NON-UNIQUENESS OF FORCED ACTIVE SCALAR EQUATIONS WITH EVEN DRIFT OPERATORS

MIMI DAI AND SUSAN FRIEDLANDER

Abstract. We consider forced active scalar equations with even and homogeneous degree 0 drift operator on $\mathbb{T}^d$. Inspired by the non-uniqueness construction for dyadic fluid models [17, 23], by implementing a sum-difference convex integration scheme we obtain non-unique weak solutions for the active scalar equation in space $C_t^0 C_x^\alpha$ with $\alpha < \frac{1}{2d+1}$. Without external forcing, Isett and Vicol [30] constructed non-unique weak solutions for such active scalar equations with spatial regularity C_x^α for $\alpha < \frac{1}{4d+1}$.

KEY WORDS: forced active scalar equation; non-uniqueness; convex integration sum-difference scheme.

CLASSIFICATION CODE: 35Q35, 35Q86, 76D03.

1. Introduction

1.1. **Background.**

The active scalar equation with external forcing

$$\begin{aligned} \partial_t \theta + u \cdot \nabla \theta &= -\nu \Lambda^\gamma \theta + f, \\ u &= T[\theta], \\ \nabla \cdot u &= 0 \end{aligned} \tag{1.1}$$

describes a number of physical phenomena arising in fluid dynamics. The unknown θ is a real-valued scalar function, while u is the drift velocity defined from θ through the nonlocal Zygmund operator T. The given function f denotes the external buoyancy forcing. The parameter $\nu \geq 0$ is the dissipation coefficient, and $\gamma > 0$ indicates the strength of the dissipation. The operator T has Fourier symbol $m(\xi)$ which is even, homogeneous of degree 0, and satisfies $\xi \cdot m(\xi) = 0$. We consider (1.1) on $\mathbb{T}^d \times [0, \infty)$ with $d \geq 2$.

Particular physical examples of (1.1) with even drift operators include the incompressible porous media (IPM) equation [1, 16] and the magnetogeostrophic (MG) equation [25, 33, 34]. These physical models have attracted attention due to their application in various physics contexts and their connection to hydrodynamic equations.

The class of active scalar equations with odd drift operators, including the surface quasi-geostrophic equation (SQG) [13], has also been extensively studied in the literature. The different symmetry features of the even and odd classes result in different ill/well-posedness theories. The cancellation property for the odd class of scalar equations is beneficial in establishing well-posedness, see [7, 14, 31, 32, 35]; while such cancellation structure is absent for the even class. The main objective of this paper is to investigate the ill-posedness phenomena for (1.1) with even operators T through the lens of convex integration techniques.

M.D. is partially supported by the NSF grants DMS–2009422 and DMS–2308208.

A pair (θ, u) is a weak solution of (1.1) if the equations in (1.1) are satisfied in the distributional sense. In the inviscid case $\nu = 0$, the existence of global weak solution for the active scalar equation with even operators is a challenging problem in the framework of classical energy method. Nevertheless, it was shown in [36] that there are infinitely many bounded weak solutions for (1.1) with $\nu = 0$ and $f \equiv 0$ via the application of convex integration techniques, which were developed for Euler equations in [20, 22]. In separate works [15, 37], non-unique bounded weak solutions were also constructed for the 2D IPM equation based on the convex integration method. In the recent work [8] for the 2D IPM, the authors obtained infinitely many mixing solutions in Sobolev space by combining convex integration, contour dynamics and pseudodifferential operators techniques. Similar result for the 2D IPM with curved interfaces was established in [10]. We emphasize that the convex integration method in the aforementioned works is rooted in the Tartar framework through the concept of subsolution. In particular, the solutions constructed in [15, 36, 37] are in the space $L^\infty_{t,x}$.

In the time line of the progress toward solving Onsager's conjecture (verifying $\frac{1}{3}$ Hölder regularity threshold for energy conservation) for the Euler equation, it was first shown the existence of bounded weak solutions violating the energy conservation in [22], then improvements were obtained in [12, 20, 21, 27, 2, 3] by producing continuous and C^α dissipative solutions for $\alpha < \frac{1}{5}$. Eventually dissipative solutions with spatial regularity C^α for any $\alpha < \frac{1}{3}$ were constructed in [28]. In this development, the improvement from constructing bounded weak solutions to Hölder continuous solutions relies on a crucial cancellation property which involves the use of stationary plane wave solutions for the Euler equation. The benefit of taking such plane waves as building blocks is that interference terms between different waves can be controlled.

Coming back to active scalar equations with even drift operators, an analogous Onsager's conjecture is that $\frac{1}{3}$ spatial regularity is expected to guarantee energy conservation for the solution. However, it is not trivial to adapt the ideas for the Euler equations in the hope of obtaining wild weak solutions that are continuous or C^α for $\alpha > 0$. The obstacle is that interference terms in different waves for active scalar equations can not be controlled efficiently due to the lack of a similar cancellation structure as that for the Euler equation. Remarkably the authors of [30] discovered a new mechanism for producing cancellations between overlapping waves for active scalar equations with even operators, and constructed non-unique dissipative solutions in C^α_x with $\alpha < \frac{1}{4d+1}$ for (1.1) with $\nu = 0$ and $f \equiv 0$. The cancellation comes from the vanishing of self-interaction terms which is due to the property $\xi \cdot m(\xi) = 0$ (the divergence free condition). Such a cancellation determines that the iteration convex integration scheme is essentially based on one dimensional oscillations. Hence considering the problem in d-dimension, i.e. on $\mathbb{T}^d$, requires d stages to correct the stress error in the iteration step from R_q to R_{q+1}. This explains why the spatial regularity obtained in [30] depends on dimension, which appears to be counter intuitive in some sense since one expects to have more flexibility in the higher dimensional case.

We note that the forced surface quasi-geostrophic (SQG) equation (an active scalar equation with odd drift operator) was recently studied in [6, 18, 19]. For the forced stationary SQG, it was shown that there are more than one solutions in the space C^α_x with $\alpha < -\frac{1}{4}$ in [18]. As a contrast, a non-trivial weak solution was constructed in C^α_x with $\alpha < -\frac{1}{3}$ for stationary SQG without external forcing in [11]. For the evolutionary SQG without external forcing, non-unique weak solutions with spatial regularity C^α_x for $\alpha < -\frac{1}{5}$ were constructed independently in [4] and [29]. While for the forced evolutionary SQG, the authors of [6] and [19] recently constructed non-unique weak solutions with spatial regularity C^α_x for $\alpha < 0$. The construction of [6] is in the framework of [4] and the construction of [19] is in the framework of [29]. Moreover in both [18] and [19], the authors exploit the flexibility due to the presence of forcing through the sum-difference formulation of two solutions for the underlying system, which was inspired by [17, 23]. Such sum-difference formulation will be adapted in the current paper as well. More details will be provided in Section 3.

Comparing the active scalar equations with even and odd drift operators, we observe that it seems much harder to construct weak solutions with higher regularity that violate uniqueness and the energy law in the odd case. The reason is that the cancellation property in the odd case presents an obstruction, see [4].

Among the active scalar equations with even operator, the MG equation is of particular interest since the operator T has an unbounded region in the Fourier space (c.f. [26]). Due to the unboundedness and evenness of T, ill-posedness for (1.1) with $\nu = 0$, $f \equiv 0$ and the MG operator T was shown in [26] in the sense that there is no Lipschitz solutions map at the initial time. While when $\nu > 0$, the unforced MG equation was shown well-posed in [25] since the diffusion term plays a dominant role. In the case of fractional diffusion for the MG equation, i.e. (1.1) with $\nu > 0$ and $f \equiv 0$, the authors of [24] identified a threshold value $\gamma = 1$ by proving that: the MG equation with $\gamma \in (1, 2)$ is locally well-posed, the MG equation with $\gamma \in (0, 1)$ is ill-posed and the MG equation with $\gamma = 1$ is globally well-posed for small initial data.

1.2. **Main result.** In this current paper we study the active scalar equation (1.1) with $\nu \geq 0$ and $f \not\equiv 0$. The purpose is to test whether the flexibility of allowing an external forcing can lead to the construction of wild solutions that reach the critical spatial regularity $\frac{1}{3}$ for the energy law. We adapt the cancellation mechanism discovered in [30] in our construction, and thus our result also depends on dimension. In particular, the wild solutions we obtain have spatial regularity C^α_x with $\alpha < \frac{1}{2d+1}$ and violate uniqueness. It is clear that this regularity is higher than $\frac{1}{4d+1}$ for the non-forced case in [30]. The improvement on the regularity for the forced equation is explained in the item (ii) of Subsection 3.2 below. The key idea is to design a particular increment in the iterative convex integration scheme such that $\theta_{q+1} = \theta_q$, which leads to improvement in the estimates of the stress errors. Such improvement can be seen by comparing the heuristic estimates (3.11) for the forced case and the error estimates on the stress error for the unforced case in Section 2.

The main result is stated below.

Theorem 1.1. *Let $\alpha < \frac{1}{2d+1}$, $0 \leq \gamma < 1-\alpha$ and $\zeta < \frac{1}{2d}$. There exists $f \in C^0_t C^{2\alpha-1}_x$ such that there are more than one solutions $\theta \in C^0_t C^\alpha_x(\mathbb{T}^d) \cap C^\zeta_t C^0_x(\mathbb{T}^d)$ to (1.1) with external forcing f.*

Theorem 1.1 implies that the forced MG equation with $\gamma < 1 - \frac{1}{2d+1} = \frac{6}{7}$ (since the dimension is $d = 3$) is ill-posed due to the lack of uniqueness. This result is complementary to the ill-posedness result shown in [24] for the unforced MG with $\gamma < 1$, where the solution map is not Lipschitz continuous.

The paper is organized as follows. Section 2 gives a heuristic analysis for the result of non-forced active scalar equation with even drift operators which was proven in [30]; Section 3 provides a heuristic analysis for the forced active scalar equation with an even operator. In Section 4 we lay out technical preparations and the main iterative process. Section 5 is devoted to the proof of the main iteration statement. Section 6 concludes the proof of Theorem 1.1.

2. Heuristics of non-uniqueness for (1.1) with $\nu = 0$ and $f \equiv 0$

We provide an outline of heuristics for the earlier result of Isett and Vicol in their article [30] concerning unforced active scalar equations. We present the analysis in the latter notation of [5] for the Navier-Stokes equation.

Theorem 2.1. *Let $\alpha < \frac{1}{4d+1}$ and I be an open time interval. There exist non-trivial solutions $\theta \in C^{\alpha}_{x,t}(\mathbb{T}^d \times \mathbb{R})$ to (1.1) with $\nu = 0$ and $f \equiv 0$ such that $\theta(x,t) = 0$ for $t \notin I$.*

First, consider the approximating systems in 2D

$$\begin{aligned} \partial_t \theta_q + \nabla \cdot (u_q \theta_q) &= \nabla \cdot \widetilde{R}_q, \\ u_q &= T[\theta_q]. \end{aligned} \tag{2.1}$$

Assume the image of the even part of the multiplier m contains d linearly independent vectors given by

$$A_j = m(\xi_j) + m(-\xi_j), \quad j = 1, 2, ..., d, \quad |\xi_j| = 1.$$

The stress vector $\widetilde{R}_q$ can be decomposed as

$$\widetilde{R}_q = c_{1,q} A_1 + c_{2,q} A_2 =: c_{1,q} A_1 + R_q.$$

Without loss of generality, we assume $|c_{1,q}| \geq |c_{2,q}|$. The goal is to construct a new solution such that the principal part $c_{1,q} A_1$ in the stress error gets reduced.

We specify the index $I = (k, \pm) \in \mathbb{Z} \times \{\pm\} := \Omega$. Denote $\bar{I} = (k, \mp)$. For $I \in \Omega$, let $\theta_{I,q+1}$ and ξ_I be the amplitude and phase functions respectively, satisfying

$$\theta_{\bar{I}} = \bar{\theta}_I, \quad \xi_{\bar{I}} = -\xi_I.$$

Moreover, ξ_I is advected by u_q on a short time interval τ_q with initial state $\widehat{\xi}_I$. The increment $\Theta_{q+1} = \theta_{q+1} - \theta_q$ is constructed to take the form

$$\begin{aligned} \Theta_{I,q+1} &= \mathbb{P}_{I,\lambda_{q+1}}(e^{i\lambda_{q+1}\xi_I}\theta_{I,q+1}) = e^{i\lambda_{q+1}\xi_I}(\theta_{I,q+1} + \delta\theta_{I,q+1}), \\ \Theta_{q+1} &= \sum_{I \in \Omega} \Theta_{I,q+1} \end{aligned}$$

where the error term $\delta\theta_{I,q+1}$ comes from the application of a microlocal lemma. Applying the microlocal lemma again yields

$$\begin{aligned} U_{I,q+1} &= T[\Theta_{I,q+1}] = e^{i\lambda_{q+1}\xi_I}(u_{I,q+1} + \delta u_I) \quad \text{with} \quad u_{I,q+1} = m(\nabla\xi_I)\theta_{I,q+1} \\ U_{q+1} &= T[\Theta_{q+1}] = \sum_{I \in \Omega} T[\Theta_{I,q+1}]. \end{aligned}$$

We also have

$$u_{q+1} = u_q + U_{q+1} = T[\theta_q] + T[\Theta_{q+1}].$$

The tuple $(\theta_{q+1}, u_{q+1}, \widetilde{R}_{q+1})$ is a solution of (2.1) with q replaced by $q+1$ and with the new stress error $\widetilde{R}_{q+1}$ satisfying

$$\begin{aligned}\nabla\cdot\widetilde{R}_{q+1} =&(\partial_t+u_q\cdot\nabla)\,\Theta_{q+1}+\nabla\cdot(U_{q+1}\theta_q)\\&+\nabla\cdot\sum_{J\neq\bar{I}}U_{J,q+1}\Theta_{I,q+1}\\&+\nabla\cdot\sum_{I\in\Omega}\left(U_{I,q+1}\Theta_{\bar{I},q+1}+c_{1,q}A_1+R_q\right)\\=:&\,\nabla\cdot R_T+\nabla\cdot R_N+\nabla\cdot R_H+\nabla\cdot R_S.\end{aligned}\tag{2.2}$$

For parameters $\lambda_0 \gg 1$, $b > 1$ and $0 < \beta < 1$, define

$$\lambda_q = \left\lceil \lambda_0^{b^q} \right\rceil, \quad q \in \mathbb{N}\cup\{0\}$$

and let $\delta_q = \lambda_q^{-\beta}$. Assume θ_q and u_q are localized to frequency $\sim \lambda_q$. The cancellation $U_{I,q+1}\Theta_{\bar{I},q+1} + c_{1,q}A_1$ suggests the scaling $|c_{1,q}| \sim |\theta_{I,q+1}|^2$. We make the inductive assumptions:

$$\|\nabla^k u_q\|_{C^0} + \|\nabla^k\theta_q\|_{C^0} \lesssim \lambda_q^k\delta_{q-1}^{\frac12}, \quad k=1,2,...,L, \tag{2.3}$$

$$\|\nabla^k(\partial_t+u_q\cdot\nabla)\theta_q\|_{C^0} \lesssim \lambda_q^{k+1}\delta_{q-1}, \quad k=0,1,2,...,L-1, \tag{2.4}$$

$$\|\nabla^k c_{1,q}\|_{C^0} \lesssim \lambda_q^k\delta_q, \quad k=1,2,...,L, \tag{2.5}$$

$$\|\nabla^k(\partial_t+u_q\cdot\nabla)c_{1,q}\|_{C^0} \lesssim \lambda_q^{k+1}\delta_{q-1}^{\frac12}\delta_q, \quad k=0,1,2,...,L-1, \tag{2.6}$$

$$\|\nabla^k R_q\|_{C^0} \lesssim \lambda_q^k\delta_{q+1}, \quad k=1,2,...,L, \tag{2.7}$$

$$\|\nabla^k(\partial_t+u_q\cdot\nabla)R_q\|_{C^0} \lesssim \lambda_q^{k+1}\delta_{q-1}^{\frac12}\delta_{q+1}, \quad k=0,1,2,...,L-1. \tag{2.8}$$

The increments Θ_{q+1} and $U_{q+1} = T[\Theta_{q+1}]$ satisfy

$$\|\nabla^k\Theta_{q+1}\|_{C^0} + \|\nabla^k U_{q+1}\|_{C^0} \lesssim \lambda_{q+1}^k\delta_q^{\frac12}, \quad k=0,1, \tag{2.9}$$

$$\|(\partial_t+u_q\cdot\nabla)\Theta_{q+1}\|_{C^0} + \|(\partial_t+u_q\cdot\nabla)U_{q+1}\|_{C^0} \lesssim \tau_q^{-1}\delta_q^{\frac12} \tag{2.10}$$

where the time scale τ_q is to be determined in the following.

Since we can find R_T and R_N such that

$$R_T = \nabla\Delta^{-1}\mathbb{P}_{\sim\lambda_{q+1}}[(\partial_t+u_q\cdot\nabla)\Theta_{q+1}]$$

$$R_N = \nabla\Delta^{-1}(U_{q+1}\cdot\nabla\theta_q),$$

it follows from (2.3), (2.9) and (2.10)

$$\begin{aligned}\|R_T\|_{C^0} &\lesssim \lambda_{q+1}^{-1}\|(\partial_t+u_q\cdot\nabla)\Theta_{q+1}\|_{C^0} \lesssim \lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac12},\\ \|R_N\|_{C^0} &\lesssim \lambda_{q+1}^{-1}\|U_{q+1}\cdot\nabla\theta_q\|_{C^0} \lesssim \lambda_{q+1}^{-1}\delta_q^{\frac12}\lambda_q\delta_{q-1}^{\frac12}.\end{aligned}$$

On the other hand, we have

$$
\begin{aligned}
\|R_H\|_{C^0} &= \left\| \sum_{J\neq \bar{I}} U_{J,q+1}\Theta_{I,q+1} \right\|_{C^0} \\
&\lesssim \sum_I \|\theta_{I,q+1}\|_{C^0}^2 \left(\|m(\nabla\xi_I) - m(\nabla\widehat{\xi}_I)\|_{C^0} + \|\nabla\xi_I - \nabla\widehat{\xi}_I\|_{C^0} \right) \\
&\lesssim \sum_I \|\theta_{I,q+1}\|_{C^0}^2 \|\nabla\xi_I - \nabla\widehat{\xi}_I\|_{C^0} \\
&\lesssim \sum_I \|\theta_{I,q+1}\|_{C^0}^2 \lambda_q\tau_q\|u_q\|_{C^0} \\
&\lesssim \delta_q\lambda_q\delta_{q-1}^{\frac{1}{2}}\tau_q.
\end{aligned}
$$

To balance the error R_T and R_H, we choose $\tau_q = \delta_{q-1}^{-\frac14}\delta_q^{-\frac14}\lambda_q^{-\frac12}\lambda_{q+1}^{-\frac12}$ such that

$$
\lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac12} \sim \delta_q\lambda_q\delta_{q-1}^{\frac12}\tau_q.
$$

In the end, we observe (up to small errors)

$$
R_S = c_{2,q+1}A_2
$$

for some coefficient $c_{2,q+1}$ with $|c_{2,q+1}| \leq \delta_{q+1}$. We then denote $R_{q+1} = R_T + R_N + R_H$. Combining the estimates above gives

$$
\begin{aligned}
\|R_{q+1}\|_{C^0} &\lesssim \lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac12} + \lambda_{q+1}^{-1}\delta_q^{\frac12}\lambda_q\delta_{q-1}^{\frac12} \\
&\lesssim \lambda_{q+1}^{-\frac12}\lambda_q^{\frac12}\delta_q^{\frac34}\delta_{q-1}^{\frac14} + \lambda_{q+1}^{-1}\delta_q^{\frac12}\lambda_q\delta_{q-1}^{\frac12} \\
&\lesssim \lambda_q^{-\frac12 b+\frac12-\frac34\beta-\frac{\beta}{4b}} + \lambda_q^{-b+1-\frac12\beta-\frac{\beta}{2b}}.
\end{aligned}
$$

To make sure $\|R_{q+1}\|_{C^0} \lesssim \delta_{q+2}$, we require

$$
\begin{cases}
-\frac12 b + \frac12 - \frac34\beta - \frac{\beta}{4b} < -b^2\beta, \\
-b + 1 - \frac12\beta - \frac{\beta}{2b} < -b^2\beta.
\end{cases}
$$

Thus we solve, by recalling $b > 1$

$$
\begin{aligned}
& b^2\beta - \frac12 b + \frac12 - \frac34\beta - \frac{\beta}{4b} < 0 \\
\iff & b\beta(b-1) + \beta(b-1) + \frac{\beta}{4b}(b-1) - \frac12(b-1) < 0 \\
\iff & b\beta + \beta + \frac{\beta}{4b} - \frac12 < 0 \\
\iff & \beta < \frac{1}{2b+2+\frac{1}{2b}}.
\end{aligned}
$$

When $b = 1^+$, the inequality above implies $\beta < \frac{2}{9}$. Similarly, the other inequality gives

$$\begin{aligned} & b^2\beta - b + 1 - \frac{1}{2}\beta - \frac{\beta}{2b} < 0 \\ \Longleftrightarrow\ & b\beta(b-1) + \beta(b-1) + \frac{\beta}{2b}(b-1) - \frac{1}{2}(b-1) < 0 \\ \Longleftrightarrow\ & b\beta + \beta + \frac{\beta}{2b} - 1 < 0 \\ \Longleftrightarrow\ & \beta < \frac{1}{b + 1 + \frac{1}{2b}} \end{aligned}$$

which indicates $\beta < \frac{2}{5}$ for $b = 1^+$. The C^α of Θ_{q+1} requires

$$\|\Theta_{q+1}\|_{C^\alpha} \lesssim \lambda_{q+1}^\alpha \|\Theta_{q+1}\|_{C^0} \lesssim \lambda_{q+1}^\alpha \delta_q^{\frac{1}{2}} \lesssim \lambda_q^{b\alpha - \frac{1}{2}\beta} \lesssim 1$$

which leads to $\alpha < \frac{\beta}{2b} < \frac{1}{9}$.

In d-dimension, we need to make sure $\|R_{q+1}\|_{C^0} \lesssim \delta_{q+d}$ in order to carry on the iteration, and hence require

$$\begin{cases} -\frac{1}{2}b + \frac{1}{2} - \frac{3}{4}\beta - \frac{\beta}{4b} < -b^d\beta, \\ -b + 1 - \frac{1}{2}\beta - \frac{\beta}{2b} < -b^d\beta. \end{cases}$$

The first inequality is equivalent to

$$\begin{aligned} & b^d\beta - b + 1 - \frac{1}{2}\beta - \frac{\beta}{2b} < 0 \\ \Longleftrightarrow\ & b^{d-1}\beta(b-1) + b^{d-2}\beta(b-1) + ... + \beta(b-1) + \frac{\beta}{4b}(b-1) - \frac{1}{2}(b-1) < 0 \\ \Longleftrightarrow\ & b^{d-1}\beta + b^{d-2}\beta + ... + \beta + \frac{\beta}{4b} - \frac{1}{2} < 0 \\ \Longleftrightarrow\ & \beta < \frac{1}{2(b^{d-1} + b^{d-2} + ... + 1 + \frac{1}{4b})}, \end{aligned}$$

following which we have $\beta < 1/(2d + \frac{1}{2})$ for $b > 1$. Similarly the second inequality is equivalent to

$$\beta < \frac{1}{b^{d-1} + b^{d-2} + ... + 1 + \frac{1}{2b}}$$

and hence $\beta < 1/(d + \frac{1}{2})$ for $b > 1$. Combining the two conditions yields $\beta < \frac{2}{4d+1}$ and hence $\alpha < \frac{\beta}{2b} < \frac{1}{4d+1}$.

3. Outline of non-uniqueness constructions for forced equation (1.1)

In this section we sketch a generic convex integration scheme for forced active scalar equations with even operators. We will explore the flexibility in the convex integration construction due to the presence of an external forcing. Such flexibility was exploited in the previous works [18, 19] through the sum-difference formulation of two distinct solutions for SQG. We note an alternating formulation of convex integration techniques was used in [6] for forced SQG.

3.1. **Sum-difference system of two solutions.** Assume (θ, u) and $(\widetilde{\theta}, \widetilde{u})$ are two distinct solutions of (1.1). The new variables

$$P = \frac{1}{2}(\theta + \widetilde{\theta}), \quad M = \frac{1}{2}(\theta - \widetilde{\theta})$$

satisfy the system

$$\begin{aligned} P_t + T[P] \cdot \nabla P + T[M] \cdot \nabla M &= -\nu \Lambda^\gamma P + f, \\ M_t + T[P] \cdot \nabla M + T[M] \cdot \nabla P &= -\nu \Lambda^\gamma M, \\ \nabla \cdot T[P] = 0, \quad \nabla \cdot T[M] &= 0. \end{aligned} \tag{3.1}$$

Allowing forcing in the equation of M, we have the flexibility to find a pair (θ, u) and $(\widetilde{\theta}, \widetilde{u})$ with $\theta - \widetilde{\theta} \not\equiv 0$, satisfying the relaxed system

$$\begin{aligned} P_t + T[P] \cdot \nabla P + T[M] \cdot \nabla M &= -\nu \Lambda^\gamma P + f_1, \\ M_t + T[P] \cdot \nabla M + T[M] \cdot \nabla P &= -\nu \Lambda^\gamma M + f_2, \\ \nabla \cdot T[P] = 0, \quad \nabla \cdot T[M] &= 0 \end{aligned} \tag{3.2}$$

for some external forcing functions f_1 and f_2. We then apply a convex integration scheme to the equation of M with the aim to erase the forcing f_2 iteratively, and eventually arrive at the system (3.1).

3.2. **The convex integration scheme.** For $f_2 \not\equiv 0$, we will apply a convex integration scheme to system (3.2) with the aim of reducing the forcing f_2 in the second equation. We thus consider the approximating system

$$\begin{aligned} \partial_t P_q + T[P_q] \cdot \nabla P_q + T[M_q] \cdot \nabla M_q &= -\nu \Lambda^\gamma P_q + \nabla \cdot \bar{R}_q, \\ \partial_t M_q + T[P_q] \cdot \nabla M_q + T[M_q] \cdot \nabla P_q &= -\nu \Lambda^\gamma M_q + \nabla \cdot \widetilde{R}_q \end{aligned} \tag{3.3}$$

inductively. Consistent with notation, we have

$$\theta_q = P_q + M_q, \quad \widetilde{\theta}_q = P_q - M_q,$$

$$T[\theta_q] = T[P_q] + T[M_q] = u_q, \quad T[\widetilde{\theta}_q] = T[P_q] - T[M_q] = \widetilde{u}_q.$$

Due to the presence of forcing terms in both equations, we have the abundance to find an initial tuple $(P_0, M_0, \bar{R}_0, \widetilde{R}_0)$ with $M_0 \not\equiv 0$ satisfying (3.3). Starting from this tuple, we construct another solution $(P_1, M_1, \bar{R}_1, \widetilde{R}_1)$ of (3.3) with $\widetilde{R}_1$ smaller than $\widetilde{R}_0$ in an appropriate way. Without loss of generality, assume $(P_q, M_q, \bar{R}_q, \widetilde{R}_q)$ satisfies (3.3) for an even integer q. To take the advantage of the flexibility of having two unknown variables, each stage of the construction consists two steps: from $(P_q, M_q, \bar{R}_q, \widetilde{R}_q)$ to $(P_{q+1}, M_{q+1}, \bar{R}_{q+1}, \widetilde{R}_{q+1})$ and from $(P_{q+1}, M_{q+1}, \bar{R}_{q+1}, \widetilde{R}_{q+1})$ to $(P_{q+2}, M_{q+2}, \bar{R}_{q+2}, \widetilde{R}_{q+2})$. In particular, we construct W_{q+1} and W_{q+2} such that

$$M_{q+1} = M_q + W_{q+1}, \quad P_{q+1} = P_q - W_{q+1}$$

and

$$M_{q+2} = M_{q+1} + W_{q+2}, \quad P_{q+2} = P_{q+1} + W_{q+2}.$$

Consequently we note, for even q

$$\begin{aligned} \theta_{q+1} &= P_{q+1} + M_{q+1} = P_q + M_q = \theta_q, \\ \widetilde{\theta}_{q+1} &= P_{q+1} - M_{q+1} = P_q - M_q - 2W_{q+1} = \widetilde{\theta}_q - 2W_{q+1} \end{aligned} \tag{3.4}$$

and

$$\begin{aligned} \theta_{q+2} &= P_{q+2} + M_{q+2} = P_{q+1} + M_{q+1} + 2W_{q+2} = \theta_{q+1} + 2W_{q+2}, \\ \widetilde{\theta}_{q+2} &= P_{q+2} - M_{q+2} = P_{q+1} - M_{q+1} = \widetilde{\theta}_{q+1}. \end{aligned} \tag{3.5}$$

The "pause" reflected in $\theta_{q+1} = \theta_q$ and $\widetilde{\theta}_{q+2} = \widetilde{\theta}_{q+1}$ will play a key role to gain better estimates in stress errors.

Since the three tuples $(P_{q+j}, M_{q+j}, \bar{R}_{q+j}, \widetilde{R}_{q+j})$ with $j = 0, 1, 2$ all satisfy (3.3), the stress terms can be expressed as

$$\begin{aligned} \nabla \cdot \widetilde{R}_{q+1} =& \left(\partial_t + T[\widetilde{\theta}_q] \cdot \nabla\right) W_{q+1} + \nu\Lambda^\gamma W_{q+1} + T[W_{q+1}] \cdot \nabla\widetilde{\theta}_q \\ &+ \left(\nabla \cdot \widetilde{R}_q - 2T[W_{q+1}] \cdot \nabla W_{q+1}\right), \end{aligned} \tag{3.6}$$

$$\begin{aligned} \nabla \cdot \widetilde{R}_{q+2} =& \left(\partial_t + T[\theta_{q+1}] \cdot \nabla\right) W_{q+2} + \nu\Lambda^\gamma W_{q+2} + T[W_{q+2}] \cdot \nabla\theta_{q+1} \\ &+ \left(\nabla \cdot \widetilde{R}_{q+1} + 2T[W_{q+2}] \cdot \nabla W_{q+2}\right), \end{aligned} \tag{3.7}$$

$$\begin{aligned} \nabla \cdot \bar{R}_{q+1} =& -\left(\partial_t + T[\widetilde{\theta}_q] \cdot \nabla\right) W_{q+1} - \nu\Lambda^\gamma W_{q+1} - T[W_{q+1}] \cdot \nabla\widetilde{\theta}_q \\ &+ \left(\nabla \cdot \bar{R}_q + 2T[W_{q+1}] \cdot \nabla W_{q+1}\right), \end{aligned} \tag{3.8}$$

$$\begin{aligned} \nabla \cdot \bar{R}_{q+2} =& \left(\partial_t + T[\theta_{q+1}] \cdot \nabla\right) W_{q+2} + \nu\Lambda^\gamma W_{q+2} + T[W_{q+2}] \cdot \nabla\theta_{q+1} \\ &+ \left(\nabla \cdot \bar{R}_{q+1} + 2T[W_{q+2}] \cdot \nabla W_{q+2}\right). \end{aligned} \tag{3.9}$$

The forms of the stress terms above provide some insights on the construction of the increments in the two steps $q \to q+1$ and $q+1 \to q+2$:
(i) W_{q+j} with $j = 1, 2$ will be designed such that

$$\nabla \cdot \widetilde{R}_q - 2T[W_{q+1}] \cdot \nabla W_{q+1} \quad \text{and} \quad \nabla \cdot \widetilde{R}_{q+1} + 2T[W_{q+2}] \cdot \nabla W_{q+2} \tag{3.10}$$

are small;
(ii) in view of (3.5), we have $\widetilde{\theta}_q = \widetilde{\theta}_{q-1}$ in the iteration process and hence expect to have better estimates for the terms containing $\widetilde{\theta}_q$ in (3.6); similarly, thanks to $\theta_{q+1} = \theta_q$ in (3.4), better estimates may be achieved for $\widetilde{R}_{q+2}$ as in (3.7);
(iii) comparing (3.6) and (3.8), the reduced stress error in the process $\widetilde{R}_q \to \widetilde{R}_{q+1}$ is gained in the process $R_q \to R_{q+1}$; while according to (3.7) and (3.9), both processes $\widetilde{R}_{q+1} \to \widetilde{R}_{q+2}$ and $R_{q+1} \to R_{q+2}$ have stress error reduced by the same amount. It indicates that this scheme is likely to reduce the forcing in one equation, but not in both equations.

3.3. **Heuristics.** Now we estimate $\widetilde{R}_{q+1}$ given in (3.6) using the "pause" feature $T[\widetilde{\theta}_q] = T[\widetilde{\theta}_{q-1}]$ described in item (ii) above. As before, we write

$$\nabla \cdot \widetilde{R}_{q+1} = \nabla \cdot R_T + \nabla \cdot R_N + \nabla \cdot R_H + \nabla \cdot R_S + \nabla \cdot R_D$$

with

$$\begin{aligned} \nabla \cdot R_T &= \left(\partial_t + T[\widetilde{\theta}_{q-1}] \cdot \nabla\right) W_{q+1} \\ \nabla \cdot R_N &= T[W_{q+1}] \cdot \nabla\widetilde{\theta}_{q-1} \\ \nabla \cdot R_H &= \nabla \cdot \sum_{J \neq \bar{I}} U_{J,q+1}\Theta_{I,q+1} \\ \nabla \cdot R_D &= \nu\Lambda^\gamma W_{q+1} \end{aligned}$$

and R_S similar as in (2.2). We point out that the forms of the stress errors here are not exactly the same as the stress errors in the rigorous analysis to be carried out in Section 5, since we need to include an additional step of mollification. Nevertheless, the scalings of the stress errors here and those in Section 5 are consistent. Applying (2.3), (2.9) and (2.10) gives

$$\begin{aligned}
\|R_T\|_{C^0} &\lesssim \lambda_{q+1}^{-1}\|(\partial_t + \widetilde{u}_{q-1}\cdot\nabla)\Theta_{q+1}\|_{C^0} \lesssim \lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac{1}{2}},\\
\|R_N\|_{C^0} &\lesssim \lambda_{q+1}^{-1}\|U_{q+1}\cdot\nabla\widetilde{\theta}_{q-1}\|_{C^0} \lesssim \lambda_{q+1}^{-1}\delta_q^{\frac{1}{2}}\lambda_{q-1}\delta_{q-2}^{\frac{1}{2}},\\
\|R_D\|_{C^0} &\lesssim \lambda_{q+1}^{-1+\gamma}\|W_{q+1}\|_{C^0} \lesssim \lambda_{q+1}^{-1+\gamma}\delta_q^{\frac{1}{2}},\\
\|R_H\|_{C^0} &\lesssim \sum_I \|\theta_{I,q+1}\|_{C^0}^2\lambda_{q-1}\tau_q\|\widetilde{u}_{q-1}\|_{C^0} \lesssim \delta_q\lambda_{q-1}\delta_{q-2}^{\frac{1}{2}}\tau_q
\end{aligned} \tag{3.11}$$

where in the last inequality we used the fact that ξ_I is now advected by $\widetilde{u}_{q-1}$. To balance the estimates of $\|R_T\|_{C^0}$ and $\|R_H\|_{C^0}$, we set

$$\lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac{1}{2}} = \delta_q\lambda_{q-1}\delta_{q-2}^{\frac{1}{2}}\tau_q$$

which implies

$$\tau_q = \lambda_{q-1}^{-\frac{1}{2}}\lambda_{q+1}^{-\frac{1}{2}}\delta_{q-2}^{-\frac{1}{4}}\delta_q^{-\frac{1}{4}}.$$

In the end, to obtain

$$\|R_{q+1}\|_{C^0} \lesssim \delta_{q+2}$$

we need to require

$$\begin{cases}
\lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac{1}{2}} \lesssim \delta_{q+2},\\
\lambda_{q+1}^{-1}\delta_q^{\frac{1}{2}}\lambda_{q-1}\delta_{q-2}^{\frac{1}{2}} \lesssim \delta_{q+2},\\
\lambda_{q+1}^{-1+\gamma}\delta_q^{\frac{1}{2}} \lesssim \delta_{q+2}
\end{cases}$$

which are valid provided

$$\begin{aligned}
&b^2\beta - \frac{3}{4}\beta - \frac{\beta}{4b^2} + \frac{1}{2b} - \frac{b}{2} < 0\\
\Longleftrightarrow\ & b\beta(b-1) + \beta(b-1) + \frac{\beta}{4b^2}(b^2-1) - \frac{1}{2b}(b^2-1) < 0\\
\Longleftrightarrow\ & b\beta + \beta + \frac{\beta}{4b^2}(b+1) - \frac{1}{2b}(b+1) < 0\\
\Longleftrightarrow\ & \beta < \frac{\frac{1}{2b}(b+1)}{b+1+\frac{b+1}{4b^2}}
\end{aligned}$$

and

$$b(-1+\gamma) - \frac{1}{2}\beta + b^2\beta < 0 \Longleftrightarrow \beta < \frac{2b(1-\gamma)}{2b^2-1}.$$

So for $b = 1^+$, we have $\beta < \frac{2}{5}$, $\alpha < \frac{\beta}{2b} < \frac{1}{5}$, and $\gamma < 1-\alpha$.

In general for d-dimension, we need to impose

$$\|R_{q+1}\|_{C^0} \lesssim \delta_{q+d}$$

and hence

$$\begin{cases}
\lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac{1}{2}} \lesssim \delta_{q+d},\\
\lambda_{q+1}^{-1}\delta_q^{\frac{1}{2}}\lambda_{q-1}\delta_{q-2}^{\frac{1}{2}} \lesssim \delta_{q+d},\\
\lambda_{q+1}^{-1+\gamma}\delta_q^{\frac{1}{2}} \lesssim \delta_{q+d}.
\end{cases}$$

Thus we have

$$\begin{aligned} & b^d\beta - \frac{3}{4}\beta - \frac{\beta}{4b^2} + \frac{1}{2b} - \frac{b}{2} < 0 \\ \iff & b^{d-1}\beta(b-1) + ... + \beta(b-1) + \frac{\beta}{4b^2}(b^2-1) - \frac{1}{2b}(b^2-1) < 0 \\ \iff & b^{d-1}\beta + ... + b\beta + \beta + \frac{\beta}{4b^2}(b+1) - \frac{1}{2b}(b+1) < 0 \\ \iff & \beta < \frac{\frac{1}{2b}(b+1)}{b^{d-1} + ... + b + 1 + \frac{b+1}{4b^2}} \end{aligned}$$

and

$$b(-1+\gamma) - \frac{1}{2}\beta + b^d\beta < 0 \iff \beta < \frac{2b(1-\gamma)}{2b^d - 1}.$$

For $b = 1^+$, $\beta < \frac{1}{d+\frac{1}{2}}$, $\alpha < \frac{\beta}{2b} < \frac{1}{2d+1}$, and $\gamma < 1 - \alpha$.

3.4. **Key idea to reduce the stress error.** We discuss how to achieve the error reduction in (3.10). The assumptions on the Fourier symbol m of the operator T imply that there are two linearly independent vectors in the image of the even part of m

$$A_1 = m(\xi^{(1)}) + m(-\xi^{(1)}), \quad A_2 = m(\xi^{(2)}) + m(-\xi^{(2)}) \tag{3.12}$$

with $\xi^{(1)}, \xi^{(2)} \in \mathbb{Z}^2$.

Consider the increment ansatz

$$\begin{aligned} W_{I,q+1} &= \mathbb{P}_{I,\lambda_{q+1}}\left(a_{I,q+1}(x,t)e^{i\lambda_{q+1}\xi_I}\right), \\ W_{q+1} &= \sum_{I\in\Omega} W_{I,q+1}. \end{aligned}$$

By the Microlocal Lemma 4.1 and zero degree of homogeneity of m, the drift term takes the form

$$T[W_{I,q+1}] = m(\nabla\xi_I)W_{I,q+1} + \delta T[W_{I,q+1}]$$

with a small error term $\delta T[W_{I,q+1}]$. A straightforward computation shows that

$$\begin{aligned} T[W_{q+1}]W_{q+1} =& \frac{1}{2}\sum_{I\in\Omega}\left(T[W_{I,q+1}]W_{\bar{I},q+1} + T[W_{\bar{I},q+1}]W_{I,q+1}\right) \\ &+ \sum_{J\neq\bar{I}} T[W_{I,q+1}]W_{J,q+1} \\ =& \frac{1}{2}\sum_{I\in\Omega}|a_{I,q+1}|^2\left(m(\nabla\xi_I) + m(-\nabla\xi_I)\right) + \text{error} \\ &+ \sum_{J\neq\bar{I}} T[W_{I,q+1}]W_{J,q+1}. \end{aligned}$$

Since m is not odd, the leading order (low frequency) term is non zero,

$$|a_{I,q+1}|^2\left(m(\nabla\xi_I) + m(-\nabla\xi_I)\right) \neq 0.$$

The goal is to construct coefficient functions $a_{I,q+1}$ and phase functions ξ_I such that

$$\nabla\cdot\left(\widetilde{R}_q - \sum_{I\in\Omega}|a_{I,q+1}|^2\left(m(\nabla\xi_I) + m(-\nabla\xi_I)\right)\right) \tag{3.13}$$

is small. As A_1 and A_2 defined in (3.12) span $\mathbb{R}^2$, we expect to choose ξ_I to guarantee

$$m(\nabla\xi_I) + m(-\nabla\xi_I) \approx A_1 \quad \text{or} \quad m(\nabla\xi_I) + m(-\nabla\xi_I) \approx A_2.$$

Then with an appropriate choice of $a_{I,q+1}$ we hope the principal part of $\widetilde{R}_q$ can be canceled through (3.13). The reduction of $\widetilde{R}_{q+1}$ is achieved analogously.

To ensure other terms in the stress fields given in (3.6) and (3.7) can be controlled appropriately, we need several technical tools which will be provided in Section 4.

4. Construction of the increment

We describe the construction of the highly oscillatory correction (increment) in this section. We start with some technical preparations.

4.1. **Microlocal Lemma.** The drift operator T is a nonlocal differential operator. When acting T on plane waves, we need the following lemma from [30] to extract the leading order term.

Lemma 4.1 (Microlocal Lemma)**.** *Let $K : \mathbb{R}^2 \to \mathbb{C}$ be a Schwartz function and*

$$T[\Theta](x) = \int_{\mathbb{R}^2} \Theta(x-h)K(h)\,dh$$

for $\Theta : \mathbb{T}^2 \to \mathbb{C}$. For any $\Theta = e^{i\lambda\xi(x)}\theta(x)$ with $\lambda \in \mathbb{Z}$ and smooth functions $\xi : \mathbb{T}^2 \to \mathbb{R}$ and $\theta : \mathbb{T}^2 \to \mathbb{C}$, we have

$$T[\Theta](x) = e^{i\lambda\xi(x)}\left(\theta(x)\widehat{K}(\lambda\nabla\xi) + \delta[T\Theta](x)\right)$$

with the error term given by

$$\delta[T\Theta](x) = \int_0^1 \frac{d}{dr}\int_{\mathbb{R}^2} e^{-i\lambda\nabla\xi\cdot h} e^{iZ(r,x,h)}\theta(x-rh)K(h)\,dr\,dh$$

$$Z(r,x,h) = r\lambda\int_0^1 h^j h^l \partial_j\partial_l\xi(x-sh)(1-s)\,ds.$$

4.2. **Mollification.** As is standard in convex integration method, to avoid loss of derivative, we need to regularize the solution $(P_q, M_q, \bar{R}_q, \widetilde{R}_q)$ before adding increments to produce $(P_{q+1}, M_{q+1}, \bar{R}_{q+1}, \widetilde{R}_{q+1})$. We first regularize $P_q, M_q, T[P_q]$ and $T[M_q]$. Fix some $L \geq 1$. Choose $\mu_q = \lambda_{q+1}^{\frac{1}{L}}\lambda_q^{1-\frac{1}{L}}$. Denote $\mathbb{P}_{\leq\mu_q}$ by the Littlewood-Paley projection onto frequency $\leq \mu_q$. Define

$$P_{\epsilon,q} = \mathbb{P}^2_{\leq\mu_q}P_q, \quad M_{\epsilon,q} = \mathbb{P}^2_{\leq\mu_q}M_q,$$
$$T[P_{\epsilon,q}] = \mathbb{P}^2_{\leq\mu_q}T[P_q], \quad T[M_{\epsilon,q}] = \mathbb{P}^2_{\leq\mu_q}T[M_q].$$

Lemma 4.2. *The estimates*

$$\|P_q - P_{\epsilon,q}\|_{C^0} \lesssim \mu_q^{-j}\|\nabla^j P_q\|_{C^0},$$
$$\|M_q - M_{\epsilon,q}\|_{C^0} \lesssim \mu_q^{-j}\|\nabla^j M_q\|_{C^0},$$
$$\|T[P_q] - T[P_{\epsilon,q}]\|_{C^0} \lesssim \mu_q^{-j}\|\nabla^j T[P_q]\|_{C^0},$$
$$\|T[M_q] - T[M_{\epsilon,q}]\|_{C^0} \lesssim \mu_q^{-j}\|\nabla^j T[M_q]\|_{C^0}$$

hold for $0 \leq j \leq L$.

The regularization of $\widetilde{R}_q$ is slightly more involved. We first define the coarse scale flow $\Phi_q(x,s;t) : \mathbb{T}^2 \times \mathbb{R} \times \mathbb{R} \to \mathbb{T}^2 \times \mathbb{R}$ to be the solution of

$$\begin{cases} \partial_s \Phi_q(x,s;t) = \widetilde{u}_q(\Phi_q(x,s;t),s) \\ \Phi_q(x,t;t) = x \end{cases}$$

with $\widetilde{u}_q = T[M_q]$. Let η be a standard mollifier in space with $supp\, \eta \subset B(0,1)$ and ρ be a standard mollifier in time with $supp\, \rho \subset (-1,1)$. Denote

$$\eta_\delta(x) = \delta^{-d}\eta(\delta^{-1}x), \quad \rho_\delta(s) = \delta^{-1}\rho(\delta^{-1}s).$$

Without loss of generality, assume $\widetilde{R}_q$ can be decomposed as

$$\widetilde{R}_q = c_{1,q}A_1 + R_q^*.$$

For the spatial scale ϵ_x and time scale ϵ_t to be determined later, we define the regularized stress and component

$$\begin{aligned} R^*_{q,\epsilon_x} &= \eta_{\epsilon_x} * R_q^*, \\ R^*_{q,\epsilon}(x,t) &= \int_{\mathbb{R}} R^*_{q,\epsilon_x}(\Phi_q(x,t+s;t),t+s)\rho_{\epsilon_t}(s)\,ds, \\ c_{1,q,\epsilon_x} &= \eta_{\epsilon_x} * c_{1,q}, \\ c_{1,q,\epsilon}(x,t) &= \int_{\mathbb{R}} c_{1,q,\epsilon_x}(\Phi_q(x,t+s;t),t+s)\rho_{\epsilon_t}(s)\,ds, \end{aligned}$$

and

$$\widetilde{R}_{q,\epsilon} = c_{1,q,\epsilon}A_1 + R^*_{q,\epsilon}.$$

The purpose of such regularization is to have better estimates on the advective derivatives, since the advective derivative commutes with the flow map Φ. See [27] for more details.

Choose the length and time scales

$$\epsilon_x = (\lambda_{q+1}^{-1}\lambda_q)^{\frac{1}{L}}\lambda_q^{-1}, \quad \epsilon_t = \lambda_{q+1}^{-1}\delta_q^{-\frac{1}{2}}. \tag{4.1}$$

Lemma 4.3. *The regularized stress field satisfies the following estimates*

$$\begin{aligned} \|(c_{1,q,\epsilon} - c_{1,q})A_1\|_{C^0} + \|R^*_{q,\epsilon} - R^*_q\|_{C^0} &\lesssim \delta_{q-1}^{\frac{1}{2}}\delta_q^{\frac{1}{2}}\lambda_{q+1}^{-1}\lambda_q, \\ \|\nabla^k c_{1,q,\epsilon}\|_{C^0} &\lesssim \lambda_q^k \delta_q (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}, \\ \|\nabla^k R^*_{q,\epsilon}\|_{C^0} &\lesssim \lambda_q^k \delta_{q+1} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}, \\ \|\nabla^k \frac{D}{Dt} c_{1,q,\epsilon}\|_{C^0} &\lesssim \lambda_q^k \delta_q \lambda_q \delta_{q-1}^{\frac{1}{2}} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}, \\ \|\nabla^k \frac{D}{Dt} R^*_{q,\epsilon}\|_{C^0} &\lesssim \lambda_q^k \delta_{q+1} \lambda_q \delta_{q-1}^{\frac{1}{2}} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}, \\ \|\nabla^k (\frac{D}{Dt})^2 c_{1,q,\epsilon}\|_{C^0} &\lesssim \lambda_q^k \delta_q \lambda_q \delta_{q-1}^{\frac{1}{2}} \lambda_{q+1}\delta_q^{\frac{1}{2}} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}, \\ \|\nabla^k (\frac{D}{Dt})^2 R^*_{q,\epsilon}\|_{C^0} &\lesssim \lambda_q^k \delta_{q+1} \lambda_q \delta_{q-1}^{\frac{1}{2}} \lambda_{q+1}\delta_q^{\frac{1}{2}} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}. \end{aligned}$$

4.3. **Lifting function.** To achieve the cancellation in (3.13), we further decompose $R^*_{q,\epsilon}$ as

$$R^*_{q,\epsilon} = c_{1,q,r}A_1 + c_{2,q+1}A_2 \tag{4.2}$$

and hence

$$\widetilde{R}_{q,\epsilon} = (c_{1,q,\epsilon} + c_{1,q,r})\, A_1 + c_{2,q+1}A_2.$$

Then we require

$$\sum_I |a_{I,q+1}|^2 A_1 = e_q(t)A_1 + (c_{1,q,\epsilon} + c_{1,q,r})\, A_1$$

for some function $e_q(t)$ such that

$$e_q(t) + (c_{1,q,\epsilon} + c_{1,q,r}) > 0, \quad e_q(t) > 2\,(c_{1,q,\epsilon} + c_{1,q,r})\,.$$

The function e_q is referred to be the lifting function.

4.4. **Time cutoff.** To optimize the control of main error terms, we choose the life span τ_q of the increment $W_{I,q+1}$ to be an appropriate short time interval. The size of τ_q will be specified later. Let ϕ satisfy the condition of partition of unity in time

$$\sum_{n\in\mathbb{Z}} \phi^2(t-n) = 1.$$

Define

$$\phi_k(t) = \phi\left(\frac{t-k\tau_q}{\tau_q}\right)$$

and consider the amplitude function

$$a_{I,q+1} = e^{\frac12}(t)\phi_k(t)(1 + (c_{1,q,\epsilon} + c_{1,q,r})e^{-1}(t))^{\frac12}, \quad I = (k,\pm).$$

The choice of ϕ_k indicates that the amplitude function $a_{I,q+1}$ with $I = (k,f)$ has support $[k\tau_q - \frac23\tau_q, k\tau_q + \frac23\tau_q]$ in time.

4.5. **Phase functions.** Denote $\widetilde{u}_{q,\epsilon} = T[P_{q,\epsilon}] - T[M_{q,\epsilon}]$. We identify the phase functions ξ_I to be solutions of the transport equation

$$\begin{aligned}(\partial_t + \widetilde{u}_{q,\epsilon}\cdot\nabla)\xi_I &= 0\\ \xi_I(k\tau_q, x) &= \widehat{\xi}_I(x),\end{aligned} \tag{4.3}$$

where the initial data for the phase function is given by

$$\widehat{\xi}_I(x) = \widehat{\xi}_{(k,\pm)}(x) = \pm 10^{[k]}\xi^{(1)}\cdot x$$

with

$$[k] = \begin{cases} 0, & \text{if} \quad k \quad \text{is even,}\\ 1, & \text{if} \quad k \quad \text{is odd.}\end{cases}$$

In the iteration step of $q+1 \to q+2$, the initial data is chosen as

$$\widehat{\xi}_I(x) = \widehat{\xi}_{(k,\pm)}(x) = \pm 10^{[k]}\xi^{(2)}\cdot x.$$

The time scale τ_q will be chosen such

$$|\nabla\xi_I - \nabla\widehat{\xi}_I| \le \frac14|\nabla\widehat{\xi}_I|.$$

4.6. **Construction of the increment W_{q+1}.** We are ready to introduce the increment W_{q+1},

$$W_{q+1} = \sum_{I\in\Omega} W_{I,q+1}, \quad W_{I,q+1} = \mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \tag{4.4}$$

with the amplitude function

$$a_{I,q+1}(x,t) = e^{\frac{1}{2}}(t)\phi_k(t)\left(1+(c_{1,q,\epsilon}+c_{1,q,r})(x,t)e^{-1}(t)\right)^{\frac{1}{2}}, \quad I=(k,\pm). \tag{4.5}$$

The projection operator $\mathbb{P}_{\approx\lambda_{q+1}}$ is defined through its Fourier multiplier as

$$\widehat{\mathbb{P}_{\approx\lambda_{q+1}}f}(\xi) = \eta_{\lambda_{q+1}}(\xi)\hat{f}(\xi).$$

We specify the multiplier $\eta_{\lambda_{q+1}}$ in the following. Let η be a smooth bump function in Fourier space with frequency support on $B_{|\xi^{(1)}|/2}(\xi^{(1)})$, satisfying

$$\eta(\xi) = 1, \quad \text{if} \quad |\xi-\xi^{(1)}| \le \frac{1}{4}|\xi^{(1)}|.$$

Then we define

$$\eta_{\lambda_{q+1}}(\xi) = \eta(\pm 10^{-[k]}\lambda_{q+1}^{-1}\xi).$$

Applying the Microlocal Lemma 4.1 gives

$$W_{I,q+1} = e^{i\lambda_{q+1}\xi_I}(a_{I,q+1}+\delta a_{I,q+1}),$$
$$T[W_{I,q+1}] = e^{i\lambda_{q+1}\xi_I}(u_{I,q+1}+\delta u_{I,q+1})$$

with $u_{I,q+1} = m(\lambda_{q+1}\nabla\xi_I)a_{I,q+1} = m(\nabla\xi_I)a_{I,q+1}$ since m is homogeneous of order 0, where $\delta a_{I,q+1}$ and $\delta u_{I,q+1}$ are error terms.

4.7. **Main iteration process.** For the tuple $(P_q, M_q, \bar{R}_q, \widetilde{R}_q)$ satisfying (3.3), we make the following inductive assumptions. Assume $\widetilde{R}_q$ can be written as

$$\widetilde{R}_q = \begin{cases} c_{1,q}A_1 + R_q^*, & \text{if } q \text{ is even} \\ c_{2,q}A_2 + R_q^*, & \text{if } q \text{ is odd.} \end{cases} \tag{4.6}$$

Assume the estimates below hold:

$$\begin{aligned}&\|\nabla^k P_q\|_{C^0} + \|\nabla^k M_q\|_{C^0} \\ &+ \|\nabla^k T[P_q]\|_{C^0} + \|\nabla^k T[M_q]\|_{C^0} \lesssim \lambda_q^k \delta_{q-1}^{\frac{1}{2}}, \quad k=1,...,L\end{aligned} \tag{4.7}$$

$$\begin{aligned}&\|\nabla^k(\partial_t + T[P_q]\cdot\nabla)T[M_q]\|_{C^0} \\ &+ \|\nabla^k(\partial_t + T[M_q]\cdot\nabla)T[M_q]\|_{C^0} \lesssim \lambda_q^{k+1}\delta_{q-1}, \quad k=0,1,...,L-1\end{aligned} \tag{4.8}$$

$$\|\nabla^k c_{1,q}\|_{C^0} \lesssim \lambda_q^k \delta_q, \quad k=0,1,...,L \tag{4.9}$$

$$\begin{aligned}&\|\nabla^k(\partial_t + T[P_q]\cdot\nabla)c_{1,q}\|_{C^0} \\ &+ \|\nabla^k(\partial_t + T[M_q]\cdot\nabla)c_{1,q}\|_{C^0} \lesssim \lambda_q^{k+1}\delta_{q-1}^{\frac{1}{2}}\delta_q, \quad k=0,1,...,L-1\end{aligned} \tag{4.10}$$

$$\|\nabla^k R_q^*\|_{C^0} \lesssim \lambda_q^k \delta_{q+1}, \quad k=0,1,...,L \tag{4.11}$$

$$\begin{aligned}&\|\nabla^k(\partial_t + T[P_q]\cdot\nabla)R_q^*\|_{C^0} \\ &+ \|\nabla^k(\partial_t + T[M_q]\cdot\nabla)R_q^*\|_{C^0} \lesssim \lambda_q^{k+1}\delta_{q-1}^{\frac{1}{2}}\delta_{q+1}, \quad k=0,1,...,L-1.\end{aligned} \tag{4.12}$$

In view of the crucial feature (3.4)-(3.5) of the two-step scheme, we have for even q and $k = 0, 1, ..., L$

$$\begin{cases} \widetilde{\theta}_q = \widetilde{\theta}_{q-1}, & \|\nabla^k \widetilde{\theta}_q\|_{C^0} \lesssim \lambda_{q-1}^k \delta_{q-2}^{\frac{1}{2}}, \\ \theta_{q+1} = \theta_q, & \|\nabla^k \theta_q\|_{C^0} \lesssim \lambda_q^k \delta_{q-1}^{\frac{1}{2}}. \end{cases} \tag{4.13}$$

Proposition 4.4 (Main Iteration)**.** *Let $L \geq 2$ and $K, C \geq 4$. Assume $(P_q, M_q, \bar{R}_q, \widetilde{R}_q)$ satisfies (3.3) and (4.6)-(4.13). Let $I_t \subset \mathbb{R}$ be a nonempty closed interval such that*

$$supp\, R_q^* \cup supp\, c_{1,q} \subset I_t \times \mathbb{T}^2. \tag{4.14}$$

Let e be a function of time satisfying

$$e(t) \geq K\delta_q \quad \forall \quad t \in I_t \pm \hat{\tau}_q \tag{4.15}$$

with the natural time scale $\hat{\tau}_q = \lambda_q^{-1}\delta_{q-1}^{-\frac{1}{2}}$, and

$$\|\frac{d^r}{dt^r} e^{\frac{1}{2}}(t)\|_{C^0} \leq C(\lambda_q \delta_{q-1}^{\frac{1}{2}})^r \delta_q^{\frac{1}{2}}, \quad 0 \leq r \leq 2. \tag{4.16}$$

There exists another tuple $(P_{q+1}, M_{q+1}, \bar{R}_{q+1}, \widetilde{R}_{q+1})$ satisfying (3.3) with q replaced by $q+1$ in the form

$$\begin{cases} M_{q+1} = M_q + W_{q+1}, \quad P_{q+1} = P_q - W_{q+1}, \quad \text{if } q \text{ is even}, \\ M_{q+1} = M_q + W_{q+1}, \quad P_{q+1} = P_q + W_{q+1}, \quad \text{if } q \text{ is odd} \end{cases}$$

and $W_{q+1} = \nabla \cdot \widetilde{W}_{q+1}$. Moreover, $\widetilde{R}_{q+1}$ can be written as

$$\widetilde{R}_{q+1} = \begin{cases} c_{2,q+1} A_2 + R_{q+1}^*, & \text{if } q \text{ is even}, \\ c_{1,q+1} A_1 + R_{q+1}^*, & \text{if } q \text{ is odd} \end{cases} \tag{4.17}$$

with

$$(supp\, R_{q+1}^* \cup supp\, c_{j,q+1}) \subset supp\, e \times \mathbb{T}^2, \quad j = 1, 2. \tag{4.18}$$

The estimates (4.7)-(4.12) are satisfied with q replaced by $q+1$ and in particular $c_{1,q}$ replaced by $c_{2,q+1}$. The correction W_{q+1} satisfies the estimates

$$\|\nabla^k W_{q+1}\|_{C^0} + \|\nabla^k T[W_{q+1}]\|_{C^0} \lesssim \lambda_{q+1}^k \delta_q^{\frac{1}{2}}, \quad k = 0, 1, \tag{4.19}$$

$$\begin{aligned} &\|(\partial_t + T[P_{\epsilon,q}] \cdot \nabla) W_{q+1}\|_{C^0} + \|(\partial_t + T[P_{\epsilon,q}] \cdot \nabla) T[W_{q+1}]\|_{C^0} \\ &+ \|(\partial_t + T[M_{\epsilon,q}] \cdot \nabla) W_{q+1}\|_{C^0} + \|(\partial_t + T[M_{\epsilon,q}] \cdot \nabla) T[W_{q+1}]\|_{C^0} \\ \lesssim\ & \tau_q^{-1} \delta_q^{\frac{1}{2}}, \end{aligned} \tag{4.20}$$

$$\|\nabla^k \widetilde{W}_{q+1}\|_{C^0} \lesssim \lambda_{q+1}^{k-1} \delta_q^{\frac{1}{2}}, \quad k = 0, 1, \tag{4.21}$$

$$\begin{aligned} &\|(\partial_t + T[P_{\epsilon,q}] \cdot \nabla) \widetilde{W}_{q+1}\|_{C^0} \\ &+ \|(\partial_t + T[M_{\epsilon,q}] \cdot \nabla) \widetilde{W}_{q+1}\|_{C^0} \lesssim \tau_q^{-1} \lambda_{q+1}^{-1} \delta_q^{\frac{1}{2}}. \end{aligned} \tag{4.22}$$

5. Proof of the main iteration argument

5.1. Basic estimates of increments.

Lemma 5.1. *Let $L \geq 2$ be the integer in Proposition 4.4. The regularizations $T[P_{\epsilon,q}]$ and $T[M_{\epsilon,q}]$ satisfy*

$$\begin{aligned}
&\|\nabla^k T[P_{\epsilon,q}]\|_{C^0} + \|\nabla^k T[M_{\epsilon,q}]\|_{C^0}\\
&\lesssim \lambda_q^k \delta_{q-1}^{\frac{1}{2}} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k-L)_+}{L}}, \quad k \geq 1,\\
&\|\nabla^k (\partial_t + T[P_{\epsilon,q}]\cdot\nabla) T[M_{\epsilon,q}]\|_{C^0}\\
&+ \|\nabla^k (\partial_t + T[M_{\epsilon,q}]\cdot\nabla) T[M_{\epsilon,q}]\|_{C^0}\\
&\lesssim \lambda_q^{k+1} \delta_{q-1} (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1-L)_+}{L}}, \quad k \geq 0.
\end{aligned}$$

See [30] (Lemma 7.1) for a proof.

Lemma 5.2. *Let h be a kernel function satisfying*

$$\| |x|^a |\nabla^b h|(x)\|_{L^1(\mathbb{R}^2)} \leq \lambda^{b-a}, \quad \lambda \geq \lambda_{q+1}, \quad 0 \leq a \leq b \leq N.$$

Denote

$$\frac{D_{P,q,\epsilon}}{Dt} = \partial_t + T[P_{\epsilon,q}]\cdot\nabla, \quad \frac{D_{M,q,\epsilon}}{Dt} = \partial_t + T[M_{\epsilon,q}]\cdot\nabla.$$

For the convolution operator

$$Qf(x) = \int_{\mathbb{R}^2} f(x-y) h(y)\, dy,$$

the commutators $\left[\frac{D_{P,q,\epsilon}}{Dt}, Q\right]$ and $\left[\frac{D_{M,q,\epsilon}}{Dt}, Q\right]$ are bounded operators on $C^0(\mathbb{T}^2\times\mathbb{R})$ and satisfy

$$\left\|\nabla^k [\frac{D_{P,q,\epsilon}}{Dt}, Q]\right\| + \left\|\nabla^k [\frac{D_{M,q,\epsilon}}{Dt}, Q]\right\| \lesssim \lambda_q \delta_{q-1}^{\frac{1}{2}} \lambda^k, \quad 0 \leq k \leq N-1.$$

Lemma 5.3. *Let $L \geq 2$ be the integer in Proposition 4.4. For $\frac{D_q}{Dt} \in \{\frac{D_{P,q}}{Dt}, \frac{D_{M,q}}{Dt}\}$, define*

$$D_q^{(k,r)} = \nabla^{k_1} (\frac{D_q}{Dt})^{r_1} \nabla^{k_2} (\frac{D_q}{Dt})^{r_2} \nabla^{k_3}, \ k_1+k_2+k_3 = k, \ r_1+r_2 = r.$$

The phase function ξ_I satisfies

$$\begin{aligned}
&\|\nabla^k (\frac{D_q}{Dt})^r \nabla\xi_I\|_{C^0} + \|D_q^{(k,r)}\nabla\xi_I\|_{C^0}\\
&\lesssim \lambda_q^k (\lambda_q \delta_{q-1}^{\frac{1}{2}})^r (\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+1+(r-1)_+-L)_+}{L}}, \quad k \geq 1, \ r = 0,1,2.
\end{aligned}$$

Moreover, we have

$$|\nabla\xi_I(\Phi_q(x,s;t)) - \nabla\widehat{\xi}_I(x)| \leq C\tau_q \lambda_q \delta_{q-1}^{\frac{1}{2}}, \quad |s| \leq \tau_q.$$

Lemma 5.4. *Let $L \geq 2$. The principal part of the amplitude function satisfies the estimate*

$$\|D_{q'}^{(k,r)} a_{I,q+1}\|_{C^0} + \|D_{q'}^{(k,r)} u_{I,q+1}\|_{C^0} \lesssim \lambda_{q'}^k \delta_q^{\frac{1}{2}} \tau_{q'}^{-r} (\lambda_{q'+1}\lambda_{q'}^{-1})^{\frac{(k+1-L)_+}{L}}$$

for $q' \geq 0$, $k \geq 0$ and $r = 0,1,2$.

Lemma 5.5. *Let $L \geq 2$. The amplitude error terms $\delta a_{I,q+1}$ and $\delta u_{I,q+1}$ satisfy*

$$\|D_{q'}^{(k,r)}\delta a_{I,q+1}\|_{C^0} + \|D_{q'}^{(k,r)}\delta u_{I,q+1}\|_{C^0} \lesssim (\lambda_{q'+1}^{-1}\lambda_{q'})\lambda_q^k\delta_q^{\frac{1}{2}}\tau_{q'}^{-r}(\lambda_{q+1}\lambda_q^{-1})^{\frac{(k+2-L)_+}{L}}$$

for $q' \geq 0$, $k \geq 0$ and $r = 0, 1, 2$.

Lemma 5.6. *The estimates for the corrections W_{q+1} and $T[W_{q+1}]$*

$$\|D_{q'}^{(k,r)}W_{I,q+1}\|_{C^0} + \|D_{q'}^{(k,r)}T[W_{I,q+1}]\|_{C^0} \lesssim \lambda_{q'+1}^k\tau_{q'}^{-r}\delta_q^{\frac{1}{2}}$$

hold for $q' \geq 0$, $k \geq 0$ and $r = 0, 1, 2$.

The lemmas above can be proved analogously as in [30].

5.2. **Proof of Proposition 4.4.** We only prove the statements for even q; the statements for odd q can be established by minor modifications of the proof. Let W_{q+1} be the correction term constructed in Subsection 4.6 and define

$$M_{q+1} = M_q + W_{q+1}, \quad P_{q+1} = P_q - W_{q+1}.$$

For $\bar{R}_{q+1}$ and $\widetilde{R}_{q+1}$ defined respectively through (3.8) and (3.6), the tuple

$$(P_{q+1}, M_{q+1}, \bar{R}_{q+1}, \widetilde{R}_{q+1})$$

satisfies (3.3) with q replaced by $q + 1$.

The estimates (4.19) and (4.20) follow from Lemma 5.6. Since $W_{I,q+1}$ is localized near frequency λ_{q+1} in phase space, we can define

$$\widetilde{W}_{I,q+1} = \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}(a_{I,q+1}e^{i\lambda_{q+1}\xi_I})$$

and hence $W_{q+1} = \nabla \cdot \widetilde{W}_{q+1}$ with $\widetilde{W}_{q+1} = \sum_I \widetilde{W}_{I,q+1}$. Then the estimate (4.21) follows from Lemma 5.4 with $k = r = 0$. Regarding the advective derivative, we can rewrite

$$\begin{aligned}&(\partial_t + T[M_{\epsilon,q}] \cdot \nabla)\widetilde{W}_{I,q+1}\\ =&\left[\frac{D_{M,q,\epsilon}}{Dt}, \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\right](a_{I,q+1}e^{i\lambda_{q+1}\xi_I}) + \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\left(e^{i\lambda_{q+1}\xi_I}\frac{D_{M,q,\epsilon}}{Dt}a_{I,q+1}\right).\end{aligned}$$

As a consequence, it follows from Lemma 5.2 with a suitable rescaling and Lemma 5.4 with $k = 0$ and $r = 1$

$$\begin{aligned}\|(\partial_t + T[M_{\epsilon,q}] \cdot \nabla)\widetilde{W}_{I,q+1}\|_{C^0} &\lesssim \lambda_q\delta_{q-1}^{\frac{1}{2}}\lambda_{q+1}^{-1}\delta_q^{\frac{1}{2}} + \lambda_{q+1}^{-1}\delta_q^{\frac{1}{2}}\tau_q^{-1}\\ &\lesssim \lambda_{q+1}^{-1}\delta_q^{\frac{1}{2}}\tau_q^{-1}\end{aligned}$$

provided

$$\tau_q \lesssim \lambda_q^{-1}\delta_{q-1}^{-\frac{1}{2}}. \tag{5.1}$$

Other terms in (4.22) can be estimated similarly.

By (4.7) and (4.19), we have

$$\begin{aligned}\|\nabla^k P_{q+1}\|_{C^0} &\leq \|\nabla^k P_q\|_{C^0} + \|\nabla^k W_{q+1}\|_{C^0}\\ &\lesssim \lambda_q^k\delta_{q-1}^{\frac{1}{2}} + \lambda_{q+1}^k\delta_q^{\frac{1}{2}}\\ &\lesssim \lambda_{q+1}^k\delta_q^{\frac{1}{2}}\end{aligned}$$

since $b > 1$ and $k \geq 1$. The estimate of $\nabla^k M_{q+1}, \nabla^k T[P_{q+1}], \nabla^k T[M_{q+1}]$ and hence (4.7) with q replaced by $q + 1$ follows analogously.

Next we show (4.8) with q replaced by $q+1$. We write

$$
\begin{aligned}
&(\partial_t + T[M_{q+1}]\cdot\nabla)T[M_{q+1}] \\
&= (\partial_t + T[M_q]\cdot\nabla)T[M_q] + T[W_{q+1}]\cdot\nabla T[M_q] + T[W_{q+1}]\cdot\nabla T[W_{q+1}] \\
&\quad + (\partial_t + T[M_q]\cdot\nabla)T[W_{q+1}]
\end{aligned}
$$

and further decompose

$$
\begin{aligned}
&(\partial_t + T[M_q]\cdot\nabla)T[W_{q+1}] \\
=&(\partial_t + T[M_{\epsilon,q}]\cdot\nabla)T[W_{q+1}] + (T[M_q] - T[M_{\epsilon,q}])\cdot\nabla T[W_{q+1}].
\end{aligned}
$$

Immediately it follows from the induction assumption (4.8)

$$\|(\partial_t + T[M_q]\cdot\nabla)T[M_q]\|_{C^0} \lesssim \lambda_q\delta_{q-1},$$

and the estimate (4.20)

$$\|(\partial_t + T[M_{\epsilon,q}]\cdot\nabla)T[W_{q+1}]\|_{C^0} \lesssim \tau_q^{-1}\delta_q^{\frac{1}{2}}.$$

The assumption (4.7) and estimate (4.19) together yield

$$\|T[W_{q+1}]\cdot\nabla T[M_q]\|_{C^0} \le \|T[W_{q+1}]\|_{C^0}\|\nabla T[M_q]\|_{C^0} \lesssim \delta_q^{\frac{1}{2}}\lambda_q\delta_{q-1}^{\frac{1}{2}}.$$

The estimate (4.19) also implies

$$\|T[W_{q+1}]\cdot\nabla T[W_{q+1}]\|_{C^0} \le \|T[W_{q+1}]\|_{C^0}\|\nabla T[W_{q+1}]\|_{C^0} \lesssim \delta_q^{\frac{1}{2}}\lambda_{q+1}\delta_q^{\frac{1}{2}}.$$

In view of Lemma 4.2 and (4.19) we have

$$
\begin{aligned}
\|(T[M_q] - T[M_{\epsilon,q}])\cdot\nabla T[W_{q+1}]\|_{C^0} \le& \|T[M_q] - T[M_{\epsilon,q}]\|_{C^0}\|\nabla T[W_{q+1}]\|_{C^0} \\
\lesssim& \mu_q^{-L}\lambda_q^L\delta_{q-1}^{\frac{1}{2}}\lambda_{q+1}\delta_q^{\frac{1}{2}}.
\end{aligned}
$$

Summarizing the estimates above we obtain for $b>1$ and $0<\beta<1$

$$
\begin{aligned}
\|(\partial_t + T[M_{q+1}]\cdot\nabla)T[M_{q+1}]\|_{C^0} &\lesssim \lambda_q\delta_{q-1} + \tau_q^{-1}\delta_q^{\frac{1}{2}} + \delta_q^{\frac{1}{2}}\lambda_q\delta_{q-1}^{\frac{1}{2}} \\
&\quad + \lambda_{q+1}\delta_q + \mu_q^{-L}\lambda_q^L\delta_{q-1}^{\frac{1}{2}}\lambda_{q+1}\delta_q^{\frac{1}{2}} \\
&\lesssim \lambda_{q+1}\delta_q
\end{aligned}
$$

where we used $\mu_q = \lambda_{q+1}^{\frac{1}{L}}\lambda_q^{1-\frac{1}{L}}$ and supposed

$$\tau_q^{-1} \le \lambda_{q+1}\delta_q^{\frac{1}{2}}. \tag{5.2}$$

For $1 \le k \le L-1$, higher order derivatives in (4.8) with q replaced by $q+1$ can be estimated similarly.

Next we establish the estimates for the new stress field. Invoking $\widetilde{\theta}_q = \widetilde{\theta}_{q-1}$ and recalling (3.6), we have

$$
\begin{aligned}
\nabla\cdot\widetilde{R}_{q+1} =& \Big(\partial_t + T[\widetilde{\theta}_{\epsilon,q-1}]\cdot\nabla\Big)W_{q+1} + \nu\Lambda^\gamma W_{q+1} + T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{\epsilon,q-1} \\
&+ \nabla\cdot\big(c_{1,q}A_1 + R_q^* - 2T[W_{q+1}]W_{q+1}\big) \\
=& \Big(\partial_t + T[\widetilde{\theta}_{\epsilon,q-1}]\cdot\nabla\Big)W_{q+1} + \nu\Lambda^\gamma W_{q+1} + T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{\epsilon,q-1} \\
&+ \nabla\cdot\big(c_{\epsilon,1,q}A_1 + R_{\epsilon,q}^* - 2T[W_{q+1}]W_{q+1}\big) \\
&+ \nabla\cdot\Big((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{\epsilon,q-1}])W_{q+1} + T[W_{q+1}](\widetilde{\theta}_{q-1} - \widetilde{\theta}_{\epsilon,q-1}) \\
&+(c_{1,q} - c_{\epsilon,1,q})A_1 + (R_q^* - R_{\epsilon,q}^*)\Big)
\end{aligned}
$$

with $\widetilde{\theta}_{\epsilon,q-1}=P_{\epsilon,q-1}-M_{\epsilon,q-1}$. We denote

$$\begin{aligned}
\nabla\cdot R_T &= \left(\partial_t+T[\widetilde{\theta}_{\epsilon,q-1}]\cdot\nabla\right)W_{q+1},\\
\nabla\cdot R_D &= \nu\Lambda^\gamma W_{q+1},\\
\nabla\cdot R_N &= T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{\epsilon,q-1},\\
\nabla\cdot R_O &= \nabla\cdot\left(c_{\epsilon,1,q}A_1+R^*_{\epsilon,q}-2T[W_{q+1}]W_{q+1}\right),\\
\nabla\cdot R_M &= \nabla\cdot\Big((T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{\epsilon,q-1}])W_{q+1}+T[W_{q+1}](\widetilde{\theta}_{q-1}-\widetilde{\theta}_{\epsilon,q-1})\\
&\quad+(c_{1,q}-c_{\epsilon,1,q})A_1+(R^*_q-R^*_{\epsilon,q})\Big).
\end{aligned}$$

Estimates of R_T: Note W_{q+1} is localized to frequency $\approx\lambda_{q+1}$ in Fourier space. Therefore we can find R_T such that

$$R_T=\nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\left[(\partial_t+T[\widetilde{\theta}_{\epsilon,q-1}]\cdot\nabla)W_{q+1}\right].$$

As a consequence we obtain

$$\|R_T\|_{C^0}\lesssim\lambda_{q+1}^{-1}\|(\partial_t+T[\widetilde{\theta}_{\epsilon,q-1}]\cdot\nabla)W_{q+1}\|_{C^0}.$$

Since

$$T[\widetilde{\theta}_{\epsilon,q-1}]=T[P_{\epsilon,q-1}]-T[M_{\epsilon,q-1}],$$

applying Lemma 5.6 with $q'=q$, $r=1$ and $k=0$ we have

$$\begin{aligned}
&\|(\partial_t+T[\widetilde{\theta}_{\epsilon,q-1}]\cdot\nabla)W_{q+1}\|_{C^0}\\
\le&\|\left(\partial_t+T[P_{\epsilon,q-1}]\cdot\nabla\right)W_{q+1}\|_{C^0}+\|\left(\partial_t+T[M_{\epsilon,q-1}]\cdot\nabla\right)W_{q+1}\|_{C^0}\\
\lesssim&\,\tau_q^{-1}\delta_q^{\frac12}.
\end{aligned}$$

Therefore, we conclude

$$\|R_T\|_{C^0}\lesssim\lambda_{q+1}^{-1}\tau_q^{-1}\delta_q^{\frac12}. \tag{5.3}$$

Estimates of R_D: It is obvious that there exists R_D satisfying

$$R_D=\nu\nabla\Delta^{-1}\Lambda^\gamma W_{q+1}.$$

It follows from (4.19) that

$$\|R_D\|_{C^0}\lesssim\lambda_{q+1}^{-1+\gamma}\|W_{q+1}\|_{C^0}\lesssim\lambda_{q+1}^{-1+\gamma}\delta_q^{\frac12}. \tag{5.4}$$

Estimates of R_N: Again, due to the frequency localization property of $\widetilde{\theta}_{\epsilon,q-1}$ and W_{q+1}, we can define

$$R_N=\nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}[T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{\epsilon,q-1}].$$

Applying (4.13) and (4.19) gives

$$\begin{aligned}
\|R_N\|_{C^0}&\lesssim\lambda_{q+1}^{-1}\|T[W_{q+1}]\|_{C^0}\|\nabla\widetilde{\theta}_{\epsilon,q-1}\|_{C^0}\\
&\lesssim\lambda_{q+1}^{-1}\delta_q^{\frac12}\lambda_{q-1}\delta_{q-2}^{\frac12}.
\end{aligned} \tag{5.5}$$

Estimates of R_O: It follows from Microlocal Lemma 4.1 that

$$\begin{aligned}
W_{I,q+1}&=\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}]\\
&=a_{I,q+1}\eta_{\lambda_{q+1}}(\lambda_{q+1}\nabla\xi_I)e^{i\lambda_{q+1}\xi_I}+\delta a_{I,q+1}e^{i\lambda_{q+1}\xi_I}\\
&=a_{I,q+1}e^{i\lambda_{q+1}\xi_I}+\delta a_{I,q+1}e^{i\lambda_{q+1}\xi_I}
\end{aligned}$$

and

$$\begin{aligned} T[W_{I,q+1}] &= T\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &= a_{I,q+1}m(\lambda_{q+1}\nabla\xi_I)\eta_{\lambda_{q+1}}(\lambda_{q+1}\nabla\xi_I)e^{i\lambda_{q+1}\xi_I} + \delta u_{I,q+1}e^{i\lambda_{q+1}\xi_I} \\ &= a_{I,q+1}m(\nabla\xi_I)e^{i\lambda_{q+1}\xi_I} + \delta u_{I,q+1}e^{i\lambda_{q+1}\xi_I}. \end{aligned}$$

Consequently we compute

$$\begin{aligned} T[W_{q+1}]\cdot\nabla W_{q+1} = \frac{1}{2}\cdot\nabla\sum_{I\in\Omega}\big(T[W_{I,q+1}]W_{\bar{I},q+1} + T[W_{\bar{I},q+1}]W_{I,q+1}\big) \\ + \sum_{J\neq\bar{I}} T[W_{J,q+1}]\cdot\nabla W_{I,q+1} \end{aligned}$$

with

$$\begin{aligned} &\frac{1}{2}\sum_{I\in\Omega}\big(T[W_{I,q+1}]W_{\bar{I},q+1} + T[W_{\bar{I},q+1}]W_{I,q+1}\big) \\ &= \frac{1}{2}\sum_{I\in\Omega} a_{I,q+1}^2\Big(m(\nabla\widehat{\xi}_I) + m(-\nabla\widehat{\xi}_I)\Big) \\ &\quad + \frac{1}{2}\sum_{I\in\Omega} a_{I,q+1}^2\Big(m(\nabla\xi_I) - m(\nabla\widehat{\xi}_I) + m(-\nabla\xi_I) - m(-\nabla\widehat{\xi}_I)\Big) \\ &\quad + \frac{1}{2}\sum_{I\in\Omega}\big(a_{I,q+1}\delta u_{\bar{I},q+1} + a_{I,q+1}m(\nabla\xi_I)\delta a_{\bar{I},q+1} + a_{\bar{I},q+1}\delta u_{I,q+1} \\ &\quad + a_{\bar{I},q+1}m(\nabla\xi_{\bar{I}})\delta a_{I,q+1} - \delta a_{\bar{I},q+1}\delta u_{I,q+1} - \delta a_{I,q+1}\delta u_{\bar{I},q+1}\big) \\ &=: \frac{1}{2}\sum_{I\in\Omega} a_{I,q+1}^2\Big(m(\nabla\widehat{\xi}_I) + m(-\nabla\widehat{\xi}_I)\Big) + R_{O,1} + R_{O,2}. \end{aligned}$$

Since

$$\nabla W_{I,q+1} = i\lambda_{q+1}\nabla\xi_I\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] + \mathbb{P}_{\approx\lambda_{q+1}}[\nabla a_{I,q+1}e^{i\lambda_{q+1}\xi_I}],$$

we further compute

$$\begin{aligned} &\sum_{J\neq\bar{I}} T[W_{J,q+1}]\cdot\nabla W_{I,q+1} \\ &= \sum_{J\neq\bar{I}} i\lambda_{q+1}T[W_{J,q+1}]\cdot\nabla\xi_I\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &\quad + \sum_{J\neq\bar{I}} T[W_{J,q+1}]\cdot\mathbb{P}_{\approx\lambda_{q+1}}[\nabla a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &= \sum_{J\neq\bar{I}} i\lambda_{q+1}a_{J,q+1}m(\nabla\xi_J)e^{i\lambda_{q+1}\xi_J}\cdot\nabla\xi_I\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &\quad + \sum_{J\neq\bar{I}} i\lambda_{q+1}\delta u_{J,q+1}e^{i\lambda_{q+1}\xi_J}\cdot\nabla\xi_I\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &\quad + \sum_{J\neq\bar{I}} T[W_{J,q+1}]\cdot\mathbb{P}_{\approx\lambda_{q+1}}[\nabla a_{I,q+1}e^{i\lambda_{q+1}\xi_I}]. \end{aligned}$$

Note that

$$\begin{aligned} m(\nabla\xi_J)\cdot\nabla\xi_I &= m(\nabla\xi_{J,in})\cdot\nabla\xi_{I,in} + m(\nabla\xi_{J,in})\cdot(\nabla\xi_I - \nabla\xi_{I,in}) \\ &\quad + (m(\nabla\xi_J) - m(\nabla\xi_{J,in}))\cdot\nabla\xi_I \\ &= m(\nabla\xi_{J,in})\cdot(\nabla\xi_I - \nabla\xi_{I,in}) \\ &\quad + (m(\nabla\xi_J) - m(\nabla\xi_{J,in}))\cdot\nabla\xi_I \end{aligned}$$

since

$$m(\nabla\xi_{J,in})\cdot\nabla\xi_{I,in} = m(\pm\nabla\xi_{I,in})\cdot\nabla\xi_{I,in} = 0.$$

Therefore we have

$$\begin{aligned} &\sum_{J\neq\bar{I}} T[W_{J,q+1}]\cdot\nabla W_{I,q+1} \\ =&\sum_{J\neq\bar{I}} i\lambda_{q+1}a_{J,q+1}m(\nabla\xi_{J,in})\cdot(\nabla\xi_I - \nabla\xi_{I,in})e^{i\lambda_{q+1}\xi_J}\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &+ \sum_{J\neq\bar{I}} i\lambda_{q+1}a_{J,q+1}(m(\nabla\xi_J) - m(\nabla\xi_{J,in}))\cdot\nabla\xi_I e^{i\lambda_{q+1}\xi_J}\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &+ \sum_{J\neq\bar{I}} i\lambda_{q+1}\delta u_{J,q+1}e^{i\lambda_{q+1}\xi_J}\cdot\nabla\xi_I\mathbb{P}_{\approx\lambda_{q+1}}[a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ &+ \sum_{J\neq\bar{I}} T[W_{J,q+1}]\cdot\mathbb{P}_{\approx\lambda_{q+1}}[\nabla a_{I,q+1}e^{i\lambda_{q+1}\xi_I}] \\ =&: \nabla\cdot R_{O,3} + \nabla\cdot R_{O,4} + \nabla\cdot R_{O,5} + \nabla\cdot R_{O,6}. \end{aligned}$$

Summarizing the analysis above we obtain

$$\begin{aligned} T[W_{q+1}]\cdot\nabla W_{q+1} = &\frac{1}{2}\nabla\cdot\sum_{I\in\Omega} a_{I,q+1}^2\left(m(\nabla\xi_{I,in}) + m(-\nabla\xi_{I,in})\right) + \nabla\cdot R_{O,1} \\ &+ \nabla\cdot R_{O,2} + \nabla\cdot R_{O,3} + \nabla\cdot R_{O,4} + \nabla\cdot R_{O,5} + \nabla\cdot R_{O,6}. \end{aligned}$$

According to the choice of $a_{I,q+1}$ in (4.5), we have that

$$\begin{aligned} \nabla\cdot R_O &= \nabla\cdot\left(c_{\epsilon,1,q}A_1 + R_{\epsilon,q}^* - 2T[W_{q+1}]W_{q+1}\right) \\ &= \nabla\cdot\left(c_{2,q+1}A_2 - 2R_{O,1} - 2R_{O,2} - 2R_{O,3} - 2R_{O,4} - 2R_{O,5} - 2R_{O,6}\right). \end{aligned}$$

Now we estimate the error terms above. It follows from the definition of $c_{2,q+1}$ and the assumption (4.11) that

$$\|c_{2,q+1}\|_{C^0} \lesssim \|R_{\epsilon,q}^*\|_{C^0} \lesssim \|R_q^*\|_{C^0} \lesssim \delta_{q+1}.$$

Applying Lemma 5.3 and 5.4, noticing that ξ_I is advected by $\widetilde{u}_q = \widetilde{u}_{q-1}$, leads to

$$\begin{aligned} \|R_{O,1}\|_{C^0} &\lesssim \sum_{I\in\Omega}\|a_{I,q+1}\|_{C^0}^2|\nabla\xi_I - \nabla\widehat{\xi}_I| \\ &\lesssim \sum_{I\in\Omega}\|a_{I,q+1}\|_{C^0}^2\lambda_{q-1}\tau_q\|\widetilde{u}_{q-1}\|_{C^0} \\ &\lesssim \lambda_{q-1}\tau_q\delta_{q-2}^{\frac{1}{2}}\delta_q. \end{aligned}$$

Applying Lemma 5.4 and Lemma 5.5

$$\begin{aligned}\|R_{O,2}\|_{C^0} \lesssim &\sum_{I\in\Omega}\|a_{I,q+1}\|_{C^0}\|\delta u_{\bar I,q+1}\|_{C^0}+\sum_{I\in\Omega}\|a_{I,q+1}\|_{C^0}\|\delta a_{\bar I,q+1}\|_{C^0}\\ &+\sum_{I\in\Omega}\|\delta a_{I,q+1}\|_{C^0}\|\delta u_{\bar I,q+1}\|_{C^0}\\ \lesssim &\ \delta_q^{\frac12}\lambda_{q+1}^{-1}\lambda_q\delta_q^{\frac12}.\end{aligned}$$

We observe that due to the frequency support of W_{q+1} and $T[W_{q+1}]$, we can define

$$R_{O,3}+R_{O,4}+R_{O,5}+R_{O,6}=\sum_{J\neq\bar I}\nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\left[T[W_{J,q+1}]\cdot\nabla W_{I,q+1}\right].$$

Therefore we deduce from Lemma 5.3 and Lemma 5.4

$$\begin{aligned}\|R_{O,3}\|_{C^0} &\lesssim\sum_{J\neq\bar I}\|a_{J,q+1}\|_{C^0}\|a_{I,q+1}\|_{C^0}\|m(\nabla\xi_{I,in})\|_{C^0}|\nabla\xi_I-\nabla\xi_{I,in}|\\ &\lesssim\sum_{J\neq\bar I}\|a_{J,q+1}\|_{C^0}\|a_{I,q+1}\|_{C^0}\lambda_{q-1}\tau_q\|\widetilde u_{q-1}\|_{C^0}\\ &\lesssim\lambda_{q-1}\tau_q\delta_{q-2}^{\frac12}\delta_q\end{aligned}$$

$$\begin{aligned}\|R_{O,4}\|_{C^0} &\lesssim\sum_{J\neq\bar I}\|a_{J,q+1}\|_{C^0}\|a_{I,q+1}\|_{C^0}\|m(\nabla\xi_J)-m(\nabla\xi_{J,in})\|_{C^0}\|\nabla\xi_I\|_{C^0}\\ &\lesssim\sum_{J\neq\bar I}\|a_{J,q+1}\|_{C^0}\|a_{I,q+1}\|_{C^0}|\nabla\xi_I-\nabla\xi_{I,in}|\\ &\lesssim\lambda_{q-1}\tau_q\delta_{q-2}^{\frac12}\delta_q\end{aligned}$$

where we used the fact $\|m(\nabla\xi_J)-m(\nabla\xi_{J,in})\|_{C^0}\lesssim\|\nabla\xi_J-\nabla\xi_{J,in}\|_{C^0}$;

$$\begin{aligned}\|R_{O,5}\|_{C^0} &\lesssim\sum_{J\neq\bar I}\|\delta u_{J,q+1}\|_{C^0}\|a_{I,q+1}\|_{C^0}\|\nabla\xi_I\|_{C^0}\\ &\lesssim\lambda_{q+1}^{-1}\lambda_q\delta_q^{\frac12}\delta_q^{\frac12}\end{aligned}$$

using (4.19) and Lemma 5.5, and

$$\begin{aligned}\|R_{O,6}\|_{C^0} &\lesssim\sum_{J\neq\bar I}\lambda_{q+1}^{-1}\|T[W_{J,q+1}]\|_{C^0}\|\nabla a_{I,q+1}\|_{C^0}\\ &\lesssim\lambda_{q+1}^{-1}\delta_q^{\frac12}\lambda_q\delta_q^{\frac12}.\end{aligned}$$

Summarizing the estimates above gives

$$\|R_{O,1}\|_{C^0}+\|R_{O,3}\|_{C^0}+\|R_{O,4}\|_{C^0}\lesssim\lambda_{q-1}\tau_q\delta_{q-2}^{\frac12}\delta_q,\tag{5.6}$$

$$\|R_{O,2}\|_{C^0}+\|R_{O,5}\|_{C^0}+\|R_{O,6}\|_{C^0}\lesssim\lambda_{q+1}^{-1}\lambda_q\delta_q.\tag{5.7}$$

Estimates of R_M: It follows from the fact $\widetilde{\theta}_q=\widetilde{\theta}_{q-1}$, Lemma 4.2 and estimate (4.19) that

$$\begin{aligned}&\|(T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{\epsilon,q-1}])W_{q+1}\|_{C^0}+\|T[W_{q+1}](\widetilde{\theta}_{q-1}-\widetilde{\theta}_{\epsilon,q-1})\|_{C^0}\\ &=\|(T[\widetilde{\theta}_q]-T[\widetilde{\theta}_{\epsilon,q}])W_{q+1}\|_{C^0}+\|T[W_{q+1}](\widetilde{\theta}_q-\widetilde{\theta}_{\epsilon,q})\|_{C^0}\\ &\lesssim \delta_q^{\frac12}\mu_q^{-L}\lambda_q^L\delta_{q-1}^{\frac12}\\ &\lesssim \delta_q^{\frac12}\delta_{q-1}^{\frac12}\lambda_{q+1}^{-1}\lambda_q.\end{aligned}$$

By Lemma 4.3, we have

$$\|(c_{1,q,\epsilon}-c_{1,q})A_1\|_{C^0}+\|R^*_{q,\epsilon}-R^*_q\|_{C^0}\lesssim \delta_q^{\frac12}\delta_{q-1}^{\frac12}\lambda_{q+1}^{-1}\lambda_q.$$

Therefore

$$\|R_M\|_{C^0}\lesssim \delta_q^{\frac12}\delta_{q-1}^{\frac12}\lambda_{q+1}^{-1}\lambda_q. \tag{5.8}$$

In summary, the new stress error can be written as

$$\widetilde{R}_{q+1}=c_{2,q+1}A_2+R^*_{q+1}$$

with

$$\|c_{2,q+1}\|_{C^0}\lesssim \delta_{q+1}$$

and

$$R^*_{q+1}=R_T+R_N+R_D+R_M-2(R_{O,1}+...+R_{O,6}).$$

Thus (4.9) is satisfied with $c_{1,q}$ replaced by $c_{2,q+1}$. To show (4.11) with q replaced by $q+1$, we just need to show $\|R^*_{q+1}\|_{C^0}\le\delta_{q+2}$. We choose $\tau_q=\lambda_{q+1}^{-\frac12}\lambda_{q-1}^{-\frac12}\delta_q^{-\frac14}\delta_{q-2}^{-\frac14}$ to optimize the two estimates (5.3) and (5.6) such that

$$\|R_T\|_{C^0}+\|R_{O,1}\|_{C^0}+\|R_{O,3}\|_{C^0}+\|R_{O,4}\|_{C^0}\lesssim \lambda_{q+1}^{-\frac12}\lambda_{q-1}^{\frac12}\delta_q^{\frac34}\delta_{q-2}^{\frac14}.$$

In view of the last estimate, (5.5), (5.4), (5.7) and (5.8), we impose

$$\begin{cases}\lambda_{q+1}^{-\frac12}\lambda_{q-1}^{\frac12}\delta_q^{\frac34}\delta_{q-2}^{\frac14}\lesssim\delta_{q+2}\\ \lambda_{q+1}^{-1}\lambda_{q-1}\delta_q^{\frac12}\delta_{q-2}^{\frac12}\lesssim\delta_{q+2}\\ \lambda_{q+1}^{-1+\gamma}\delta_q^{\frac12}\lesssim\delta_{q+2}\\ \lambda_{q+1}^{-1}\lambda_q\delta_q\lesssim\delta_{q+2}\\ \lambda_{q+1}^{-1}\lambda_q\delta_q^{\frac12}\delta_{q-1}^{\frac12}\lesssim\delta_{q+2}\end{cases}$$

which are satisfied provided $b=1^+$ (close enough to 1 from the right), $\beta<\frac25$ and for $\gamma\ge0$ satisfying $\beta<\frac{2b(1-\gamma)}{2b^2-1}$. Recall θ is in $C^0_tC^\alpha_x$ with $\alpha<\frac{\beta}{2b}<\frac15$. The conditions $\beta<\frac{2b(1-\gamma)}{2b^2-1}$ and $\alpha<\frac{\beta}{2b}$ together imply $0\le\gamma<1-\alpha$.

Estimates of higher order spatial derivatives: For $c_{2,q+1}$, each derivative cost is $\lesssim\lambda_{q+1}$. Combining with the C^0 estimate of $c_{2,q+1}$, we have

$$\|\nabla^k c_{2,q+1}\|_{C^0}\lesssim\lambda_{q+1}^k\delta_{q+1}.$$

We observe that R_T, R_D, R_N, $R_{O,3}$, $R_{O,4}$,$R_{O,5}$ and $R_{O,6}$ are localized in Fourier space near frequency λ_{q+1}. Hence each spatial derivative of them costs at most $\approx\lambda_{q+1}$. From Lemma 4.2, we know

$$\|\widetilde{\theta}_{q-1}-\widetilde{\theta}_{q-1,\epsilon}\|_{C^0}\lesssim\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}.$$

We also have the estimate from (4.13)

$$\|\nabla\widetilde{\theta}_{q-1}\|_{C^0}+\|\nabla\widetilde{\theta}_{q-1,\epsilon}\|_{C^0}\lesssim\lambda_{q-1}\delta_{q-2}^{\frac{1}{2}}.$$

Therefore one spatial derivative cost of $\widetilde{\theta}_{q-1}-\widetilde{\theta}_{q-1,\epsilon}$ and $T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}]$ is at most $\approx\lambda_q$. On the other hand, one spatial derivative cost of W_{q+1} and $T[W_{q+1}]$ is at most $\approx\lambda_{q+1}$. Hence one spatial derivative cost of R_M is at most $\approx\lambda_{q+1}$. Regarding $R_{O,1}$, we compare the estimates from Lemma 5.3

$$\|\nabla^2\xi_I(\Phi_q(x,s,t))\|_{C^0}\lesssim\lambda_q,$$
$$\|\nabla\xi_I(\Phi_q(x,s,t))-\nabla\xi_{I,in}\|_{C^0}\lesssim\tau_q\lambda_q\delta_{q-1}^{\frac{1}{2}}.$$

Thus one spatial derivative of $\nabla\xi_I(\Phi_q(x,s,t))-\nabla\xi_{I,in}$ is at most $\approx\tau_q^{-1}\delta_{q-1}^{-\frac{1}{2}}\lesssim\lambda_{q+1}$. The analysis shows that

$$\|\nabla^kR^*_{q+1}\|_{C^0}\lesssim\lambda_{q+1}^k\delta_{q+2},\quad 0\le k\le L.$$

Estimates of advective derivative: First we rewrite

$$\partial_t+T[M_{q+1}]\cdot\nabla=\partial_t+T[M_{q,\epsilon}]\cdot\nabla+(T[M_q]-T[M_{q,\epsilon}])\cdot\nabla+T[W_{q+1}]\cdot\nabla$$

and

$$\partial_t+T[\theta_{q+1}]\cdot\nabla=\partial_t+T[\theta_{q,\epsilon}]\cdot\nabla+(T[\theta_q]-T[\theta_{q,\epsilon}])\cdot\nabla+T[W_{q+1}]\cdot\nabla.$$

For R_D:

$$\begin{aligned}&\|\nabla^k(\partial_t+T[M_{q+1}]\cdot\nabla)R_D\|_{C^0}\\ \lesssim&\|\nabla^k(\partial_t+T[M_{q,\epsilon}]\cdot\nabla)R_D\|_{C^0}+\|\nabla^k\left((T[M_q]-T[M_{q,\epsilon}])\cdot\nabla\right)R_D\|_{C^0}\\ &+\|\nabla^k(T[W_{q+1}]\cdot\nabla)R_D\|_{C^0}\\ \lesssim&\lambda_{q+1}^{k+1}\delta_q^{\frac{1}{2}}\delta_{q+1}.\end{aligned}$$

In view of the definition of $c_{2,q+1}$ in (4.2), we observe

$$\begin{aligned}&\|\nabla^k(\partial_t+T[M_{q+1}]\cdot\nabla)c_{2,q+1}\|_{C^0}\\ \lesssim&\|\nabla^k(\partial_t+T[M_{q+1}]\cdot\nabla)R^*_{q,\epsilon}\|_{C^0}\\ \lesssim&\|\nabla^k(\partial_t+T[M_{q+1}]\cdot\nabla)R^*_q\|_{C^0}\\ \lesssim&\|\nabla^k(\partial_t+T[M_{q,\epsilon}]\cdot\nabla)R^*_q\|_{C^0}+\|\nabla^k\left((T[M_q]-T[M_{q,\epsilon}])\cdot\nabla\right)R^*_q\|_{C^0}\\ &+\|\nabla^k(T[W_{q+1}]\cdot\nabla)R^*_q\|_{C^0}\\ \lesssim&\lambda_q^{k+1}\delta_{q-1}^{\frac{1}{2}}\delta_{q+1}+\lambda_{q+1}^k\delta_q^{\frac{1}{2}}\lambda_q\delta_{q+1}\\ \lesssim&\lambda_{q+1}^{k+1}\delta_q^{\frac{1}{2}}\delta_{q+1}.\end{aligned}$$

Lemma 5.7. *Let $k\ge0$ and $0\le r\le2$, we have*

$$\|\nabla^k\left(\frac{D_q}{Dt}\right)^r(\nabla\xi_I-\nabla\xi_{I,in})\|_{C^0}\lesssim\lambda_{q+1}^k\tau_q^{-r}\tau_{q+1}\lambda_q\delta_{q-1}^{\frac{1}{2}}.$$

Lemma 5.8. *Let $k\ge0$ and $0\le r\le2$, we also have*

$$\|\nabla^k\left(\frac{D_q}{Dt}\right)^r(m(\nabla\xi_I)-m(\nabla\xi_{I,in}))\|_{C^0}\lesssim\lambda_{q+1}^k\tau_q^{-r}\tau_q\lambda_q\delta_{q-1}^{\frac{1}{2}}.$$

Proof: Note

$$m(\nabla\xi_I) - m(\nabla\xi_{I,in}) = (\nabla\xi_I - \nabla\xi_{I,in}) \int_0^1 \partial_a m\left((1-s)\nabla\xi_{I,in} + s\nabla\xi_I\right) ds.$$

The estimate follows from Lemma 5.7.

□

Combining Lemma 5.4, Lemma 5.5, Lemma 5.7, Lemma 5.8, we have

$$\begin{aligned}
&\|\nabla^k(\partial_t + T[M_{q+1}]\cdot\nabla)R_{O,1}\|_{C^0} + \|\nabla^k(\partial_t + T[M_{q+1}]\cdot\nabla)R_{O,2}\|_{C^0}\\
\lesssim\ &\|\nabla^k(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)R_{O,1}\|_{C^0} + \|\nabla^k\left((T[M_q] - T[M_{q,\epsilon}])\cdot\nabla\right)R_{O,1}\|_{C^0}\\
&+ \|\nabla^k(T[W_{q+1}]\cdot\nabla)R_{O,1}\|_{C^0} + \|\nabla^k(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)R_{O,2}\|_{C^0}\\
&+ \|\nabla^k\left((T[M_q] - T[M_{q,\epsilon}])\cdot\nabla\right)R_{O,2}\|_{C^0} + \|\nabla^k(T[W_{q+1}]\cdot\nabla)R_{O,2}\|_{C^0}\\
\lesssim\ &\lambda_{q+1}^{k+1}\delta_q^{\frac{1}{2}}\delta_{q+2}.
\end{aligned}$$

For R_T, recall

$$R_T = \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}[(\partial_t + T[\widetilde{\theta}_{q-1}]\cdot\nabla)W_{q+1}].$$

Then

$$\begin{aligned}
&\|(\partial_t + T[\theta_{q+1}]\cdot\nabla)R_T\|_{C^0}\\
\lesssim\ &\|(\partial_t + T[\theta_{q,\epsilon}]\cdot\nabla)R_T\|_{C^0} + \|(T[\theta_q] - T[\theta_{q,\epsilon}])\cdot\nabla R_T\|_{C^0}\\
&+ \|T[W_{q+1}]\cdot\nabla R_T\|_{C^0}.
\end{aligned}$$

Denote

$$\frac{D_{q,\theta,\epsilon}}{Dt} = \partial_t + T[\theta_{q,\epsilon}]\cdot\nabla, \quad \frac{\widetilde{D}_{q,\theta}}{Dt} = \partial_t + T[\widetilde{\theta}_q]\cdot\nabla$$

For the first term, we apply the commutator

$$\begin{aligned}
(\partial_t + T[\theta_{q,\epsilon}]\cdot\nabla)R_T = &\left[\frac{D_{q,\theta,\epsilon}}{Dt}, \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\right]\frac{\widetilde{D}_{q-1,\theta}}{Dt}W_{q+1}\\
&+ \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\frac{D_{q,\theta,\epsilon}}{Dt}\frac{\widetilde{D}_{q-1,\theta}}{Dt}W_{q+1}.
\end{aligned}$$

Hence applying Lemma 5.2, Lemma 5.6

$$\begin{aligned}
\|(\partial_t + T[\theta_{q,\epsilon}]\cdot\nabla)R_T\|_{C^0} \le &\left\|\left[\frac{D_{q,\theta,\epsilon}}{Dt}, \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\right]\frac{\widetilde{D}_{q-1,\theta}}{Dt}W_{q+1}\right\|_{C^0}\\
&+ \left\|\nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\frac{D_{q,\theta,\epsilon}}{Dt}\frac{\widetilde{D}_{q-1,\theta}}{Dt}W_{q+1}\right\|_{C^0}\\
\lesssim\ &\lambda_q\delta_{q-1}^{\frac{1}{2}}\tau_{q-1}^{-1}\delta_q^{\frac{1}{2}} + \lambda_{q+1}^{-1}\tau_q^{-1}\tau_{q-1}^{-1}\delta_q^{\frac{1}{2}}.
\end{aligned}$$

Applying Lemma 4.2, estimate (4.19) and the spatial derivative established earlier

$$\begin{aligned}
\|(T[\theta_q] - T[\theta_{q,\epsilon}])\cdot\nabla R_T\|_{C^0} &\lesssim \|T[\theta_q] - T[\theta_{q,\epsilon}]\|_{C^0}\|\nabla R_T\|_{C^0}\\
&\lesssim \lambda_{q+1}^{-1}\lambda_q\delta_{q-1}^{\frac{1}{2}}\lambda_{q+1}\delta_{q+2},\\
\|T[W_{q+1}]\cdot\nabla R_T\|_{C^0} &\lesssim \|T[W_{q+1}]\|_{C^0}\|\nabla R_T\|_{C^0}\\
&\lesssim \delta_q^{\frac{1}{2}}\lambda_{q+1}\delta_{q+2}.
\end{aligned}$$

Summarizing we have

$$\begin{aligned}\|(\partial_t + T[\theta_{q+1}]\cdot\nabla)R_T\|_{C^0} &\lesssim \lambda_q\delta_{q-1}^{\frac12}\tau_{q-1}^{-1}\delta_q^{\frac12} + \lambda_{q+1}^{-1}\tau_q^{-1}\tau_{q-1}^{-1}\delta_q^{\frac12}\\ &\quad + \lambda_{q+1}^{-1}\lambda_q\delta_{q-1}^{\frac12}\lambda_{q+1}\delta_{q+2} + \delta_q^{\frac12}\lambda_{q+1}\delta_{q+2}.\end{aligned}$$

The advective derivative of R_N can be handled similarly. Recall

$$R_N = \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}[T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{q-1}].$$

$$\begin{aligned}&\|(\partial_t + T[\theta_{q+1}]\cdot\nabla)R_N\|_{C^0}\\ &\lesssim \|(\partial_t + T[\theta_{q,\epsilon}]\cdot\nabla)R_N\|_{C^0} + \|(T[\theta_q]-T[\theta_{q,\epsilon}])\cdot\nabla R_N\|_{C^0}\\ &\quad + \|T[W_{q+1}]\cdot\nabla R_N\|_{C^0}.\end{aligned}$$

$$\begin{aligned}(\partial_t + T[\theta_{q,\epsilon}]\cdot\nabla)R_N &= \left[\frac{D_{q,\theta,\epsilon}}{Dt}, \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\right](T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{q-1})\\ &\quad + \nabla\Delta^{-1}\mathbb{P}_{\approx\lambda_{q+1}}\frac{D_{q,\theta,\epsilon}}{Dt}(T[W_{q+1}]\cdot\nabla\widetilde{\theta}_{q-1}).\end{aligned}$$

The advective derivative of $R_{O,3},...,R_{O,6}$ can be estimated similarly and the details are omitted.

To estimate the material derivative of R_M, we first consider the term $(T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])W_{q+1}$,

$$\begin{aligned}&\|(\partial_t + T[M_{q+1}]\cdot\nabla)\left((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])W_{q+1}\right)\|_{C^0}\\ &\lesssim \|(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)\left((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\right)W_{q+1}\|_{C^0}\\ &\quad + \|(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)W_{q+1}(T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\\ &\quad + \|(T[M_q] - T[M_{q,\epsilon}])\cdot\nabla\left((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\right)W_{q+1}\|_{C^0}\\ &\quad + \|(T[M_q] - T[M_{q,\epsilon}])\cdot\nabla W_{q+1}(T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\\ &\quad + \|(T[W_{q+1}]\cdot\nabla)\left((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\right)W_{q+1}\|_{C^0}\\ &\quad + \|(T[W_{q+1}]\cdot\nabla)W_{q+1}(T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}.\end{aligned}$$

We see from Lemma 5.1 that the cost of $(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)$ is $\lambda_q\delta_{q-1}^{\frac12}$. Combining with Lemma 4.2 and estimate (4.19) we deduce

$$\begin{aligned}&\|(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)\left((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\right)W_{q+1}\|_{C^0}\\ &\lesssim \|(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)\left((T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\right)\|_{C^0}\|W_{q+1}\|_{C^0}\\ &\lesssim \lambda_q\delta_{q-1}^{\frac12}\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}\delta_q^{\frac12},\end{aligned}$$

$$\begin{aligned}&\|(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)W_{q+1}(T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\\ &\lesssim \|(\partial_t + T[M_{q,\epsilon}]\cdot\nabla)W_{q+1}\|_{C^0}\|(T[\widetilde{\theta}_{q-1}] - T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\\ &\lesssim \lambda_q\delta_{q-1}^{\frac12}\delta_q^{\frac12}\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}.\end{aligned}$$

Using Lemma 4.2, (4.13) and estimate (4.19),

$$\begin{aligned}
&\|(T[M_q]-T[M_{q,\epsilon}])\cdot\nabla\left((T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}])\right)W_{q+1}\|_{C^0}\\
\lesssim&\|T[M_q]-T[M_{q,\epsilon}]\|_{C^0}\|\nabla(T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\|W_{q+1}\|_{C^0}\\
\lesssim&\lambda_{q+1}^{-1}\lambda_q\delta_{q-1}^{\frac12}\lambda_{q-1}\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}\delta_q^{\frac12},
\end{aligned}$$

$$\begin{aligned}
&\|(T[M_q]-T[M_{q,\epsilon}])\cdot\nabla W_{q+1}(T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\\
\lesssim&\|T[M_q]-T[M_{q,\epsilon}]\|_{C^0}\|T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}]\|_{C^0}\|\nabla W_{q+1}\|_{C^0}\\
\lesssim&\lambda_{q+1}^{-1}\lambda_q\delta_{q-1}^{\frac12}\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}\lambda_q\delta_q^{\frac12},
\end{aligned}$$

Applying Lemma 4.2 and estimate (4.19),

$$\begin{aligned}
&\|(T[W_{q+1}]\cdot\nabla)\left((T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}])\right)W_{q+1}\|_{C^0}\\
\lesssim&\|T[W_{q+1}]\|_{C^0}\|\nabla(T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\|W_{q+1}\|_{C^0}\\
\lesssim&\delta_q^{\frac12}\lambda_{q-1}\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}\delta_q^{\frac12},
\end{aligned}$$

$$\begin{aligned}
&\|(T[W_{q+1}]\cdot\nabla)W_{q+1}(T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}])\|_{C^0}\\
\lesssim&\|T[W_{q+1}]\|_{C^0}\|T[\widetilde{\theta}_{q-1}]-T[\widetilde{\theta}_{q-1,\epsilon}]\|_{C^0}\|\nabla W_{q+1}\|_{C^0}\\
\lesssim&\delta_q^{\frac12}\lambda_q^{-1}\lambda_{q-1}\delta_{q-2}^{\frac12}\lambda_q\delta_q^{\frac12}.
\end{aligned}$$

Other terms in R_M can be estimated analogously.

Hölder estimates in time: It follows from

$$\partial_t W_{q+1}=D_{t,q}W_{q+1}-u_q\cdot\nabla W_{q+1}$$

that

$$\|W_{q+1}\|_{C_t^1C_x^0}\lesssim\delta_{q-1}^{\frac12}\lambda_{q+1}\delta_q^{\frac12},$$

and hence by interpolation

$$\|W_{q+1}\|_{C_t^\zeta C_x^0}\lesssim(\lambda_{q+1}\delta_{q-1}^{\frac12})^\zeta\delta_q^{\frac12}\sim\lambda_q^{(b-\frac{1}{2b}\beta)\zeta-\frac12\beta}.$$

The C^ζ regularity in time is assured if

$$(b-\frac{1}{2b}\beta)\zeta-\frac12\beta<0$$

i.e.

$$\zeta<\frac{\beta}{2b-\frac1b\beta}<\frac{1}{2d}$$

when choosing $b=1^+$ and $\beta<\frac{2}{2d+1}$.

6. Proof of Theorem 1.1

Due to the presence of an external forcing in each equation of (3.3), there is redundancy to find an initial tuple $(P_0, M_0, \bar{R}_0, \widetilde{R}_0)$ satisfying the system (3.3) and (4.6)-(4.12) at level 0, such that $M_0 \not\equiv 0$. For instance, we can choose $P_0 \equiv 0$ and $M_0 \not\equiv 0$ with $M_0 = \mathbb{P}_{\le \lambda_0} M_0$. We then define $\bar{R}_0$ and $\widetilde{R}_0$ such that

$$\nabla \cdot \bar{R}_0 = \nabla \cdot (T[M_0]M_0)\,,$$
$$\nabla \cdot \widetilde{R}_0 = \partial_t M_0 + \nu \Lambda^\gamma M_0.$$

Obviously, such tuple $(P_0, M_0, \bar{R}_0, \widetilde{R}_0)$ satisfies the system (3.3) at level $q = 0$. It is easy to see that, choosing $\lambda_0 > 0$ large enough and $\delta_{-1} = 1$, the estimates (4.6)-(4.12) at level 0 can be satisfied. Applying the inductive Proposition 4.4 iteratively we obtain a sequence $\{(P_q, M_q, \bar{R}_q, \widetilde{R}_q\}$ satisfying (3.3) and (4.6)-(4.12). Thanks to the estimates (4.7), (4.9) and (4.11), there exists a subsequence such that P_q converges to a function P, M_q converges to a function M and $\widetilde{R}_q$ converges to 0 as $q \to \infty$. Note

$$M = M_0 + \sum_{q=0}^{\infty} (M_{q+1} - M_q) = M_0 + \sum_{q=0}^{\infty} W_{q+1}$$

and W_{q+1} are supported on frequencies near λ_{q+1}. Hence $M \not\equiv 0$. Regarding the convergence of $\bar{R}_q$, it follows from (3.8) that

$$\begin{aligned}\nabla \cdot (\bar{R}_{q+1} - \bar{R}_q) = &- \left(\partial_t + T[\widetilde{\theta}_q] \cdot \nabla\right) W_{q+1} - \nu \Lambda^\gamma W_{q+1} - T[W_{q+1}] \cdot \nabla \widetilde{\theta}_q \\ &+ 2\nabla \cdot (T[W_{q+1}] W_{q+1})\end{aligned}$$

which implies

$$\bar{R}_{q+1} - \bar{R}_q = -R_T - R_D - R_N + 2T[W_{q+1}]W_{q+1}.$$

We then obtain from the estimates (5.3), (5.4), (5.5) and Lemma 5.6 that

$$\|\bar{R}_{q+1} - \bar{R}_q\|_{C^0} \lesssim \lambda_{q+1}^{-1} \tau_q^{-1} \delta_q^{\frac12} + \lambda_{q+1}^{-1+\gamma} \delta_q^{\frac12} + \lambda_{q+1}^{-1} \delta_q^{\frac12} \lambda_{q-1} \delta_{q-2}^{\frac12} + \delta_q \lesssim \delta_q \lesssim \lambda_q^{-\beta}$$

for $\beta > 0$. Therefore there is a subsequence of $\bar{R}_q$ which converges to a vector field $\bar{R}$. The limit tuple $(P, M, \bar{R}, 0)$ is a weak solution of (3.1) with $f = \nabla \cdot \bar{R}$. Hence there are at least two weak solutions to (1.1) with external forcing $f = \nabla \cdot \bar{R}$. It follows from (3.8), (3.9) and the estimates in Proposition 4.4 that $f \in C_t^0 C_x^{2\alpha - 1}$.

Acknowledgement

The authors would like to express their gratitude to the anonymous referee for his/her valuable comments and constructive suggestions which have helped improve the manuscript.

References

[1] J. Bear. *Dynamics of Fluids in Porous Media*. Courier Dover Publications, New York, 1972.

[2] T. Buckmaster, C. De Lellis, P. Isett, and L. Székelyhidi. *Anomalous dissipation for* 1/5-*Hölder Euler flows*. Ann. of Math., Vol. 182, No. 1: 127-172, 2015.

[3] T. Buckmaster, C. De Lellis, and L. Székelyhidi. *Dissipative Euler flows with Onsager-critical spatial regularity*. Comm. Pure Appl. Math., Vol. 69 No. 9, 16131670, 2016.

[4] T. Buckmaster, S. Shkoller, and V. Vicol. *Nonuniqueness of weak solutions to the SQG equation*. Comm. Pure Appl. Math., Vol. LXXII, 1809–1874, 2019.

[5] T. Buckmaster, and V. Vicol. *Nonuniqueness of weak solutions to the Navier-Stokes equation*. Ann. of Math., (2) 189(1): 101–144, 2019.

[6] A. Bulut, M.K. Huynh, and S. Palasek. *Non-uniqueness up to the Onsager threshold for the forced SQG equation*. arXiv:2310.12947, 2023.

[7] L.A. Caffarelli and A. Vasseur. *Drift diffusion equations with fractional diffusion and the quasi-geostrophic equation*. Ann. of Math. (2), 171(3):1903–1930, 2010.

[8] A. Castro, D. Córdoba, and D. Faraco. *Mixing solutions for the Muskat problem*. Inventiones mathematicae, 226:251–348, 2021.

[9] A. Castro, D. Faraco, and B. Gebhard. *Entropy solutions to macroscopic IPM*. arXiv:2309.03637, 2023.

[10] A. Castro, D. Faraco, and F. Mengual. *Localized mixing zone for Muskat bubbles and turned interfaces*. Ann. PDE, 8(1), 7, 50, 2022.

[11] X. Cheng, H. Kwon, and D. Li. *Non-uniqueness of steady-state weak solutions to the surface quasi-geostrophic equations*. Commun. Math. Phys. 388, 1281–1295, 2021.

[12] A. Choffrut, C. De Lellis, and L. Székelyhidi Jr. *Dissipative continuous Euler flows in two and three dimensions*. arXiv:1205.1226, 2012.

[13] P. Constantin, A. J. Majda, and E. Tabak. *Formation of strong fronts in the 2-D quasi-geostrophic thermal active scalar*. Nonlinearity, 7(6): 1495–1533, 1994.

[14] P. Constantin and V. Vicol. *Nonlinear maximum principles for dissipative linear nonlocal operators and applications*. Geometric and Functional Analysis, 22(5): 1289–1321, 2012.

[15] A. Córdoba, D. Faraco and F. Gancedo. *Lack of uniqueness for weak solutions of the incompressible porous media equation*. Arch. Ration. Mech. Anal., 200(3):725–746, 2011.

[16] D. Córdoba, F. Gancedo and R. Orive. *Analytical behavior of two-dimensional incompressible flow in porous media*. J. Math. Phys., 48(6), 065206, 2007.

[17] M. Dai and S. Friedlander. *Uniqueness and non-uniqueness results for dyadic MHD models*. Journal of Nonlinear Science, https://doi.org/10.1007/s00332-022-09868-9, 2022.

[18] M. Dai and Q. Peng. *Non-unique stationary solutions of forced SQG*. arXiv:2302.03283, 2023.

[19] M. Dai and Q. Peng. *Non-unique weak solutions of forced SQG*. arXiv:2310.13537, 2023.

[20] C. De Lellis, and L. Székelyhidi. *Dissipative continuous Euler flows*. Invent. Math., Vol.193 No. 2: 377–407, 2013.

[21] C. De Lellis, and L. Székelyhidi. *Dissipative Euler flows and Onsager's conjecture*. JEMS, Vol.16, No. 7: 1467–1505, 2014.

[22] C. De Lellis, and L. Székelyhidi. *The Euler equations as a differential inclusion*. Ann. of Math., Vol.170 No.3: 1417–1436, 2009.

[23] N. Filonov and P. Khodunov. *Non-uniqueness of Leray-Hopf solutions for a dyadic model*. St. Petersburg Math. J., Vol. 32: 371–387, 2021.

[24] S. Friedlander, W. Rusin and V. Vicol. *On the supercritically diffusive magneto-geostrophic equations*. Nonlinearity, 25(11): 3071-3097, 2012.

[25] S. Friedlander and V. Vicol. *Global well-posedness for an advection-diffusion equation arising in magneto-geostrophic dynamics*. Ann. Inst. H. Poincaré Anal. Non Linéaire, 28(2): 283-301, 2011.

[26] S. Friedlander and V. Vicol. *On the ill/well-posedness and nonlinear instability of the magneto-geostrophic equations*. Nonlinearity, 24(11): 3019-3042, 2011.

[27] P. Isett. *Holder continuous Euler flows with compact support in time*. Ph. D. Thesis, Princeton University, 2013.

[28] P. Isett. *A Proof of Onsager's Conjecture*. Ann. of Math., Vol.188 No.3: 1–93, 2018.

[29] P. Isett and A. Ma. *A direct approach to non-uniqueness and failure of compactness for the SQG equation*. Nonlinearity, 34(5): 3122–3162, 2021.

[30] P. Isett and V. Vicol. *Hölder continuous solutions of active scalar equations*. Ann. PDE., 1(1): 1–77, 2015.

[31] A. Kiselev, F. Nazarov, and A. Volberg. *Global well-posedness for the critical 2D dissipative quasi-geostrophic equation*. Invent. Math., 167(3):445–453, 2007.

[32] F. Marchand. *Existence and regularity of weak solutions to the quasi-geostrophic equations in the spaces L^p or $\dot{H}^{-\frac{1}{2}}$*. Commun. Math. Phys., 277(1):45–67, 2008.

[33] H. K. Moffatt. *Magnetostrophic turbulence and the geodynamo*. In: IUTAM Symposium on Computational Physics and New Perspectives in Turbulence, vol. 4 of IUTAM Bookser. Springer, Dordrecht, pp 339-346, 2008.

[34] H. K. Moffatt and D. E. Loper. *The magnetostrophic rise of a buoyant parcel in the Earth's core*. Geophys. J. Int., 117(2):394-402, 1994.

[35] S.G. Resnick. *Dynamical problems in non-linear adjective partial differential equations*. Ph. D. Thesis, University of Chicago, 1995.

[36] R. Shvydkoy. *Convex integration for a class of active scalar equations*. J. Am. Math. Soc., 24(4): 1159–1174, 2011.

[37] L. Székelyhidi Jr. *Relaxation of the incompressible porous media equation*. Ann. Sci. Éc. Norm. Supér., (4) 45(3): 491–509, 2012.

Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Chicago, IL 60607, USA

Email address: mdai@uic.edu

Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA

Email address: susanfri@usc.edu