|^2\, \mathrm{d}s + C_\eta \int_0^t \|w^n\|^2\, \mathrm{d}s.$$

Next, we use the estimate (2.7) and the boundedness of U^n in $C([0, T], V)$ to deduce

$$\begin{aligned} &\left|2\int_0^t \left(A^{1/2}\left[B(U^n) - B(U_h)\right], A^{1/2}w^n\right) \mathrm{d}s\right| \\ &\quad \leq \int_0^t \left|\left(B(U^n, U^n - U_h), Aw^n\right)\right| \mathrm{d}s + \int_0^t \left|\left(B(U^n - U_h, U_h), Aw^n\right)\right| \mathrm{d}s \end{aligned}$$

$$\leq \frac{\eta}{3}\int_0^t |Aw^n|^2\,\mathrm{d}s + C_\eta \int_0^t |AU^n|^2\|w^n\|^2\,\mathrm{d}s.$$

The final term is treated using the idea from Dong et al. [12, Theorem 5.8] and the additional regularity of the solution of the skeleton Eq. (3.2). By the Lipschitz continuity of σ (2.16) in $L_2(\mathcal{U}, H)$ and the growth estimate of σ (2.15) in $L_2(\mathcal{U}, V)$, we obtain

$$\begin{aligned}
&\left|\int_0^t \left(A^{1/2}\left[\sigma(U^n)h_n - \sigma(U_h)h\right], A^{1/2}w^n\right)\mathrm{d}s\right| \\
&\quad\leq \left|\int_0^t \left(A^{1/2}\left[\sigma(U^n) - \sigma(U_h)\right]h_n, A^{1/2}w^n\right)\mathrm{d}s\right| \\
&\qquad+ \left|\int_0^t \left(A^{1/2}\sigma(U_h)\left[h_n - h\right], A^{1/2}w^n\right)\mathrm{d}s\right| \\
&\quad\leq \frac{\eta}{3}\int_0^t |Aw^n|^2\,\mathrm{d}s + C_\eta \int_0^t |h_n|_{\mathcal{U}}^2\|w^n\|^2\,\mathrm{d}s \\
&\qquad+ C\sup_{s\in[0,t]}\left(1 + \|U_h\|_{H^2}\right)\left(\int_0^t |h_n - h|_{\mathcal{U}}^2\,\mathrm{d}s\right)^{1/2}\left(\int_0^t \|w^n\|^2\,\mathrm{d}s\right)^{1/2}.
\end{aligned}$$

Collecting the estimates above and choosing $\eta > 0$ sufficiently small, we get

$$\begin{aligned}
\|w^n(t)\|^2 + \int_0^t |Aw^n|^2\,\mathrm{d}s &\leq C_h\left(\int_0^t \|w^n\|^2\,\mathrm{d}s\right)^{1/2} \\
&\quad+ C_\eta \int_0^t \|w^n\|^2\left(1 + |AU^n|^2 + |AU_h|^2 + |h_n|_{\mathcal{U}}^2\right)\mathrm{d}s.
\end{aligned} \tag{3.8}$$

for all $t \in [0, T]$. By the Gronwall lemma and the bounds (3.7), we get

$$\|w^n(t)\|^2 \leq C_{h,\{h_k\},u_0}\left(\int_0^t \|w^n\|^2\,\mathrm{d}s\right)^{1/2}, \tag{3.9}$$

where the constant C_h depends only on h, the bound of $\{h_k\}_{k=1}^\infty$ in $L^2(0, T; \mathcal{U})$ and u_0, and is independent of n. Recalling that $w^n \to 0$ in $L^2(0, T; V)$ by (3.7), we may pass to the limit in (3.9) and obtain strong convergence in $C([0, T]; V)$. Strong convergence in $L^2(0, T; H^2)$ then follows immediately from (3.8).

The measurable map $\mathcal{G}^0 : C([0, T], \mathcal{U}_0) \to C([0, T], V) \cap L^2(0, T; D(A))$ is defined by

$$\mathcal{G}^0(h) = \begin{cases} U_{\tilde{h}}, & \text{if } \tilde{h} = \int_0^\cdot h(t)\,\mathrm{d}t \text{ for some } h \in L^2(0, T; \mathcal{U}), \\ 0, & \text{otherwise,} \end{cases} \tag{3.10}$$

as in [12]. □

3.2 Preliminary Estimate

Let $h_\varepsilon \in L^2(0,T;\mathcal{U})$. The equation

$$\begin{aligned} dU^\varepsilon &+ \left[AU^\varepsilon + B(U^\varepsilon, U^\varepsilon) + F(U^\varepsilon)\right] dt \\ &= \sigma(U^\varepsilon) h_\varepsilon \, dt + \sqrt{\varepsilon}\sigma(U^\varepsilon)\, dW, \quad U^\varepsilon(0) = u_0. \end{aligned} \tag{3.11}$$

will play an important role in the proof of Theorem 1.1. In [12], the existence and uniqueness of (3.11) is established by the means of the Girsanov theorem. However, for the argument in the proof of [12, Theorem 1.1, Step 2] to go through as described, one needs additional information on the behaviour of AU_{h_ε}.

Existence of global pathwise solutions U^ε essentially follows from Brzeźniak and Slavík [2], which is inspired by the arguments of Cao and Titi [6] and Debussche et al. [9,10]. In order to avoid repetition, we will only state the necessary preliminary result. The following proposition can be established by a series of estimates similar to those in [2, Section 4], see also Sect. 4.2 for estimates of the same kind for a more complicated equation.

Proposition 3.4 *Let U_ε be the solution of (3.11) and let $\tau_K^{U,\varepsilon}$ be the stopping time defined by*

$$\tau_K^{U,\varepsilon,p} = \inf\left\{ t \geq 0 \mid \int_0^t \|U_\varepsilon\|^{p-2}|AU_\varepsilon|^2 \,\mathrm{d}s \geq K \right\}., \quad K > 0, \varepsilon \in (0, \varepsilon_0].$$

Then, $\tau_K^{U,\varepsilon} \to \infty$ $\mathbb{P}$-almost surely. Moreover, for all $t > 0$, one has

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{U,\varepsilon} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0].$$

3.3 Proof of Theorem 1.1

We follow the proof presented in [12, Theorem 1.1]. There are two minor technical issues in the original proof that deserve clarification. First, in [12], it is assumed that the solution of the stochastic 3D primitive equations has the regularity

$$U \in L^2\left(\Omega; C\left([0,T], V\right)\right) \cap L^2\left(\Omega; L^2\left(0,T; D(A)\right)\right) \text{ for all } T \geq 0,$$

but the stronger integrability in Ω compared to Definition 2.2 has not been established. This regularity allows the authors to prove the convergence

$$Z^\varepsilon \to 0 \quad \text{in} \quad L^2\left(\Omega; C\left([0,T], V\right) \cap L^2\left(0,T; D(A)\right)\right),$$

see (3.13) for the definition of Z^ε. We can resolve the issue by establishing the discussed convergence in probability, see (3.15). Second, after using the Skorokhod theorem, the existence of the solution of the equation for $X_{\tilde{h}_\varepsilon}$, in the notation used here denoted by R^ε, see (3.16), is not discussed. As we will see in Sect. 4.2, if we were to establish the existence using the decomposition into the barotropic and baroclinic mode and similar estimates in [6], we would need additional regularity of Z^ε which we do not have. Here, the existence of R^ε is immediate by definition. After using the Skorokhod theorem, the same problem, arising for $\tilde{R}^\varepsilon$, is resolved by the Bensoussan trick from Bensoussan [1].

By Theorem 2.6, it suffices to check that the conditions specified in (2.31) and (2.32) are satisfied. The second condition (2.32) holds by Proposition 3.3. Regarding the first condition, let $M > 0$ and let $h_\varepsilon \in \mathcal{A}_M$ be such that $h_\varepsilon \to h$ as $\mathcal{T}_M$-valued random variables in law. Let

$$Y^\varepsilon = \mathcal{G}^\varepsilon \left(W(\cdot) + \frac{1}{\sqrt{\varepsilon}} \int_0^\cdot h_\varepsilon(s)\,\mathrm{d}s \right).$$

Our goal is to show that $\mathrm{law}(Y^\varepsilon) \to \mathrm{law}(U_h)$, where U_h is the solution of (3.2). By the Girsanov theorem, Y^ε solves

$$dY^\varepsilon + \left[AY^\varepsilon + B(Y^\varepsilon) + F(Y^\varepsilon)\right] dt = \sqrt{\varepsilon}\sigma(Y^\varepsilon)\,dW + \sigma(Y^\varepsilon) h_\varepsilon\,dt, \quad Y^\varepsilon(0) = u_0.$$

Let $\tau_K^{Y,\varepsilon}$ be the stopping time defined for $K > 0$ by

$$\tau_K^{Y,\varepsilon} = \inf \left\{ t \geq 0 \mid \int_0^t \|Y^\varepsilon\|_{H^2}^2\,\mathrm{d}s \geq K \right\}.$$

An argument similar to the one in [2, Section 4] or Sect. 4.2 in this manuscript shows that

$$\tau_K^{Y,\varepsilon} \to \infty \text{ as } K \to \infty \text{ in probability uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0]. \tag{3.12}$$

Let Z^ε be the solution of

$$dZ^\varepsilon + AZ^\varepsilon\,dt = \sqrt{\varepsilon}\sigma(Y^\varepsilon)\,dW, \quad Z^\varepsilon(0) = 0. \tag{3.13}$$

By the Itô lemma, see, e.g. [2, Theorem A.1], and the bound (2.15) on σ in $L_2(\mathcal{U}, V)$, we get

$$\mathbb{E}\left[\sup_{t\in[0,T\wedge\tau_K^{U,\varepsilon}]} \|Z^\varepsilon\|^2 + \int_0^{T\wedge\tau_K^{U,\varepsilon}} |AZ^\varepsilon|^2\,\mathrm{d}t \right] \leq C_{T,M}\varepsilon\mathbb{E}\left[\int_0^{T\wedge\tau_K^{U,\varepsilon}} 1 + |AY^\varepsilon|^2\,\mathrm{d}t \right]. \tag{3.14}$$

The convergence (3.12) implies

$$Z^\varepsilon \in C([0,T]; V) \cap L^2(0,T; D(A)), \quad \mathbb{P}\text{-a.s. for all } \varepsilon \in (0, \varepsilon_0] \text{ and } K \geq 0.$$

However, we do not know whether $Z^\varepsilon \in L^2(\Omega; C([0,T], V) \cap L^2(0,T; D(A)))$. Using the estimate (3.14) and the bounds on AY^ε from Proposition 3.4 we may employ an argument similar to the one in the proof of the uniform stochastic Gronwall lemma in Proposition A.1 to show that $Z^\varepsilon \to 0$ in X in probability as $\varepsilon \to 0+$, that is

$$\lim_{\varepsilon \to 0+} \mathbb{P}\left(\left\{\sup_{t\in[0,T]} \|Z^\varepsilon\|^2 + \int_0^T |AZ^\varepsilon|^2 \, \mathrm{d}t \ge \eta\right\}\right) = 0 \quad \text{for all } \eta > 0. \tag{3.15}$$

Let us also define $R^\varepsilon = Y^\varepsilon - Z^\varepsilon$, that is R^ε solves

$$dR^\varepsilon + \left[AR^\varepsilon + B(R^\varepsilon + Z^\varepsilon) + F(R^\varepsilon + Z^\varepsilon)\right] dt = \sigma(R^\varepsilon + Z^\varepsilon)h_\varepsilon \, dt, \quad R^\varepsilon(0) = u. \tag{3.16}$$

Since $h_\varepsilon \to h$ in law on the space $\mathcal{T}_m$ and $Z_\varepsilon \to 0$ in X in probability by (3.15), one can use the Portmanteau lemma in Polish spaces, see, e.g. [14, Theorem 18.2.6], to establish a version of van der Vaart [46, Theorem 2.17(v)]. In particular, the convergences above imply $(h_\varepsilon, Z_\varepsilon) \to (h, 0)$ in law on the space $\mathcal{T}_M \times X$. By the Skorokhod representation theorem, see, e.g. [8, Theorem 2.4], there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ and $\mathcal{T}_M \times \mathcal{T}_M \times X \times C([0,T], \mathcal{U}_0) \times C([0,T], \mathcal{U}_0)$-valued random variables $(\tilde{h}_\varepsilon, \tilde{h}, \tilde{Z}_\varepsilon, \tilde{W}_\varepsilon, \tilde{W})$ such that $\tilde{W}_\varepsilon$ and $\tilde{W}$ are cylindrical Wiener processes with reproducing kernel Hilbert space $\mathcal{U}$ and

$$\begin{aligned} \text{law}\,(h_\varepsilon, Z_\varepsilon, W) &= \text{law}\left(\tilde{h}_\varepsilon, \tilde{Z}_\varepsilon, \tilde{W}_\varepsilon\right), \quad \left(\tilde{Z}_\varepsilon, \tilde{h}_\varepsilon, \tilde{W}_\varepsilon\right) \\ &\to \left(0, \tilde{h}, \tilde{W}\right) \ \mathbb{P}\text{-a.s. in } \left(X, \mathcal{T}_M, C\left([0,T], \mathcal{U}_0\right)\right). \end{aligned} \tag{3.17}$$

Let

$$\tilde{Y}^\varepsilon = \mathcal{G}^\varepsilon\left(\tilde{W}(\cdot) + \frac{1}{\sqrt{\varepsilon}} \int_0^\cdot \tilde{h}_\varepsilon(s)\, \mathrm{d}s\right).$$

Using the Girsanov theorem, we observe that $\tilde{Y}^\varepsilon$ solves

$$d\tilde{Y}^\varepsilon + \left[A\tilde{Y}^\varepsilon + B(\tilde{Y}^\varepsilon) + F(\tilde{Y}^\varepsilon)\right] dt = \sqrt{\varepsilon}\sigma(\tilde{Y}^\varepsilon)\, d\tilde{W}_\varepsilon + \sigma(\tilde{Y}^\varepsilon)\tilde{h}_\varepsilon \, dt, \quad \tilde{Y}^\varepsilon(0) = u.$$

By [41, Theorem 1.6], we have

$$\text{law}\left(Y^\varepsilon, W\right) = \text{law}\left(\tilde{Y}^\varepsilon, \tilde{W}_\varepsilon\right). \tag{3.18}$$

The equality of laws (3.18) allows us to employ a stopped version of the Bensoussan trick to show that $\tilde{Z}^\varepsilon$ solves

$$d\tilde{Z}^\varepsilon + A\tilde{Z}^\varepsilon \, dt = \sqrt{\varepsilon}\sigma(\tilde{Y}^\varepsilon)\, d\tilde{W}_\varepsilon, \qquad \tilde{Z}^\varepsilon(0) = 0.$$

Indeed, defining $\tilde{\mathcal{F}}_t^\varepsilon = \sigma(\mathbb{1}_{[0,t]}\tilde{W}_\varepsilon, \mathbb{1}_{[0,t]}\tilde{Y}^\varepsilon, \mathbb{1}_{[0,t]}\tilde{Z}^\varepsilon)$, $\tilde{\mathcal{F}}_t = \sigma(\mathbb{1}_{[0,t]}\tilde{W}, \mathbb{1}_{[0,t]}\tilde{Y})$, $\tilde{\mathbb{F}}^\varepsilon = (\tilde{\mathcal{F}}_t^\varepsilon)_{t\geq 0}$ and $\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_t)_{t\geq 0}$, we observe that $\tilde{W}_\varepsilon$, resp. $\tilde{W}$, is a cylindrical $\tilde{\mathbb{F}}^\varepsilon$-Wiener, resp. $\tilde{\mathbb{F}}$-Wiener, process. Let

$$\tilde{\tau}_K^{Y,\varepsilon} = \inf\left\{t \geq 0 \mid \int_0^t \|\tilde{Y}^\varepsilon\|_{H^2}^2\, \mathrm{d}s \geq K\right\}, \quad K \geq 0, \varepsilon \in (0, \varepsilon_0]. \tag{3.19}$$

Then, since the map that maps a single trajectory in $L^2(0, T; H^2)$ to the hitting time defined in (3.19) is Borel measurable, $\tilde{\tau}_K^{Y,\varepsilon}$ is a random variable. By the debut theorem [11, Theorem 50, p. 116] $\tilde{\tau}_K^{Y,\varepsilon}$ is a $\tilde{\mathbb{F}}^\varepsilon$-stopping time. It remains to apply the Bensoussan trick from [1, Section 4.3.4] to the equation for $Z_K^\varepsilon = Z^\varepsilon(\cdot \wedge \tau_K^{Y,\varepsilon})$

$$dZ_K^\varepsilon + \mathbb{1}_{[0,\tau_K^{Y,\varepsilon}]} A Z_K^\varepsilon\, dt = \mathbb{1}_{[0,\tau_K^{Y,\varepsilon}]}\sqrt{\varepsilon}\sigma(Y^\varepsilon)\, dW, \quad Z^\varepsilon(0) = 0,$$

to show that $\tilde{Z}_K^\varepsilon = \tilde{Z}^\varepsilon(\cdot \wedge \tilde{\tau}_K^{Y,\varepsilon})$ is a solution of

$$d\tilde{Z}_K^\varepsilon + \mathbb{1}_{[0,\tilde{\tau}_K^{Y,\varepsilon}]} A \tilde{Z}_K^\varepsilon\, dt = \mathbb{1}_{[0,\tilde{\tau}_K^{Y,\varepsilon}]}\sqrt{\varepsilon}\sigma(\tilde{Y}^\varepsilon)\, d\tilde{W}_\varepsilon, \quad \tilde{Z}^\varepsilon(0) = 0,$$

and take the limit $K \to \infty$.

Defining $\tilde{R}^\varepsilon = \tilde{Y}^\varepsilon - \tilde{Z}^\varepsilon$, we observe that $\tilde{R}^\varepsilon$ satisfies

$$d\tilde{R}^\varepsilon + \left[A\tilde{R}^\varepsilon + B(\tilde{R}^\varepsilon + \tilde{Z}^\varepsilon) + F(\tilde{R}^\varepsilon + \tilde{Z}^\varepsilon)\right] dt = \sigma(\tilde{R}^\varepsilon + \tilde{Z}^\varepsilon)\tilde{h}_\varepsilon\, dt, \qquad \tilde{R}^\varepsilon(0) = u.$$

Referring to the equality in laws in (3.17) and (3.18), we obtain

$$\mathrm{law}(\tilde{R}^\varepsilon) = \mathrm{law}(R^\varepsilon). \tag{3.20}$$

By the $\tilde{\mathbb{P}}$-a.s. convergences in (3.17), we deduce that $\tilde{R}^\varepsilon \to \tilde{R}^0$ $\tilde{\mathbb{P}}$-almost surely, where $\tilde{R}^0$ satisfies

$$d\tilde{R}^0 + \left[A\tilde{R}^0 + B(\tilde{R}^0) + F(\tilde{R}^0)\right] dt = \sigma(\tilde{R}^0)\tilde{h}\, dt, \quad \tilde{R}^0(0) = u,$$

and therefore, recalling (3.17) again, we have

$$\mathrm{law}(\hat{R}^0) = \mathrm{law}(U_{\tilde{h}}) = \mathrm{law}(U_h). \tag{3.21}$$

We are now ready to finalize the proof. By (3.20), we have $\mathrm{law}(Y^\varepsilon - Z_\varepsilon) = \mathrm{law}(\tilde{R}^\varepsilon)$. Using the convergence $\tilde{R}^\varepsilon \to \tilde{R}^0$ $\tilde{\mathbb{P}}$-almost surely and (3.21), we have $\mathrm{law}(Y^\varepsilon - Z^\varepsilon) \to \mathrm{law}(U_h)$. Recalling $Z^\varepsilon \to 0$ in probability, we finally have $\mathrm{law}(Y^\varepsilon) \to \mathrm{law}(U_h)$. By Theorem 2.6, the LDP holds with the good rate function $I : C([0, T], V) \cap L^2(0, T; D(A)) \to \mathbb{R}$ given by

$$I(U) = \inf\left\{\frac{1}{2}\int_0^T |h|_{\mathcal{U}}^2\, \mathrm{d}t \mid h \in L^2(0, T; \mathcal{U}) \text{ s.t. } U = \mathcal{G}^0\left(\int_0^\cdot h\, \mathrm{d}t\right)\right\} \tag{3.22}$$

where $\mathcal{G}^0$ has been defined in (3.10). □

4 Moderate Deviations Principle

Let $u_0 \in V \cap H^2$ and let U^0 be the solution of the deterministic equation

$$dU^0 + \left[AU^0 + B\left(U^0\right) + F\left(U^0\right)\right] dt = 0, \quad U^0(0) = u_0. \tag{4.1}$$

Let R^ε be defined by

$$R^\varepsilon = \frac{U^\varepsilon - U^0}{\sqrt{\varepsilon}\lambda(\varepsilon)}, \quad \varepsilon \in (0, \varepsilon_0], \tag{4.2}$$

where U^ε is the solution of (3.1) with the same initial condition u_0. It is straightforward to check that R^ε is well defined and satisfies

$$\begin{aligned} &dR^\varepsilon + \left[AR^\varepsilon + B\left(R^\varepsilon, U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right) + B\left(U^0, R^\varepsilon\right) + A_{\mathrm{pr}}R^\varepsilon + ER^\varepsilon\right] dt \\ &= \lambda^{-1}(\varepsilon)\sigma\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right) dW, \quad R^\varepsilon(0) = 0. \end{aligned} \tag{4.3}$$

Similarly as for the LDP, we need to establish certain estimates on R^ε before we proceed to the proof of the main theorem. The process itself is more technically involved than the respective estimates for U^ε and seems to require additional regularity of the deterministic solution U^0. The additional regularity will be discussed in detail in Sect. 4.2.

4.1 Skeleton Equation

Let $h \in L^2(0, T; \mathcal{U})$. The *skeleton equation* corresponding to (4.3) is the equation

$$\begin{aligned} &dR_h + \left[AR_h + B\left(R_h, U^0\right) + B\left(U^0, R_h\right) + A_{\mathrm{pr}}R_h + ER_h\right] dt \\ &= \sigma\left(U^0\right) h\, dt, \quad R_h(0) = 0. \end{aligned} \tag{4.4}$$

Following a similar argument as in Sect. 4.1, we may establish the following equivalent of Proposition 3.3. The additional regularity of R_h is not required in this case since the potentially problematic term with σ does not play any role in the compactness argument.

Proposition 4.1 *Let $\{h_n\}_{n=1}^\infty \subseteq \mathcal{T}_M$ and let $R^n = R_{h_n}$ be the respective solutions of (4.4). Then, there exist $h \in L^2(0, T; \mathcal{U})$ and a subsequence (for simplicity not relabelled) such that $h_n \rightharpoonup h$ in $L^2(0, T; \mathcal{U})$ and $R^n \to R_h$ in $L^\infty(0, T; V) \cap L^2(0, T; D(A))$.*

Moreover, there exists a measurable map $\mathcal{G}_R^0 : C([0,T],\mathcal{U}_0) \to C([0,T],V) \cap L^2(0,T;D(A))$ *such that* $R_h = \mathcal{G}_R^0\Big(\int_0^\cdot h_t\,\mathrm{d}t\Big)$.

4.2 Preliminary Estimates

In the estimates in this subsection, we require additional regularity of the solution U^0 of the deterministic Eq. (4.1). Even though this is restrictive, it is not unusual to assume more regular initial data to obtain qualitative results on the primitive equations, see, e.g. the mathematical justification of the hydrostatic limit in [31]. To be more precise, it suffices to assume

$$U^0 \in C\left([0,T],H^1\right) \cap L^2\left(0,T;H^2\right) \cap L^\infty\left(0,T;L^\infty\right), \quad \partial_z U^0 \in L^\infty\left(0,T;H^1\right) \cap L^2\left(0,T;H^2\right).$$

Let $u_0 \in V \cap H^2$. By [5, Proposition 2.1], there exists a unique global solution of (4.1) satisfying

$$U^0 \in C\left([0,T],H^2\right) \cap L^2\left(0,T;H^3\right) \text{ for all } T \geq 0.$$

Let τ_K^0 be the stopping time defined by

$$\tau_K^0 = \inf\left\{t \geq 0 \mid \sup_{s\in[0,t]} \|U^0\| + \sup_{s\in[0,t]} |U^0|_{L^\infty} + \sup_{s\in[0,t]} \|\partial_z U^0\| + \int_0^t \|U^0\|_{H^2}^2 + \|\partial_z U^0\|_{H^2}^2\,\mathrm{d}s \geq K\right\}.$$

Since τ_K^0 is essentially deterministic and U^0 is a global solution, we observe

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^0 \leq t\right\}\right) = 0 \text{ for all } t \geq 0. \tag{4.5}$$

Proposition 4.2 *Let* $\tau_K^{w,\varepsilon}$ *be the stopping time defined for* $K \geq 0$ *and* $\varepsilon \in (0,\varepsilon_0]$ *by*

$$\tau_K^{w,\varepsilon} = \inf\left\{t \geq 0 \mid \int_0^t |R^\varepsilon|^4 + \|R^\varepsilon\|^2 + |R^\varepsilon|^4\|R^\varepsilon\|^2\,\mathrm{d}s \geq K\right\}.$$

Then, $\tau_K^{w,\varepsilon} \to \infty$ $\mathbb{P}$*-almost surely as* $K \to \infty$ *for all* $\varepsilon \in (0,\varepsilon_0]$ *and, for all* $t \geq 0$, *one has*

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{w,\varepsilon} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0,\varepsilon_0].$$

Proof Let $p \geq 2$. Applying the Itô lemma [2, Theorem A.1], we obtain

$$
\begin{aligned}
& d|R^\varepsilon|^p + p|R^\varepsilon|^{p-2}\left[\|R^\varepsilon\|^2 + \left(B\left(R^\varepsilon, U^0\right), R^\varepsilon\right) + \left(A_{\rm pr}R^\varepsilon + ER^\varepsilon, R^\varepsilon\right)\right] dt \\
& \quad \leq p\lambda^{-1}(\varepsilon)|R^\varepsilon|^{p-2}\left(\sigma(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\, dW, R^\varepsilon\right) + p|R^\varepsilon|^p\, dt \\
& \qquad + \frac{p(p-1)}{2}\lambda^{-2}(\varepsilon)|R^\varepsilon|^{p-2}\|\sigma(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\|^2_{L_2(\mathcal{U},H)}\, dt.
\end{aligned}
$$

By the estimate (2.8) on B, we have

$$
\begin{aligned}
p|R^\varepsilon|^{p-2}\left|\left(B\left(R^\varepsilon, U^0\right), R^\varepsilon\right)\right| & \leq Cp|R^\varepsilon|^{p-2}\|R^\varepsilon\|\|U^0\|^{1/2}\|U^0\|^{1/2}_{H^2}|R^\varepsilon|^{1/2}\|R^\varepsilon\|^{1/2} \\
& \leq \eta|R^\varepsilon|^{p-2}\|R^\varepsilon\|^2 + C_\eta|R^\varepsilon|^p\|U^0\|^2\|U^0\|^2_{H^2}
\end{aligned}
$$

for some $\eta > 0$ precisely determined later. Using the boundedness of the operators $A_{\rm pr}$ and E, we deduce

$$
p|R^\varepsilon|^{p-2}\left|\left(A_{\rm pr}R^\varepsilon + ER^\varepsilon, R^\varepsilon\right)\right| \leq C|R^\varepsilon|^{p-1}\|R^\varepsilon\| \leq \eta|R^\varepsilon|^{p-2}\|R^\varepsilon\|^2 + C_\eta|R^\varepsilon|^p.
$$

By the bound (2.14) on σ in $L_2(\mathcal{U}, H)$, we have

$$
\begin{aligned}
& \frac{p(p-1)}{2}\lambda^{-2}(\varepsilon)|R^\varepsilon|^{p-2}\|\sigma(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\|^2_{L_2(\mathcal{U},H)} \\
& \quad \leq C\lambda^{-2}(\varepsilon)|R^\varepsilon|^{p-2}\left(1 + |U^0|^2 + \|U^0\|^2\right) \\
& \qquad + C\varepsilon|R^\varepsilon|^p + p(p-1)\varepsilon\eta_0|R^\varepsilon|^{p-2}|\nabla_3 R^\varepsilon|^2.
\end{aligned}
$$

Let $K \geq 0$, $N \in \mathbb{N}$ and let τ_a and τ_b be stopping times such that $0 \leq \tau_a \leq \tau_b \leq T \wedge \tau_K^0$. We estimate the stochastic integral using the Burkholder–Davis–Gundy inequality (2.28) and the bound (2.14) on σ in $L_2(\mathcal{U}, H)$ as

$$
\begin{aligned}
& p\lambda^{-1}(\varepsilon)\mathbb{E}\sup_{t\in[\tau_a,\tau_b]}\left|\int_{\tau_a}^t |R^\varepsilon|^{p-2}\left(\sigma(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\,\mathrm{d}W, R^\varepsilon\right)\right| \\
& \quad \leq C_{BDG}\, p\lambda^{-1}\mathbb{E}\left(\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{2p-2}\|\sigma(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\|^2_{L_2(\mathcal{U},H)}\,\mathrm{d}s\right)^{1/2} \\
& \quad \leq C\lambda^{-1}(\varepsilon)\mathbb{E}\left(\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{2p-2}\left(1 + \|U^0\|^2 + \varepsilon\lambda^2(\varepsilon)|R^\varepsilon|^2\right)\mathrm{d}s\right)^{1/2} \\
& \qquad + C_{BDG}\, p\sqrt{2\varepsilon\eta_0}\,\mathbb{E}\left(\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{2p-2}\|R^\varepsilon\|^2\,\mathrm{d}s\right)^{1/2} \\
& \quad = I_1 + I_2.
\end{aligned}
$$

Using the Young inequality, we obtain

$$\begin{aligned} I_1 &\leq C\lambda^{-1}(\varepsilon)\mathbb{E}\left[\left(\sup_{s\in[\tau_a,\tau_b]}|R^\varepsilon|^{p/2}\right)\left(\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{p-2}\left(1\right.\right.\right. \\ &\quad \left.\left.\left.+\|U^0\|^2+\varepsilon\lambda^2(\varepsilon)|R^\varepsilon|^2\right)\,\mathrm{d}s\right)^{1/2}\right] \\ &\leq \eta\mathbb{E}\sup_{s\in[\tau_a,\tau_b]}|R^\varepsilon|^p + C_\eta\lambda^{-2}(\varepsilon)\mathbb{E}\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{p-2}\left(1+\|U^0\|^2\right)\mathrm{d}r + C_\eta\varepsilon\int_{\tau_a}^{\tau_b}|R^\varepsilon|^p\,\mathrm{d}s, \end{aligned} \tag{4.6}$$

$$\begin{aligned} I_2 &\leq C_{BDG}\,p\sqrt{2\varepsilon\eta_0}\,\mathbb{E}\left[\sup_{s\in[\tau_a,\tau_b]}|R^\varepsilon|^{p/2}\left(\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{p-2}\|R^\varepsilon\|^2\,\mathrm{d}s\right)^{1/2}\right] \\ &\leq (1-\delta)\mathbb{E}\sup_{s\in[\tau_a,\tau_b]}|R^\varepsilon|^p + \frac{C_{BDG}^2p^2\varepsilon\eta_0}{2(1-\delta)}\mathbb{E}\int_{\tau_a}^{\tau_b}|R^\varepsilon|^{p-2}\|R^\varepsilon\|^2\,\mathrm{d}s, \end{aligned} \tag{4.7}$$

for some $\delta\in(0,1)$. Collecting the estimates above, choosing δ and η sufficiently small and assuming that η_0 or ε_0 are sufficiently small, we observe

$$\mathbb{E}\left[\sup_{s\in[\tau_a,\tau_b]}|R^\varepsilon|^p + \int_{\tau_a}^{\tau_b}|R^\varepsilon|^{p-2}\|R^\varepsilon\|^2\,\mathrm{d}s\right] \leq C\mathbb{E}\int_{\tau_a}^{\tau_b}\left(1+|R^\varepsilon|^p\right)\left(1+\|U^0\|^2\right)\mathrm{d}s$$

with the constant C independent of ε. The claim then follows by the means of the uniform stochastic Gronwall lemma from Proposition A.1. □

In the rest of this Section, let us denote the velocity and temperature part of R^ε by $R^\varepsilon=(\nu^\varepsilon,\Upsilon^\varepsilon)$. Following the argument by Cao and Titi [6], we decompose the velocity part ν^ε into its barotropic and baroclinic modes $\nu^\varepsilon=\bar\nu^\varepsilon+\tilde\nu^\varepsilon$, where

$$\bar\nu^\varepsilon=\mathcal{A}_2\nu^\varepsilon,\quad \tilde\nu^\varepsilon=\mathcal{R}\nu^\varepsilon.$$

Similarly as in [6], one can establish that the barotropic mode $\bar\nu^\varepsilon$ satisfies the equation in $\mathcal{M}_0$

$$\begin{aligned} &d\bar\nu^\varepsilon+\Big[-\mu\Delta\bar\nu^\varepsilon+\sqrt{\varepsilon}\lambda(\varepsilon)\left[\left(\bar\nu^\varepsilon\cdot\nabla\right)\bar\nu^\varepsilon+\mathcal{J}_1(\nu^\varepsilon,\nu^\varepsilon)\right]+\left(\bar\nu^\varepsilon\cdot\nabla\right)\bar v^0+\mathcal{J}_1(\nu^\varepsilon,v^0) \\ &\quad+\left(\bar v^0\cdot\nabla\right)\bar\nu^\varepsilon+\mathcal{J}_1(v^0,\nu^\varepsilon)+\nabla p_S^\varepsilon-\beta_T g\nabla\mathcal{A}_2\int_z^0\Upsilon^\varepsilon\,\mathrm{d}z'+f\vec{k}\times\bar\nu^\varepsilon\Big]\mathrm{d}t \\ &=\lambda^{-1}(\varepsilon)\mathcal{A}_2\sigma_1(U^0+\sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\,\mathrm{d}W_1, \end{aligned} \tag{4.8}$$

where, for sufficiently regular u, v, we denote

$$\begin{aligned} B_1(u,v)&=(u\cdot\nabla)\,v, \\ \mathcal{J}_1(u,v)&=\mathcal{A}_2\left[(\tilde u\cdot\nabla)\,\tilde v+(\operatorname{div}\tilde u)\,\tilde v\right], \end{aligned}$$

with the periodic boundary conditions on $\mathcal{M}_0$, and

$$\operatorname{div}\bar{\mathrm{v}}^\varepsilon = 0, \qquad \bar{\mathrm{v}}^\varepsilon(0) = 0.$$

The baroclinic mode $\tilde{\mathrm{v}}^\varepsilon$ solves the equation in $\mathcal{M}$

$$\begin{aligned} d\tilde{\mathrm{v}}^\varepsilon + \Big[&- \mu\Delta\tilde{\mathrm{v}}^\varepsilon - \nu\partial_{zz}\tilde{\mathrm{v}}^\varepsilon + \sqrt{\varepsilon}\lambda(\varepsilon)\left[B_1(\tilde{\mathrm{v}}^\varepsilon, \tilde{\mathrm{v}}^\varepsilon) + \mathcal{J}_2(\mathrm{v}^\varepsilon, \mathrm{v}^\varepsilon)\right] + B_1(\tilde{\mathrm{v}}^\varepsilon, \tilde{v}^0) \\ &+ \mathcal{J}_2(\mathrm{v}^\varepsilon, v^0) + B_1(\tilde{v}^0, \tilde{\mathrm{v}}^\varepsilon) + \mathcal{J}_2(v^0, \mathrm{v}^\varepsilon) - \beta_T g\nabla\mathcal{R}\int_z^0 \Upsilon^\varepsilon\, dz' + f\vec{k}\times\tilde{\mathrm{v}}^\varepsilon\Big]\mathrm{d}t \\ = \lambda^{-1}&(\varepsilon)\mathcal{R}\sigma_1(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon)\, dW_1, \end{aligned} \tag{4.9}$$

where for sufficiently regular u, v we set

$$\begin{aligned} B_2(u, v) &= B_1(u, v) + w(u)\partial_z v = (u\cdot\nabla)\, v - \left(\int_{-h}^z \operatorname{div} u\, \mathrm{d}z'\right)\partial_z v, \\ \mathcal{J}_2(u, v) &= (\tilde{u}\cdot\nabla)\,\bar{v} - \mathcal{J}_1(u, v) = (\tilde{u}\cdot\nabla)\,\bar{v} - \mathcal{A}_3\left[(\tilde{u}\cdot\nabla)\,\tilde{v} + (\operatorname{div}\tilde{u})\,\tilde{v}\right] \end{aligned}$$

with the periodic boundary conditions on Γ_l and

$$\partial_z\tilde{\mathrm{v}}^\varepsilon = 0 \text{ on } \Gamma_i\cup\Gamma_b, \quad \tilde{\mathrm{v}}^\varepsilon(0) = 0.$$

Proposition 4.3 *Let $K \geq 0$ and let $\tau_K^{6,\varepsilon}$ be the stopping time defined by*

$$\tau_K^{6,\varepsilon} = \inf\left\{t \geq 0 \mid \int_0^t |\tilde{\mathrm{v}}^\varepsilon|_{L^6}^6 + \left(\int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 |\nabla_3\tilde{\mathrm{v}}^\varepsilon|^2\, \mathrm{d}\mathcal{M}\right) \mathrm{d}s \geq K\right\}. \tag{4.10}$$

Then, $\tau_K^{6,\varepsilon} \to \infty$ $\mathbb{P}$-a.s. as $K \to \infty$ for all $\varepsilon \in (0, \varepsilon_0]$ and, for all $t \geq 0$, one has

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{6,\varepsilon} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0].$$

Proof We will need the following estimates from Cao and Titi [6, p 256]. There exists $C > 0$ such that, for all $f \in L^6$ satisfying $\nabla_3(f^3) \in L^2$, one has

$$|f|_{L_x^{12}L_z^6}^6 \leq C|f|_{L^6}^3 |\nabla_3 f^3|_{L^2} + |f|_{L^6}^6, \tag{4.11}$$

$$|f|_{L_x^8 L_z^4}^4 \leq C|f|_{L^6}^3 \|f\|, \tag{4.12}$$

$$|f|_{L_x^8 L_z^2}^2 \leq C|f|_{L^6}^{3/2} \|f\|^{1/2}. \tag{4.13}$$

The estimate (4.11) follows directly from (2.23) with $q = 4$ and the others can be established using the Minkowski inequality similarly as in Lemma 2.1.

Employing the Itô lemma [2, Theorem A.1], integration by parts and the cancellation property

$$\big(B_2(u, v), |v|^q v\big) = 0, \qquad u \in V_1, v \in V_1 \cap H^2,$$

where $q \geq 0$ and $|v|$ denotes the modulus of v, we have

$$\begin{aligned}
& d|\tilde{\mathrm{v}}^\varepsilon|_{L^6}^6 + 6\,(\mu \wedge \nu) \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 |\nabla_3 \tilde{\mathrm{v}}^\varepsilon|^2 \,\mathrm{d}\mathcal{M}\,\mathrm{d}t \\
&\quad \leq 6 \left| \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 \tilde{\mathrm{v}}^\varepsilon \cdot \left(\beta_T g \nabla \mathcal{R} \int_z^0 \Upsilon^\varepsilon \, dz' + \sqrt{\varepsilon}\lambda(\varepsilon) \mathcal{J}_2(\mathrm{v}^\varepsilon, \mathrm{v}^\varepsilon) \right) \mathrm{d}\mathcal{M} \right| \mathrm{d}t \\
&\qquad + 6 \left| \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 \tilde{\mathrm{v}}^\varepsilon \cdot \left(B_2(\tilde{\mathrm{v}}^\varepsilon, \tilde{v}^0) + \mathcal{J}_2(\mathrm{v}^\varepsilon, v^0) + \mathcal{J}_2(v^0, \mathrm{v}^\varepsilon) \right) \mathrm{d}\mathcal{M} \right| \mathrm{d}t \\
&\qquad + 6\lambda^{-1} \sum_{k=1}^{\infty} \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 \tilde{\mathrm{v}}^\varepsilon \cdot \left(\mathcal{R}\sigma_1(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon) R^\varepsilon) e_k \right) d\mathcal{M}\, dW_1^k \\
&\qquad + 15\lambda^{-2}(\varepsilon) \sum_{k=1}^{\infty} \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 \left| \mathcal{R}\sigma_1(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon) R^\varepsilon) \right|^2 \mathrm{d}\mathcal{M}\,\mathrm{d}t \\
&\quad = I_1\, dt + I_2\, dt + \sum_{k=1}^{\infty} I_3^k \, dW_1^k + \sum_{k=1}^{\infty} I_4^k \, dt.
\end{aligned}$$

The estimates of I_1 follow from Cao and Titi [6, Section 3.2]. Estimates of the stochastic terms I_3^k and the Itô correction terms I_4^k are essentially the same as in [2, Proposition 4.3]. The remaining term I_2 seems to require more work. We use and slightly refine ideas from Debussche et al. [10, Proposition 4.3] with the anisotropic estimates from Guillén-González et al. [19].

Before we proceed to the estimates of the integrals above, let us make a short auxiliary estimate. Let $u, v, w \in D(A)$. Following the argument from Cao and Titi [6, Section 3.2], one may use integration by parts to establish

$$\begin{aligned}
& - \int_{\mathcal{M}} |\tilde{v}|^4 \tilde{v} \cdot \mathcal{J}_2(u, w) \,\mathrm{d}\mathcal{M} \\
&\quad = \int_{\mathcal{M}} (\operatorname{div} \tilde{u})\, |\tilde{v}|^4 \tilde{v} \cdot \bar{w} + (\tilde{u} \cdot \nabla) \left(|\tilde{v}|^4 \tilde{v} \right) \cdot \bar{w} \\
&\qquad - \partial_j \left(|\tilde{v}|^4 \tilde{v} \right) \overline{\tilde{u}_j \tilde{w}_k} \,\mathrm{d}\mathcal{M},
\end{aligned}$$

which in turn leads to the estimate

$$\begin{aligned}
& \left| \int_{\mathcal{M}} |\tilde{v}|^4 \tilde{v} \cdot \mathcal{J}_2(u, w) \,\mathrm{d}\mathcal{M} \right| \\
&\quad \leq \int_{\mathcal{M}_0} |\bar{w}| \left(\int_{-h}^0 |\nabla \tilde{u}| |\tilde{v}|^5 \,\mathrm{d}z \right) d\mathcal{M}_0 + \int_{\mathcal{M}_0} |\bar{w}| \left(\int_{-h}^0 |\tilde{u}| |\nabla \tilde{v}| |\tilde{v}|^4 \,\mathrm{d}z \right) \mathrm{d}\mathcal{M}_0
\end{aligned}$$

$$
\begin{aligned}
&+ \int_{\mathcal{M}_0} \left(\int_{-h}^{0} |\tilde{u}||\tilde{w}| \, \mathrm{d}z \right) \left(\int_{-h}^{0} |\nabla \tilde{v}||\tilde{v}|^4 \, \mathrm{d}z \right) d\mathcal{M}_0 \\
&= \mathcal{J}_2^1(v,u,w) + \mathcal{J}_2^2(v,u,w) + \mathcal{J}_2^3(v,u,w).
\end{aligned} \tag{4.14}
$$

We split the integral I_1 into I_1^1 and I_1^2 as

$$
\begin{aligned}
I_1 &\leq 6\beta_T g \left| \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 \tilde{\mathrm{v}}^\varepsilon \cdot \nabla \mathcal{R} \left(\int_z^0 \Upsilon^\varepsilon \, \mathrm{d}z' \right) \mathrm{d}\mathcal{M} \right| \\
&\quad + 6\sqrt{\varepsilon}\lambda(\varepsilon) \left| \int_{\mathcal{M}} |\tilde{\mathrm{v}}^\varepsilon|^4 \tilde{\mathrm{v}}^\varepsilon \cdot \mathcal{J}_2(\mathrm{v}^\varepsilon, \mathrm{v}^\varepsilon) \, d\mathcal{M} \right| \\
&= I_1^1 + I_1^2.
\end{aligned}
$$

Regarding I_1^1, we use integration by parts, the Ladyzhenskaya inequality and the estimate (4.12) to obtain

$$
\begin{aligned}
I_1^1 &= 6\beta_T g \left| \int_{\mathcal{M}} \operatorname{div} \left(|\tilde{\mathrm{v}}^\varepsilon|^4 \tilde{\mathrm{v}}^\varepsilon \right) \mathcal{R} \left(\int_z^0 \Upsilon^\varepsilon \, \mathrm{d}z' \right) \mathrm{d}\mathcal{M} \right| \\
&\leq C \int_{\mathcal{M}_0} \overline{|\Upsilon^\varepsilon|} \left(\int_{-h}^0 |\nabla \tilde{\mathrm{v}}^\varepsilon||\tilde{\mathrm{v}}^\varepsilon|^4 \, \mathrm{d}z \right) \mathrm{d}\mathcal{M}_0 \\
&\leq C \int_{\mathcal{M}_0} \overline{|\Upsilon^\varepsilon|} |\nabla (\tilde{\mathrm{v}}^\varepsilon)^3|_{L_z^2} |\tilde{\mathrm{v}}^\varepsilon|_{L_z^4}^2 \, \mathrm{d}\mathcal{M}_0 \\
&\leq C |\overline{|\Upsilon^\varepsilon|}|_{L^4(\mathcal{M}_0)} |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3| |\tilde{\mathrm{v}}^\varepsilon|_{L_x^8 L_z^4}^2 \\
&\leq C \left(|\overline{|\Upsilon^\varepsilon|}|_{L^2(\mathcal{M}_0)}^{1/2} |\nabla \overline{|\Upsilon^\varepsilon|}|_{L^2(\mathcal{M}_0)}^{1/2} + |\overline{|\Upsilon^\varepsilon|}|_{L^2(\mathcal{M}_0)} \right) |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3| |\tilde{\mathrm{v}}^\varepsilon|_{L^6}^{3/2} \|\tilde{\mathrm{v}}^\varepsilon\|^{1/2} \\
&\leq \eta |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3|^2 + C_\eta |\tilde{\mathrm{v}}^\varepsilon|_{L^6}^6 \left(|R^\varepsilon|^2 \|R^\varepsilon\|^2 + |R^\varepsilon|^4 \right) + C_\eta \|R^\varepsilon\|^2.
\end{aligned} \tag{4.15}
$$

We split the integral I_1^2 using (4.14). The first term can be estimated by the estimate (4.11) and the Ladyzhenskaya inequality by

$$
\begin{aligned}
\mathcal{J}_2^1(\mathrm{v}^\varepsilon, \mathrm{v}^\varepsilon, \mathrm{v}^\varepsilon) &\leq \int_{\mathcal{M}_0} \overline{|\tilde{\mathrm{v}}^\varepsilon|} |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3|_{L_z^2} |\tilde{\mathrm{v}}^\varepsilon|_{L_z^6}^3 \, \mathrm{d}\mathcal{M}_0 \\
&\leq |\overline{|\tilde{\mathrm{v}}^\varepsilon|}|_{L^4(\mathcal{M}_0)} |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3| |\tilde{\mathrm{v}}^\varepsilon|_{L_x^{12} L_z^6}^3 \\
&\leq C \left(|\overline{|\tilde{\mathrm{v}}^\varepsilon|}|_{L^2(\mathcal{M}_0)}^{1/2} |\nabla \overline{|\tilde{\mathrm{v}}^\varepsilon|}|_{L^2(\mathcal{M}_0)}^{1/2} + |\overline{|\tilde{\mathrm{v}}^\varepsilon|}|_{L^2(\mathcal{M}_0)} \right) |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3| \\
&\quad \cdot \left(|\tilde{\mathrm{v}}^\varepsilon|_{L^6}^{3/2} |\nabla_3 (\tilde{\mathrm{v}}^\varepsilon)^3|^{1/2} + |\tilde{\mathrm{v}}^\varepsilon|_{L^6}^3 \right)
\end{aligned}
$$

and therefore, by the Young inequality, we get

$$
\begin{aligned}
&6\sqrt{\varepsilon}\lambda(\varepsilon)\mathcal{J}_2^1(v^\varepsilon, v^\varepsilon, v^\varepsilon)\\
&\quad\le \eta|\nabla_3(\tilde{v}^\varepsilon)^3|^2 + C_{\eta,\varepsilon}|\tilde{v}^\varepsilon|_{L^6}^6\left(|R^\varepsilon|^2\|R^\varepsilon\|^2 + |R^\varepsilon|\|R^\varepsilon\| + |R^\varepsilon|^4 + |R^\varepsilon|^2\right),
\end{aligned}
\tag{4.16}
$$

where for fixed η we have $C_{\eta,\varepsilon} \to 0$ as $\varepsilon \to 0+$. The second term in the splitting of I_1^2 is estimated in exactly the same way as the first one. Hence, we obtain

$$
\mathcal{J}_2^2(v^\varepsilon, v^\varepsilon, v^\varepsilon) \le \eta|\nabla_3(\tilde{v}^\varepsilon)^3|^2 + C_{\eta,\varepsilon}|\tilde{v}^\varepsilon|_{L^6}^6\left(|R^\varepsilon|^2\|R^\varepsilon\|^2 + |R^\varepsilon|\|R^\varepsilon\| + |R^\varepsilon|^4 + |R^\varepsilon|^2\right)
\tag{4.17}
$$

with $C_{\eta,\varepsilon}$ as above. The third term in the splitting of I_1^2 is bounded using the estimates (4.12) and (4.13) as

$$
\begin{aligned}
\mathcal{J}_2^3(v^\varepsilon, v^\varepsilon, v^\varepsilon) &\le \int_{\mathcal{M}_0} |\tilde{v}^\varepsilon|_{L_z^2}^2 |\nabla_3(\tilde{v}^\varepsilon)^3|_{L_z^2} |\tilde{v}^\varepsilon|_{L_4^2}^2 \, \mathrm{d}\mathcal{M}_0\\
&\le |\tilde{v}^\varepsilon|_{L_x^8 L_z^2}^2 |\nabla_3(\tilde{v}^\varepsilon)^3| |\tilde{v}^\varepsilon|_{L_x^8 L_z^4}^2\\
&\le C|\tilde{v}^\varepsilon|_{L^6}^{3/2}\|\tilde{v}^\varepsilon\|^{1/2}|\nabla_3(\tilde{v}^\varepsilon)^3||\tilde{v}^\varepsilon|_{L^6}^{3/2}\|\tilde{v}^\varepsilon\|^{1/2}.
\end{aligned}
$$

Employing the Young inequality, we get

$$
\begin{aligned}
&6\sqrt{\varepsilon}\lambda(\varepsilon)\mathcal{J}_2^1(v^\varepsilon, v^\varepsilon, v^\varepsilon)\\
&\quad\le \eta|\nabla_3(\tilde{v}^\varepsilon)^3|^2 + C_{\eta,\varepsilon}|\tilde{v}^\varepsilon|_{L^6}^6\|R^\varepsilon\|^2,
\end{aligned}
\tag{4.18}
$$

where for fixed η we have $C_{\eta,\varepsilon} \to 0$ as $\varepsilon \to 0+$. Collecting [(4.15)–(4.18)] and applying the Young inequality, we get

$$
I_1 \le \eta|\nabla_3(\tilde{v}^\varepsilon)^3|^2 + C_\eta|\tilde{v}^\varepsilon|_{L^6}^6\left(|R^\varepsilon|^2\|R^\varepsilon\|^2 + |R^\varepsilon|^4 + \|R^\varepsilon\|^2\right) + C_\eta\|R^\varepsilon\|^2.
\tag{4.19}
$$

Let us proceed to I_2. We split the integral as

$$
\begin{aligned}
I_2 &\le 6\left|\int_{\mathcal{M}} |\tilde{v}^\varepsilon|^4\tilde{v}^\varepsilon \cdot B_2(\tilde{v}^\varepsilon, \tilde{v}^0)\, d\mathcal{M}\right| + 6\left|\int_{\mathcal{M}} |\tilde{v}^\varepsilon|^4\tilde{v}^\varepsilon \cdot \mathcal{J}_2(v^\varepsilon, v^0)\,\mathrm{d}\mathcal{M}\right|\\
&\quad + 6\left|\int_{\mathcal{M}} |\tilde{v}^\varepsilon|^4\tilde{v}^\varepsilon \cdot \mathcal{J}_2(v^0, v^\varepsilon)\,\mathrm{d}\mathcal{M}\right|\\
&= I_2^1 + I_2^2 + I_2^3.
\end{aligned}
$$

From the definition of B_2, we have

$$I_2^1 \le 6\left|\int_{\mathcal{M}} |\tilde{v}^\varepsilon|^4 \tilde{v}^\varepsilon \cdot \left[\left(\tilde{v}^\varepsilon \cdot \nabla\right) v^0\right] \mathrm{d}\mathcal{M}\right| \\ + 6\left|\int_{\mathcal{M}} |\tilde{v}^\varepsilon|^4 \tilde{v}^\varepsilon \cdot \left[w(\tilde{v}^\varepsilon)\partial_z v^0\right] \mathrm{d}\mathcal{M}\right| = I_2^{11} + I_2^{12}.$$

We estimate I_2^{11} using the Gagliardo–Nirenberg inequality as

$$\begin{aligned} I_2^{11} &\le C\int_{\mathcal{M}} |\nabla v^0||\tilde{v}^\varepsilon|^6 \,\mathrm{d}\mathcal{M} \le C|\nabla v^0|_{L^6}|\tilde{v}^\varepsilon|^6_{L^{36/5}} = C|\nabla v^0|_{L^6}|(\tilde{v}^\varepsilon)^3|^2_{L^{12/5}} \\ &\le C|\nabla v^0|_{L^6}\left(|\tilde{v}^\varepsilon|^{3/2}_{L^6}|\nabla_3(\tilde{v}^\varepsilon)^3|^{1/2} + |\tilde{v}^\varepsilon|^2_{L^6}\right) \\ &\le \eta|\nabla_3(\tilde{v}^\varepsilon)^3|^2 + C_\eta|\tilde{v}^\varepsilon|^2_{L^6}\left(\|v^0\|^{4/3}_{H^2} + \|v^0\|_{H^2}\right). \end{aligned} \tag{4.20}$$

Regarding I_2^{12}, we proceed similarly as in the proof of [10, Proposition 4.3] and use integration by parts to obtain

$$\begin{aligned} &-\int_{\mathcal{M}}\left(\int_{-h}^{z} \partial_j \tilde{v}^\varepsilon_j \,\mathrm{d}z'\right)\partial_z v^0_k (\tilde{v}^\varepsilon_k)^5 \,\mathrm{d}\mathcal{M} \\ &\quad= \int_{\mathcal{M}}\left(\int_{-h}^{z} \tilde{v}^\varepsilon_j \,\mathrm{d}z'\right)\partial_j\partial_z v^0_k(\tilde{v}^\varepsilon_k)^5 + 5\int_{\mathcal{M}}\left(\int_{-h}^{z} \tilde{v}^\varepsilon_j \,\mathrm{d}z'\right)\partial_z v^0_k \partial_j \tilde{v}^\varepsilon_k (\tilde{v}^\varepsilon_k)^4 \,\mathrm{d}\mathcal{M} \\ &\quad= I_2^{121} + I_2^{122}. \end{aligned}$$

By the anisotropic estimate (2.26) and the embedding $L^q \hookrightarrow L^q_x L^2_z$ with $q = 6$, we get

$$\begin{aligned} I_2^{121} &\le C\int_{\mathcal{M}_0}\left(\int_{-h}^{0} |\tilde{v}^\varepsilon_j| \,\mathrm{d}z'\right)\left(\int_{-h}^{0} |\partial_j\partial_z v^0_k||\tilde{v}^\varepsilon_k|^5 \,dz\right)\mathrm{d}\mathcal{M}_0 \\ &\le C\int_{\mathcal{M}_0} |\tilde{v}^\varepsilon_j|_{L^2_z}|\partial_j\partial_z v^0_k|_{L^2_z}|(\tilde{v}^\varepsilon_k)^5|_{L^2_z}\,\mathrm{d}\mathcal{M}_0 \le C|\tilde{v}^\varepsilon|_{L^6_x L^2_z}\|\partial_z v^0\||(\tilde{v}^\varepsilon)^5|_{L^3_x L^2_z} \\ &\le C|\tilde{v}^\varepsilon|^2_{L^6}\|\partial_z v^0\|\left(|\nabla_3(\tilde{v}^\varepsilon)^3|^{4/3} + |\tilde{v}^\varepsilon|^4_{L^6}\right) \\ &\le \eta|\nabla_3(\tilde{v}^\varepsilon)^3|^2 + C_\eta|\tilde{v}^\varepsilon|_{L^6}\left(\|\partial_z v^0\|^3 + \|\partial_z v^0\|\right). \end{aligned} \tag{4.21}$$

Employing the anisotropic estimates (2.23) and (2.27), we have

$$\begin{aligned} I_2^{122} &\le C\int_{\mathcal{M}_0}\left(\int_{-h}^{0} |\tilde{v}^\varepsilon_j| \,\mathrm{d}z'\right)\left(\int_{-h}^{0} |\partial_z v^0_k||(\tilde{v}^\varepsilon_k)^4\partial_j\tilde{v}^\varepsilon_k| \,\mathrm{d}z\right)\mathrm{d}\mathcal{M}_0 \\ &\le C\int_{\mathcal{M}_0} |\tilde{v}^\varepsilon_j|_{L^2_z}|\partial_z v^0_k|_{L^6_z}|(\tilde{v}^\varepsilon_k)^2\partial_j\tilde{v}^\varepsilon_k|_{L^2_z}|(\tilde{v}^\varepsilon_k)^2|_{L^3_z}\,\mathrm{d}\mathcal{M}_0 \\ &\le C|\tilde{v}^\varepsilon|_{L^{12}_x L^2_z}|\partial_z v^0|_{L^6}|\nabla_3(\tilde{v}^\varepsilon)^3||(\tilde{v}^\varepsilon)^2|_{L^4_x L^3_z} \end{aligned}$$

$$\leq C\left(|\tilde{\nu}^\varepsilon|_{L^6}^2\|\partial_z v^0\||\nabla_3(\tilde{\nu}^\varepsilon)^3|^{4/3}+|\tilde{\nu}^\varepsilon|_{L^6}^{5/2}\|\partial_z v^0\||\nabla_3(\tilde{\nu}^\varepsilon)^3|^{7/6}\right.$$
$$\left.+|\tilde{\nu}^\varepsilon|_{L^6}^3\|\partial_z v^0\||\nabla_3(\tilde{\nu}^\varepsilon)^3|\right)$$
$$\leq \eta|\nabla_3(\tilde{\nu}^\varepsilon)^3|^2+C_\eta|\tilde{\nu}^\varepsilon|_{L^6}^6\left(\|\partial_z v^0\|^3+\|\partial_z v^0\|^2\right)+C_\eta|\tilde{\nu}^\varepsilon|_{L^6}^{35/8}\|\partial_z v^0\|^{7/4}. \tag{4.22}$$

Collecting [(4.20)–(4.22)], we obtain

$$I_2^1\leq \eta|\nabla_3(\tilde{\nu}^\varepsilon)^3|^2+C_\eta\left(1+|\tilde{\nu}^\varepsilon|_{L^6}^6\right)\left(1+\|\partial_z v^0\|^3\right). \tag{4.23}$$

Continuing with I_2^2, we recall the estimate (4.14) and therefore

$$I_2^2\leq 6\left(\mathcal{J}_2^1(\nu^\varepsilon,\nu^\varepsilon,v^0)+\mathcal{J}_2^2(\nu^\varepsilon,\nu^\varepsilon,v^0)+\mathcal{J}_2^3(\nu^\varepsilon,\nu^\varepsilon,v^0)\right).$$

By the Hölder inequality, the Ladyzhenskaya inequality, the anisotropic estimate (4.11) and the Young inequality, we have

$$\begin{aligned}\mathcal{J}_2^1(\tilde{\nu}^\varepsilon,\tilde{\nu}^\varepsilon,v^0)&\leq C\int_{\mathcal{M}_0}|\bar{v}^0||\nabla_3(\tilde{\nu}^\varepsilon)^3|_{L_z^2}|\tilde{\nu}^\varepsilon|_{L_z^6}^3\,\mathrm{d}\mathcal{M}_0\\&\leq C|\bar{v}^0|_{L^4}|\nabla_3(\tilde{\nu}^\varepsilon)^3||v|_{L_x^{12}L_z^6}^3\\&\leq C|\bar{v}^0|_{L^2(\mathcal{M}_0)}^{1/2}\|\bar{v}^0\|_{H^1(\mathcal{M}_0)}^{1/2}|\nabla_3(\tilde{\nu}^\varepsilon)^3||\tilde{\nu}^\varepsilon|_{L^6}^{3/2}\|(\tilde{\nu}^\varepsilon)^3\|^{1/2}\\&\leq \eta|\nabla_3(\tilde{\nu}^\varepsilon)^3|^2+C_\eta|\tilde{\nu}^\varepsilon|_{L^6}^6\left(|U^0|^2\|U^0\|^2+|U^0|\|U^0\|\right).\end{aligned} \tag{4.24}$$

The next term can be estimated in exactly the same manner. Therefore, we get

$$\mathcal{J}_2^2(\nu^\varepsilon,\nu^\varepsilon,v^0)\leq \eta|\nabla_3(\tilde{\nu}^\varepsilon)^3|^2+C_\eta|\tilde{\nu}^\varepsilon|_{L^6}^6\left(|U^0|^2\|U^0\|^2+|U^0|\|U^0\|\right). \tag{4.25}$$

Regarding the final part of I_2^2, we employ the Hölder inequality, the anisotropic estimates (4.12) and (4.13) and the Sobolev embedding $W^{1,2}\hookrightarrow L^6$ and deduce

$$\begin{aligned}\mathcal{J}_2^3(\tilde{\nu}^\varepsilon,\tilde{\nu}^\varepsilon,v^0)&\leq \int_{\mathcal{M}_0}|\tilde{\nu}^\varepsilon|_{L_z^2}|\tilde{v}^0|_{L_z^2}|\nabla_3(\tilde{\nu}^\varepsilon)^3|_{L_z^2}|\tilde{\nu}^\varepsilon|_{L_z^4}^2\,\mathrm{d}\mathcal{M}_0\\&\leq |\tilde{\nu}^\varepsilon|_{L_x^8L_z^2}|\tilde{v}^0|_{L_x^8L_z^2}|\nabla_3(\tilde{\nu}^\varepsilon)^3||\tilde{\nu}^\varepsilon|_{L_x^8L_z^4}^2\\&\leq C\|U^0\|\|R^\varepsilon\|^{3/4}|\tilde{\nu}^\varepsilon|_{L^6}^{9/4}|\nabla_3(\tilde{\nu}^\varepsilon)^3|\\&\leq \eta|\nabla_3(\tilde{\nu}^\varepsilon)^3|^2+C_\eta|\tilde{\nu}^\varepsilon|_{L^6}^{9/2}\|U^0\|^2\|R^\varepsilon\|^{3/2}.\end{aligned} \tag{4.26}$$

Collecting [(4.24)–(4.26)], we obtain

$$I_2^2\leq \eta|\nabla_3(\tilde{\nu}^\varepsilon)^3|^2+C_\eta|\tilde{\nu}^\varepsilon|_{L^6}^6\left(1+|U^0|^2\|U^0\|^2+\|R^\varepsilon\|^2+\|U^0\|^8\right)$$

$$+ C_\eta \left(1 + \|R^\varepsilon\|^2 + \|U^0\|^8\right). \tag{4.27}$$

The remaining part of the term I_2, that is I_2^3, can be estimated using (4.14) as

$$I_2^3 \le 6\left(\mathcal{J}_2^1(\tilde{v}^\varepsilon, v^0, \tilde{v}^\varepsilon) + \mathcal{J}_2^2(\tilde{v}^\varepsilon, v^0, \tilde{v}^\varepsilon) + \mathcal{J}_2^3(\tilde{v}^\varepsilon, v^0, \tilde{v}^\varepsilon)\right)$$

By the Hölder inequality, the Gagliardo–Nirenberg inequality both in 2D and 3D and the anisotropic estimate (2.23) for $\nabla_3 \tilde{v}^0$, we obtain

$$\begin{aligned} \mathcal{J}_2^1(v^\varepsilon, v^0, v^\varepsilon) &\le \int_{\mathcal{M}_0} |\bar{v}^\varepsilon| |\nabla_3 \tilde{v}^0|_{L^2_z} |\tilde{v}^\varepsilon|_{L^{10}_z} \, \mathrm{d}\mathcal{M}_0 \le |\bar{v}^\varepsilon|_{L^3(\mathcal{M}_0)} |\nabla_3 \tilde{v}^0|_{L^6_x L^2_z} |\tilde{v}^\varepsilon|^3_{L^{10}} \\ &\le C|R^\varepsilon|^{2/3} \|R^\varepsilon\|^{1/3} \|U^0\|^{1/3} \|U^0\|^{2/3}_{H^2} |\tilde{v}^\varepsilon|^2_{L^6} \|(\tilde{v}^\varepsilon)^3\| \\ &\le \eta |\nabla_3 (\tilde{v}^\varepsilon)^3|^2 + C_\eta |\tilde{v}^\varepsilon|^4_{L^6} \left(|R^\varepsilon|^4 \|R^\varepsilon\|^2 + \|U^0\| \|U^0\|^2_{H^2}\right) \\ &\quad + C|\tilde{v}^\varepsilon|^5_{L^6} \left(|R^\varepsilon| \|R^\varepsilon\|^{1/2} + \|U^0\| \|U^0\|^2_{H^2}\right). \end{aligned} \tag{4.28}$$

Next, we employ the Hölder inequality, the Gagliardo–Nierenberg inequality in 2D and the anisotropic estimate (4.11) and get

$$\begin{aligned} \mathcal{J}_2^2(v^\varepsilon, v^0, v^\varepsilon) &\le C \int_{\mathcal{M}_0} |\bar{v}^\varepsilon| |\nabla_3 (\tilde{v}^\varepsilon)^3|_{L^2_z} |\tilde{v}^\varepsilon|_{L^6_z} |\tilde{v}^0|_{L^6_z} \, \mathrm{d}\mathcal{M}_0 \\ &\le C |\bar{v}^\varepsilon|_{L^6(\mathcal{M}_0)} |\nabla_3 (\tilde{v}^\varepsilon)^3| |\tilde{v}^\varepsilon|^2_{L^{12}_x L^6_z} |\tilde{v}^0|_{L^6} \\ &\le C|R^\varepsilon|^{1/3} \|R^\varepsilon\|^{2/3} |\nabla_3 (\tilde{v}^\varepsilon)^3| |\tilde{v}^\varepsilon|_{L^6} \|(\tilde{v}^\varepsilon)^3\|^{1/3} \|U^0\| \\ &\le \eta |\nabla_3 (\tilde{v}^\varepsilon)^3|^2 + C_\eta |\tilde{v}^\varepsilon|^3_{L^6} |R^\varepsilon| \|R^\varepsilon\|^2 \|U^0\|^3 \\ &\quad + C_\eta |\tilde{v}^\varepsilon|^4_{L^6} |R^\varepsilon|^{2/3} \|R^\varepsilon\|^{4/3} \|U^0\|^2. \end{aligned} \tag{4.29}$$

The final part of I_2^3 is estimated in exactly the same way as $\mathcal{J}_2^3(\tilde{v}^\varepsilon, \tilde{v}^\varepsilon, v^0)$. Therefore, we have

$$\mathcal{J}_2^3(v^\varepsilon, v^0, v^\varepsilon) \le \eta |\nabla_3 (\tilde{v}^\varepsilon)^3|^2 + C_\eta |\tilde{v}^\varepsilon|^{9/2}_{L^6} \|U^0\|^2 \|R^\varepsilon\|^{3/2}. \tag{4.30}$$

Collecting [(4.28)–(4.30)] and using the Young inequality, we deduce

$$\begin{aligned} I_2^3 &\le \eta |\nabla_3 (\tilde{v}^\varepsilon)^3|^2 + C_\eta \left(1 + |\tilde{v}^\varepsilon|^6_{L^6}\right) \left(1 + |R^\varepsilon|^4 \|R^\varepsilon\|^2 + \|U^0\| \|U^0\|^2_{H^2}\right) \\ &\quad + C_\eta \left(1 + |\tilde{v}^\varepsilon|^6_{L^6}\right) \left(|R^\varepsilon| \|R^\varepsilon\|^2 \|U^0\|^3 + \|R^\varepsilon\|^2 + \|U^0\|^6\right) \end{aligned} \tag{4.31}$$

This finishes the estimates of I_2.

Let $K \ge 0$ and let τ_a and τ_b be stopping times such that $0 \le \tau_a \le \tau_b \le T \wedge \tau^w_{K,\varepsilon} \wedge \tau^0_K$. Using the Burkholder–Davis–Gundy inequality (2.28) and the bound (2.18) on

$\mathcal{R}\sigma_1$ in $L_2(\mathcal{U}, L^6)$, we get

$$
\begin{aligned}
&\mathbb{E} \sup_{t\in[\tau_a,\tau_b]} \left| \int_{\tau_a}^{t} \sum_{k=1}^{\infty} I_3^k \, \mathrm{d}W_1^k \right| \\
&\leq 6\lambda^{-1}(\varepsilon) C_{BDG} \mathbb{E} \left(\int_{\tau_a}^{\tau_b} \sum_{k=1}^{\infty} \left[\int_{\mathcal{M}} |\tilde{v}^\varepsilon|^5 |\mathcal{R}\sigma_1 \left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon) R^\varepsilon \right) e_k| \, d\mathcal{M} \right] \mathrm{d}s \right)^{1/2} \\
&\leq C\lambda^{-1}(\varepsilon) \mathbb{E} \left(\int_{\tau_a}^{\tau_b} |\tilde{v}^\varepsilon|_{L^6}^{10} \sum_{k=1}^{\infty} |\mathcal{R}\sigma_1 \left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon) R^\varepsilon \right) e_k|_{L^6}^2 \, ds \right)^{1/2} \\
&\leq C\lambda^{-1}(\varepsilon) \mathbb{E} \left(\int_{\tau_a}^{\tau_b} |\tilde{v}^\varepsilon|_{L^6}^{10} \left(1 + |\tilde{v}^0|_{L^6}^2 + \varepsilon\lambda^{-2}(\varepsilon) |\tilde{v}^\varepsilon|_{L^6}^2 \right) \mathrm{d}s \right)^{1/2} \\
&\leq \eta \mathbb{E} \sup_{s\in[\tau_a,\tau_b]} |\tilde{v}^\varepsilon|_{L^6}^6 + C_\eta \lambda^{-2}(\varepsilon) \mathbb{E} \int_{\tau_a}^{\tau_b} \left(1 + |\tilde{v}^\varepsilon|_{L^6}^6 \right) \left(1 + \|U^0\|^2 \right) \mathrm{d}s. \qquad (4.32)
\end{aligned}
$$

Regarding the Itô correction term, from the bound (2.18) on $\mathcal{R}\sigma_1$ in $L_2(\mathcal{U}, L^6)$ we deduce

$$
\begin{aligned}
\sum_{k=1}^{\infty} I_4^k &\leq C\lambda^{-2}(\varepsilon) |\tilde{v}^\varepsilon|_{L^6}^4 \sum_{k=1}^{\infty} |\mathcal{R}\sigma_1 \left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon) R^\varepsilon \right) e_k|_{L^6}^2 \\
&\leq C\lambda^{-2}(\varepsilon) \left(1 + |\tilde{v}^\varepsilon|_{L^6}^6 \right) \left(1 + \|U^0\|^2 \right). \qquad (4.33)
\end{aligned}
$$

Finally, choosing $\eta > 0$ sufficiently small, we use the estimates (4.19), (4.23), (4.27), (4.31), (4.32) and (4.33) to infer

$$
\mathbb{E} \left[\sup_{s\in[\tau_a,\tau_b]} |\tilde{v}^\varepsilon|^6 + \int_{\tau_a}^{\tau_b} |\nabla_3 (\tilde{v}^\varepsilon)^3|^2 \, \mathrm{d}s \right] \leq C\mathbb{E} \int_{\tau_a}^{\tau_b} \left(1 + |\tilde{v}^\varepsilon|_{L^6}^6 \right) \Phi(s) \, \mathrm{d}s,
$$

where

$$
\Phi(s) = 1 + |R^\varepsilon|^4 \|R^\varepsilon\|^2 + |R^\varepsilon|^4 + \|R^\varepsilon\|^2 + \|\partial_z U^0\|^3 + \|U^0\|^8 + |R^\varepsilon| \|R^\varepsilon\|^2 \|U^0\|^3.
$$

We observe that the definitions of the stopping times τ_K^0 and $\tau_K^{w,\varepsilon}$ and the "uniform" convergence (4.5) justify the use of the uniform stochastic Gronwall lemma from Proposition A.1, which finishes the proof. □

Proposition 4.4 *Let $K \geq 0$ and let $\tau_K^{\nabla,\varepsilon}$ be the stopping time defined by*

$$
\tau_K^{\nabla,\varepsilon} = \inf \left\{ t \geq 0 \mid \int_0^t \|\tilde{v}^\varepsilon\|_{H^1(\mathcal{M}_0)}^4 \, \mathrm{d}s \geq K \right\}.
$$

Then, $\tau_K^{\nabla,\varepsilon} \to \infty$ $\mathbb{P}$*-almost surely as* $K \to \infty$ *for all* $\varepsilon \in (0, \varepsilon_0]$ *and, for all* $t \geq 0$*, one has*

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{\nabla,\varepsilon} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0].$$

Proof Applying the Itô lemma [2, Theorem A.1] to (4.8) with the function $|A_S^{1/2} P_{\bar{H}} \cdot|^4$ and using the cancellation property

$$\int_{\mathcal{M}_0} (u \cdot \nabla)\, v \cdot \Delta v \, \mathrm{d}\mathcal{M}_0 = 0$$

holding for $u \in H^2(\mathcal{M}_0)$ with $\operatorname{div} u = \operatorname{div} v = 0$ and periodic boundary conditions, we obtain

$$\begin{aligned}
& d|\nabla\bar{\mathrm{v}}^\varepsilon|^4 + 4\mu|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon|^2\, dt \\
&\quad \leq 4|\nabla\bar{\mathrm{v}}^\varepsilon|^2\Bigg[\beta_T g \left|\left(\Delta\bar{\mathrm{v}}^\varepsilon, P_{\bar{H}}\mathcal{A}_2 \int_z^0 \nabla\Upsilon^\varepsilon\right)\right| + \sqrt{\varepsilon}\lambda(\varepsilon)\left|\left(P_{\bar{H}}\mathcal{J}_1(\mathrm{v}^\varepsilon, \mathrm{v}^\varepsilon), \Delta\bar{\mathrm{v}}^\varepsilon\right)\right| \\
&\qquad + \left|\left(P_{\bar{H}}\mathcal{J}_1(v^0, \mathrm{v}^\varepsilon), \Delta\bar{\mathrm{v}}^\varepsilon\right)\right| + \left|\left(P_{\bar{H}}B_1(\bar{\mathrm{v}}^\varepsilon, v^0), \Delta\bar{\mathrm{v}}^\varepsilon\right)\right| + \left|\left(P_{\bar{H}}\mathcal{J}_1(\mathrm{v}^\varepsilon, v^0), \Delta\bar{\mathrm{v}}^\varepsilon\right)\right|\Bigg]\mathrm{d}t \\
&\qquad + 6\lambda^{-2}(\varepsilon)|\nabla\bar{\mathrm{v}}^\varepsilon|^2\|\nabla\mathcal{A}_2\sigma_1\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right)\|^2_{L_2(\mathcal{U}, L^2(\mathcal{M}_0))}\, dt \\
&\qquad + 4\lambda^{-1}(\varepsilon)|\nabla\bar{\mathrm{v}}^\varepsilon|^2\left(\nabla\bar{\mathrm{v}}^\varepsilon, \nabla\mathcal{A}_2\sigma_1\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right) dW_1\right) \\
&\quad = \sum_{k=1}^6 I_k\, dt + I_7\, dW_1.
\end{aligned}$$

The term I_1 can be estimated by the Hölder inequality. We have

$$I_1 \leq \eta|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon|^2 + C_\eta|\nabla\bar{\mathrm{v}}^\varepsilon|^2\|R^\varepsilon\|^2 \tag{4.34}$$

for some $\eta > 0$ fixed precisely determined later. Next, we argue as in [6, Section 3.3] and use the Hölder inequality to obtain

$$\begin{aligned}
I_2 &\leq C\sqrt{\varepsilon}\lambda(\varepsilon)|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon||\nabla_3(\tilde{\mathrm{v}}^\varepsilon)^3|^{1/2}\|R^\varepsilon\|^{1/2} \\
&\leq \eta|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon|^2 + C_\eta\varepsilon\lambda^2(\varepsilon)\left(|\nabla\bar{\mathrm{v}}^\varepsilon|^4|\nabla_3(\tilde{\mathrm{v}}^\varepsilon)^3|^2 + \|R^\varepsilon\|^2\right).
\end{aligned} \tag{4.35}$$

We deal with I_3 using the Hölder inequality and Ladyzhenskaya inequality. We get

$$\begin{aligned}
I_4 &\leq C|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon||(\bar{\mathrm{v}}^\varepsilon\cdot\nabla)v^0| \leq C|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon||\bar{\mathrm{v}}^\varepsilon|_{L^4}|\nabla\bar{v}^0|_{L^4} \\
&\leq C|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon||\bar{\mathrm{v}}^\varepsilon|^{1/2}\|\bar{\mathrm{v}}^\varepsilon\|^{1/2}|\nabla\bar{v}^0|^{1/2}\|\nabla\bar{v}^0\|^{1/2} \\
&\leq \eta|\nabla\bar{\mathrm{v}}^\varepsilon|^2|\Delta\bar{\mathrm{v}}^\varepsilon|^2 + C_\eta|\nabla\bar{\mathrm{v}}^\varepsilon|^2\left(|R^\varepsilon|^2\|R^\varepsilon\|^2 + \|U^0\|^2\|U^0\|^2_{H^2}\right).
\end{aligned} \tag{4.36}$$

The term I_4 can be estimated in exactly the same way as I_3, therefore

$$I_4 \le \eta |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon|^2 + C_\eta |\nabla \bar{v}^\varepsilon|^2 \left(|R^\varepsilon|^2 \|R^\varepsilon\|^2 + \|U^0\|^2 \|U^0\|^2_{H^2} \right) \tag{4.37}$$

Regarding I_5, we split the term by the Hölder inequality as follows:

$$I_5 \le C |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon| \left(|\mathcal{A}_2 \left[(\tilde{v}^\varepsilon \cdot \nabla)\tilde{v}^0 \right]| + |\mathcal{A}_2 \left[(\operatorname{div} \tilde{v}^\varepsilon)\tilde{v}^0 \right]| \right) = I_5^1 + I_5^2.$$

By the Hölder inequality, the Young inequality and the anisotropic estimate (2.23) applied to $\nabla \tilde{v}^0 \in H^1(\mathcal{M}; \mathbb{R}^2)$, we get

$$\begin{aligned} I_5^1 &\le C |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon| \left(\int_{\mathcal{M}_0} |\tilde{v}^\varepsilon|^2_{L^2_z} |\nabla \tilde{v}^0|^2_{L^2_z} \, d\mathcal{M}_0 \right)^{1/2} \le C |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon| |\tilde{v}^\varepsilon|_{L^6_x L^2_z} |\nabla \tilde{v}^0|_{L^3_x L^2_z} \\ &\le C |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon| |\tilde{v}^\varepsilon|_{L^6} |\nabla \tilde{v}^0|^{2/3} \|\nabla \tilde{v}^0\|^{1/3} \\ &\le \eta |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon|^2 + C_\eta |\nabla \bar{v}^\varepsilon|^2 \left(|\tilde{v}^\varepsilon|^4_{L^6} + \|\tilde{v}^0\|^{8/3} \|\tilde{v}^0\|^{4/3}_{H^2} \right). \end{aligned} \tag{4.38}$$

Regarding the term I_5^2, we use the boundedness of $\mathcal{R}$ in L^∞ to deduce

$$I_5^2 \le C |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon| |\tilde{v}^0|_{L^\infty} |\nabla \tilde{v}^\varepsilon| \le \eta |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon|^2 + C_\eta |\nabla \bar{v}^\varepsilon|^2 \|R^\varepsilon\|^2 |v^0|^2_{L^\infty}. \tag{4.39}$$

We estimate the Itô correction term using the estimate (2.20) and the Young inequality by

$$I_6 \le C \lambda^{-2}(\varepsilon) |\nabla \bar{v}^\varepsilon|^2 \left(1 + \|U^0\|^2 + |\Delta \bar{v}^0|^2 \right) + C\varepsilon |\nabla \bar{v}^\varepsilon|^2 \|R^\varepsilon\|^2 + C\varepsilon \eta_2 |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon|^2. \tag{4.40}$$

Let $K \ge 0$ and let τ_a and τ_b be stopping times such that $0 \le \tau_a \le \tau_b \le T \wedge \tau_K^{6,\varepsilon} \wedge \tau_K^{w,\varepsilon} \wedge \tau_K^0$. Using the Burkholder–Davis–Gundy inequality and the bound (2.20), we may estimate the stochastic term similarly as in (4.6) and (4.7) by

$$\begin{aligned} &4\lambda^{-1}(\varepsilon) \mathbb{E} \sup_{t \in [\tau_a, \tau_b]} \left| \int_{\tau_a}^t |\nabla \bar{v}^\varepsilon|^2 \left(\nabla \bar{v}^\varepsilon, \nabla \mathcal{A}_2 \sigma_1 \left(U^0 + \sqrt{\varepsilon} \lambda(\varepsilon) R^\varepsilon \right) dW_1 \right) \right| \\ &\quad \le (1 - \delta + \eta) \, \mathbb{E} \sup_{s \in [\tau_a, \tau_b]} + C_\eta \lambda^{-2}(\varepsilon) \mathbb{E} \int_{\tau_a}^{\tau_b} |\nabla \bar{v}^\varepsilon|^2 \left(1 + \|U^0\|^2 + \|U^0\|^2_{H^2} \right) ds \\ &\qquad + C_\eta \varepsilon \mathbb{E} \int_{\tau_a}^{\tau_b} |\nabla \bar{v}^\varepsilon|^4 \, ds + \frac{C_\eta \eta_2 \varepsilon}{1 - \delta} \mathbb{E} \int_{\tau_a}^{\tau_b} |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon|^2 \, ds \end{aligned} \tag{4.41}$$

for some $\delta \in [0, 1)$.

Collecting [(4.34)–(4.41)], choosing η and δ appropriately small and assuming that ε_0 is sufficiently small, we infer

$$\mathbb{E} \left[\sup_{s \in [\tau_a, \tau_b]} |\nabla \bar{v}^\varepsilon|^4 + \int_{\tau_a}^{\tau_b} |\nabla \bar{v}^\varepsilon|^2 |\Delta \bar{v}^\varepsilon|^2 \, ds \right] \le C \mathbb{E} \int_{\tau_a}^{\tau_b} \left(1 + |\nabla \bar{v}^\varepsilon|^4 \right) \Phi(s) \, ds,$$

where

$$\Phi(s) = 1 + \|R^\varepsilon\|^2 + |\nabla_3(\tilde{v}^\varepsilon)^3|^2 + |R^\varepsilon|^2\|R^\varepsilon\|^2 + \|U^0\|^4\|U^0\|_{H^2}^2 + |\tilde{v}^\varepsilon|_{L^6}^6 \\ + \|R^\varepsilon\|^2|v^0|_{L^\infty}^2 + \|U^0\|^2 + \|U^0\|_{H^2}^2.$$

The rest follow from the uniform stochastic Gronwall lemma, see Proposition A.1. □

Proposition 4.5 *Let $K \geq 0$ and let $\tau_K^{z,\varepsilon}$ be the stopping time defined by*

$$\tau_K^{z,\varepsilon} = \inf\left\{t \geq 0 \mid \int_0^t |\partial_z R^\varepsilon|^2\|\partial_z R^\varepsilon\|^2 \,\mathrm{d}s \geq K\right\}. \tag{4.42}$$

Then, $\tau_K^{z,\varepsilon} \to \infty$ $\mathbb{P}$-almost surely as $K \to \infty$ for all $\varepsilon \in (0, \varepsilon_0]$ and, for all $t \geq 0$, one has

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{z,\varepsilon} \leq t\right\}\right) = 0 \,\text{uniformly w.r.t.}\, \varepsilon \in (0, \varepsilon_0].$$

Proof The proof uses elements from the proof [10, Proposition 5.2]. Let $p \geq 2$. By the Itô lemma [2, Theorem A.1], integration by parts and the cancellation property (2.9), we get

$$\begin{aligned} &d|\partial_z v^\varepsilon|^p + p(\mu\wedge\nu)|\partial_z v^\varepsilon|^{p-2}|\nabla_3\partial_z v^\varepsilon|^2\,dt \\ &\leq p|\partial_z v^\varepsilon|^{p-2}\Big[\left|\left(\nabla\Upsilon^\varepsilon, \partial_z v^\varepsilon\right)\right| + \sqrt{\varepsilon}\lambda(\varepsilon)\left|\left(\partial_z B_2(v^\varepsilon, v^\varepsilon), \partial_z v^\varepsilon\right)\right| \\ &\quad + \left|\left(\partial_z B_2(v^0, v^\varepsilon), \partial_z v^\varepsilon\right)\right| \\ &\quad + \left|\left(\partial_z B_2(v^\varepsilon, v^0), \partial_z v^\varepsilon\right)\right| + \frac{p-2}{2}\|\partial_z\sigma_1\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right)\|_{L_2(\mathcal{U},L^2)}^2\Big]\,dt \\ &\quad + p|\partial_z v^\varepsilon|^{p-2}\left(\partial_z\sigma_1\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right)dW_1, \partial_z v^\varepsilon\right) \\ &= \sum_{k=1}^5 I_k + I_6\,dW_1. \end{aligned}$$

From the Cauchy–Schwartz inequality, we immediately obtain

$$I_1 \leq C|\partial_z v^\varepsilon|^{p-1}\|R^\varepsilon\|. \tag{4.43}$$

Before we proceed to I_2, let $u, v \in D(A)$. By the cancellation property (2.9) and integration by parts, we have

$$\begin{aligned} &(\partial_z B_2(u, v), \partial_z v) \\ &\quad = ((\partial_z u\cdot\nabla)\,v - (\operatorname{div} u)\,\partial_z v, \partial_z v) \\ &\quad = \int_{\mathcal{M}} \partial_z u_j\partial_j v_k\partial_z v_k - \partial_j u_j\partial_z v_k\partial_z v_k\,\mathrm{d}\mathcal{M} \end{aligned}$$

$$= -\int_{\mathcal{M}} \partial_z \partial_j u_j v_k \partial_z v_k + \partial_z u_j v_k \partial_j \partial_z v_k - 2 u_j \partial_j \partial_z v_k \partial_z v_k \, \mathrm{d}\mathcal{M}. \tag{4.44}$$

Using the Hölder inequality, (4.44), the Gagliardo–Nirenberg inequality and the Young inequality, we deduce

$$\begin{aligned} I_2 &\le C\sqrt{\varepsilon}\lambda(\varepsilon)|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\nabla \partial_z \mathrm{v}^\varepsilon||\mathrm{v}^\varepsilon|_{L^6}|\partial_z \mathrm{v}^\varepsilon|_{L^3} \\ &\le C\sqrt{\varepsilon}\lambda(\varepsilon)\left(|\partial_z \nabla \mathrm{v}^\varepsilon|^{3/2}|\partial_z \mathrm{v}^\varepsilon|^{p-3/2}|\mathrm{v}^\varepsilon|_{L^6} + |\partial_z \nabla \mathrm{v}^\varepsilon|^{1/2}|\partial_z \mathrm{v}^\varepsilon|^{p-1}|\mathrm{v}^\varepsilon|_{L^6}\right) \\ &\le \eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta|\partial_z \mathrm{v}^\varepsilon|^p\left(|\mathrm{v}^\varepsilon|_{L^6}^4 + 1\right) \end{aligned} \tag{4.45}$$

for some $\eta > 0$ small precisely determined later. Using the Hölder inequality and (4.44), we obtain

$$\begin{aligned} I_3 &\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2}\left(\|v^0\|_{H^2}|\mathrm{v}^\varepsilon|_{L^6}|\partial_z \mathrm{v}^\varepsilon|_{L^3} + |\partial_z v^0|_{L^3}|\mathrm{v}^\varepsilon|_{L^6}|\partial_z \nabla \mathrm{v}^\varepsilon| + |v^0|_{L^\infty}|\partial_z \nabla \mathrm{v}^\varepsilon||\partial_z \mathrm{v}^\varepsilon|\right) \\ &= I_3^1 + I_3^2 + I_3^3. \end{aligned}$$

By the Gagliardo–Nierenberg inequality and the Young inequality, we get

$$\begin{aligned} I_3^1 &\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2}\|v^0\|_{H^2}|\mathrm{v}^\varepsilon|_{L^6}\left(|\partial_z \mathrm{v}^\varepsilon|^{1/2}|\nabla \partial_z \mathrm{v}^\varepsilon|^{1/2} + |\partial_z \mathrm{v}^\varepsilon|\right) \\ &\le \eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta|\partial_z \mathrm{v}^\varepsilon|^{p-4/3}\|v^0\|_{H^2}^{4/3}|\mathrm{v}^\varepsilon|_{L^6}^{4/3} + C|\partial_z \mathrm{v}^\varepsilon|^{p-1}\|v^0\|_{H^2}|\mathrm{v}^\varepsilon|_{L^6}. \end{aligned} \tag{4.46}$$

Using the Young inequality and the Gagliardo–Nirenberg inequality for the term I_3^2, we obtain

$$I_3^2 + I_3^3 \le \eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\mathrm{v}^\varepsilon|_{L^6}^2\|v^0\|\|v^0\|_{H^2} + C_\eta|\partial_z \mathrm{v}^\varepsilon|^p|v^0|_{L^\infty}^2. \tag{4.47}$$

Proceeding to I_4, we use integration by parts to deduce

$$\begin{aligned} I_4 \le C|\partial_z \mathrm{v}^\varepsilon|^{p-2}\Big[&\left|\left(B_2(\mathrm{v}^\varepsilon, v^0), \partial_{zz}\mathrm{v}^\varepsilon\right)\right| + \left|\left((\operatorname{div}\mathrm{v}^\varepsilon)\partial_z v^0, \partial_z \mathrm{v}^\varepsilon\right)\right| \\ &+ \left|\left(w(\mathrm{v}^\varepsilon)\partial_{zz} v^0, \partial_z \mathrm{v}^\varepsilon\right)\right|\Big] \\ = I_4^1 + I_4^2 + I_4^3&. \end{aligned}$$

Similarly as above, we use the Hölder inequality, the Gagliardo–Nirenberg inequality and the Young inequality to get

$$I_4^1 \le \eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\mathrm{v}^\varepsilon|_{L^6}\|v^0\|\|v^0\|_{H^2}. \tag{4.48}$$

The term I_4^2 can be estimated in exactly the same way as I_3^1; therefore, we have

$$I_4^2 \le \eta|\partial_z \mathrm{v}^\varepsilon|^{p-2}|\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta |\partial_z \mathrm{v}^\varepsilon|^{p-4/3}\|v^0\|_{H^2}^{4/3}|\mathrm{v}^\varepsilon|_{L^6}^{4/3} + C|\partial_z \mathrm{v}^\varepsilon|^{p-1}\|v^0\|_{H^2}|\mathrm{v}^\varepsilon|_{L^6}. \tag{4.49}$$

Next, we employ integration by parts to get

$$\begin{aligned}
&-\int_{\mathcal{M}} \left(\int_{-h}^0 \operatorname{div} \mathrm{v}^\varepsilon \,\mathrm{d}z'\right) \partial_{zz} v^0 \partial_z \mathrm{v}^\varepsilon \,\mathrm{d}\mathcal{M}_0 \\
&\quad = \int_{\mathcal{M}_0} \left(\int_{-h}^z \mathrm{v}_j^\varepsilon \,\mathrm{d}z'\right) \partial_j \partial_{zz} v_k^0 \partial_z \mathrm{v}_k^\varepsilon + \left(\int_{-h}^z \mathrm{v}_j^\varepsilon \,\mathrm{d}z'\right) \partial_{zz} v_k^0 \partial_j \partial_z \mathrm{v}_k^\varepsilon \,\mathrm{d}\mathcal{M}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
I_4^3 &\le C|\partial_z \tilde{\mathrm{v}}^\varepsilon|^{p-2} \left|\int_{\mathcal{M}_0} \left(\int_{-h}^z \mathrm{v}_j^\varepsilon \,\mathrm{d}z'\right) \partial_j \partial_{zz} v_k^0 \partial_z \mathrm{v}_k^\varepsilon \,\mathrm{d}\mathcal{M}\right| \\
&\quad + C|\partial_z \tilde{\mathrm{v}}^\varepsilon|^{p-2} \left|\int_{\mathcal{M}_0} \left(\int_{-h}^z \mathrm{v}_j^\varepsilon \,dz'\right) \partial_{zz} v_k^0 \partial_j \partial_z \mathrm{v}_k^\varepsilon \,\mathrm{d}\mathcal{M}\right| \\
&= I_4^{31} + I_4^{32}.
\end{aligned}$$

By the Hölder inequality, the embedding $L^6 \hookrightarrow L_x^6 L_z^2$, the anisotropic estimate (2.23) and the Young inequality, we get

$$\begin{aligned}
I_4^{31} &\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2} \int_{\mathcal{M}} |\mathrm{v}^\varepsilon|_{L_z^2} |\partial_{zz} v^0|_{L_z^2} |\nabla \partial_z \mathrm{v}^\varepsilon|_{L_z^2} \,\mathrm{d}\mathcal{M}_0 \\
&\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2} |\mathrm{v}^\varepsilon|_{L_x^6 L_z^2} |\partial_{zz} v^0|_{L_x^3 L_z^2} |\partial_z \nabla \mathrm{v}^\varepsilon| \\
&\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2} |\partial_z \nabla \mathrm{v}^\varepsilon| |\mathrm{v}^\varepsilon|_{L^6} \|\partial_z v^0\|^{2/3} \|\partial_z v^0\|_{H^2}^{1/3} \\
&\le \eta |\partial_z \mathrm{v}^\varepsilon|^{p-2} |\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta |\partial_z \mathrm{v}^\varepsilon|^{p-2} |\mathrm{v}^\varepsilon|_{L^6}^2 \|\partial_z v^0\|^{4/3} \|\partial_z v^0\|_{H^2}^{2/3}. \qquad (4.50)
\end{aligned}$$

Similarly, we deduce

$$\begin{aligned}
I_4^{32} &\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2} \int_{\mathcal{M}} |\mathrm{v}^\varepsilon|_{L_z^2} |\nabla \partial_{zz} v^0|_{L_z^2} |\partial_z \mathrm{v}^\varepsilon|_{L^2} \,\mathrm{d}\mathcal{M} \\
&\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2} |\mathrm{v}^\varepsilon|_{L_x^6 L_z^2} \|\partial_z v^0\|_{H^2} |\partial_z \mathrm{v}^\varepsilon|_{L_x^3 L_z^2} \\
&\le C|\partial_z \mathrm{v}^\varepsilon|^{p-2} |\mathrm{v}^\varepsilon|_{L^6} \|\partial_z v^0\|_{H^2} \left(|\partial_z \mathrm{v}^\varepsilon|^{2/3} |\nabla_3 \partial_z \mathrm{v}^\varepsilon|^{1/3} + |\partial_z \mathrm{v}^\varepsilon|\right) \\
&\le \eta |\partial_z \mathrm{v}^\varepsilon|^{p-2} |\partial_z \nabla \mathrm{v}^\varepsilon|^2 + C_\eta |\partial_z \mathrm{v}^\varepsilon|^{p-6/5} \|\partial_z v^0\|_{H^2}^{6/5} |\mathrm{v}^\varepsilon|_{L^6}^{6/5}. \qquad (4.51)
\end{aligned}$$

The Itô correction term is estimated using the bound (2.21) and the Young inequality by

$$\begin{aligned}
I_5 &\le C\lambda^{-2}(\varepsilon) |\partial_z \mathrm{v}^\varepsilon|^{p-2} \left(1 + \|U^0\|^2 + \|U^0\|_{H^2}^2\right) \\
&\quad + C\varepsilon |\partial_z \mathrm{v}^\varepsilon|^p + p(p-1)\varepsilon\eta_0 |\partial_z \mathrm{v}^\varepsilon|^{p-2} |\nabla_3 \partial_z \mathrm{v}^\varepsilon|^2. \qquad (4.52)
\end{aligned}$$

Let $K \geq 0$ and let τ_a and τ_b be stopping times such that

$$0 \leq \tau_a \leq \tau_b \leq T \wedge \tau_K^{\nabla,\varepsilon} \wedge \tau_K^{6,\varepsilon} \wedge \tau_K^{w,\varepsilon} \wedge \tau_K^0.$$

Proceeding similarly as in (4.6) and (4.7), we employ the Burkholder–Davis–Gundy inequality and the bound (2.21) on $\partial_z \sigma_1$ in $L_2(\mathcal{U}, L^2)$ and deduce

$$\begin{aligned} &p\lambda^{-1}(\varepsilon)\mathbb{E} \sup_{t\in[\tau_a,\tau_b]} \left| \int_{\tau_a}^t |\partial_z v^\varepsilon|^{p-2} \left(\partial_z v^\varepsilon, \partial_z \sigma_1 \left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon) R^\varepsilon \right) \mathrm{d}W_1 \right) \right| \\ &\quad \leq (1-\delta+\eta)\, \mathbb{E} \sup_{s\in[\tau_a,\tau_b]} |\partial_z v^\varepsilon|^p + C_\eta \lambda^{-2}(\varepsilon) \mathbb{E} \int_{\tau_a}^{\tau_b} |\partial_z v^\varepsilon|^{p-2} \left(1 + \|U^0\|_{H^2}^2 \right) \mathrm{d}s \\ &\qquad + C_\eta \varepsilon \mathbb{E} \int_{\tau_a}^{\tau_b} |\partial_z v^\varepsilon|^p \,\mathrm{d}s + \frac{C_\eta \eta_3 \varepsilon}{1-\delta} \mathbb{E} \int_{\tau_a}^{\tau_b} |\partial_z v^\varepsilon|^{p-2} |\nabla_3 \partial_z v^\varepsilon|^2 \,\mathrm{d}s \end{aligned} \tag{4.53}$$

for some $\delta \in (\eta, 1)$.

Collecting the above estimates, choosing δ and η sufficiently small and assuming that η_3 or ε_0 are sufficiently small, we get

$$\mathbb{E}\left[\sup_{s\in[\tau_a,\tau_b]} |\partial_z v^\varepsilon|^p + \int_{\tau_a}^{\tau_b} |\partial_z v^\varepsilon|^{p-2} |\nabla_3 \partial_z v^\varepsilon|^2 \, ds \right] \leq C\mathbb{E} \int_{\tau_a}^{\tau_b} \left(1 + |\partial_z v^\varepsilon|^p\right) \Phi(s) \,\mathrm{d}s,$$

where

$$\Phi(s) = 1 + \|R^\varepsilon\|^2 + |v^\varepsilon|_{L^6}^4 + \|U^0\|_{H^2}^2 + \|U^0\|^2 \|U^0\|_{H^2}^2 + \|\partial_z U^0\|^4 \|\partial_z U^0\|_{H^2}^2 + \|\partial_z U^0\|_{H^2}^2.$$

The claim follows from the uniform stochastic Gronwall lemma from Proposition A.1. □

The proof of the following proposition is similar to the proof of Proposition 4.5 and [2, Proposition 4.6] and is therefore omitted.

Proposition 4.6 *Let $K \geq 0$ and let $\tau_K^{T,\varepsilon}$ be the stopping time defined by*

$$\tau_K^{T,\varepsilon} = \inf \left\{ t \geq 0 \mid \sup_{s\in[0,t]} |\partial_z \Upsilon^\varepsilon|^4 + \int_0^t |\partial_z \Upsilon^\varepsilon|^2 \|\partial_z \Upsilon^\varepsilon\|^2 \,\mathrm{d}s \geq K \right\}. \tag{4.54}$$

Then, $\tau_K^{T,\varepsilon} \to \infty$ $\mathbb{P}$-almost surely as $K \to \infty$ for all $\varepsilon \in (0, \varepsilon_0]$ and, for all $t \geq 0$, one has

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{ \tau_K^{T,\varepsilon} \leq t \right\} \right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0].$$

Proposition 4.7 *Let $p \geq 2$, $K \geq 0$ and let $\tau_K^{U,\varepsilon,p}$ be the stopping time defined by*

$$\tau_K^{R,\varepsilon,p} = \inf \left\{ t \geq 0 \mid \sup_{s\in[0,t]} \|R^\varepsilon\|^p + \int_0^t \|R^\varepsilon\|^{p-2} \|R^\varepsilon\|_{H^2}^2 \,\mathrm{d}s \geq K \right\}. \tag{4.55}$$

Then, $\tau_K^{R,\varepsilon,p} \to \infty$ $\mathbb{P}$-almost surely as $K \to \infty$ for all $\varepsilon \in (0, \varepsilon_0]$ and, for all $t \geq 0$, one has

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{R,\varepsilon,p} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0].$$

Proof By the Itô lemma [2, Theorem A.1] and integrating by parts, we have

$$\begin{aligned}
& d|A^{1/2}R^\varepsilon|^p + p\,(\mu \wedge \nu)\,|A^{1/2}R^\varepsilon|^{p-2}|AR^\varepsilon|^2\,dt \\
&\quad \leq p|A^{1/2}R^\varepsilon|^{p-2}\Bigg[\sqrt{\varepsilon}\lambda(\varepsilon)\left|\left(B(R^\varepsilon, R^\varepsilon), AR^\varepsilon\right)\right| + \left|\left(B(U^0, R^\varepsilon), AR^\varepsilon\right)\right| \\
&\qquad + \left|\left(B(R^\varepsilon, U^0), AR^\varepsilon\right)\right| \\
&\qquad + \left|\left(F_d(R^\varepsilon), AR^\varepsilon\right)\right| + \tfrac{p-1}{2}\lambda^{-2}(\varepsilon)\|\sigma\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right)\|^2_{L_2(\mathcal{U},V)}\Bigg]dt \\
&\qquad + p\lambda^{-1}(\varepsilon)|A^{1/2}R^\varepsilon|^{p-2}\left(A^{1/2}\sigma\left(U^0 + \sqrt{\varepsilon}\lambda(\varepsilon)R^\varepsilon\right)dW, A^{1/2}R^\varepsilon\right) \\
&\quad = \sum_{k=1}^{5} I_k\,dt + I_6\,dW.
\end{aligned}$$

The estimate (2.6) and the Young inequality imply

$$I_1 \leq \eta\|R^\varepsilon\|^{p-2}\|R^\varepsilon\|^2_{H^2} + C_\eta\varepsilon\lambda^2(\varepsilon)\|R^\varepsilon\|^p\left(|v^\varepsilon|^4_{L^6} + |\partial_z R^\varepsilon|^2\|\partial_z R^\varepsilon\|^2\right), \tag{4.56}$$

where $\eta > 0$ will be determined later. From the estimate (2.7) and the Young inequality, we get

$$I_2 + I_3 + I_4 \leq \eta\|R^\varepsilon\|^{p-2}\|R^\varepsilon\|^2_{H^2} + C_\eta\|R^\varepsilon\|^p\left(1 + \|U^0\|^2\|U^0\|^2_{H^2}\right). \tag{4.57}$$

For the Itô correction term, we have

$$I_5 \leq C\lambda^{-2}(\varepsilon)\|R^\varepsilon\|^{p-2}\left(1 + \|U^0\|^2_{H^2}\right) + C\varepsilon\|R^\varepsilon\|^p + C\varepsilon\eta_1\|R^\varepsilon\|^{p-2}\|R^\varepsilon\|^2_{H^2}. \tag{4.58}$$

Similarly as in the above proofs, let τ_a and τ_b be stopping times such that

$$0 \leq \tau_a \leq \tau_b \leq T \wedge \tau_K^{T,\varepsilon} \wedge \tau_K^{z,\varepsilon} \wedge \tau_K^{\nabla,\varepsilon} \wedge \tau_K^{6,\varepsilon} \wedge \tau_K^{w,\varepsilon} \wedge \tau_K^0.$$

Using the Burkholder–Davis–Gundy inequality (2.28) and similar estimates as in (4.6) and (4.7), we deduce

$$\begin{aligned}
& p\lambda^{-1}(\varepsilon)\mathbb{E}\sup_{t\in[\tau_a,\tau_b]}\left|\int_{\tau_a}^{t} I_6\,\mathrm{d}W\right| \\
&\quad \leq (1 - \delta + \eta)\,\mathbb{E}\sup_{s\in[\tau_a,\tau_b]}\|R^\varepsilon\|^p + C_\eta\lambda^{-2}(\varepsilon)\mathbb{E}\int_{\tau_a}^{\tau_b}\|R^\varepsilon\|^{p-2}\left(1 + \|U^0\|^2_{H^2}\right)\mathrm{d}s
\end{aligned}$$

$$+ C_\eta \varepsilon \mathbb{E} \int_{\tau_a}^{\tau_b} \|R^\varepsilon\|^p \, \mathrm{d}s + \frac{C_\eta \eta_3 \varepsilon}{1-\delta} \mathbb{E} \int_{\tau_a}^{\tau_b} \|R^\varepsilon\|^{p-2} \|R^\varepsilon\|_{H^2}^2 \, \mathrm{d}s. \tag{4.59}$$

Collecting the above estimates and choosing $\eta, \delta > 0$ and ε_0 sufficiently small, we obtain

$$\mathbb{E}\left[\sup_{s\in[\tau_a,\tau_b]} \|R^\varepsilon\|^p + \int_{\tau_a}^{\tau_b} \|R^\varepsilon\|^{p-2}\|R^\varepsilon\|_{H^2}^2 \, \mathrm{d}s\right] \le C_p \mathbb{E} \int_{\tau_a}^{\tau_b} \left(1 + \|R^\varepsilon\|^p\right)(1+\Phi(s)) \, \mathrm{d}s,$$

where the constant C_p depends on p but is independent of τ_a and τ_b and

$$\Phi(s) = 1 + |v^\varepsilon|_{L^6}^4 + |\partial_z R^\varepsilon|^2 \|\partial_z R^\varepsilon\|^2 + \|U^0\|^2\|U^0\|_{H^2}^2 + \|U^0\|_{H^2}^2.$$

The proof is closed using the Gronwall lemma from Proposition A.1 similarly as in the proofs above. □

4.3 Proof of Theorem 1.2

The proof of Theorem 1.2 closely follows the argument of the proof of Theorem 1.1 in Sect. 3.3 using the results of Sects. 4.1 and 4.2 . For this reason, we omit the details. The respective good rate function $I : C([0,T], V) \cap L^2(0,T; D(A)) \to \mathbb{R}$ is given by

$$I(U) = \inf\left\{\frac{1}{2}\int_0^T |h|_{\mathcal{U}}^2 \, \mathrm{d}t \mid h \in L^2(0,T;\mathcal{U}) \text{ s.t. } U = \mathcal{G}_R^0\left(\int_0^\cdot h \, dt\right)\right\} \tag{4.60}$$

where $\mathcal{G}_R^0$ has been defined in Proposition 4.1. □

5 Proof of Theorem 1.3

Most of the estimates in this section are straightforward adaptations of the estimates from Sect. 4.2. Thus, we only go through the main steps of the proof, which closely follows the one from Zhang et al. [49].

Let U^ε and U^0 be the solutions of (3.1) and (4.1), respectively, and recall that, for $K \ge 0$, the stopping time

$$\tau_K^0 = \inf\left\{t \ge 0 \mid \int_0^t \|U^0\|_{H^2}^2 + \|U^0\|^2\|U^0\|_{H^2}^2 \, \mathrm{d}s \ge K\right\}$$

satisfies $\tau_K^0 \to \infty$ as $K \to \infty$. Let $R^\varepsilon = (U^\varepsilon - U^0)/\sqrt{\varepsilon}$. Clearly, R^ε satisfies

$$dR^\varepsilon + \left[AR^\varepsilon + B\left(R^\varepsilon, U^0 + \sqrt{\varepsilon}R^\varepsilon\right) + B\left(U^0, R^\varepsilon\right) + A_{\mathrm{pr}}R^\varepsilon + ER^\varepsilon\right] dt = \sigma\left(U^\varepsilon\right) dW, \tag{5.1}$$

with $R^\varepsilon(0) = 0$, which is essentially (4.3) with $\lambda(\varepsilon) \equiv 1$. Let $\tau_K^{R,\varepsilon}$ be the stopping time defined for $K \geq 0$ by

$$\tau_K^{R,\varepsilon} = \inf\left\{ t \geq 0 \mid \sup_{s\in[0,t]} \|R^\varepsilon\|^2 + \int_0^t \|R^\varepsilon\|_{H^2}^2 + \|R^\varepsilon\|^2 \|R^\varepsilon\|_{H^2}^2 \,\mathrm{d}s \geq K \right\}.$$

By the result of Sect. 4.2, we have $\tau_K^{R,\varepsilon} \to \infty$ $\mathbb{P}$-a.s. as $K \to \infty$. Let $\hat{U}$ be the solution of

$$d\hat{U} + \left[A\hat{U} + B(U^0, \hat{U}) + B(\hat{U}, U^0) + A_{\mathrm{pr}}\hat{U} + E\hat{U} \right] ds = \sigma(U^0)\, dW, \quad \hat{U}(0) = 0. \tag{5.2}$$

Let $Y^\varepsilon = R^\varepsilon - \hat{U}$. Using bilinearity of B, we observe that Y^ε satisfies

$$\begin{aligned} dY^\varepsilon &+ \left[AY^\varepsilon + \sqrt{\varepsilon} B(R^\varepsilon, R^\varepsilon) + B(U^0, Y^\varepsilon) + B(Y^\varepsilon, U^0) + F_d(Y^\varepsilon) \right] dt \\ &= \left[\sigma(U^\varepsilon) - \sigma(U^0) \right] dW, \quad Y^\varepsilon(0) = 0. \end{aligned} \tag{5.3}$$

Let τ_a and τ_b be stopping times such that $0 \leq \tau_a \leq \tau_b \leq T \wedge \tau_K^{R,\varepsilon} \wedge \tau_K^0$. By the Itô lemma and similar estimates as in Proposition 4.7 with the estimate (2.7) on B, the Lipschitz continuity (2.17) of σ in $L_2(\mathcal{U}, V)$ and the boundedness of the operators A_{pr} and E, we obtain

$$\begin{aligned} \mathbb{E}\left[\sup_{s\in[\tau_a,\tau_b]} \|Y^\varepsilon\|^2 + \int_{\tau_a}^{\tau_b} \|Y^\varepsilon\|_{H^2}^2 \,\mathrm{d}s \right] &\leq C\mathbb{E}\int_{\tau_a}^{\tau_a} \|Y^\varepsilon\|^2 \left(1 + \|U^0\|^2 \|U^0\|_{H^2}^2 \right) \mathrm{d}s \\ + C\mathbb{E}\|Y^\varepsilon(\tau_a)\|^2 + C\varepsilon\mathbb{E}\int_{\tau_a}^{\tau_b} &\|R^\varepsilon\|^2 + \|R^\varepsilon\|_{H^2}^2 + \|R^\varepsilon\|^2 \|R^\varepsilon\|_{H^2}^2 \,\mathrm{d}s. \end{aligned}$$

Using the stochastic Gronwall lemma from Glatt-Holtz and Ziane [18, Lemma 5.3], we get

$$\begin{aligned} &\mathbb{E}\left[\sup_{t\in[0,T\wedge\tau_K^{R,\varepsilon}]} \|Y^\varepsilon\|^2 + \int_0^{T\wedge\tau_K^{R,\varepsilon}} \|Y^\varepsilon\|_{H^2}^2 \,\mathrm{d}t \right] \\ &\quad\leq C_K \varepsilon \mathbb{E}\int_0^{T\wedge\tau_K^{R,\varepsilon}} \|R^\varepsilon\|^2 + \|R^\varepsilon\|_{H^2}^2 + \|R^\varepsilon\|^2 \|R^\varepsilon\|_{H^2}^2 \,\mathrm{d}t. \end{aligned} \tag{5.4}$$

If we define

$$\Omega_K = \left\{ \omega \in \Omega \mid \tau_K^0 \wedge \tau_K^{R,\varepsilon} \geq T \right\}$$

and if we consider the process $\mathbb{1}_{\Omega_K} Y^\varepsilon$ instead of Y^ε, keep K fixed in (5.4) and take the limit $\varepsilon \to 0+$, we get $Y^\varepsilon \to 0$ in $C([0,T], V) \cap L^2(0,T; H^2)$ $\mathbb{P}$-a.s. on Ω_K.

Since $\tau_K^0 \wedge \tau_K^{R,\varepsilon} \to \infty$ as $K \to \infty$ $\mathbb{P}$-almost surely, we have $\mathbb{P}(\Omega \setminus \bigcup_{K\in\mathbb{N}} \Omega_K) = 0$, which gives the convergence $\mathbb{P}$-a.s. on Ω. □

Acknowledgements The author wishes to thank Prof. Z. Brzeźniak and P. Razafimandimby for fruitful discussions, the University of York for their kind hospitality and the anonymous referee for a careful reading of the manuscript and suggestions improving the readability of the manuscript.

A Uniform Version of the Stochastic Gronwall Lemma

The following result is not new; in fact, it is a combination of the stochastic Gronwall lemma from Glatt-Holtz and Ziane [18, Lemma 5.3] and a part of the proof of global existence of the strong solutions of 2D stochastic Navier–Stokes equation from Glatt-Holtz and Ziane [18, Theorem 4.2]. We include the proof for the sake of completeness.

Proposition A.1 *Let $\varepsilon_0 > 0$. Let $X^\varepsilon, Y^\varepsilon, Z^\varepsilon, R^\varepsilon : [0,\infty) \times \Omega \to [0,\infty)$ be stochastic processes on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\tau_K^{R,\varepsilon}$ be the stopping time defined by*

$$\tau_K^{R,\varepsilon} = \inf\left\{t \geq 0 \mid \int_0^t R^\varepsilon \, \mathrm{d}s \geq K\right\}, \quad K > 0, \varepsilon \in (0, \varepsilon_0].$$

Let for all $t > 0$

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{R,\varepsilon} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0,\varepsilon_0]. \tag{A.1}$$

Let $T > 0$ and let for all $K > 0$ and $\varepsilon \in (0, \varepsilon_0]$

$$\mathbb{E}\int_0^{T\wedge\tau_K^{R,\varepsilon}} R^\varepsilon X^\varepsilon + Z^\varepsilon \, \mathrm{d}s \leq C_{T,K} < \infty.$$

Let there exist a constant $C_0 = C_0(T)$ such that, for all $\varepsilon \in (0,\varepsilon_0]$ and all stopping times τ_a and τ_b satisfying $0 \leq \tau_a \leq \tau_b \leq T \wedge \tau_K^{R,\varepsilon}$, one has

$$\mathbb{E}\left[\sup_{s\in[\tau_a,\tau_b]} X^\varepsilon + \int_{\tau_a}^{\tau_b} Y^\varepsilon \, \mathrm{d}s\right] \leq C_0\left[X(\tau_a) + \int_{\tau_a}^{\tau_b} R^\varepsilon X^\varepsilon + Z^\varepsilon \, \mathrm{d}s\right].$$

Then, for all $\varepsilon \in (0,\varepsilon_0]$ and $K > 0$, we have

$$\mathbb{E}\left[\sup_{s\in\left[0,T\wedge\tau_K^{R,\varepsilon}\right]} X^\varepsilon + \int_0^{T\wedge\tau_K^{R,\varepsilon}} Y^\varepsilon \, \mathrm{d}s\right] \leq C_{C_0,T,K}\mathbb{E}\left[X(0) + \int_0^{T\wedge\tau_K^{R,\varepsilon}} Z^\varepsilon \, \mathrm{d}s\right]. \tag{A.2}$$

Moreover, if we define the stopping time $\tau_K^{X,\varepsilon}$ *by*

$$\tau_K^{X,\varepsilon} = \inf\left\{ t \geq 0 \mid \sup_{s\in[0,t]} X^\varepsilon + \int_0^t Y^\varepsilon \, \mathrm{d}s \geq K \right\}, \quad K > 0, \varepsilon \in (0, \varepsilon_0],$$

then for all $t > 0$

$$\lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{X,\varepsilon} \leq t\right\}\right) = 0 \text{ uniformly w.r.t. } \varepsilon \in (0, \varepsilon_0], \tag{A.3}$$

and $\tau_K^{X,\varepsilon} \to +\infty$ *as* $K \to \infty$ $\mathbb{P}$*-a.s. for all* $\varepsilon \in (0, \varepsilon_0]$.

Proof The inequality (A.2) is the stochastic Gronwall lemma from Glatt-Holtz and Ziane [18, Lemma 5.3]. Following the argument from Glatt-Holtz and Ziane [18, Theorem 4.2], we use the Chebyshev theorem and (A.2) to estimate

$$\begin{aligned}
\mathbb{P}\left(\left\{\tau_K^{X,\varepsilon} \leq t\right\}\right) &\leq \mathbb{P}\left(\left\{\tau_K^{X,\varepsilon} \leq t\right\} \cap \left\{\tau_M^{R,\varepsilon} > t\right\}\right) + \mathbb{P}\left(\left\{\tau_M^{R,\varepsilon} \leq t\right\}\right) \\
&\leq \mathbb{P}\left(\left\{\sup_{s\in\left[0,t\wedge\tau_M^{R,\varepsilon}\right]} X^\varepsilon + \int_0^{t\wedge\tau_M^{R,\varepsilon}} Y^\varepsilon \, \mathrm{d}s \geq K\right\}\right) + \mathbb{P}\left(\left\{\tau_M^{R,\varepsilon} \leq t\right\}\right) \\
&\leq \frac{1}{K}\mathbb{E}\left[\sup_{s\in\left[0,t\wedge\tau_M^{R,\varepsilon}\right]} X^\varepsilon + \int_0^{t\wedge\tau_M^{R,\varepsilon}} Y^\varepsilon \, \mathrm{d}s\right] + \mathbb{P}\left(\left\{\tau_M^{R,\varepsilon} \leq t\right\}\right) \\
&\leq \frac{C_{t,M}}{K} + \mathbb{P}\left(\left\{\tau_M^{R,\varepsilon} \leq t\right\}\right).
\end{aligned}$$

Let $\delta > 0$ be arbitrary. By the uniform convergence (A.1), we find $M \in \mathbb{N}$ such that, for all $\varepsilon \in (0, \varepsilon_0]$, we have

$$\mathbb{P}\left(\left\{\tau_M^{R,\varepsilon} \leq t\right\}\right) < \frac{\delta}{2}.$$

Let $K_0 \in \mathbb{N}$ be such that $C_{t,M}/K < \delta/2$ for all $K \in \mathbb{N}$, $K \geq K_0$. Collecting the above, we deduce that

$$\mathbb{P}\left(\left\{\tau_K^{X,\varepsilon} \leq t\right\}\right) < \delta,$$

for all $\varepsilon \in (0, \varepsilon_0]$ and all $K \in \mathbb{N}$, $K \geq 0$, which finishes the proof of (A.3).

To establish the $\mathbb{P}$-a.s. convergence, we argue by contradiction. Assume that $\mathbb{P}(\{\lim_{K\to\infty} \tau_K^{X,\varepsilon} < +\infty\}) > 0$ for some $\varepsilon \in (0, \varepsilon_0]$. Then, since $\{\lim_{K\to\infty} \tau_K^{X,\varepsilon} < +\infty\} = \bigcup_{N\in\mathbb{N}}\{\lim_{K\to\infty} \tau_K^{X,\varepsilon} \leq N\}$, there exists $N_0 \in \mathbb{N}$ such that $\mathbb{P}(\{\lim_{K\to\infty} \tau_K^{X,\varepsilon} \leq N_0\}) > 0$. On the other hand, since $\tau_K^{X,\varepsilon}$ is monotone, i.e. $\{\tau_K^{X,\varepsilon} \leq N_0\} \supseteq \{\tau_L^{X,\varepsilon} \leq N_0\}$

for $K \leq L$, we observe $\{\lim_{K\to\infty} \tau_K^{X,\varepsilon} \leq N_0\} = \bigcap_{K\in\mathbb{N}} \{\tau_K^{X,\varepsilon} \leq N_0\}$. However, (A.3) implies

$$0 < \mathbb{P}\left(\left\{\lim_{K\to\infty} \tau_K^{X,\varepsilon} \leq N_0\right\}\right) = \mathbb{P}\left(\bigcap_{K\in\mathbb{N}} \left\{\tau_K^{X,\varepsilon} \leq N_0\right\}\right) = \lim_{K\to\infty} \mathbb{P}\left(\left\{\tau_K^{X,\varepsilon} \leq N_0\right\}\right) = 0,$$

a contradiction. □

References

1. Bensoussan, A.: Stochastic Navier–Stokes equations. Acta Appl. Math. **38**(3), 267–304 (1995)
2. Brzeźniak, Z., Slavík, J.: Well-posedness of the 3D stochastic primitive equations with transport noise. J. Differ. Equ. (2021). https://doi.org/10.1016/j.jde.2021.05.049
3. Budhiraja, A., Dupuis, P.: A variational representation for positive functionals of infinite dimensional Brownian motion. Probab. Math. Stat. **20**(1, Acta Univ. Wratislav. No. 2246), 39–61 (2000)
4. Budhiraja, A., Dupuis, P., Maroulas, V.: Large deviations for infinite dimensional stochastic dynamical systems. Ann. Probab. **36**(4), 1390–1420 (2008)
5. Cao, C., Li, J., Titi, E.S.: Local and global well-posedness of strong solutions to the 3D primitive equations with vertical eddy diffusivity. Arch. Ration. Mech. Anal. **214**(1), 35–76 (2014)
6. Cao, C., Titi, E.S.: Global well-posedness of the three-dimensional viscous primitive equations of large scale ocean and atmosphere dynamics. Ann. Math. (2) **166**(1), 245–267 (2007)
7. Chueshov, I., Millet, A.: Stochastic 2D hydrodynamical type systems: well posedness and large deviations. Appl. Math. Optim. **61**(3), 379–420 (2010)
8. Da Prato, G., Zabczyk, J.: In: Stochastic Equations in Infinite Dimensions, Volume 152 of Encyclopedia of Mathematics and its Applications, 2nd edn. Cambridge University Press, Cambridge (2014)
9. Debussche, A., Glatt-Holtz, N., Temam, R.: Local martingale and pathwise solutions for an abstract fluids model. Phys. D **240**(14–15), 1123–1144 (2011)
10. Debussche, A., Glatt-Holtz, N., Temam, R., Ziane, M.: Global existence and regularity for the 3D stochastic primitive equations of the ocean and atmosphere with multiplicative white noise. Nonlinearity **25**(7), 2093–2118 (2012)
11. Dellacherie, C., Meyer, P.-A.: Probabilities and Potential, Volume 29 of North-Holland Mathematics Studies. North-Holland, Amsterdam, New York (1978)
12. Dong, Z., Zhai, J., Zhang, R.: Large deviation principles for 3D stochastic primitive equations. J. Differ. Equ. **263**(5), 3110–3146 (2017)
13. Freidlin, M.I., Wentzell, A.D.: In: Random Perturbations of Dynamical Systems, Volume 260 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer, New York (1984)
14. Fristedt, B., Gray, L.: A Modern Approach to Probability Theory, Probability and its Applications. Birkhäuser, Boston (1997)
15. Gao, H., Sun, C.: Well-posedness and large deviations for the stochastic primitive equations in two space dimensions. Commun. Math. Sci. **10**(2), 575–593 (2012)
16. Glatt-Holtz, N., Kukavica, I., Vicol, V., Ziane, M.: Existence and regularity of invariant measures for the three dimensional stochastic primitive equations. J. Math. Phys. **55(5):051504**, 34 (2014)
17. Glatt-Holtz, N., Temam, R., Wang, C.: Time discrete approximation of weak solutions to stochastic equations of geophysical fluid dynamics and applications. Chin. Ann. Math. Ser. B **38**(2), 425–472 (2017)
18. Glatt-Holtz, N., Ziane, M.: Strong pathwise solutions of the stochastic Navier–Stokes system. Adv. Differ. Equ. **14**(5–6), 567–600 (2009)
19. Guillén-González, F., Masmoudi, N., Rodríguez-Bellido, M.A.: Anisotropic estimates and strong solutions of the primitive equations. Differ. Integral Equ. **14**(11), 1381–1408 (2001)
20. Guo, B., Huang, D.: On the 3D viscous primitive equations of the large-scale atmosphere. Acta Math. Sci. Ser. B (Engl. Ed.) **29**(4), 846–866 (2009)
21. Hieber, M., Kashiwabara, T.: Global strong well-posedness of the three dimensional primitive equations in L^p-spaces. Arch. Ration. Mech. Anal. **221**(3), 1077–1115 (2016)

22. Hu, C., Temam, R., Ziane, M.: The primitive equations on the large scale ocean under the small depth hypothesis. Discrete Contin. Dyn. Syst. **9**(1), 97–131 (2003)
23. Hussein, A., Saal, M., Wrona, M.: Primitive equations with horizontal viscosity: the initial value and the time-periodic problem for physical boundary conditions. AIMS. (2021). https://doi.org/10.3934/dcds.2020398
24. Ju, N.: The global attractor for the solutions to the 3D viscous primitive equations. Discrete Contin. Dyn. Syst. **17**(1), 159–179 (2007)
25. Ju, N.: Global uniform boundedness of solutions to viscous 3D primitive equations with physical boundary conditions. Indiana Univ. Math. J. (2020). https://doi.org/10.1512/iumj.2020.69.7976
26. Ju, N.: On H^2 solutions and z-weak solutions of the 3D primitive equations. Indiana Univ. Math. J. **66**(3), 973–996 (2017)
27. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus, Volume 113 of Graduate Texts in Mathematics, 2nd edn. Springer, New York (1991)
28. Kobelkov, G.M.: Existence of a solution “in the large” for ocean dynamics equations. J. Math. Fluid Mech. **9**(4), 588–610 (2007)
29. Kukavica, I., Ziane, M.: On the regularity of the primitive equations of the ocean. Nonlinearity **20**(12), 2739–2753 (2007)
30. Li, J., Titi, E.S.: Recent advances concerning certain class of geophysical flows. In: Handbook of Mathematical Analysis in Mechanics of Viscous Fluids, pp. 933–971. Springer, Cham (2018)
31. Li, J., Titi, E.S.: The primitive equations as the small aspect ratio limit of the Navier–Stokes equations: rigorous justification of the hydrostatic approximation. J. Math. Pures Appl. **124**, 30–58 (2019)
32. Lions, J.-L., Temam, R., Wang, S.: Models for the coupled atmosphere and ocean. (CAO I, II). Comput. Mech. Adv. **1**(1), 120 (1993)
33. Lions, J.-L., Temam, R., Wang, S.H.: New formulations of the primitive equations of atmosphere and applications. Nonlinearity **5**(2), 237–288 (1992)
34. Lions, J.-L., Temam, R., Wang, S.H.: On the equations of the large-scale ocean. Nonlinearity **5**(5), 1007–1053 (1992)
35. Liu, W.: Large deviations for stochastic evolution equations with small multiplicative noise. Appl. Math. Optim. **61**(1), 27–56 (2010)
36. Liu, W., Röckner, M., Zhu, X.-C.: Large deviation principles for the stochastic quasi-geostrophic equations. Stoch. Process. Appl. **123**(8), 3299–3327 (2013)
37. Mikulevicius, R., Rozovskii, B.: Stochastic Navier–Stokes equations. Propagation of chaos and statistical moments. In: Optimal Control and Partial Differential Equations, pp. 258–267. IOS, Amsterdam (2001)
38. Mikulevicius, R., Rozovskii, B.L.: Stochastic Navier–Stokes equations for turbulent flows. SIAM J. Math. Anal. **35**(5), 1250–1310 (2004)
39. Pedlosky, J.: Geophysical Fluid Dynamics, Springer Study Edition. Springer, New York (1992)
40. Petcu, M., Temam, R. M., Ziane, M.: Some mathematical problems in geophysical fluid dynamics. In: Handbook of Numerical Analysis. Vol. XIV. Special Volume: Computational Methods for the Atmosphere and the Oceans, Volume 14 of Handb. Numer. Anal., pp. 577–750. Elsevier/North-Holland, Amsterdam (2009)
41. Qiao, H.: A theorem dual to Yamada-Watanabe theorem for stochastic evolution equations. Stoch. Dyn. **10**(3), 367–374 (2010)
42. Röckner, M., Schmuland, B., Zhang, X.: Yamada–Watanabe theorem for stochastic evolution equations in infinite dimensions. Condens. Matt. Phys. (2008)
43. Samelson, R., Temam, R., Wang, S.: Some mathematical properties of the planetary geostrophic equations for large-scale ocean circulation. Appl. Anal. **70**(1–2), 147–173 (1998)
44. Temam, R.: Navier–Stokes Equations, Volume 2 of Studies in Mathematics and its Applications, Revised Edition. North-Holland, Amsterdam, New York (1979). Theory and numerical analysis, With an appendix by F. Thomasset
45. Vallis, G.K.: Atmospheric and Oceanic Fluid Dynamics. Cambridge University Press, Cambridge (2017)
46. van der Vaart, A.W.: Asymptotic Statistics, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 3. Cambridge University Press, Cambridge (1998)
47. Varadhan, S.R.S.: Asymptotic probabilities and differential equations. Commun. Pure Appl. Math. **19**, 261–286 (1966)

48. Wang, R., Zhai, J., Zhang, T.: A moderate deviation principle for 2-D stochastic Navier–Stokes equations. J. Differ. Equ. **258**(10), 3363–3390 (2015)
49. Zhang, R., Zhou, G., Guo, B.: Stochastic 2D primitive equations: central limit theorem and moderate deviation principle. Comput. Math. Appl. **77**(4), 928–946 (2019)