Global Infinite Energy Solutions for the 2D Gravity Water Waves System

XUECHENG WANG
Princeton University

Abstract

We prove global existence and a modified scattering property for the solutions of the 2D gravity water waves system in the infinite depth setting for a class of initial data, which is only required to be small above the level $\dot{H}^{1/5} \times \dot{H}^{1/5+1/2}$. No assumption is made below this level. Therefore, the nonlinear solution can have infinite energy. As a direct consequence, the momentum condition assumed on the physical velocity in all previous small energy results by Ionescu-Pusateri, Alazard-Delort, and Ifrim-Tataru is removed. © 2017 Wiley Periodicals, Inc.

1 Introduction

1.1 The Gravity Water Waves System

We consider an incompressible irrotational inviscid fluid with density 1 that occupies a time-dependent domain $\Omega(t)$ with the free interface $\Gamma(t)$ and without a bottom. Above $\Gamma(t)$, there is a vacuum. Assume that the interface $\Gamma(t)$ of the domain $\Omega(t)$ is given as $\Gamma(t) = \partial\Omega(t) = \{(x, h(t,x)) : x \in \mathbb{R}\}$; then $\Omega(t) = \{(x,y) : y \leq h(t,x), x \in \mathbb{R}\}$.

Let v and p denote the velocity and the pressure of the fluid, respectively. As the evolution of the fluid is described by the Euler equations with a free boundary, then v and p satisfy the following system of equations:

$$(1.1)\quad \begin{cases} \partial_t v(t,X) + v(t,X)\cdot\nabla v(t,X) = -\nabla p(t,X) - g(0,1), & X \in \Omega(t), \\ \nabla\cdot v(t,X) = 0,\ \nabla\times v(t,X) = 0, & X \in \Omega(t), \\ \partial_t + v\cdot\nabla \text{ is tangent to} \bigcup_t \Gamma(t),\ \ p(t,x) = 0, & x \in \Gamma(t), \\ v(0,X) = v_0(X), & X \in \Omega(0), \end{cases}$$

where g is the gravitational constant, which is assumed to be 1 throughout this paper, and $v_0(X)$ is the initial physical velocity field.

As the fluid is irrotational, we assume that the velocity field is given by the gradient of a potential function ϕ. Let $\psi(t,x) := \phi(t,x,h(t,x))$ be the restriction of potential ϕ to the boundary $\Gamma(t)$. From the incompressible condition, we have

Communications on Pure and Applied Mathematics, Vol. LXXI, 0090–0162 (2018)

the following Laplace equation with a Dirichlet-type boundary condition at the interface $\Gamma(t)$,

$$v(t,X) = \nabla\phi(t,X), \quad \Delta\phi(t,X) = 0, X \in \Omega(t), \quad \phi(t,X)\big|_{\Gamma(t)} = \psi(t,x). \tag{1.2}$$

Therefore, it is sufficient to study the evolution of the system at the interface $\Gamma(t)$, i.e., the evolution of the height $h(t,x)$ and the restricted velocity potential $\psi(t,x)$ at the interface.

Following Zakharov [37] and Craig-Sulem-Sulem [11], we can derive the system of equations satisfied by $(h,\psi) : \mathbb{R}_t \times \mathbb{R}_x \to \mathbb{R}$ as follows:

$$\begin{cases} \partial_t h = G(h)\psi \\ \partial_t \psi = -h - \frac{1}{2}|\partial_x\psi|^2 + \dfrac{(G(h)\psi + \partial_x h \partial_x \psi)^2}{2(1+|\partial_x h|^2)}, \end{cases} \tag{1.3}$$

where $G(h)\psi = \sqrt{1+|\partial_x h|^2}\mathcal{N}(h)\psi$ and $\mathcal{N}(h)\psi$ is the Dirichlet-Neumann operator associated with the domain $\Omega(t)$. The system (1.3) is generally referred to as the gravity water waves system. It has the conserved Hamiltonian as follows:

$$\mathcal{H}(h,\psi) := \int_{\mathbb{R}} \frac{1}{2}\psi G(h)\psi + \frac{1}{2}|h|^2\, dx \approx \frac{1}{2}\big[\|h\|_{L^2}^2 + \||\nabla|^{\frac{1}{2}}\psi\|_{L^2}^2\big]. \tag{1.4}$$

1.2 Previous Results

There is an extensive literature on the water waves problems. Without trying to be exhaustive, we only list some representative references on the Cauchy problem. For blowup behavior and splash singularity, please refer to [7] and references therein for more details.

We mention that the motion of the boundary is subject to the Taylor instability when the surface tension effect is neglected. As proved by Wu [32, 33], she showed that, as long as the interface is non-self-intersecting, the Taylor sign condition $-\partial p/\partial \mathbf{n} \geq c_0 > 0$ always holds for the n-dimensional infinite-depth gravity water waves system, where $n \geq 2$ ($n = 2, 3$ for physical relevance). Hence the Taylor instability is not an issue.

On the local theory side of the water waves system, we have the work of Nalimov [28] and the work of Yosihara [36] for the small initial data, the works of Wu [32, 33] for the general initial data in Sobolev spaces, and subsequent works by Christodoulou-Lindblad [8], Lannes [25], Lindblad [26], Coutand-Shkoller [9], Shatah-Zeng [30], and Alazard-Burq-Zuily [2]. Local well-posedness also holds when the surface tension effect is considered. Please refer to the works by Beyer-Gunther [6], Ambrose-Masmoudi [5], Coutand-Shkoller [9], Shatah-Zeng [30], and Alazard-Burq-Zuily [1] for more details.

On the long-time behavior side, we start with the breakthrough work of Wu [34], where she proved almost global existence for the 2D gravity water waves system for small initial data; then Germain-Masmoudi-Shatah [14] and Wu [35] proved global existence for the 3D gravity water waves system for small initial data. When the surface tension effect is considered but the gravity effect is neglected (the so-called capillary waves system), Germain-Masmoudi-Shatah [15] proved global existence of the 3D capillary waves system for small initial data.

Due to the slower decay rate in 2D, it is considerably more difficult to prove global existence than the 3D case. Until very recently, we had several results. Global existence of the 2D gravity water waves for small initial data was first proved by Ionescu-Pusateri [21], and a similar result was proved independently by Alazard-Delort [3] in Eulerian coordinates. More recently Hunter-Ifrim-Tataru [16] used holomorphic coordinates to give a different proof of the almost global existence result; then Ifrim-Tataru [19] extended it to global existence in holomorphic coordinates. Ionescu-Pusateri [24] proved the global existence of the 2D capillary waves system for small initial data without the momentum condition on the associated profile in the Eulerian coordinates. Ifrim-Tataru [17] proved the global existence of the same system for small initial data with the momentum condition on the associated profile in holomorphic coordinates.

1.3 Momentum Condition

We mentioned that all initial data are assumed to be small at the level of $L^2 \times \dot{H}^{1/2}$ in all previous results [3, 19, 21]. That is to say, $\||\nabla|^{1/2}\psi\|_{L^2}$ is not only finite but also small. Intuitively speaking, the following holds for the physical velocity:

$$\||\nabla|^{\frac{1}{2}}\psi\|_{L^2}^2 \sim \||\nabla|^{-1/2}v\|_{L^2}^2 = \int_{\mathbb{R}} \frac{1}{|\xi|}|\widehat{v}(\xi)|^2\, d\xi < \infty. \tag{1.5}$$

Hence, the finiteness of $\||\nabla|^{1/2}\psi\|_{L^2}^2$ implies that the physical velocity is neutral (i.e., $\widehat{v}(0) = \int v(x)dx = 0$) and behaves very nicely at the very low frequency. However, this assumption does not generally hold, even for a Schwartz function.

So a natural question is what if the physical velocity is not neutral. Do we still have a global solution for the gravity water waves system (1.3)? The main goal of this paper is to show that we do have global solutions, which can have infinite energy. The L^2 norm of $|\nabla|^{1/2}\psi$ is not necessary to be finite. Hence, the physical velocity at the interface is not necessary to be neutral.

1.4 Main Result

Before stating our main theorem, we define the main function spaces that we will use very often. We use k_+ to denote $\max\{k, 0\}$ and k_- to denote $\min\{0, k\}$ throughout this paper. We define function spaces as follows:

$$\begin{aligned}\|f\|_{H^{N,p}} &:= \Big[\sum_{k\in\mathbb{Z}}(2^{Nk} + 2^{pk})^2\|P_k f\|_{L^2}^2\Big]^{\frac{1}{2}} \\ &\approx \|(|\partial_x|^p + |\partial_x|^N)f\|_{L^2}, \quad 0 \le p \le N,\end{aligned} \tag{1.6}$$

$$\begin{aligned} \|f\|_{W^{\gamma,b}} &:= \sum_{k\in\mathbb{Z}} \big(2^{\gamma k_+} + 2^{bk_-}\big)\|P_k f\|_{L^\infty}, \\ \|f\|_{W^{\gamma}} &:= \sum_{k\in\mathbb{Z}} (2^{N_2 k} + 1)\|P_k f\|_{L^\infty}, \end{aligned} \tag{1.7}$$

$$\|f\|_{\widetilde{W^{\gamma}}} := \|P_{\le 0} f\|_{L^\infty} + \sum_{k\ge 0} 2^{\gamma k}\|P_k f\|_{L^\infty}, \quad \gamma \ge 0. \tag{1.8}$$

Our main result is stated as follows:

THEOREM 1.1. *Let* $N_0 = 8$, $N_1 = 1$, $N_2 = \frac{61}{20}$, $p = \frac{1}{5}$, *and* $p_0 \in (0, 10^{-10}]$. *Assume that* $(h_0, \psi_0) \in H^{N_0+1/2,p} \times (H^{N_0+1/2,1/2+p} \cap L^\infty)$ *satisfies the following smallness condition*:

$$\|h_0\|_{H^{N_0+1/2,p}} + \|\psi_0\|_{H^{N_0+1/2,1/2+p}} + \|x\partial_x h_0\|_{H^{N_1+1/2,p}} + \|x\partial_x \psi_0\|_{H^{N_1+1/2,1/2+p}} \le \epsilon_0, \tag{1.9}$$

where ϵ_0 *is a sufficiently small constant. Then there exists a unique global solution* (h, ψ) *of the system* (1.3) *with initial data* (h_0, ψ_0). *Moreover, the following estimates hold*:

$$\sup_{t\in[0,\infty)} (1+t)^{-p_0}\big[\|h\|_{H^{N_0,p}} + \|\psi\|_{H^{N_0,1/2+p}} + \|Sh\|_{H^{N_1,p}} + \|S\psi\|_{H^{N_1,1/2+p}}\big] \lesssim \epsilon_0, \tag{1.10}$$

$$\sup_{t\in[0,\infty)} (1+t)^{\frac{1}{2}}\|(h, |\nabla|^{\frac{1}{2}}\psi)\|_{W^{N_2}} \lesssim \epsilon_0, \tag{1.11}$$

where $S := t\partial_t + 2x\partial_x$ *is the scaling vector field associated with the system* (1.3). *Furthermore, the solution possesses the modified scattering property as* $t \to +\infty$.

Remark 1.2. Note that $p = \frac{1}{5}$ in the above theorem is not optimal; we can improve it after being more careful with the argument used here. However, given the methods used in this paper, p should be strictly less than $\frac{1}{4}$. See Remark 4.3 for more details.

Remark 1.3. Note that the function space L^∞ is used because the $H^{N,p}$-norm does not define a natural space of distribution when $p > \frac{1}{2}$. We remark that the assumption $\psi \in L^\infty$ does not imply the neutrality of the physical velocity. For a nonneutral Schwartz physical velocity v, the restricted velocity potential behaves like the antiderivative of v, which belongs to L^∞.

Remark 1.4. The property of the modified scattering in the infinite energy setting is the same as the small energy setting in [21, 23]. We first describe the modified scattering property and then give a comment.

Define $f(t) := h(t) + i|\nabla|^{1/2}\psi(t)$ and

$$G(t,\xi) := \frac{|\xi|^4}{\pi}\int_0^t |\widehat{f}(s,\xi)|^2 \frac{ds}{s+1}. \tag{1.12}$$

Then, there exist ω_∞ and $p_1 < p_0$ such that

$$\sup_{t\in[0,\infty)} (1+t)^{p_1} \big\| |\xi|^{1/4}(1+|\xi|)^{N_2} \cdot \big(e^{iG(t,\xi)}\widehat{f}(t,\xi) - e^{-it|\xi|^{1/2}}\omega_\infty(\xi)\big)\big\|_{L^2_\xi} \lesssim \epsilon_0. \tag{1.13}$$

We can see that after rotating the profile $\widehat{f}(t,\xi)$ through the angle of $G(t,\xi)$, it approaches a linear solution. We can also write this modified scattering behavior in the physical space (t,x). Following the argument in [23], one can show that there exists a uniformly bounded function f_∞ such that

$$\Big| f(t,x) - \frac{e^{-it|t/4|x||}}{\sqrt{1+|t|}} f_\infty\Big(\frac{x}{t}\Big) \cdot \exp\Big[-\frac{i|f_\infty(x/t)|^2}{64|x/t|^5}\log(1+|t|)\Big]\Big| \lesssim \epsilon_0(1+|t|)^{-(1+p_1)/2}. \tag{1.14}$$

Note that the same formula has also been derived by Alazard-Delort [3] and Ifrim-Tataru [19].

We comment that the modified scattering property is only a by-product of the L^∞ decay estimate. Note that the modified scattering property happens because of the space-time resonance set, which is only a point. At this point, all inputs of the cubic terms have the same size of frequency. For this case, the symbol contributes $\frac{5}{2}$ degrees of smallness. Although the spaces we use in the infinite energy setting are weaker than spaces used in the small energy setting, the infinite energy setting has little effect on the modified scattering. We modify the profile in the same way as in the small energy setting in [23] to get the sharp decay rate and the modified scattering property.

1.5 Main Difficulties and Main Ideas of the Proof

A very essential observation, which makes it possible to consider the infinite energy solution, is that the nonlinearities of the system (1.3) only depend on the steepness, $\partial_x h$, and the physical velocity, $\partial_x \psi$, in the infinite depth setting. In Appendix B, we will show this observation by using a fixed-point-type method to analyze the Dirichlet-Neumann operator. This method is not only interesting on its own but also can be applied to other settings, for example, the flat bottom setting.

In the infinite energy setting, we also have to deal with the issues in the small energy setting, which can be summarized as the losing derivatives issue and the slow decay rate issue. Besides those difficulties, we also confront an additional difficulty, which comes from the fact that we do not make any assumptions below $\dot{H}^{1/5}$. The additional difficulty is that we lose $\frac{1}{5}$ derivatives at the low frequency in the sense that only $2^{k/5}\|P_k u\|_{L^2}$, $k \le 0$, is controlled when we put an input in L^2. The issue of losing derivatives at the low frequency is very problematic in the energy estimate of Sh and $S\psi$.

To avoid losing derivatives at the high frequency, we need to find good substitution variables for h and ψ. As the system (1.3) lacks symmetries, one fails at the beginning of the energy estimate due to the quasilinear nature. Thanks to the work of Alazard-Métivier [4], the works of Alazard-Burq-Zuily [1,2], and the work of the Alazard-Delort [3], their paralinearization method helps us to see the good structures inside the system (1.3) and to find the good substitution variables U^1 and U^2. The system of equations satisfied by U^1 and U^2 has requisite symmetries to avoid losing derivatives when doing the energy estimate.

For intuitive purposes, we postpone the detailed formulas of U^1 and U^2 to (2.16). At this moment, it is enough to know that the difference between (U^1, U^2) and $(h, |\nabla|^{1/2}\psi)$ is quadratic and higher order. Thus, (U^1, U^2) and $(h, |\nabla|^{1/2}\psi)$ have comparable sizes of the L^∞-decay rate and the total energy.

To establish the global existence, very naturally, the first step is to establish the local existence for the infinite energy initial data. Recall that the the local well-posedness for finite energy initial data was obtained first in [32] and then in [2]. We remark that, with minor modifications, local well-posedness also holds for the initial data with infinite energy, which is small in the $H^{N,p}$ space. The main idea behind this is still the observation that nonlinearities only depend on the steepness $\partial_x h$ and the physical velocity $\partial_x \psi$. We discuss this point in more detail in Section 2.3.

Therefore, to establish global well-posedness and to prove our main theorem, we use the standard bootstrap argument. To close the argument, it is sufficient to control the following quantities over time:

$$\|(U^1, U^2)(t)\|_{H^{N_0,p}}, \quad \|(SU^1, SU^2)\|_{H^{N_1,p}}, \quad \|(U^1, U^2)(t)\|_{W^{N_2}}.$$

Energy Estimate for (U^1, U^2)

We now discuss the main ideas of controlling $\|(U^1, U^2)(t)\|^2_{H^{N_0,p}}$ over time. As mentioned before, there are three difficulties: (i) losing derivatives at the high frequency, (ii) the slow decay rate, and (iii) losing derivatives at the low frequency. We will address how to get around those difficulties one by one.

Recall that the system of equations satisfied by U^1 and U^2 has requisite symmetries inside, which help us to get around the potential difficulty of losing derivatives at the high frequency. However, due to the slow $t^{-1/2}$ decay rate, the cubic terms inside the derivative of usual energy are troublesome.

To get around this issue, we will use the normal form transformation to find the cubic correction terms, which cancel out those problematic terms. A very important observation is that the normal form transformation is not singular because the symbol of quadratic terms covers the loss of doing the normal form transformation. Then, we add the cubic correction terms to the usual energy. As a result, the derivative of the modified energy is quartic and higher, which has sufficient decay rate.

Because of the cubic correction terms, due to the quasilinear nature, it is possible to lose derivatives again after taking a derivative with respect to time. To get around this issue, we also add quartic correction terms to the usual energy, which enables us to see cancellations by utilizing again the symmetries inside the system of equations satisfied by U^1 and U^2. As a result, the derivative of the modified energy does not lose derivatives and has the critical $\frac{1}{t}$ decay rate.

Lastly, we discuss the issue of losing $\frac{1}{5}$ derivatives at the low frequency. For intuitive purposes, we use the following terms from the time derivative of modified energy as an example:

$$\int_{\mathbb{R}} \partial_x^p \big(Q_1(U^1, U^2) + \mathcal{R}_1\big) \partial_x^p A(U^1, U^1) dx, \quad p = \frac{1}{5}, \tag{1.15}$$

where $Q_1(U^1, U^2)$ represents the quadratic terms of $\partial_t U^1$, $\mathcal{R}_1$ represents the cubic and higher-order remainder term of $\partial_t U^1$, and $A(\cdot,\cdot)$ represents one of the normal form transformations that we will do later.

Note that the quartic term is not problematic, since we can always put the input with larger frequency in L^2 and the input with smaller frequency in L^∞. Hence, the quartic terms do not lose derivatives at the low frequency.

The remainder term is also not problematic, as the L^2-type and the L^∞-type estimates of the Dirichlet-Neumann operator only depend on $\partial_x h$ and $\partial_x \psi$ as long as a certain smallness condition on $\partial_x h$ is satisfied. This observation guarantees that we gain $\frac{1}{5}$ derivatives at the low frequency.

Energy Estimate for (SU^1, SU^2)

Many parts of the energy estimate of SU^1 and SU^2 are the same as the energy estimate of U^1 and U^2. We still need to utilize symmetries to avoid losing derivatives at the high frequency and to add correction terms to cancel out the slow-decay cubic terms.

However, a major difference is that inputs SU^1 and SU^2 are forced to be put in L^2 even when they have relatively smaller frequencies. For the case when SU^1 and SU^2 can be safely put in L^2, we will redo the procedures that we did in the energy estimate of U^1 and U^2. Hence we only need to focus on the case when SU^1 and SU^2 have relatively smaller frequencies and cannot be safely put in L^2.

To illustrate our strategy, we use the problematic cubic terms from the derivative of usual energy of (SU^1, SU^2) as follows as an example:

$$\int_{\mathbb{R}} \partial_x^p P_k(SU^1) \partial_x^p P_k\big(Q_1(P_{k_1}(SU^1), P_{k_2}(U^1))\big), \qquad k_1 \le k_2 - 10,\ k_1, k_2, k \in \mathbb{Z}. \tag{1.16}$$

Note that the symbols of quadratic terms contribute to the smallness of $2^{k_1/2}$, which is essential.

If one uses the normal form transformation directly, then the above problematic cubic terms will vanish. However, we will have the following quartic term:

$$I := \int_{\mathbb{R}} \partial_x^p Q_1(P_{k_1'}[SU^1], P_{k_2} U^2) \partial_x^p A(P_{k_1}[SU^1], P_{k_2}[U^1]) dx, \qquad k_1, k_1' \le k_2 - 10. \tag{1.17}$$

After putting SU^1 in L^2 and U^i, $i \in \{1,2\}$, in L^∞, we have

$$|I| \lesssim 2^{-pk_1 - pk_1'} \frac{1}{t} E(t), \quad p = \frac{1}{5},$$

where $E(t)$ represents the energy of the nonlinear solution. Hence, it is not sufficient to close the argument when k_1 and k_1' are sufficiently small. This difficulty does not appear in the small energy setting, because $p = 0$ in the small energy setting.

To get around this difficulty, we first identify the most problematic term by using a more subtle Fourier method and then show that the most problematic term is not bad in the sense that the associated phase has a good lower bound. As a result, we can divide the phase again to gain an extra $t^{-1/2}$ decay rate, which covers the loss of derivatives at the low frequency. We explain our strategy in more detail as follows.

Note that the cubic term in (1.16) actually has $\frac{1}{t}$ decay if $2^{(1-p)k_1} \leq (1+t)^{-1}$. Therefore we only have to use the normal form transformation to cancel out (1.16) when $2^{(1-p)k_1} \geq (1+t)^{-1}$, i.e., $2^{k_1} \geq (1+t)^{-5/4}$.

Next, we find out all problematic quartic terms. For intuitive purposes, we use the quartic term in (1.17) as an example. Note that the symbol in (1.17) contributes to the smallness of $2^{k_1'/2}$. As a result, we know that the loss of $2^{-pk_1 - pk_1'}$ can be covered by the symbol when $k_1 + 10 \geq k_1'$ or $2^{(1/2-p)k_1'} \leq (1+t)^{-1/4}$. Hence, it is only problematic when $k_1 + 10 \leq k_1'$ and $2^{(1/2-p)k_1'} \geq (1+t)^{-1/4}$.

A very important observation for this problematic quartic term is that the size of associated phases has a good lower bound. More precisely, the associated phases for quartic terms are given as follows:

$$\Phi^{\mu,\nu,\tau}(\xi,\eta,\sigma) = |\xi|^{1/2} - \mu|\xi-\eta|^{1/2} - \nu|\eta-\sigma|^{1/2} - \tau|\sigma|^{1/2}, \quad \mu,\nu,\tau \in \{+,-\}.$$

Recall that $k_2 - 10 \geq k_1' \geq k_1 + 10$ in the problematic case. For this case, the following estimate holds:

$$|\Phi^{\mu,\nu,\tau}(\xi,\eta,\sigma)| \psi_{k_1'}(\xi) \psi_{k_2}(\xi-\eta) \psi_{k_1}(\eta-\sigma) \psi_{k_2}(\sigma) \gtrsim 2^{k_1'/2}. \tag{1.18}$$

Hence, the price of dividing the phases $|\Phi^{\mu,\nu,\tau}(\xi,\eta,\sigma)|$ can be paid by the smallness of the symbol in (1.17). As a result, we can gain an extra $t^{-1/2}$ decay rate, which is sufficient to cover the loss of $2^{-pk_1 - pk_1'} \leq (1+t)^{1/4+1/6}$ from putting SU^1 in $\dot{H}^p$. Note that we have used the fact that $2^{k_1} \geq (1+t)^{-5/4}$ and $2^{k_1'} \geq (1+t)^{-5/6}$.

For all other problematic quartic terms, a similar estimate as in (1.18) also holds. Hence, we can use the same argument to close the energy estimate of SU^1 and SU^2.

The Sharp L^∞-Decay Estimate for (U^1, U^2)

From the energy estimates and the linear dispersion estimate in Lemma 4.2, to prove the sharp decay rate, it is sufficient to prove that the L^∞_ξ-type norm (i.e., the Z-norm, which is defined in (4.3)) of the profile does not grow with respect to time.

Note that one has to be very careful when trying to define an appropriate Z-norm. We cannot let the Z-norm to be very strong, because the finiteness of the Z-norm implies the finiteness of energy. As a result, the Z-normed space that we will use in this paper is weaker than the Z-normed space used in [21]. To work in this weaker Z-normed space, we have to be more careful when doing estimates at the low frequency. It turns out that the difficulties are still manageable. The ideas of proving the L^∞-type decay estimate are mainly based on the works of Ionescu-Pusateri [20–22, 24]. To improve the understanding of the L^∞ decay estimate, many ideas are combined. As a result, the argument we present here is more concise.

1.6 Outline

In Section 2, we fix the notation, find the good unknown variables for the water waves system (1.3), state the bootstrap assumption, and then reduce the proof of the main theorem 1.1 into proofs of two main propositions. In Section 3, we prove that the energy of the nonlinear solution only grows at most like $(1+t)^{p_0}$ with respect to time. In Section 4, we prove the sharp decay rate of the L^∞-norm of the nonlinear solution. In Appendix A, we show that the requisite estimates of all remainder terms also hold in the infinite energy setting. In Appendix B, we use a fixed-point-type method to analyze the Dirichlet-Neumann operator. As a result, one can see that the Dirichlet-Neumann operator only depends on the steepness of interface and the physical velocity, which is very essential to the existence of a infinite energy solution. In Appendix C, we do paralinearization for the full system (1.3) to show that the remainder terms do not lose derivatives.

2 Preliminaries

2.1 Notation and the Multilinear Estimate

We fix an even smooth function $\widetilde{\psi} : \mathbb{R} \to [0,1]$, which is supported in $[-\frac{3}{2}, \frac{3}{2}]$ and is equal to 1 in $[-\frac{5}{4}, \frac{5}{4}]$. For any $k \in \mathbb{Z}$, define

$$\psi_k(x) := \widetilde{\psi}(x/2^k) - \widetilde{\psi}(x/2^{k-1}),$$
$$\psi_{\leq k}(x) := \widetilde{\psi}(x/2^k), \quad \psi_{\geq k}(x) := 1 - \psi_{\leq k-1}(x).$$

Denote the projection operators P_k, $P_{\leq k}$, and $P_{\geq k}$ by the Fourier multipliers ψ_k, $\psi_{\leq k}$, and $\psi_{\geq k}$, respectively. We use f_k to abbreviate $P_k f$. For a well-defined nonlinearity $\mathcal{N}$ and $p \in \mathbb{N}_+$, we will use $\Lambda_p(\mathcal{N})$ to denote the p^{th}-order terms of the nonlinearity $\mathcal{N}$ when a Taylor expansion of this nonlinearity is available. We use $\Lambda_{\geq p}(\mathcal{N}) := \sum_{q \geq p} \Lambda_p(\mathcal{N})$ ($\Lambda_{\leq p}(\mathcal{N})$) to denote the p^{th}- and higher- (lower-) order terms of the nonlinearity $\mathcal{N}$. For example, $\Lambda_3(\mathcal{N})$ denotes the cubic terms of $\mathcal{N}$ and $\Lambda_{\geq 3}(\mathcal{N})$ denotes the cubic and higher terms of $\mathcal{N}$. The cubic and lower-order terms of the Dirichlet-Neumann operator are given as follows:

$$\Lambda_{\leq 3}[G(h)\psi] = |\nabla|\psi - |\nabla|(h|\nabla|\psi),$$
$$-\partial_x(h\partial_x\psi) + |\nabla|(h|\nabla|(h|\nabla|\psi)) + \frac{|\nabla|(h^2\partial_x^2\psi) + \partial_x^2(h^2|\nabla|\psi)}{2}, \tag{2.1}$$

which can be found in [3, 21].

The Fourier transform is defined as follows:

$$\widehat{f}(\xi) = \mathcal{F}(f)(\xi) = \int e^{-ix\xi} f(x)dx.$$

For two well-defined functions f and g and a bilinear form $Q(f,g)$, the symbol $q(\cdot,\cdot)$ of $Q(\cdot,\cdot)$ is defined in the following sense throughout this paper,

$$\mathcal{F}[Q(f,g)](\xi) = \frac{1}{2\pi}\int_{\mathbb{R}^2} \widehat{f}(\xi-\eta)\widehat{g}(\eta)q(\xi-\eta,\eta)d\eta. \tag{2.2}$$

For a trilinear form $C(f,g,h)$, its symbol $c(\cdot,\cdot,\cdot)$ is defined in the following sense:

$$\mathcal{F}[C(f,g,h)](\xi) = \frac{1}{4\pi^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \widehat{f}(\xi-\eta)\widehat{g}(\eta-\sigma)\widehat{h}(\sigma)c(\xi-\eta,\eta-\sigma,\sigma)d\eta\, d\sigma.$$

For $a, f \in L^2$ and a pseudodifferential operator $\widetilde{a}(x,\xi)$, we define the operators $T_a f$ and $T_{\widetilde{a}} f$ as follows:

$$T_a f = \mathcal{F}^{-1}\left[\int_{\mathbb{R}} \widehat{a}(\xi-\eta)\theta(\xi-\eta,\eta)\widehat{f}(\eta)d\eta\right],$$
$$T_{\widetilde{a}} f = \mathcal{F}^{-1}\left[\int_{\mathbb{R}} \mathcal{F}_x(\widetilde{a})(\xi-\eta,\eta)\theta(\xi-\eta,\eta)\widehat{f}(\eta)d\eta\right],$$

where the cutoff function is defined as follows:

$$\theta(\xi-\eta,\eta) = \begin{cases} 1 & \text{when } |\xi-\eta| \leq 2^{-10}|\eta|, \\ 0 & \text{when } |\xi-\eta| \geq 2^{10}|\eta|. \end{cases}$$

We also use $\widetilde{\theta}(\xi-\eta,\eta)$ to denote the cutoff function such that $\xi-\eta$ and η have comparable size inside the support. More precisely,

$$\widetilde{\theta}(\xi-\eta,\eta) := 1 - \theta(\xi-\eta,\eta) - \theta(\eta,\xi-\eta). \tag{2.3}$$

Let the Fourier multiplier $\widetilde{\theta}(\xi-\eta,\eta)$ define the bilinear operator $R_{\mathcal{B}}(\cdot,\cdot)$. Therefore, the following paraproduct decomposition holds for two well-defined functions a and b:

$$ab = T_a b + T_b a + R_{\mathcal{B}}(a,b). \tag{2.4}$$

DEFINITION 2.1. Given $\rho \in \mathbb{N}_+$, $\rho \geq 0$, and $m \in \mathbb{R}$, we use $\Gamma^m_\rho(\mathbb{R}^n)$ to denote the space of locally bounded functions $a(x,\xi)$ on $\mathbb{R}^n \times (\mathbb{R}^n/\{0\})$, which are C^∞ with respect to ξ for $\xi \neq 0$. Moreover, they satisfy the estimate

$$\forall |\xi| \geq \frac{1}{2} \quad \left\| \partial^\alpha_\xi a(\cdot,\xi) \right\|_{W^{\rho,\infty}} \lesssim_\alpha (1+|\xi|)^{m-|\alpha|}, \quad \alpha \in \mathbb{N}^n,$$

where $W^{\rho,\infty}$ is the usual Sobolev space.

For a symbol $a \in \Gamma^m_\rho$, we can define its norm as follows:

$$M^m_\rho(a) := \sup_{|\alpha| \leq 2+\rho} \sup_{|\xi| \geq 1/2} \left\| (1+|\xi|)^{|\alpha|-m} \partial^\alpha_\xi a(\cdot,\xi) \right\|_{W^{\rho,\infty}}.$$

We will use the following composition lemma for paradifferential operators very often. It can be found, for example, in [1, 27].

LEMMA 2.2 (Composition lemma). *Let $m \in \mathbb{R}$ and $\rho > 0$. If given symbols $a \in \Gamma^m_\rho(\mathbb{R}^d)$ and $b \in \Gamma^{m'}_\rho(\mathbb{R}^d)$, we define*

$$a \sharp b = \sum_{|\alpha| < \rho} \frac{1}{i^{|\alpha|}\alpha!} \partial^\alpha_\xi a \partial^\alpha_x b;$$

then for all $\mu \in \mathbb{R}$, there exists a constant K such that

$$\|T_a T_b - T_{a\sharp b}\|_{H^\mu \to H^{\mu-m-m'+\rho}} \leq K M^m_\rho(a) M^{m'}_\rho(b). \tag{2.5}$$

We will also use the following equality very often. Recall that $S := t\partial_t + 2x\partial_x$. For a bilinear operator $A(\cdot,\cdot)$ with symbol $a(\cdot,\cdot)$, we have

$$SA(f,g) = A(Sf,g) + A(f,Sg) + \widetilde{A}(f,g),$$
$$\widetilde{a}(\xi,\eta) = -2(\xi\partial_x \xi a(\xi,\eta) + \eta\partial_\eta a(\xi,\eta)).$$

In particular, when $a(\cdot,\cdot)$ is homogeneous of degree λ, we have

$$SA(f,g) = A(Sf,g) + A(f,Sg) - 2\lambda A(f,g),$$

which can be derived from the equality $\xi\partial_\xi a(\xi,\eta) + \eta\partial_\eta a(\xi,\eta) = \lambda a(\xi,\eta)$.

Define a class of symbols and its associated norms as follows:

$$\mathcal{S}^\infty := \{m : \mathbb{R}^2 \text{ or } \mathbb{R}^3 \to \mathbb{C},\ m \text{ is continuous, and } \|\mathcal{F}^{-1}(m)\|_{L^1} < \infty\},$$
$$\|m\|_{\mathcal{S}^\infty} := \|\mathcal{F}^{-1}(m)\|_{L^1},$$
$$\|m(\xi,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} := \|m(\xi,\eta)\psi_k(\xi)\psi_{k_1}(\xi-\eta)\psi_{k_2}(\eta)\|_{\mathcal{S}^\infty},$$
$$\|m(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} := \|m(\xi,\eta,\sigma)\psi_k(\xi)\psi_{k_1}(\xi-\eta)\psi_{k_2}(\eta-\sigma)\psi_{k_3}(\sigma)\|_{\mathcal{S}^\infty}.$$

LEMMA 2.3. *Assume that m, $m' \in \mathcal{S}^\infty$, $p,q,r,s \in [1,\infty]$, and we have smooth well-defined functions f, g, h, and $\widetilde{f}$. Then the following estimates hold:*

$$\|m \cdot m'\|_{\mathcal{S}^\infty} \lesssim \|m\|_{\mathcal{S}^\infty} \|m'\|_{\mathcal{S}^\infty}, \tag{2.6}$$

$$\Big\|\mathcal{F}^{-1}\Big[\int_{\mathbb{R}} m(\xi,\eta)\widehat{g}(\eta)\widehat{h}(\xi-\eta)d\eta\Big]\Big\|_{L^r} \lesssim \|m\|_{\mathcal{S}^\infty}\|g\|_{L^p}\|h\|_{L^q} \quad \text{if } \frac{1}{r}=\frac{1}{p}+\frac{1}{q}, \tag{2.7}$$

$$\Big\|\mathcal{F}^{-1}\Big[\int_{\mathbb{R}^2} m'(\xi,\eta,\sigma)\widehat{f}(\sigma)\widehat{g}(\eta-\sigma)\widehat{h}(\xi-\eta)d\eta\, d\sigma\Big]\Big\|_{L^s} \lesssim \|m'\|_{\mathcal{S}^\infty}\|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r} \quad \text{if } \frac{1}{s}=\frac{1}{p}+\frac{1}{q}+\frac{1}{r}. \tag{2.8}$$

PROOF. The proof is standard; see [20, 21] for details. □

To estimate the $\mathcal{S}^\infty_{k,k_1,k_2}$-norm and the $\mathcal{S}^\infty_{k,k_1,k_2,k_3}$-norm of symbols, we will use the following lemma very often.

LEMMA 2.4. *For $i \in \{2,3\}$, if $f:\mathbb{R}^i \to \mathbb{C}$ is a smooth function and $k_1,\dots,k_i \in \mathbb{Z}$, then the following estimate holds*:

$$\Big\|\int_{\mathbb{R}^i} f(\xi_1,\dots,\xi_i)\prod_{j=1}^{i} e^{ix_j\xi_j}\psi_{k_j}(\xi_j)d\xi_1\cdots d\xi_i\Big\|_{L^1_{x_1,\dots,x_i}} \lesssim \sum_{m=0}^{i}\sum_{j=1}^{i} 2^{mk_j}\|\partial^m_{\xi_j} f\|_{L^\infty}. \tag{2.9}$$

PROOF. We only consider the case when $i=2$, because the proof of the case when $i=3$ is very similar. Through scaling, it is sufficient to prove the above estimate for the case when $k_1=k_2=0$. From the Plancherel theorem, we have the following two estimates:

$$\Big\|\int_{\mathbb{R}^i} f(\xi_1,\xi_2)e^{i(x_1\xi_1+x_2\xi_2)}\psi_0(\xi_1)\psi_0(\xi_2)d\xi_1 d\xi_2\Big\|_{L^2_{x_1,x_2}} \lesssim \|f(\xi_1,\xi_2)\|_{L^\infty_{\xi_1,\xi_2}},$$

$$\|(|x_1|+|x_2|)^2\int_{\mathbb{R}^i} f(\xi_1,\xi_2)e^{i(x_1\xi_1+x_2\xi_2)}\psi_0(\xi_1)\psi_0(\xi_2)d\xi_1 d\xi_2\|_{L^2_{x_1,x_2}} \lesssim \sum_{m=0}^{2}\big[\|\partial^m_{\xi_1} f\|_{L^\infty}+\|\partial^m_{\xi_2} f\|_{L^\infty}\big],$$

which are sufficient to finish the proof of (2.9). □

2.2 The Good Substitution Variables

At the beginning, we mention that the content of this subsection is not novel. We only briefly introduce and explain the main ideas behind this subsection. More details can be found in Appendix C. Interested readers can refer to [3] for more elaborated details.

The main process of deriving the good substitution variables can be summarized as the paralinearization process and the symmetrization process. Before introducing the main ideas of doing these processes, we define the following variables:

$$B := B(h)\psi := \frac{G(h)\psi + \partial_x h \partial_x \psi}{1 + |\partial_x h|^2}, \quad V := V(h)\psi := \partial_x \psi - \partial_x h B(h)\psi, \tag{2.10}$$

$$a = 1 + \partial_t B + V\partial_x B, \quad \alpha = \sqrt{a} - 1, \tag{2.11}$$

where B and V represent the vertical derivative and the horizontal derivative of the velocity potential at the interface, respectively, and a is the so-called Taylor coefficient.

The main goal of the paralinearization process is to identify the cancellations and highlight the quasilinear structures inside the nonlinearities of the system (1.3).

As a result of paralinearization (one can also see Lemma C.1), we have a good decomposition of the Dirichlet-Neumann operator as follows:

$$G(h)\psi = |\nabla|\omega - \partial_x(T_{V(h)\psi}\eta) + F(h)\psi, \quad \omega := \psi - T_{B(h)\psi}h, \tag{2.12}$$

where $F(h)\psi$ is a quadratic and higher-order good error term, which does not lose derivatives. Recall (1.3). After doing paralinearization for the nonlinearity of $\partial_t\psi$, we have

$$\begin{aligned} &-\frac{1}{2}|\partial_x\psi|^2 + \frac{(G(h)\psi + \partial_x h\partial_x\psi)^2}{2(1+|\partial_x h|^2)} \\ &= -\frac{1}{2}|\partial_x\psi|^2 + \frac{1}{2}(1+|\partial_x h|^2)(B(h)\psi)^2 \\ &= -T_V\partial_x(\psi - T_B h) - T_{V\partial_x B}h + T_B G(h)\psi + \mathcal{R}, \end{aligned} \tag{2.13}$$

where $\mathcal{R}$ is also a quadratic and higher-order good error term, which does not lose derivatives.

Recall the definition of ω in (2.12). From (2.13), we have

$$\begin{aligned} \partial_t\omega &= -h - T_V\partial_x\omega - T_{V\partial_x B}h + T_B\partial_t h - T_{\partial_t B}h - T_B\partial_t h + \mathcal{R} \\ &= -T_a h - T_V\partial_x\omega + \mathcal{R}. \end{aligned}$$

To sum up, we have

$$\begin{cases} \partial_t h = |\nabla|\omega - T_V\partial_x h + F(h)\psi, \\ \partial_t\omega = -T_a h - T_V\partial_x\omega + \mathcal{R}. \end{cases} \tag{2.14}$$

One might find that we cannot use the system (2.14) to do the energy estimate, as the following term loses derivatives and cannot be treated simply:

$$\int \partial_x^k h\partial_x^k|\nabla|\omega - \partial_x^k|\nabla|^{\frac{1}{2}}\omega\partial_x^j|\nabla|^{\frac{1}{2}}(T_a h). \tag{2.15}$$

To get around this issue, we need to symmetrizing the system (2.14) by using the following good substitution variables,

$$U^1 = T_{\sqrt{a}}h = h + T_\alpha h, \quad U^2 = |\nabla|^{\frac{1}{2}}\omega. \tag{2.16}$$

As a result, from (2.14), the system of equations satisfied by U^1 and U^2 can be formulated as follows:

$$\begin{cases} \partial_t U^1 - |\nabla|^{\frac{1}{2}} U^2 = T_\alpha |\nabla|^{\frac{1}{2}} U^2 - T_V \partial_x U^1 + \mathcal{R}_1, \\ \partial_t U^2 + |\nabla|^{\frac{1}{2}} U^1 = -T_\alpha |\nabla|^{\frac{1}{2}} U^1 - T_V \partial_x U^2 + \mathcal{R}_2, \end{cases} \tag{2.17}$$

where $\mathcal{R}_1$ and $\mathcal{R}_2$ are quadratic and higher-order good error terms that do not lose derivatives. Their detailed formulas can be found in (A.1) and (A.2).

As a result of symmetrization, the bulk term in (2.15) essentially becomes

$$\int \partial_x^k U^1 \partial_x^k (T_{\sqrt{a}} |\nabla|^{\frac{1}{2}} U^2) - \partial_x^k U^2 \partial_x^k (T_{\sqrt{a}} |\nabla|^{\frac{1}{2}} U^1) dx,$$

which does not lose derivatives. One can verify this fact by utilizing the symmetries on the Fourier side.

Because we will cancel out the quadratic terms inside the system (2.17), it is important to know exactly what quadratic terms are. Until the end of this section, *we assume that the expansion of the projection operator* $\Lambda_{\geq 3}[\cdot]$ *is in terms of* U^1 *and* U^2.

From (2.1) and (2.16), we can rewrite (2.17) as follows:

$$\begin{cases} \partial_t U^1 - |\nabla|^{\frac{1}{2}} U^2 = Q_1(U^1, U^2) + C_1 + \Lambda_{\geq 3}[\mathcal{R}_1], \\ \partial_t U^2 + |\nabla|^{\frac{1}{2}} U^1 = Q_2(U^1, U^1) + Q_3(U^2, U^2) + C_2 + \Lambda_{\geq 3}[\mathcal{R}_2], \end{cases} \tag{2.18}$$

where $Q_1(\cdot,\cdot)$, $Q_2(\cdot,\cdot)$, and $Q_3(\cdot,\cdot)$ denote quadratic terms and C_1 and C_2 denote the cubic and higher-order terms that at most lose one derivative. More precisely,

$$\begin{aligned} Q_1(U^1, U^2) &= -T_{\partial_x |\nabla|^{-1/2} U^2} \partial_x U^1 - \frac{1}{2} T_{\partial_x^2 |\nabla|^{-1/2} U^2} U^1 \\ &\quad - \frac{1}{2} T_{|\nabla| U^1} |\nabla|^{1/2} U^2 - |\nabla| (U^1 |\nabla|^{1/2} U^2) + |\nabla| (T_{|\nabla|^{1/2} U^2} U^1) \\ &\quad - \partial_x (U^1 \partial_x |\nabla|^{-1/2} U^2) + \partial_x (T_{\partial_x |\nabla|^{-1/2} U^2} U^1), \end{aligned}$$

$$\begin{aligned} Q_2(U^1, U^1) &= \frac{1}{2} \big(|\nabla|^{1/2} \big(T_{|\nabla| U^1} U^1 \big), \\ Q_3(U^2, U^2) &= -|\nabla|^{1/2} \big(T_{\partial_x |\nabla|^{-1/2} U^2} \partial_x |\nabla|^{-1/2} U^2 \big), \\ C_1 &= T_{\Lambda_{\geq 2}[\alpha]} |\nabla|^{1/2} U^2 - T_{\Lambda_{\geq 2}[V]} \partial_x U^1, \\ C_2 &= -T_{\Lambda_{\geq 2}[\alpha]} |\nabla|^{1/2} U^1 - T_{\Lambda_{\geq 2}[V]} \partial_x U^2. \end{aligned} \tag{2.19}$$

We remark that the good error terms $\mathcal{R}_1$ and $\mathcal{R}_2$ contribute the following quadratic terms:

$$\Lambda_2[\mathcal{R}_1](U^1, U^2) = Q_1(U^1, U^2) + \frac{1}{2}T_{|\nabla|U^1}|\nabla|^{\frac{1}{2}}U^2 + T_{\partial_x|\nabla|^{-1/2}U^2}\partial_x U^1,$$
$$\Lambda_2[\mathcal{R}_2](U^1, U^1) = Q_2(U^1, U^1) - \frac{1}{2}T_{|\nabla|U^1}|\nabla|^{\frac{1}{2}}U^1,$$
$$\Lambda_2[\mathcal{R}_2](U^2, U^2) = Q_3(U^2, U^2) + T_{\partial_x|\nabla|^{-1/2}U^2}\partial_x U^2.$$

The symbols $q_i(\cdot,\cdot)$ of quadratic term $Q_i(\cdot,\cdot)$, $i \in \{1,2,3\}$, are given as follows:

$$q_1(\xi-\eta,\eta) = \sum_{i=1,2,3} q_1^i(\xi-\eta,\eta), \tag{2.20}$$

$$q_2(\xi-\eta,\eta) = \frac{1}{2}|\xi|^{1/2}|\eta|\theta(\eta,\xi-\eta), \tag{2.21}$$

$$q_3(\xi-\eta,\eta) = |\xi|^{1/2}(\xi-\eta)\eta|\xi-\eta|^{-1/2}|\eta|^{-1/2}\theta(\eta,\ \xi-\eta), \tag{2.22}$$

where

$$\begin{aligned} q_1^1(\xi-\eta,\eta) &= \left((\xi-\eta)\eta|\eta|^{-1/2} + \frac{|\eta|^{3/2}}{2}\right)\theta(\eta,\xi-\eta),\\ q_1^2(\xi-\eta,\eta) &= -\frac{1}{2}|\xi-\eta||\eta|^{1/2}\theta(\xi-\eta,\eta),\\ q_1^3(\xi-\eta,\eta) &= (-|\xi||\eta|^{1/2} + \xi\eta|\eta|^{-1/2})\widetilde{\theta}(\eta,\xi-\eta). \end{aligned} \tag{2.23}$$

LEMMA 2.5. *If $|k_1-k_2| \geq 5$, the following estimates hold*:

$$\sum_{i=1,2,3} \|q_i(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{\min\{k_1,k_2\}/2+\max\{k_1,k_2\}}. \tag{2.24}$$

$$\begin{aligned} &\left\|q_1^1(\xi-\eta,\eta) + q_1^1(-\xi,\eta)\right\|_{\mathcal{S}^\infty_{k,k_1,k_2}}\\ &\quad + \left\|q_2(\eta,\xi-\eta) + q_1^2(\eta-\xi,\xi)\right\|_{\mathcal{S}^\infty_{k,k_1,k_2}}\\ &\quad + \|q_3(\xi-\eta,\eta) + q_3(-\xi,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{3\min\{k_1,k_2\}/2}. \end{aligned} \tag{2.25}$$

If $|k_1-k_2| \leq 5$, the following estimate holds:

$$\sum_{i=1,2,3} \|q_i(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{k+k_1/2}. \tag{2.26}$$

PROOF. From Lemma 2.4 and the explicit formulas in (2.20), (2.23), (2.21), and (2.22), it is easy to see that our stated estimates hold. □

2.3 The Bootstrap Assumption and Proof of the Main Theorem

Before proceeding to the bootstrap argument, we briefly discuss the proof of local well-posedness in the finite energy setting and explain why the local well-posedness also holds for small initial data in $H^{N,p}$ space, which has infinite energy. Although the formulations of the problem and the main purposes in [1, 2, 32] are different, the main ideas behind the proofs of the local well-posedness are very similar. Since this paper also works in the Eulerian coordinate system, we mainly introduce the main ideas used in [2].

To prove the local well-posedness in the infinite energy setting, we pay attention to the low frequency and redo the argument in [2], which can be summarized in three main steps as follows:

(i) Establish the local well-posedness and the uniform boundedness in a time interval $[0, T]$ for the ϵ-approximation system (parabolic regularization) as follows:

$$\begin{cases} \partial_t h - G(h)\psi = \epsilon\Delta h \\ \partial_t \psi + h + \frac{1}{2}|\partial_x\psi|^2 - \dfrac{(G(h)\psi + \partial_x h\partial_x\psi)^2}{2(1+|\partial_x h|^2)} = \epsilon\Delta\psi \\ (h,\psi)\big|_{t=0} = \big(\mathcal{F}^{-1}[\widetilde{\psi}(\epsilon\xi)\widehat{h_0}(\xi)], \mathcal{F}^{-1}[\widetilde{\psi}(\epsilon\xi)\widehat{\psi_0}(\xi)]\big), \end{cases} \tag{2.27}$$

where $\widetilde{\psi}(\cdot)$ is defined at the beginning of Section 2.1. Note that the time T is determined by the size of initial data and is independent of ϵ.

(ii) Show the convergence of $(h_\epsilon, \psi_\epsilon)$ to (h, ψ), which is unique, in a lower-level regularity space. Due to the quasilinear nature, generally speaking, it is not possible to prove the convergence directly in the top regularity space.

(iii) Show that the limit (h, ψ) is a solution of the water waves system (1.3). Moreover, the limit lies in the space with top regularity and is continuous with respect to time.

With minor modifications in the argument in [2], we can show the local well-posedness for the infinite energy initial data because of three observations as follows:

(i) From the estimates of the Dirichlet-Neumann operator in Lemma B.3, we know that all estimates only depend on $\partial_x h$ and $\partial_x\psi$.

(ii) From the fact that the Dirichlet-Neumann operator is linear with respect to ψ, we know that $G(h)\psi_1 - G(h)\psi_2$ only depends on $\partial_x(\psi_1 - \psi_2)$ and $\partial_x h$.

(iii) From the fixed-point-type structure inside the Dirichlet-Neumann operator (B.7), it is not difficult to know that $G(h_1)\psi - G(h_2)\psi$ only depends on $(\partial_x\psi, \partial_x h_1, \partial_x h_2)$ and $\partial_x(h_1 - h_2)$. We remark that the small-data regime is necessary to guarantee the validity of the fixed-point-type argument in the estimates of the Dirichlet-Neumann operator.

Given the local well-posedness, we will use the bootstrap argument to show that the solution of the system (2.18) for small initial data extends globally, which further gives us the global existence of (h, ψ).

We will first discuss the initial data. Since the sizes of (U^1, U^2) and $(h, |\nabla|^{1/2}\psi)$ are comparable, from the smallness assumption of initial data (h_0, ψ_0) in (1.9), it is not difficult to see that the following estimate holds:

$$\|(U^1, U^2)(0)\|_{H^{N_0,p}} + \|x\partial_x(U^1, U^2)(0)\|_{H^{N_1,p}} \lesssim \epsilon_0. \tag{2.28}$$

Next, we will state our bootstrap assumption. As in the small energy setting, we expect that the energy grows appropriately and the L^∞-type norm of the nonlinear solution decays sharply. This expectation leads to the bootstrap assumption as follows:

$$\begin{aligned}\sup_{t\in[0,T]} (1+t)^{-p_0}\big[\|(U^1, U^2)(t)\|_{H^{N_0,p}} + \|S(U^1, U^2)(t)\|_{H^{N_1,p}}\big] \\ + (1+t)^{1/2}\|(U^1, U^2)\|_{W^{N_2}} \lesssim \epsilon_1 := \epsilon_0^{5/6} \ll 1.\end{aligned} \tag{2.29}$$

Remark 2.6. Recall that $U^2 = |\nabla|^{1/2}(\psi - T_{B(h)\psi}h)$. To put U^2 in a Sobolev space, the high-frequency parts of ψ and h should have the same level of regularity. To make sure that the initial condition (2.28) holds, we required the top regularities of the initial data to be $N_0 + \frac{1}{2}$ and $N_1 + \frac{1}{2}$ instead of N_0 and N_1; see (1.9). The global well-posedness of (U^1, U^2) provides a global $\dot{H}^{N_0}$-estimate of (U^1, U^2), which further gives us the $\dot{H}^{N_0}$-estimate of the height h. Hence, we can only recover the $\dot{H}^{N_0}$-estimate of ψ from the estimate of U^2; see (1.10).

In Section 3, we will prove the following proposition, which is sufficient to show that the total energy appropriately grows:

PROPOSITION 2.7. *Under the bootstrap assumption* (2.29)*, we can define modified energies* $E_{\text{mod}}(t) \approx \|(U^1, U^2)(t)\|^2_{H^{N_0,p}}$ *and* $E^S_{\text{mod}}(t) \approx \|(SU^1, SU^2)(t)\|^2_{H^{N_1,p}}$ *and have the following energy estimate*:

$$\sup_{t\in[0,T]} (1+t)^{1-2p_0}\left[\left|\frac{d}{dt}E_{\text{mod}}(t)\right| + \left|\frac{d}{dt}E^S_{\text{mod}}(t)\right|\right] \lesssim \epsilon_0. \tag{2.30}$$

Therefore,

$$\sup_{t\in[0,T]} (1+t)^{-p_0}\big[\|(U^1, U^2)\|_{H^{N_0,p}} + \|S(U^1, U^2)\|_{H^{N_1,p}}\big] \lesssim \epsilon_0. \tag{2.31}$$

In Section 4, we will now prove the following decay estimate for the L^∞-type norm.

PROPOSITION 2.8. *Under the bootstrap assumption* (2.29) *and the improved energy estimate* (2.31)*, we can derive the following improved decay estimate*:

$$\sup_{t\in[0,T]} (1+t)^{1/2}\|(U^1, U^2)\|_{W^{N_2}} \lesssim \epsilon_0. \tag{2.32}$$

With above two propositions, it is easy to see that our main theorem holds.

3 Energy Estimate

3.1 Normal Form Transformation

We first find out the normal form transformations, which aim to cancel out the quadratic terms. Let

$$V_1 := U^1 + A_1(U^1, U^1) + A_2(U^2, U^2), \quad V_2 := U^2 + B(U^1, U^2), \tag{3.1}$$

where $A_1(\cdot,\cdot)$ and $A_2(\cdot,\cdot)$ are two symmetric bilinear operators. Recall (2.18). Direct computations give us the equations satisfied by V_1 and V_2 as follows:

$$\begin{cases} \partial_t V_1 - |\nabla|^{1/2} V_2 = 2A_1(|\nabla|^{1/2} U^2, U^1) - 2A_2(|\nabla|^{1/2} U^1, U^2) \\ \qquad\qquad - |\nabla|^{1/2} B(U^1, U^2) + Q_1(U^1, U^2) \\ \qquad\qquad + \text{cubic and higher-order terms} \\ \partial_t V_2 + |\nabla|^{1/2} V_1 = B(|\nabla|^{1/2} U^2, U^2) - B(U^1, |\nabla|^{1/2} U^1) \\ \qquad\qquad + |\nabla|^{1/2}(A_1(U^1, U^1) + A_2(U^2, U^2)) \\ \qquad\qquad + Q_2(U^1, U^1) + Q_3(U^2, U^2) \\ \qquad\qquad + \text{cubic and higher-order terms.} \end{cases} \tag{3.2}$$

To make the quadratic terms in (3.2) vanish, it is sufficient for the symbols of the bilinear operators $A_1(\cdot,\cdot)$, $A_2(\cdot,\cdot)$, and $B(\cdot,\cdot)$ to satisfy the following system of equations:

$$\begin{cases} q_1(\xi-\eta,\eta) + 2|\eta|^{1/2} a_1(\xi-\eta,\eta) - 2|\xi-\eta|^{1/2} a_2(\xi-\eta,\eta) \\ \quad - |\xi|^{1/2} b(\xi-\eta,\eta) = 0, \\ q_2(\xi-\eta,\eta) - b(\xi-\eta,\eta)|\eta|^{1/2} + q_2(\eta,\xi-\eta) \\ \quad - b(\eta,\xi-\eta)|\xi-\eta|^{1/2} + 2|\xi|^{1/2} a_1(\xi-\eta,\eta) = 0, \\ q_3(\xi-\eta,\eta) + q_3(\eta,\xi-\eta) + b(\xi-\eta,\eta)|\xi-\eta|^{1/2} \\ \quad + b(\eta,\xi-\eta)|\eta|^{1/2} + 2|\xi|^{1/2} a_2(\xi-\eta,\eta) = 0. \end{cases} \tag{3.3}$$

The solution of the above system of equations is as follows:

$$a_1(\xi-\eta,\eta) = \frac{b(\xi-\eta,\eta)|\eta|^{1/2} - q_2(\xi-\eta,\eta) + b(\eta,\xi-\eta)|\xi-\eta|^{1/2} - q_2(\eta,\xi-\eta)}{2|\xi|^{1/2}}, \tag{3.4}$$

$$a_2(\xi-\eta,\eta) = -\frac{b(\xi-\eta,\eta)|\xi-\eta|^{1/2} + q_3(\xi-\eta,\eta) + b(\eta,\xi-\eta)|\eta|^{1/2} + q_3(\eta,\xi-\eta)}{2|\xi|^{1/2}}, \tag{3.5}$$

$$b(\xi-\eta,\eta) = \frac{(|\xi-\eta| + |\eta| - |\xi|)A(\xi-\eta,\eta) - 2A(\eta,\xi-\eta)|\xi-\eta|^{1/2}|\eta|^{1/2}}{-(|\xi-\eta| + |\eta| - |\xi|)^2 + 4|\xi-\eta||\eta|}, \tag{3.6}$$

where

$$A(\xi-\eta,\eta) := |\xi|^{1/2} q_1(\xi-\eta,\eta) - (q_2(\xi-\eta,\eta) + q_2(\eta,\xi-\eta))|\eta|^{1/2} \\ + (q_3(\xi-\eta,\eta) + q_3(\eta,\xi-\eta))|\xi-\eta|^{1/2}.$$

LEMMA 3.1. *The following estimate holds*:

$$\|a_1(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|a_2(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|b(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{\max\{k_1,k_2\}}. \tag{3.7}$$

PROOF. If $|k_1 - k_2| \geq 10$, then the estimate (2.24) in Lemma 2.5 holds. From (3.4), (3.5), (3.6), and Lemma 2.4, the following estimate holds:

$$\|a_1(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|a_2(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|b(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{\max\{k_1,k_2\}}. \tag{3.8}$$

If $|k_1 - k_2| \leq 10$, then the estimate (2.26) in Lemma 2.5 holds. Note that for this case $q_2(\xi-\eta,\eta) = q_3(\xi-\eta,\eta) = 0$. Hence, from (2.6) in Lemma 2.3, the following estimate holds:

$$\|A(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|A(\eta,\xi-\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{3k/2+k_1/2}, \tag{3.9}$$

which further gives us the estimate

$$\|a_1(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|a_2(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|b(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{k_1}. \quad \square \tag{3.10}$$

3.2 Energy Estimate of U^1 and U^2

As the cubic terms inside the time derivative of energy decay slowly over time, they are problematic when doing the energy estimate. Recall (2.18). The cubic terms at the top regularity level are given as follows:

$$\mathcal{I}_1^{N_0} = \mathfrak{Re}\Big[\int_{\mathbb{R}} \overline{\partial_x^{N_0} U^1} \partial_x^{N_0}\big[Q_1(U^1,U^2)\big] + \overline{\partial_x^{N_0} U^2} \partial_x^{N_0}\big[Q_2(U^1,U^1) + Q_3(U^2,U^2)\big]\Big]. \tag{3.11}$$

Due to the quasilinear nature, we need to utilize the symmetries inside the system (2.18) first to avoid losing a derivative again after adding the cubic correction terms. After switching the roles of inputs of the same type inside $\mathcal{I}_1^{N_0}$ on the Fourier side, we can rewrite $\mathcal{I}_1^{N_0}$ as follows:

$$\mathcal{I}_1^{N_0} = \mathfrak{Re}\Big[\int \frac{1}{2}\overline{\widehat{U^1}(\xi)}\widetilde{q}^1_{N_0}(\xi-\eta,\eta)\widehat{U^1}(\xi-\eta)\widehat{U^2}(\eta) + \overline{\widehat{U^2}(\xi)}\widetilde{q}^2_{N_0}(\xi-\eta,\eta)\widehat{U^1}(\xi-\eta)\widehat{U^1}(\eta) + \int \frac{1}{2}\overline{\widehat{U^2}(\xi)}\widetilde{q}^3_{N_0}(\xi-\eta,\eta)\widehat{U^2}(\xi-\eta)\widehat{U^2}(\eta)\Big], \tag{3.12}$$

where

$$\begin{aligned}\widetilde{q}^1_{N_0}(\xi-\eta,\eta) &= q_1^1(\xi-\eta,\eta)|\xi|^{2N_0} + q_1^1(-\xi,\eta)|\xi-\eta|^{2N_0} \\ &\quad + 2q_1^3(\xi-\eta,\eta)|\xi|^{2N_0}, \\ \widetilde{q}^2_{N_0}(\xi-\eta,\eta) &:= q_2(\eta,\xi-\eta)|\xi|^{2N_0} + q_1^2(\eta-\xi,\xi)|\eta|^{2N_0}, \\ \widetilde{q}^3_{N_0}(\xi-\eta,\eta) &:= q_3(\xi-\eta,\eta)|\xi|^{2N_0} + q_3(-\xi,\eta)|\xi-\eta|^{2N_0}.\end{aligned} \tag{3.13}$$

Recall (2.23), (2.21), and (2.22). From (2.25) in Lemma 2.5, we can see that cancellations happen in $\widetilde{q}^i_{N_0}(\cdot,\cdot)$, $i \in \{1,2,3\}$. From (2.25) in Lemma 2.5 and (2.6) in Lemma 2.3, the following estimate holds:

$$\sum_{i=1,2,3} \|\widetilde{q}^i_{N_0}(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{3\min\{k_1,k_2\}/2+2N_0k}. \tag{3.14}$$

We define bilinear operators $\widetilde{Q}_1(U^1,U^2)$, $\widetilde{Q}_2(U^1,U^1)$, and $\widetilde{Q}_3(U^2,U^2)$ by the symbols as follows:

$$\widetilde{q}_i(\xi-\eta,\eta) = \frac{\widetilde{q}^i_{N_0}(\xi-\eta,\eta)}{|\xi|^{2N_0}}, \quad i \in \{1,2,3\}. \tag{3.15}$$

From (3.14) and Lemma 2.4, the following estimate holds:

$$\sum_{i=1,2,3} \|\widetilde{q}_i(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{3\min\{k_1,k_2\}/2}. \tag{3.16}$$

Solving a similar system of equations as in (3.3) with $Q_i(\cdot,\cdot)$ replaced by $\widetilde{Q}_i(\cdot,\cdot)$, we can find bilinear operators $\widetilde{A}_1(U^1,U^1)$, $\widetilde{A}_2(U^2,U^2)$, and $\widetilde{B}(U^1,U^2)$ such that

$$2\widetilde{A}_1(|\nabla|^{\frac{1}{2}}U^2,U^1) - 2\widetilde{A}_2(|\nabla|^{\frac{1}{2}}U^1,U^2) + \widetilde{Q}_1(U^1,U^2) - |\nabla|^{\frac{1}{2}}\widetilde{B}(U^1,U^2) = 0, \tag{3.17}$$

$$\widetilde{Q}_2(U^1,U^1) + |\nabla|^{1/2}\widetilde{A}_1(U^1,U^1) - \widetilde{B}(U^1,|\nabla|^{\frac{1}{2}}U^1) = 0, \tag{3.18}$$

$$\widetilde{Q}_3(U^2,U^2) + |\nabla|^{1/2}\widetilde{A}_2(U^2,U^2) + \widetilde{B}(|\nabla|^{\frac{1}{2}}U^2,U^2) = 0. \tag{3.19}$$

Very similarly to the proof of Lemma 3.1, from (3.16) we have the estimate

$$\|\widetilde{a}_1(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|\widetilde{a}_2(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|\widetilde{b}(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{\min\{k_1,k_2\}}. \tag{3.20}$$

At this point our strategy is to use the normal form transformations $\widetilde{A}_1(U^1,U^1)$, $\widetilde{A}_2(U^2, U^2)$, and $\widetilde{B}(U^1,U^2)$ to cancel out the quadratic terms $\widetilde{Q}_i(\cdot,\cdot)$, $i \in \{1,2,3\}$. Therefore, we effectively cancel out the bulk cubic terms listed in (3.11) when doing the energy estimate. An advantage of utilizing symmetries first is that those normal form transformations do not lose derivatives; see (3.20).

Define $W^1 = U^1 + \widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2)$ and $W^2 = U^2 + \widetilde{B}(U^1,U^2)$. From the system of equations (2.18) satisfied by U^1 and U^2, we can first derive the system of equations satisfied by W^1 and W^2. Then, we substitute (U^1,U^2) for (W^1,W^2) inside the quadratic terms, C_1, and C_2 to see symmetries inside the nonlinearities. Recall (2.17), (2.18), and (2.2). As a result, we can write the system of equations satisfied by W^1 and W^2 as follows:

$$\begin{cases}\partial_t W^1 - |\nabla|^{\frac12} W^2 = \mathfrak{Q}_1 + GC_1 + \widetilde{C}_1 + \widetilde{\mathcal{R}}_1,\\ \partial_t W^2 + |\nabla|^{\frac12} W^1 = \mathfrak{Q}_2 + GC_2 + \widetilde{C}_2 + \widetilde{\mathcal{R}}_2,\end{cases} \tag{3.21}$$

where

$$\mathfrak{Q}_1 = Q_1(W^1,W^2) - \widetilde{Q}_1(W^1,W^2),$$
$$\mathfrak{Q}_2 = Q_2(W^1,W^1) + Q_3(W^2,W^2) - \widetilde{Q}_2(W^1,W^1) - \widetilde{Q}_3(W^2,W^2),$$

$$\begin{aligned}\widetilde{C}_1 &= -T_{\Lambda_{\geq 2}[V]}\partial_x W^1 + T_{\Lambda_{\geq 2}[\alpha]}|\nabla|^{1/2}W^2,\\ \widetilde{C}_2 &= -T_{\Lambda_{\geq 2}[V]}\partial_x W^2 - T_{\Lambda_{\geq 2}[\alpha]}|\nabla|^{1/2}W^1,\end{aligned} \tag{3.22}$$

$$\begin{aligned}GC_1 = {}& T_V\partial_x(\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2))\\ &- 2\widetilde{A}_1(T_V\partial_x U^1,U^1) - 2\widetilde{A}_2(T_V\partial_x U^2,U^2)\\ &- T_\alpha|\nabla|^{1/2}\widetilde{B}(U^1,U^2) + 2\widetilde{A}_1(T_\alpha|\nabla|^{\frac12}U^2,U^1)\\ &- 2\widetilde{A}_2(T_\alpha|\nabla|^{\frac12}U^1,U^2),\end{aligned} \tag{3.23}$$

$$\begin{aligned}GC_2 = {}& T_V\partial_x(\widetilde{B}(U^1,U^2)) - \widetilde{B}(T_V\partial_x U^1,U^2)\\ &- \widetilde{B}(U^1,T_V\partial_x U^2) + T_\alpha|\nabla|^{1/2}(\widetilde{A}_1(U^1,U^1)\\ &+ \widetilde{A}_2(U^2,U^2)) + \widetilde{B}(T_\alpha|\nabla|^{\frac12}U^2,U^2) - \widetilde{B}(U^1,T_\alpha|\nabla|^{\frac12}U^1).\end{aligned} \tag{3.24}$$

To improve the presentation, we postpone the detailed formulas of $\widetilde{\mathcal{R}}_1$ and $\widetilde{\mathcal{R}}_2$ to Appendix A, as their formulas are tedious and not very easy to see structures inside without detailed explanations. It is good enough to know that they have sufficient decay rates and do not lose derivatives at high frequencies or at low frequencies.

From the construction of bilinear operators $\widetilde{Q}_i(\cdot,\cdot), i \in \{1,2,3\}$, we can see that the quadratic terms $\mathfrak{Q}_1$ and $\mathfrak{Q}_2$ vanish when doing the energy estimate. We will show that there are cancellations inside the good cubic terms GC_1 and GC_2 and there are symmetries inside cubic terms $\widetilde{C}_1$ and $\widetilde{C}_2$. Therefore, W_1 and W_2 are good substitution variables of U^1 and U^2 at the top regularity. Moreover, as losing derivatives is not an issue when we estimate the growth of the $\dot{H}^p$ norm, we can cancel the quadratic terms directly by using the normal form transformation. Those intuitions motivate us to define the modified energy as follows,

$$\begin{aligned} E_{\mathrm{mod}}(t) = \frac{1}{2}\int & \big[|\partial_x^p U^1|^2 + |\partial_x^p U^2|^2\big] \\ & + \int \partial_x^p U^1 \partial_x^p\big[A_1(U^1,U^1) + A_2(U^2,U^2)\big] \\ & + \partial_x^p U^2 \partial_x^p\big[B(U^1,U^2)\big] + \frac{1}{2}\int \big[|\partial_x^{N_0} W^1|^2 + |\partial_x^{N_0} W^2|^2\big]. \end{aligned} \tag{3.25}$$

LEMMA 3.2. *Under the bootstrap condition* (2.29)*, the following estimate holds*:

$$\sup_{t\in[0,T]}\Big|E_{\mathrm{mod}}(t) - \sum_{k=p,N_0}\frac{1}{2}\int\big[\big|\partial_x^k U^1\big|^2 + \big|\partial_x^k U^2\big|^2\big]\Big| \lesssim \epsilon_0^2. \tag{3.26}$$

PROOF. From (3.7) in Lemma 3.1 and the L^2-L^∞-type estimate (2.7) in Lemma 2.3, the following estimate holds after putting the input with the larger frequency in L^2 and the input with the smaller frequency in L^∞:

$$\begin{aligned} \Big|\int \partial_x^p U^1 \partial_x^p\big[A_1(U^1,U^1) + A_2(U^2,U^2)\big] + \partial_x^p U^2 \partial_x^p\big[B(U^1,U^2)\big]\Big| \lesssim \\ \|(U^1,U^2)(t)\|_{H^{1+p,p}}^2 \|(U^1,U^2)(t)\|_{W^0} \lesssim (1+t)^{-1/2+2p_0}\epsilon_1^3. \end{aligned}$$

From (3.20) and the L^2-L^∞ estimate (2.7) in Lemma 2.3, we have

$$\begin{aligned} \big\|\partial_x^{N_0}(W^1 - U^1, W^2 - U^2)(t)\big\|_{L^2} &\lesssim \big\|(U^1,U^2)(t)\big\|_{H^{N_0,p}} \|(U^1,U^2)(t)\|_{W^1} \\ &\lesssim (1+t)^{-1/2+2p_0}\epsilon_1^2. \end{aligned}$$

Therefore, from the definition of $E_{\mathrm{mod}}(t)$ in (3.25) and the above estimates, it is easy to see that our desired estimate (3.26) holds. □

After taking a derivative with respect to time for $E_{\mathrm{mod}}(t)$, we have

$$\frac{d}{dt}E_{\mathrm{mod}}(t) = \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3,$$

where

$$\mathcal{J}_1 = \int \partial_x^p U^1 \partial_x^p\big[C_1 + \Lambda_{\geq 3}[\mathcal{R}_1]\big] + \partial_x^p U^2 \partial_x^p\big[C_2 + \Lambda_{\geq 3}[\mathcal{R}_2]\big], \tag{3.27}$$

$$\begin{aligned}\mathcal{J}_2 = \int & \partial_x^p \Lambda_{\geq 2}[\partial_t U^1]\partial_x^p\big(A_1(U^1,U^1) + A_2(U^2,U^2)\big)\\ & + \partial_x^p \Lambda_{\geq 2}[\partial_t U^2]\partial_x^p\big(B(U^1,U^2)\big)\\ & + 2\partial_x^p U^1 \partial_x^p\big(A_1(\Lambda_{\geq 2}[\partial_t U^1],U^1) + A_2(\Lambda_{\geq 2}[\partial_t U^2],U^2)\big)\\ & + \partial_x^p U^2 \partial_x^p\big(B(\Lambda_{\geq 2}[\partial_t U^1],U^2) + B(U^1,\Lambda_{\geq 2}[\partial_t U^2])\big),\\ \mathcal{J}_3 = \int & \partial_x^{N_0} W^1 \partial_x^{N_0}[\widetilde{C}_1 + GC_1 + \widetilde{\mathcal{R}}_1] + \partial_x^{N_0} W^2 \partial_x^{N_0}[\widetilde{C}_2 + GC_2 + \widetilde{\mathcal{R}}_2].\end{aligned}$$

LEMMA 3.3. *Under the bootstrap assumption* (2.29)*, we have*

$$\sup_{t\in[0,T]} \sum_{i=1,2,3} (1+t)^{1-2p_0}|\mathcal{J}_i| \lesssim \epsilon_0^2. \tag{3.28}$$

PROOF.

(i) To estimate $\mathcal{J}_1$ and $\mathcal{J}_2$, as $p \in (0,\frac{1}{4})$ is far away from N_0, we can estimate it straightforwardly by putting the input with the higher frequency in L^2 and putting the input with the lower frequency in L^∞.

From (3.7) in Lemma 3.1, the L^2-L^∞-type estimate (2.7) in Lemma 2.3, and (A.6) in Lemma A.1, the following estimate holds:

$$\begin{aligned}|\mathcal{J}_1| + |\mathcal{J}_2| \lesssim & \|\Lambda_{\geq 2}(\partial_t U^1, \partial_t U^2)\|_{H^{1+p,p}} \|(U^1,U^2)\|_{H^{1+p,p}} \|(U^1,U^2)\|_{W^0}\\ & + \|(U^1,U^2)\|_{H^{1+p,p}}^2 \|(U^1,U^2)\|_{W^2}^2\\ & + \|(U^1,U^2)\|_{H^{N_0,p}} \|\Lambda_{\geq 3}[\mathcal{R}_1,\mathcal{R}_2]\|_{H^{N_0,p}}\\ \lesssim & (1+|t|)^{-1+2p_0}\epsilon_0^2.\end{aligned}$$

(ii) Recall (3.22). We can utilize the symmetries to estimate $\widetilde{C}_1$ and $\widetilde{C}_2$ as follows:

$$\begin{aligned}&\left|\int \partial_x^{N_0} W^1 \partial_x^{N_0} \widetilde{C}_1 + \partial_x^{N_0} W^2 \partial_x^{N_0} \widetilde{C}_2\right|\\ &\quad\lesssim \sum_{i=1,2} \left|\partial_x^{N_0} W^i \partial_x^{N_0} T_{\Lambda_{\geq 2}[V]} \partial_x W^i\right|\\ &\qquad + \left|\partial_x^{N_0} W^1 \partial_x^{N_0} T_{\Lambda_{\geq 2}[\alpha]} |\nabla|^{1/2} W^2 - \partial_x^{N_0} W^2 \partial_x^{N_0} T_{\Lambda_{\geq 2}[\alpha]} |\nabla|^{1/2} W^1\right|\\ &\quad\lesssim \|(W^1,W^2)\|_{\dot{H}^{N_0}}^2 \big[\|\Lambda_{\geq 2}[V]\|_{\widetilde{W^1}} + \|\Lambda_{\geq 2}[\alpha]\|_{\widetilde{W^{1/2}}}\big]\\ &\quad\lesssim \|(U^1,U^2)\|_{H^{N_0,p}}^2 \|(U^1,U^2)\|_{W^3}^2\\ &\quad\lesssim (1+|t|)^{-1+2p_0}\epsilon_1^4 \lesssim (1+|t|)^{-1+2p_0}\epsilon_0^2.\end{aligned}$$

From (3.17) and (3.23), we can reduce GC_1 further:

$$\begin{aligned}GC_1 &= T_V\partial_x(\widetilde{A}_1(U^1,U^1)+\widetilde{A}_2(U^2,U^2)) - 2\widetilde{A}_1(T_V\partial_x U^1,U^1)\\ &\quad - 2\widetilde{A}_2(T_V\partial_x U^2,U^2) + T_\alpha[2\widetilde{A}_1(|\nabla|^{1/2}U^2,U^1) - 2\widetilde{A}_2(|\nabla|^{1/2}U^1,U^2)]\\ &\quad - 2\widetilde{A}_1(T_\alpha|\nabla|^{1/2}U^2,U^1) + 2\widetilde{A}_2(T_\alpha|\nabla|^{1/2}U^1,U^2) + T_\alpha\widetilde{Q}_1(U^1,U^2).\end{aligned}$$

Recall that $\widetilde{Q}_1(\cdot,\cdot)$ and $\widetilde{A}_i(\cdot,\cdot)$, $i\in\{1,2\}$, do not lose derivatives. See (3.16) and (3.20). Now, it is easy to see there are cancellations inside GC_1. After utilizing the symmetries on the Fourier side, the following equality holds:

$$\begin{aligned}\mathcal{F}(GC_1)(\xi) = \int_{\mathbb{R}^2} &\widehat{U^1}(\eta-\sigma)\widehat{U^1}(\sigma)\hat{V}(\xi-\eta)e_1(\xi,\eta,\sigma)\\ &+ \widehat{U^2}(\eta-\sigma)\widehat{U^2}(\sigma)\hat{V}(\xi-\eta)e_2(\xi,\eta,\sigma)\\ &+ \widehat{U^1}(\eta-\sigma)\widehat{U^2}(\sigma)\hat{\alpha}(\xi-\eta)e_3(\xi,\eta,\sigma)\,d\eta\,d\sigma,\end{aligned}$$

where, for $j\in\{1,2\}$,

$$\begin{aligned}e_j(\xi,\eta,\sigma) &= 2i(\eta-\sigma)\big[\theta(\xi-\eta,\eta)\tilde{a}_j(\eta-\sigma,\sigma) - \tilde{a}_j(\xi-\sigma,\sigma)\theta(\xi-\eta,\eta-\sigma)\big],\\ e_3(\xi,\eta,\sigma) &= 2|\sigma|^{1/2}\big[\tilde{a}_1(\sigma,\eta-\sigma)\theta(\xi-\eta,\eta) - \tilde{a}_1(\xi-\eta+\sigma,\eta-\sigma)\theta(\xi-\eta,\sigma)\big]\\ &\quad - 2|\eta-\sigma|^{1/2}\big[\tilde{a}_2(\eta-\sigma,\sigma)\theta(\xi-\eta,\eta) - \tilde{a}_2(\xi-\sigma,\sigma)\theta(\xi-\eta,\eta-\sigma)\big].\end{aligned}$$

From the above explicit formulas and Lemma 2.4, the following estimates hold:

$$\begin{aligned}\|e_1(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} + \|e_2(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} &\lesssim 2^{2\,\mathrm{med}\{k_1,k_2,k_3\}},\\ \|e_3(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} &\lesssim 2^{3\,\mathrm{med}\{k_1,k_2,k_3\}/2},\end{aligned}$$

where $\mathrm{med}\{k_1,k_2,k_3\}$ denotes the median of k_1,k_2,k_3. Hence, from a L^2-L^∞-L^∞-type estimate in Lemma 2.3, the following estimate holds:

$$\begin{aligned}\|GC_1\|_{\dot{H}^{N_0}} &\lesssim \|(U^1,U^2)\|_{H^{N_0,p}}\|\partial_x(U^1,U^2)\|_{\widetilde{W^1}}\big[\|\alpha\|_{\widetilde{W^0}} + \|V\|_{\widetilde{W^0}}\big]\\ &\lesssim \|(U^1,U^2)\|_{H^{N_0,p}}\|(U^1,U^2)\|^2_{W^3}\\ &\lesssim (1+|t|)^{-1+p_0}\epsilon_1^3 \lesssim (1+|t|)^{-1+p_0}\epsilon_0^2.\end{aligned}\tag{3.29}$$

The estimate of GC_2 is very similar; the upper bound in the right-hand side of (3.29) is still good for GC_2. We omit the details here.

From (A.6) in Lemma A.1, it is easy to see that the following estimate holds:

$$\begin{aligned}\left|\int \partial_x^{N_0}W^1\partial_x^{N_0}\widetilde{\mathcal{R}}_1 + \partial_x^{N_0}W^2\partial_x^{N_0}\widetilde{\mathcal{R}}_2\right| &\lesssim \|(W^1,W^2)\|_{H^{N_0,p}}\big\|\big(\widetilde{\mathcal{R}}_1,\widetilde{\mathcal{R}}_2\big)\big\|_{H^{N_0,p}}\\ &\lesssim (1+t)^{-1+2p_0}\epsilon_0^2.\end{aligned}$$

Recall (3.27). Now, it is easy to see that our desired estimate (3.3) holds. □

3.3 Energy Estimate of SU^1 and SU^2

Note that $[\partial_t, S] = \partial_t$ and $[\pm|\nabla|^{1/2}, S] = \pm|\nabla|^{1/2}$. On the one hand, from the system (2.17), we can derive the system of equations satisfied by SU^1 and SU^2 with the highlighted quasilinear structure as follows:

$$\begin{cases}\partial_t SU^1 - |\nabla|^{1/2} SU^2 = T_\alpha |\nabla|^{1/2} SU^2 - T_V \partial_x SU^1 + \mathcal{R}_1^S, \\ \partial_t SU^2 + |\nabla|^{1/2} SU^1 = -T_\alpha |\nabla|^{1/2} SU^1 - T_V \partial_x SU^2 + \mathcal{R}_2^S,\end{cases} \tag{3.30}$$

where $\mathcal{R}_1^S$ and $\mathcal{R}_2^S$ are quadratic and higher-order good remainder terms. Very importantly, $\mathcal{R}_1^S$ and $\mathcal{R}_2^S$ do not lose derivatives in SU^1 and SU^2. More precisely, we have

$$\mathcal{R}_1^S := (S+I)\mathcal{R}_1 + S(T_\alpha|\nabla|^{1/2}U^2 - T_V\partial_x U^1) - T_\alpha|\nabla|^{1/2}SU^2 + T_V\partial_x SU^1, \tag{3.31}$$

$$\mathcal{R}_2^S := (S+I)\mathcal{R}_2 + S(-T_\alpha|\nabla|^{1/2}U^1 - T_V\partial_x U^2) + T_\alpha|\nabla|^{1/2}SU^1 + T_V\partial_x SU^2. \tag{3.32}$$

On the other hand, from the system (2.18), we can rewrite the system (3.30) as follows to highlight the structures inside the quadratic terms:

$$\begin{cases}\partial_t SU^1 - |\nabla|^{1/2} SU^2 = \mathfrak{Q}_1^S(SU^1, SU^2) + \mathfrak{Q}_2^S + \mathcal{Q}_1 + \widetilde{C}_1 + \mathfrak{R}_1, \\ \partial_t SU^2 + |\nabla|^{1/2} SU^1 = \mathfrak{Q}_3^S(SU^1, SU^2) + \mathfrak{Q}_4^S + \mathcal{Q}_2 + \widetilde{C}_2 + \mathfrak{R}_2,\end{cases} \tag{3.33}$$

where $\mathfrak{Q}_1^S(\cdot,\cdot)$ and $\mathfrak{Q}_3^S(\cdot,\cdot)$ are quadratic terms that have quasilinear structures inside, $\mathfrak{Q}_2^S$ and $\mathfrak{Q}_4^S$ are quadratic terms that have a problematic low-frequency part, $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are commutator quadratic terms that do not depend on the scaling vector field, $\widetilde{C}_1$ and $\widetilde{C}_2$ are cubic and higher-order terms that lose at most one derivative, and $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are good cubic and higher-order terms that do not lose derivatives. Their detailed formulas are given as follows:

$$\begin{aligned}\mathfrak{Q}_1^S(SU^1, SU^2) &= Q_1(SU^1, U^2) + Q_1(U^1, SU^2) \\ &\quad - T_{\partial_x|\nabla|^{-1/2}SU^2}\partial_x U^1, \\ \mathfrak{Q}_3^S(SU^1, SU^2) &= Q_2(SU^1, U^1) + Q_2(U^1, SU^1) + Q_3(SU^2, U^2) \\ &\quad + Q_3(U^2, SU^2) + T_{\partial_x|\nabla|^{-1/2}SU^2}\partial_x U^2,\end{aligned}$$

$$\mathfrak{Q}_2^S = T_{\partial_x|\nabla|^{-1/2}SU^2}\partial_x U^1, \quad \mathfrak{Q}_4^S = -T_{\partial_x|\nabla|^{-1/2}SU^2}\partial_x U^2, \tag{3.34}$$

$$\mathcal{Q}_1 = Q_1(U^1, U^2) + SQ_1(U^1, U^2) - Q_1(SU^1, U^2) - Q_1(U^1, SU^2), \tag{3.35}$$

$$\begin{aligned}\mathcal{Q}_2 = Q_2(U^1, U^1) &+ Q_3(U^2, U^2) + SQ_2(U^1, U^1) \\ &- Q_2(SU^1, U^1) - Q_2(U^1, SU^2) + SQ_2(U^1, U^1) \\ &- Q_2(SU^1, U^1) - Q_2(U^1, SU^2),\end{aligned} \tag{3.36}$$

$$\begin{aligned}\widetilde{C}_1 &= T_{\Lambda_{\geq 2}[\alpha]}|\nabla|^{1/2}SU^2 - T_{\Lambda_{\geq 2}[V]}\partial_x SU^1, \\ \widetilde{C}_2 &= T_{\Lambda_{\geq 2}[\alpha]}|\nabla|^{1/2}SU^2 - T_{\Lambda_{\geq 2}[V]}\partial_x SU^1,\end{aligned}$$

$$\mathfrak{R}_1 = \Lambda_{\geq 3}\big[\mathcal{R}_1^S\big], \quad \mathfrak{R}_2 = \Lambda_{\geq 3}\big[\mathcal{R}_2^S\big].$$

To better see which parts of $\mathfrak{Q}_1^S(\cdot,\cdot)$ and $\mathfrak{Q}_1^S(\cdot,\cdot)$ lose derivatives in SU^1 and SU^2, which are very important in the later energy estimate part, we define the following auxiliary bilinear forms:

$$\widetilde{\mathfrak{Q}}_1^S(SU^1, SU^2) := \mathfrak{Q}_1^S(SU^1, SU^2) - T_{\Lambda_1[\alpha]}|\nabla|^{1/2}SU^2 + T_{\Lambda_1[V]}\partial_x SU^1, \tag{3.37}$$

$$\widetilde{\mathfrak{Q}}_3^S(SU^1, SU^2) := \mathfrak{Q}_3^S(SU^1, SU^2) + T_{\Lambda_1[\alpha]}|\nabla|^{1/2}SU^1 + T_{\Lambda_1[V]}\partial_x SU^2. \tag{3.38}$$

From (3.30), we can see that $\widetilde{\mathfrak{Q}}_1^S(SU^1, SU^2)$ and $\widetilde{\mathfrak{Q}}_3^S(SU^1, SU^2)$ do not lose derivatives in SU^1 and SU^2.

Handling the Bulk Quadratic Terms $\mathfrak{Q}_1^S$ and $\mathfrak{Q}_3^S$

Similar to what we did in the Section 3.2, we will also utilize the symmetries at high frequencies. For $\tau = p, N_1$, we can utilize the symmetries of the same type of inputs and change the variables on the Fourier side. As a result, the following equality holds:

$$\begin{aligned}
&\mathfrak{Re}\bigg[\int \overline{\partial_x^\tau SU^1}\partial_x^\tau \mathfrak{Q}_1^S + \overline{\partial_x^\tau SU^2}\partial_x^\tau \mathfrak{Q}_3^S\bigg] \\
&= \mathfrak{Re}\bigg[\int \overline{\widehat{SU^1}(\xi)}\hat{q}_\tau^1(\xi-\eta,\eta)\widehat{SU^1}(\xi-\eta)\widehat{U^2}(\eta) \\
&\qquad + \overline{\widehat{SU^1}(\xi)}\hat{q}_\tau^2(\xi-\eta,\eta)\widehat{SU^2}(\xi-\eta)\widehat{U^1}(\eta) \\
&\qquad + \overline{\widehat{SU^2}(\xi)}\hat{q}_\tau^3(\xi-\eta,\eta)\widehat{SU^2}(\xi-\eta)\widehat{U^2}(\eta)\bigg],
\end{aligned} \tag{3.39}$$

where

$$\begin{aligned}
\hat{q}_\tau^1(\xi-\eta,\eta) &= |\xi|^{2\tau}\big(q_1^1(\xi-\eta,\eta)/2 + q_1^2(\xi-\eta,\eta) + q_1^3(\xi-\eta,\eta)\big) \\
&\quad + q_1^1(-\xi,\eta)|\xi-\eta|^{2\tau}/2,
\end{aligned}$$

$$\begin{aligned}
\hat{q}_\tau^2(\xi-\eta,\eta) &= |\xi|^{2\tau}\big(q_1(\eta,\xi-\eta) + \eta(\xi-\eta)|\eta|^{-1/2}\theta(\xi-\eta,\eta)\big) \\
&\quad + |\xi-\eta|^{2\tau}\big(q_2(\eta,-\xi) + q_2(-\xi,\eta)\big),
\end{aligned}$$

$$\begin{aligned}
\hat{q}_\tau^3(\xi-\eta,\eta) &= |\xi|^{2\tau}\big(q_3(\xi-\eta,\eta) + q_3(\eta,\xi-\eta)\big)/2 \\
&\quad + |\xi-\eta|^{2\tau}\big(q_3(\eta,-\xi) + q_3(-\xi,\eta)\big)/2 \\
&\quad - (\xi-\eta)\eta|\xi-\eta|^{-1/2}|\eta|^{-1/2}|\xi|^{1/2+2\tau}\theta(\xi-\eta,\eta).
\end{aligned}$$

Note that $q_1(\xi-\eta,\eta)\theta(\eta,\xi-\eta) = q_1^1(\xi-\eta,\eta)$ and $q_1(\xi-\eta,\eta)\theta(\xi-\eta,\eta) = q_1^2(\xi-\eta,\eta)$. From (2.25) in Lemma 2.5, we can see that cancellations also happen when $|\eta| \ll |\xi-\eta|$. Moreover, when $|\xi-\eta| \ll |\eta|$, *the symbols contribute at least the smallness of* $|\xi-\eta|$, because we have put those exception terms into $\mathfrak{Q}_2^S$

and $\mathfrak{Q}_4^S$. For the High × High type interaction, all terms except $q_1^3(\cdot\,,\cdot)$ vanish; see (2.23), (2.21), and (2.22). As a result, the following estimate holds:

$$\sum_{i=1,2,3} \|\widehat{q}_\tau^3(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim \begin{cases} 2^{3k_2/2+2\tau k} & \text{if } k_2 \le k_1-5, \\ 2^{(2\tau+1)k+k_1/2} & \text{if } |k_1-k_2| \le 5, \\ 2^{k_1+(2\tau+1/2)k} & \text{if } k_1 \le k_2-5. \end{cases} \tag{3.40}$$

For $\tau = p$ and N_1, we define $\widetilde{Q}_{1,\tau}(SU^1, U^2)$, $\widetilde{Q}_{2,\tau}(SU^2, U^1)$, $\widetilde{Q}_{3,\tau}(SU^1, U^1)$, and $\widetilde{Q}_{4,\tau}(SU^2, U^2)$ by the symbols as follows:

$$\widetilde{q}_{1,\tau}(\xi-\eta,\eta) = \frac{\widehat{q}_\tau^1(\xi-\eta,\eta)}{|\xi|^{2\tau}}, \quad \widetilde{q}_{2,\tau}(\xi-\eta,\eta) = \frac{\widehat{q}_\tau^2(\xi-\eta,\eta)}{2|\xi|^{2\tau}},$$
$$\widetilde{q}_{3,\tau}(\xi-\eta,\eta) = \frac{\widehat{q}_\tau^2(\xi,-\eta)}{2|\xi|^{2\tau}}, \qquad \widetilde{q}_{4,\tau}(\xi-\eta,\eta) = \frac{\widehat{q}_\tau^3(\xi-\eta,\eta)}{|\xi|^{2\tau}}.$$

We define the substitution variables for SU^1 and SU^2 as follows:

$$W_{1,\tau} := SU^1 + \widetilde{C}_{1,\tau}(SU^1, U^1) + \widetilde{C}_{2,\tau}(SU^2, U^2), \tag{3.41}$$

$$W_{2,\tau} := SU^2 + \widetilde{D}_{1,\tau}(SU^1, U^2) + \widetilde{D}_{2,\tau}(SU^2, U^1). \tag{3.42}$$

Now the strategy is to use bilinear operators $\widetilde{C}_{i,\tau}(\cdot\,,\cdot)$ and $\widetilde{D}_{i,\tau}(\cdot\,,\cdot)$, $i \in \{1,2\}$, to cancel out $\widetilde{Q}_{j,\tau}(\cdot,\cdot)$, $j \in \{1,2,3,4\}$. Effectively speaking, they cancel out (3.39) in the energy estimate. To this end, it is sufficient if symbols $c_\tau^i(\cdot\,,\cdot)$ and $d_\tau^i(\cdot\,,\cdot)$ of bilinear operators $\widetilde{C}_{i,\tau}(\cdot\,,\cdot)$ and $\widetilde{D}_{i,\tau}(\cdot\,,\cdot)$ can solve the system of equations as follows:

$$\begin{cases} |\eta|^{1/2}c_\tau^1(\xi-\eta,\eta) - |\xi-\eta|^{1/2}c_\tau^2(\xi-\eta,\eta) - |\xi|^{1/2}d_\tau^1(\xi-\eta,\eta) \\ \quad + \widetilde{q}_{1,\tau}(\xi-\eta,\eta) = 0, \\ |\xi-\eta|^{1/2}c_\tau^1(\xi-\eta,\eta) - |\eta|^{1/2}c_\tau^2(\xi-\eta,\eta) - |\xi|^{1/2}d_\tau^2(\xi-\eta,\eta) \\ \quad + \widetilde{q}_{2,\tau}(\xi-\eta,\eta) = 0, \\ |\xi|^{1/2}c_\tau^1(\xi-\eta,\eta) - |\eta|^{1/2}d_\tau^1(\xi-\eta,\eta) - |\xi-\eta|^{1/2}d_\tau^2(\xi-\eta,\eta) \\ \quad + \widetilde{q}_{3,\tau}(\xi-\eta,\eta) = 0, \\ |\xi|^{1/2}c_\tau^2(\xi-\eta,\eta) + |\xi-\eta|^{1/2}d_\tau^1(\xi-\eta,\eta) + |\eta|^{1/2}d_\tau^2(\xi-\eta,\eta) \\ \quad + \widetilde{q}_{4,\tau}(\xi-\eta,\eta) = 0. \end{cases} \tag{3.43}$$

It is not difficult to solve the above system of equations and derive the solution as follows:

$$d_\tau^1(\xi-\eta,\eta) = \frac{(|\xi-\eta|+|\eta|-|\xi|)F_1(\xi-\eta,\eta) - 2F_2(\xi-\eta,\eta)|\xi-\eta|^{1/2}|\eta|^{1/2}}{-(|\xi-\eta|+|\eta|-|\xi|)^2 + 4|\xi-\eta||\eta|}, \tag{3.44}$$

$$d_\tau^2(\xi-\eta,\eta) = \frac{(|\xi-\eta|+|\eta|-|\xi|)F_2(\xi-\eta,\eta) - 2F_1(\xi-\eta,\eta)|\xi-\eta|^{1/2}|\eta|^{1/2}}{-(|\xi-\eta|+|\eta|-|\xi|)^2 + 4|\xi-\eta||\eta|}, \tag{3.45}$$

$$c_\tau^1(\xi-\eta,\eta) = \frac{|\eta|^{1/2}d_\tau^1(\xi-\eta,\eta) + |\xi-\eta|^{1/2}d_\tau^2(\xi-\eta,\eta) - \widetilde{q}_{3,\tau}(\xi-\eta,\eta)}{|\xi|^{1/2}}, \tag{3.46}$$

$$c_\tau^2(\xi-\eta,\eta) = \frac{-|\xi-\eta|^{1/2}d_\tau^1(\xi-\eta,\eta) - |\eta|^{1/2}d_\tau^2(\xi-\eta,\eta) - \widetilde{q}_{4,\tau}(\xi-\eta,\eta)}{|\xi|^{1/2}}, \tag{3.47}$$

where $F_1(\xi-\eta,\eta)$ and $F_2(\xi-\eta,\eta)$ are defined as follows:

$$F_1(\xi-\eta,\eta) = |\xi|^{1/2}\widetilde{q}_{1,\tau}(\xi-\eta,\eta) - |\eta|^{1/2}\widetilde{q}_{3,\tau}(\xi-\eta,\eta) + |\xi-\eta|^{1/2}\widetilde{q}_{4,\tau}(\xi-\eta,\eta),$$
$$F_2(\xi-\eta,\eta) = |\xi|^{1/2}\widetilde{q}_{2,\tau}(\xi-\eta,\eta) - |\xi-\eta|^{1/2}\widetilde{q}_{3,\tau}(\xi-\eta,\eta) + |\eta|^{1/2}\widetilde{q}_{4,\tau}(\xi-\eta,\eta).$$

LEMMA 3.4. *The following estimate holds for $i \in \{1,2\}$,*

$$\|c_\tau^i(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|d_\tau^i(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim \begin{cases} 2^{k_2} & \text{if } k_2 \le k_1+5, \\ 2^{k_1/2+k_2/2} & \text{if } k_1 \le k_2-5. \end{cases} \tag{3.48}$$

PROOF. Recall (3.44), (3.45), (3.46), and (3.44). From (3.40), Lemma 2.4, and (2.6) in Lemma 2.3, our desired estimate (3.48) follows straightforwardly. □

Recall (3.30) and (3.33). From (3.41) and (3.42), after substituting SU^i by $W_{i,\tau}$, $i \in \{1,2\}$, we can derive the equations satisfied by $W_{1,\tau}$ and $W_{2,\tau}$ with good structures as follows:

$$\begin{aligned} &\partial_t W_{1,\tau} - |\nabla|^{1/2}W_{2,\tau} \\ &\quad = \mathfrak{Q}_1^S(W_{1,\tau},U^2) + \mathfrak{Q}_1^S(W_{2,\tau},U^1) - \widetilde{Q}_{1,\tau}(W_{1,\tau},U^2) \\ &\qquad - \widetilde{Q}_{2,\tau}(W_{2,\tau},U^1) - T_{\Lambda_{\ge 2}[V]}\partial_x W_{1,\tau} + T_{\Lambda_{\ge 2}[\alpha]}|\nabla|^{1/2}W_{2,\tau} \\ &\qquad + \mathfrak{Q}_2^S + \mathcal{Q}_1 + \widetilde{GC}_1 + \widetilde{\mathfrak{R}}_1, \end{aligned} \tag{3.49}$$

$$\begin{aligned} &\partial_t W_{2,\tau} + |\nabla|^{1/2}W_{1,\tau} \\ &\quad = \mathfrak{Q}_3^S(W_{1,\tau},U^1) + \mathfrak{Q}_3^S(W_{2,\tau},U^2) - \widetilde{Q}_{3,\tau}(W_{1,\tau},U^1) \\ &\qquad - \widetilde{Q}_{4,\tau}(W_{2,\tau},U^2) - T_{\Lambda_{\ge 2}[V]}\partial_x W_{2,\tau} - T_{\Lambda_{\ge 2}[\alpha]}|\nabla|^{1/2}W_{1,\tau} \\ &\qquad + \mathfrak{Q}_4^S + \mathcal{Q}_2 + \widetilde{GC}_2 + \widetilde{\mathfrak{R}}_2, \end{aligned} \tag{3.50}$$

where

$$\begin{aligned}\widetilde{GC}_1 &= T_V\partial_x\big(\widetilde{C}_{1,\tau}(SU^1,U^1)+\widetilde{C}_{2,\tau}(SU^2,U^2)\big)\\&\quad - T_\alpha|\nabla|^{1/2}[\widetilde{D}_{1,\tau}(SU^1,U^2)+\widetilde{D}_{2,\tau}(SU^2,U^1)]\\&\quad - \widetilde{C}_{1,\tau}(T_V\partial_x SU^1,U^1)+\widetilde{C}_{1,\tau}(T_\alpha|\nabla|^{1/2}SU^2,U^1)\\&\quad - \widetilde{C}_{2,\tau}(T_V\partial_x SU^2,U^2)-\widetilde{C}_{2,\tau}(T_\alpha|\nabla|^{1/2}SU^1,U^2),\end{aligned}$$

$$\begin{aligned}\widetilde{GC}_2 &= T_V\partial_x[\widetilde{D}_{1,\tau}(SU^1,U^2)+\widetilde{D}_{2,\tau}(SU^2,U^1)]\\&\quad + T_\alpha|\nabla|^{1/2}[\widetilde{C}_{1,\tau}(SU^1,U^1)+\widetilde{C}_{2,\tau}(SU^2,U^2)]\\&\quad - \widetilde{D}_{1,\tau}(T_V\partial_x SU^1,U^2)+\widetilde{D}_{1,\tau}(T_\alpha|\nabla|^{1/2}SU^2,U^2)\\&\quad - \widetilde{D}_{2,\tau}(T_V\partial_x SU^2,U^1)-\widetilde{D}_{2,\tau}(T_\alpha|\nabla|^{1/2}SU^1,U^1).\end{aligned}$$

To improve the presentation, we postpone the formulas of $\widetilde{\mathfrak{R}}_1$ and $\widetilde{\mathfrak{R}}_2$ to Appendix A. It is enough to see that they are good cubic and higher-order remainder terms in the sense that they do not lose derivatives at the high frequency or the low frequency of SU^1 and SU^2. Hence the $H^{N_1,p}$-norm of those remainder terms can be estimated straightforwardly in the energy estimate.

Similar to what we did in the energy estimate of U^1 and U^2, we define the high-frequency part of the modified energy for SU^1 and SU^2 as follows:

$$\begin{aligned}E^{S,\text{high}}_{\text{mod}}(t) := \sum_{\tau=p,N_1}&\int\frac{1}{2}\big[|\partial_x^\tau W_{1,\tau}|^2+|\partial_x^\tau W_{2,\tau}|^2\big]\\&+\int\partial_x^\tau W_{1,\tau}\partial_x^\tau[E_1(U^1,U^1)+E_2(U^2,U^2)]\\&\qquad+\partial_x^\tau W_{2,\tau}\partial_x^\tau F(U^1,U^2),\end{aligned}\tag{3.51}$$

where bilinear forms $E_1(U^1,U^1)$, $E_2(U^2,U^2)$, and $F(U^1,U^2)$ are the normal form transformations, which aim to cancel out the commutator terms $\mathcal{Q}_1$ and $\mathcal{Q}_2$ in (3.35) and (3.3).

By solving a similar system of equations as in (3.3) with $Q_1(U^1,U^2)$ replaced by $\mathcal{Q}_1$ and $Q_2(U^1,U^1)+Q_3(U^2,U^2)$ replaced by $\mathcal{Q}_2$, we can explicitly solve symbols $e_1(\xi-\eta,\eta)$, $e_2(\xi-\eta,\eta)$, and $f(\xi-\eta,\eta)$ of bilinear operators $E_1(\cdot,\cdot)$, $E_2(\cdot,\cdot)$, and $F(\cdot,\cdot)$. Their precise formulas are not so important. Hence, we omit them here. Similar to the proof of estimate (3.7) in Lemma 3.1, the following estimate holds:

$$\begin{aligned}\|e_1(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}}+\|e_2(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}}&\\+\|f(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}}\lesssim 2^{\max\{k_1,k_2\}}&.\end{aligned}\tag{3.52}$$

LEMMA 3.5. *Under the bootstrap smallness assumption* (2.29)*, we have*

$$\sup_{t\in[0,T]}\left|E^{S,\mathrm{high}}_{\mathrm{mod}}(t)-\sum_{k=p,N_0}\frac{1}{2}\int\left[\left|\partial_x^k U^1\right|^2+\left|\partial_x^k U^2\right|^2\right]\right|\lesssim\epsilon_0^2. \tag{3.53}$$

PROOF. From (3.48) in Lemma 3.4 and the L^2-L^∞-type bilinear estimate (2.7) in Lemma 2.3, the following estimate holds:

$$\begin{aligned}\sum_{\tau=p,N_1}\|W_{1,\tau}(t)-SU^1(t)\|_{\dot H^\tau}+\|W_{2,\tau}(t)-SU^2(t)\|_{\dot H^\tau}\lesssim\\ \|(SU^1,SU^2)(t)\|_{H^{N_1,p}}\|(U^1,U^2)(t)\|_{W^1}\lesssim(1+|t|)^{-1/2+p_0}\epsilon_1^2.\end{aligned}\tag{3.54}$$

From (3.52), (3.54), and the L^2-L^∞-type bilinear estimate (2.7) in Lemma 2.3, our desired estimate holds as follows:

$$\begin{aligned}&\left|E^{S,\mathrm{high}}_{\mathrm{mod}}(t)-\sum_{k=p,N_1}\int\left|\partial_x^k SU^1\right|^2+\left\|\partial_x^k SU^2\right|^2\right|\\ &\quad\lesssim(1+|t|)^{-1/2+p_0}\epsilon_1^3\\ &\qquad+\|(SU^1,SU^2)(t)\|_{H^{N_1,p}}\|(U^1,U^2)\|_{H^{N_1+1,p}}\|(U^1,U^2)(t)\|_{W^1}\lesssim\epsilon_0^2.\ \square\end{aligned}$$

From direct computations, we have the estimate

$$\left|\frac{d}{dt}E^{S,\mathrm{high}}_{\mathrm{mod}}(t)+\sum_{\tau=p,N_1}\int\partial_x^\tau SU^1\partial_x^\tau\mathfrak{Q}_2^S+\partial_x^\tau SU^2\partial_x^\tau\mathfrak{Q}_4^S\right|\lesssim\sum_{i=1,2,3}\left|\mathcal{J}_i^S\right|, \tag{3.55}$$

where

$$\begin{aligned}\mathcal{J}_1^S=\sum_{\tau=p,N_1}\int&\partial_x^k W_{1,\tau}\partial_x^\tau[\widetilde{GC}_1+\widetilde{\mathfrak{R}_1}]+\partial_x^\tau W_{2,\tau}\partial_x^\tau[\widetilde{GC}_2+\widetilde{\mathfrak{R}_2}]\\ &-\partial_x^\tau\big(\widetilde{C}_{1,\tau}(SU^1,U^1),+\widetilde{C}_{2,\tau}(SU^2,U^2)\big)\partial_x^\tau\mathfrak{Q}_2^S\\ &+\partial_x^\tau\big(\widetilde{D}_{1,\tau}(SU^1,U^2)+\widetilde{D}_{2,\tau}(SU^2,U^1)\big)\partial_x^\tau\mathfrak{Q}_4^S\end{aligned}$$

$$\begin{aligned}\mathcal{J}_2^S=\sum_{\tau=p,N_1}\int&\partial_x^\tau\Lambda_{\ge2}[\partial_t W_{1,\tau}]\partial_x^\tau[E_1(U^1,U^1)+E_2(U^2,U^2)]\\ &+\partial_x^\tau\Lambda_{\ge2}[\partial_t W_{2,\tau}]\partial_x^\tau F(U^1,U^2)\\ &+2\partial_x^\tau W_{1,\tau}\partial_x^\tau[E_1(\Lambda_{\ge2}[\partial_t U^1],U^1)+E_2(\Lambda_{\ge2}[\partial_t U^2],U^2)]\\ &+\partial_x^\tau W_{2,\tau}\partial_x^\tau[F(\Lambda_{\ge2}[\partial_t U^1],U^2)+F(U^1,\Lambda_{\ge2}[\partial_t U^2])],\end{aligned}$$

$$\begin{aligned}\mathcal{J}_3^S=\sum_{\tau=p,N_1}\int&\partial_x^\tau W_{1,\tau}\partial_x^\tau[-T_{\Lambda_{\ge2}[V]}\partial_x W_{1,\tau}+T_{\Lambda_{\ge2}[\alpha]}|\nabla|^{1/2}W_{2,\tau}]\\ &-\partial_x^\tau W_{2,\tau}\partial_x^\tau[T_{\Lambda_{\ge2}[V]}\partial_x W_{2,\tau}+T_{\Lambda_{\ge2}[\alpha]}|\nabla|^{1/2}W_{1,\tau}].\end{aligned}$$

We remark that the quartic terms that depend on $\mathfrak{D}_2^S$ and $\mathfrak{D}_4^S$ inside $\mathcal{J}_1^S$ result from replacing SU^1 and SU^2 by $W_{1,\tau}$ and $W_{2,\tau}$, respectively.

LEMMA 3.6. *Under the bootstrap condition* (2.29)*, we have the estimate*

$$\sum_{i=1,2,3} \left|\mathcal{J}_i^S\right| \lesssim (1+t)^{-1+2p_0}\epsilon_0^2. \tag{3.56}$$

PROOF. We first estimate $\mathcal{J}_1^S$ and $\mathcal{J}_3^S$. From (A.9) in Lemma A.2, we have the estimate of the $H^{N_1,p}$-norm of $\widetilde{\mathfrak{R}}_1$ and $\widetilde{\mathfrak{R}}_2$. Similar to the estimate of GC_i we did in the proof of Lemma 3.3, one can see that we have cancellation after writing $\widetilde{GC}_i$ on the Fourier side and utilizing the equalities satisfied by the normal form transformations in (3.43). Also, after utilizing the symmetries, we can see the cancellations inside $\mathcal{J}_3^S$. As a result, the following estimate holds:

$$\begin{aligned}\left|\mathcal{J}_1^S\right| + \left|\mathcal{J}_3^S\right| &\lesssim \left[\|(SU^1,SU^2)\|_{H^{N_1,p}} + \|(U^1,U^2)\|_{H^{N_0,p}}\right]^2 \|(U^1,U^2)\|_{W^3}^2 \\ &\quad + (1+t)^{-1+2p_0}\epsilon_0^2 \lesssim (1+t)^{-1+2p_0}\epsilon_0^2.\end{aligned}$$

Lastly, we proceed to estimate $\mathcal{J}_2$. Note that the inputs inside the bilinear operators $E_1(\cdot,\cdot)$, $E_2(\cdot,\cdot)$, and $F(\cdot,\cdot)$ only depend on (U^1,U^2). Hence, we can always put the input with the higher frequency in L^2 to avoid losing p derivatives of smallness. As a result, we have

$$\begin{aligned}|\mathcal{J}_2| &\lesssim \left[\|(SU^1,SU^2)\|_{H^{N_1,p}} + \|(U^1,U^2)\|_{H^{N_0,p}}\right]^2 \|(U^1,U^2)\|_{W^3}^2 \\ &\lesssim (1+t)^{-1+2p_0}\epsilon_0^2.\end{aligned}$$

To sum up, we can see that the desired estimate (3.56) holds. □

Handling the Bulk Quadratic Terms $\mathfrak{Q}_2^S$ and $\mathfrak{Q}_4^S$

It remains to deal with the quadratic terms $\mathfrak{Q}_2^S$ and $\mathfrak{Q}_4^S$. As losing one derivative is not a issue here, it is not necessary to utilize the symmetries to see cancellations. We choose to work in the complex variables setting, which is more convenient in the Fourier-transform-based method. We define

$$U = U^1 + iU^2, \quad SU = SU^1 + iSU^2, \quad c_\mu = \mu/(2i), \quad \mu \in \{+,-\}.$$

From (2.18), we can write the equations satisfied by U and SU as follows:

$$\begin{aligned}\partial_t U + i|\nabla|^{1/2}U &= \sum_{\mu,\nu\in\{+,-\}} Q_{\mu,\nu}(U_\mu, U_\nu) + \Lambda_{\geq 3}[\partial_t U], \\ \partial_t SU + i|\nabla|^{1/2}SU &= \sum_{\mu,\nu\in\{+,-\}} Q^1_{\mu,\nu}((SU)_\mu, U_\nu) + Q^2_{\mu,\nu}(U_\mu, U_\nu) \\ &\quad + \Lambda_{\geq 3}[\partial_t SU],\end{aligned}$$

where $f_+ := f$ and $f_- := \overline{f}$. From (2.24) in Lemma 2.5, it is easy to see that the following estimate holds:

$$\sum_{\mu,\nu\in\{+,-\}} \|q_{\mu,\nu}(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|q^1_{\mu,\nu}(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} + \|q^2_{\mu,\nu}(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{\min\{k_1,k_2\}/2+\max\{k_1,k_2\}}. \tag{3.57}$$

Recall (3.34). We can write those problematic cubic terms on the Fourier side and in terms of SU and U as follows:

$$\sum_{\tau=p,N_1} \mathfrak{Re}\Big[\int \overline{\partial_x^\tau SU^1}\partial_x^\tau \mathfrak{Q}_2^S + \overline{\partial_x^\tau SU^2}\partial_x^\tau \mathfrak{Q}_4^S\Big] = \tag{3.58}$$
$$\sum_{\tau=p,N_1}\sum_{\mu,\nu,\kappa\in\{+,-\}} \mathfrak{Re}\Big[\int \overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_\kappa}(\eta) p^{\kappa,\tau}_{\mu,\nu}(\xi-\eta,\eta)\,d\eta\,d\xi\Big],$$

where

$$p^{\kappa,\tau}_{\mu,\nu}(\xi-\eta,\eta) = -c_\nu\Big(\frac{1}{4}+c_{-\mu}c_\kappa\Big)(\xi-\eta)\eta|\xi-\eta|^{-1/2}|\xi|^{2\tau}\theta(\xi-\eta,\eta). \tag{3.59}$$

Note that

$$\begin{aligned}&\Big|\int \overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_\kappa}(\eta) p^{\kappa,\tau}_{\mu,\nu}(\xi-\eta,\eta)\psi_{\leq 0}((1+t)^{5/4+3p_0}|\xi-\eta|)\,d\eta\,d\xi\Big|\\ &\lesssim \sum_{k_1\leq k-10}\sum_{2^{k_1}\leq(1+t)^{-5/4-3p_0}} 2^{k_1/2+k}\|P_k[SU]\|_{L^2}\|P_{k_1}[SU]\|_{L^\infty}\|P_kU\|_{L^2}\\ &\lesssim \sum_{2^{k_1}\leq(1+t)^{-5/4-3p_0}} 2^{(1-p)k_1}\|SU\|^2_{H^{N_1,p}}\|U\|_{H^{N_1,p}} \lesssim (1+t)^{-1+2p_0}\epsilon_0^2.\end{aligned} \tag{3.60}$$

From the above estimate, we can see that the low-frequency part already has a sufficient decay rate. There is no need to cancel it out. This observation motivates us to define the following cubic correction terms with time-dependent cutoff functions, which only cancel out the case when $|\xi-\eta| \gtrsim (1+t)^{-5/4-3p_0}$,

$$\begin{aligned}E^{S,\mathrm{low1}}_{\mathrm{mod}}(t) := \sum_{\tau=p,N_1}\sum_{\mu,\nu,\kappa\in\{+,-\}} \mathfrak{Re}\Big[\int \overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\\ \times \widehat{U_\kappa}(\eta) q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)\,d\eta\,d\xi\Big],\end{aligned} \tag{3.61}$$

where the superscript "low1" represents the first-level correction for the low-frequency part and the symbol $q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)$ is given as follows:

$$q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta) = \frac{p^{\kappa,\tau}_{\mu,\nu}(\xi-\eta,\eta)\psi_{\geq 0}((1+t)^{5/4+3p_0}|\xi-\eta|)}{\mu|\xi|^{1/2}-\nu|\xi-\eta|^{1/2}-\kappa|\eta|^{1/2}}. \tag{3.62}$$

Note that the following estimate holds for any $\mu, \nu, \kappa \in \{+,-\}$:

$$\begin{aligned}&\big|\mu|\xi|^{1/2} - \nu|\xi-\eta|^{1/2} - \kappa|\eta|^{1/2}\big|\theta(\xi-\eta,\eta)\\ &\quad\gtrsim \left(|\xi-\eta|^{1/2} - \Big|\frac{|\xi|-|\eta|}{|\xi|^{1/2}+|\eta|^{1/2}}\Big|\right)\theta(\xi-\eta,\eta)\\ &\quad\gtrsim |\xi-\eta|^{1/2}\theta(\xi-\eta,\eta).\end{aligned}$$

With the above estimate, from Lemma 2.4 and (3.59), the following estimates hold for $\tau \in \{p, N_1\}$:

$$\sum_{\mu,\nu,\kappa\in\{+,-\}} \|q^{\kappa,\tau}_{\mu,\nu}(\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim 2^{(2\tau+1)k}, \tag{3.63}$$

$$\sum_{\mu,\nu,\kappa\in\{+,-\}} \|\partial_t q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)\|_{\mathcal{S}^\infty_{k,k_1,k_2}} \lesssim (1+t)^{-1}2^{(2\tau+1)k}. \tag{3.64}$$

From (3.63) and (3.64), the symbols $q^{\kappa,\tau}_{\mu,\nu}(\xi-\eta,\eta)$ and $\partial_t q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)$ do not contribute any smallness when $|\xi-\eta| \ll |\eta|$. As SU is forced to be put in L^2 and the symbol cannot cover the loss of p derivatives of smallness, there is a potential problem when $|\xi-\eta|$ is the smallest inside the time derivative of $E^{S,\mathrm{low1}}_{\mathrm{mod}}(t)$.

To get around this issue, we will add quartic correction terms to cancel problematic quartic terms inside $dE^{S,\mathrm{low\,1}}_{\mathrm{mod}}(t)/dt$. We first identify those problematic quartic terms by calculating $dE^{S,\mathrm{low1}}_{\mathrm{mod}}/dt$ as follows:

$$\Lambda_{\geq 4}\left[\frac{d}{dt}E^{S,\mathrm{low1}}_{\mathrm{mod}}\right] = \sum_{\tau=p,N_1}\sum_{\mu,\nu,\kappa\in\{+,-\}} \mathfrak{G}^{\kappa,\tau}_{\mu,\nu} + \sum_{\mu,\nu,\mu',\nu'\in\{+,-\}} \mathfrak{G}^{\tau}_{\mu,\nu;\mu',\nu'},$$

where

$$\begin{aligned}\mathfrak{G}^{\kappa,\tau}_{\mu,\nu} = \mathfrak{Re}\bigg[\int \Big(&\overline{\widehat{\Lambda_{\geq 3}[\partial_t(SU)_\mu]}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_\kappa}(\eta)\\ &+ \overline{\widehat{(SU)_\mu}(\xi)}\widehat{\Lambda_{\geq 2}[(\partial_t SU)_\nu]}(\xi-\eta)\widehat{U_\kappa}(\eta)\\ &+ \overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{\Lambda_{\geq 3}[\partial_t U_\kappa]}(\eta)\Big)\\ &\times q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)\\ &+ \overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_\kappa}(\eta)\\ &\times \partial_t q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)\,d\eta\,d\xi\bigg],\end{aligned} \tag{3.65}$$

$$\mathfrak{G}^{\tau}_{\mu,\nu;\mu',\nu'} = \mathfrak{Re}\bigg[\int \overline{\widehat{(SU)_{\mu}}(\xi)}\widehat{(SU)_{\nu}}(\xi-\eta)\widehat{U_{\mu'}}(\eta-\sigma)\widehat{U_{\nu'}}(\sigma) \times e^{\tau,1}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma) + \overline{\widehat{U_{\mu}}(\xi)}\widehat{(SU)_{\nu}}(\xi-\eta)\widehat{U_{\mu'}}(\eta-\sigma)\widehat{U_{\nu'}}(\sigma) \times e^{\tau,2}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\,d\sigma\,d\eta\,d\xi\bigg],$$

where

$$e^{\tau,1}_{\mu,\nu}(t,\xi,\eta,\sigma) = \sum_{\kappa\in\{+,-\}} q^{\kappa,\tau}_{\mu,\nu}(t,\xi-\eta,\eta)P_{\kappa}[q_{\kappa\mu',\kappa\nu'}(\eta-\sigma,\sigma)] + q^{\mu',\tau}_{\kappa,\nu}(t,\xi-\eta,\eta-\sigma)P_{-\kappa}[q^{1}_{\mu,-\kappa\nu'}(\xi,-\sigma)], \tag{3.66}$$

$$e^{\tau,2}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma) = \sum_{\kappa\in\{+,-\}} q^{\mu',\tau}_{\kappa,\nu}(t,\xi-\eta,\eta-\sigma)P_{-\kappa}[q^{2}_{\mu,-\kappa\nu'}(\xi,-\sigma)], \tag{3.67}$$

and $P_{\mu}[f] := f_{\mu}$. From (3.63), (3.64), (3.57), and (2.6) in Lemma 2.3, the following estimate holds:

$$\|e^{\tau,1}_{\mu,\nu}(t,\xi,\eta,\sigma)\|_{\mathcal{S}^{\infty}_{k,k_1,k_2,k_3}} + \|e^{\tau,2}_{\mu,\nu}(t,\xi,\eta,\sigma)\|_{\mathcal{S}^{\infty}_{k,k_1,k_2,k_3}} \lesssim 2^{k/2+(2\tau+2)\max\{k_1,k_2,k_3\}}. \tag{3.68}$$

Note that $\mathfrak{G}^{\kappa,\tau}_{\mu,\nu}$ is not problematic. There are two types of terms inside $\mathfrak{G}^{\kappa,\tau}_{\mu,\nu}$:

(i) a term like $\overline{\widehat{(SU)_{\mu}}(\xi)}\ \widehat{\Lambda_{\geq 2}[(\partial_t SU)_{\nu}]}(\xi-\eta)\widehat{U_{\kappa}}(\eta)$, which is quartic and higher. As we can gain a half degree of smallness from the symbol of $\widehat{\Lambda_{\geq 2}[(\partial_t SU)_{\nu}]}(\xi-\eta)$, losing p derivatives in SU is not a issue.

(ii) all other terms inside $\mathfrak{G}^{\kappa,\tau}_{\mu,\nu}$. Note that the decay rate of those terms is at least $(1+t)^{-3/2+2p_0}$. The extra gain of $(1+t)^{-1/2}$ can cover the loss of p derivatives in SU. More precisely, we lose at most $|\xi-\eta|^{-p}$. Recall that $|\xi-\eta| \gtrsim (1+t)^{-5/4-3p_0}$ inside the support of the cutoff function. Hence, we lose at most $(1+t)^{1/4+p_0}$, which can be covered from the extra gain of $(1+t)^{-1/2}$.

Recall (3.66), (3.67), (3.57), (3.63), and (3.64). We know that the new introduced symbols $q_{\mu,\nu}(\cdot,\cdot)$ and $q^{i}_{\mu,\nu}(\cdot,\cdot)$ inside $e^{\tau,i}_{\mu,\nu}(t,\xi,\eta,\sigma)$, $i\in\{1,2\}$, contribute half degrees of smallness, which is less than or equal to the second smallest number among $\xi, \xi-\eta, \eta-\sigma$, and σ. Intuitively speaking, as $\xi-(\xi-\eta+\eta-\sigma+\sigma)=0$, we know that the largest two numbers are comparable. Note that the half degree of smallness is less than the biggest number, which implies that the half degree of smallness is less than the second largest. Therefore, it is less than or equal to the second smallest number.

As a result, after combining the estimate (3.68) with the discussion in the last paragraph, the following estimate holds:

$$\begin{aligned} &\left\| e^{\tau,1}_{\mu,\nu}(t,\xi,\eta,\sigma)\psi_k(\xi)\psi_{k_1}(\xi-\eta)\psi_{k_2}(\eta-\sigma)\psi_{k_3}(\sigma)\right\|_{\mathcal{S}^\infty} \\ &\quad + \left\| e^{\tau,2}_{\mu,\nu}(t,\xi,\eta,\sigma)\psi_k(\xi)\psi_{k_1}(\xi-\eta)\psi_{k_2}(\eta-\sigma)\psi_{k_3}(\sigma)\right\|_{\mathcal{S}^\infty} \\ &\quad \lesssim \min\{2^{k/2+(2\tau+2)\max\{k_1,k_2,k_3\}}, \\ &\qquad\qquad 2^{\mathrm{Smin}\{k,k_1,k_2,k_3\}/2+(2\tau+2)\max\{k_1,k_2,k_3\}}\}, \end{aligned} \tag{3.69}$$

where "Smin$\{k,k_1,k_2,k_3\}$" denotes the second smallest number among k, k_1, k_2, and k_3.

Recall the detailed formula of $\mathfrak{G}^\tau_{\mu,\nu;\mu',\nu'}$ in (3.65). By the L^2-L^∞-type multilinear estimate, we put all SU in L^2 and all U in L^∞. As a result, the total loss is at most of size $|\xi|^{-p}|\xi-\eta|^{-p} = 2^{-pk-pk_1}$. From (3.69), we can see that the loss of 2^{-pk-pk_1} can be covered by the symbol if $2^{k_1} \gtrsim 2^{\mathrm{Smin}\{k,k_1,k_2,k_3\}}$.

Recall that, due to the time-dependent cutoff function of $|\xi-\eta|$ in (3.62), we have $|\xi-\eta| \geq (1+t)^{-5/4-3p_0}$. Moreover, from (3.69), we know that the loss can also be covered by the symbol if $2^{(1/2-p)k}$ is less than $2^{-pk_1} \sim |\xi-\eta|^{-p} \leq (1+t)^{1/4+p_0}$. Hence, we can further rule out the case when $|\xi| \sim 2^k \leq (1+t)^{-5/6-5p_0}$.

To sum up, the quartic terms inside $\mathfrak{G}^\tau_{\mu,\nu;\mu',\nu'}$ are only problematic if $|\xi-\eta|$ is the smallest number and not comparable with the second smallest number among ξ, $\xi-\eta$, $\eta-\sigma$, and σ and $|\xi| \sim 2^k \gtrsim (1+t)^{-5/6-5p_0}$.

For this problematic scenario, a key observation is that the size of phases in this case is greater than the second smallest number instead of the smallest number among ξ, $\xi-\eta$, $\eta-\sigma$, and σ. More precisely, the following estimate holds:

$$\begin{aligned} &\big|\mu|\xi|^{1/2} - \nu|\xi-\eta|^{1/2} - \mu'|\eta-\sigma|^{1/2} - \nu'|\sigma|^{1/2}\big|\theta(\xi-\eta,\eta-\sigma)\theta(\xi-\eta,\sigma) \\ &\quad \times \theta(\xi-\eta,\xi) \gtrsim \min\{|\xi|,|\eta-\sigma|,|\sigma|\}^{1/2}, \quad \mu,\nu,\mu',\nu' \in \{+,-\}. \end{aligned} \tag{3.70}$$

The above estimate is very similar to an estimate of phases, which will be used in the L^∞-decay estimate part. Hence, we postpone the proof of (3.70) to the end of Section 4.2.

As the size of phase is not small, we can divide the phase again to gain another $t^{-1/2}$ decay rate with the price of $2^{-\mathrm{Smin}\{k,k_1,k_2,k_3\}/2}$, which can be covered by the symbol; see (3.69). This observation motivates us to define the following quartic correction terms, which cancel out the problematic quartic terms inside $\mathfrak{G}^\tau_{\mu,\nu;\mu',\nu'}$,

$$\begin{aligned} E^{S,\mathrm{low2}}_{\mathrm{mod}}(t) = \sum_{\tau=p,N_1}\sum_{\mu,\nu,\mu',\nu'\in\{+,-\}} \mathfrak{Re}\Big[\int &\overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_{\mu'}}(\eta-\sigma)\widehat{U_{\nu'}}(\sigma) \\ &\times \tilde{e}^{\tau,1}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma) \\ &+ \overline{\widehat{U_\mu}(\xi-\sigma)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_{\mu'}}(\eta)\widehat{U_{\nu'}}(\sigma) \\ &\times \tilde{e}^{\tau,2}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\,d\sigma\,d\eta\,d\xi\Big], \end{aligned} \tag{3.71}$$

where the superscript "low2" represents the second-level correction for the low-frequency part and the symbol $\widetilde{e}^{\tau,i}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)$, $i \in \{1,2\}$, is given as follows:

$$\widetilde{e}^{\tau,i}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma) = \frac{e^{\tau,i}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\theta(\xi-\eta,\min\{|\sigma|,|\eta-\sigma|,|\xi|\})}{\mu|\xi|^{1/2}-\nu|\xi-\eta|^{1/2}-\mu'|\eta-\sigma|^{1/2}-\nu'|\sigma|^{1/2}}\psi_{\geq 0}\big(\xi(1+t)^{5/6+5p_0}\big).$$

From (3.70) and (3.69), we can see that the loss from the denominator of $\widetilde{e}^{\tau,i}_{\mu,\nu;\mu',\nu'}$ can be covered by the size of $e^{\tau,i}_{\mu,\nu}$, $i \in \{1,2\}$. As a result, from (3.57), (3.66), (3.67), (3.63), (3.64), Lemma 2.4, and (2.6) in Lemma 2.3, the following estimates hold:

$$\|\widetilde{e}^{\tau,1}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} + \|\widetilde{e}^{\tau,2}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} \lesssim 2^{(2\tau+2)\max\{k_1,k_2,k_3\}}, \tag{3.72}$$

$$\|\partial_t\widetilde{e}^{\tau,1}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} + \|\partial_t\widetilde{e}^{\tau,2}_{\mu,\nu;\mu',\nu'}(t,\xi,\eta,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} \lesssim (1+t)^{-1}2^{(2\tau+2)\max\{k_1,k_2,k_3\}}. \tag{3.73}$$

Note that $|\xi|^{-p}|\xi-\eta|^{-p} \lesssim (1+t)^{5/12+2p_0} \leq (1+t)^{1/2-100p_0}$ inside the support. Hence, the extra gain of $t^{-1/2}$ is sufficient to close the argument.

From the above discussion, we define the total correction terms for the low-frequency part of energy as follows:

$$E^{S,\mathrm{low}}_{\mathrm{mod}}(t) := E^{S,\mathrm{low1}}_{\mathrm{mod}}(t) + E^{S,\mathrm{low2}}_{\mathrm{mod}}(t). \tag{3.74}$$

LEMMA 3.7. *Under the bootstrap assumption* (2.29)*, we have*

$$\sup_{t\in[0,T]}\big|E^{S,\mathrm{low}}_{\mathrm{mod}}(t)\big| \lesssim \epsilon_0^2, \tag{3.75}$$

$$\left|\frac{dE^{S,\mathrm{low}}_{\mathrm{mod}}(t)}{dt} - \sum_{\tau=p,N_1}\mathfrak{Re}\Big[\int \overline{\partial_x^\tau SU^1}\partial_x^\tau\mathfrak{Q}^S_2 + \overline{\partial_x^\tau SU^2}\partial_x^\tau\mathfrak{Q}^S_4\Big]\right| \lesssim (1+t)^{-1+2p_0}\epsilon_1^3. \tag{3.76}$$

PROOF. Recall (3.61). From (3.63) and the L^2-L^∞-type bilinear estimate (2.7) in Lemma 2.3, the following estimate holds:

$$\begin{aligned}\big|E^{S,\mathrm{low1}}_{\mathrm{mod}}(t)\big| &\lesssim \|SU(t)\|^2_{H^{N_1,p}}\|U(t)\|_{W^3}\sum_{\substack{k\in\mathbb{Z},\\ 2^k\geq(1+t)^{-5/4-3p_0}}}2^{-pk}\\ &\lesssim (1+t)^{-1/4+10p_0}\epsilon_1^3 \lesssim \epsilon_0^2.\end{aligned} \tag{3.77}$$

Recall (3.71). From (3.72) and the L^2-L^∞-L^∞-type trilinear estimate (2.8) in Lemma 2.3, we have

$$\begin{aligned}
&\left|E_{\mathrm{mod}}^{S,\mathrm{low2}}(t)\right| \\
&\lesssim \sum_{2^{k_1}\ge(1+t)^{-5/4-3p_0}} \sum_{2^k\ge(1+t)^{-5/6-5p_0}} 2^{-pk_1-pk}(\|SU\|_{H^{N_1,p}} + \|U\|_{H^{N_0,p}})^2\|U\|_{W^3} \\
&\lesssim (1+t)^{-1/2}\epsilon_1^4 \lesssim \epsilon_0^2. \qquad (3.78)
\end{aligned}$$

From (3.77) and (3.78), we can see that the desired estimate (3.75) holds. From (3.60), (3.72), (3.73), and the L^2-L^∞-type multilinear estimate in Lemma 2.3, the following estimate holds:

$$\begin{aligned}
&\text{LHS of (3.76)} \\
&\lesssim \sum_{\tau=p,N_1}\sum_{\mu,\nu,\kappa\in\{+,-\}} \left|\mathfrak{G}_{\mu,\nu}^{\kappa,\tau}\right| \\
&\quad + \left|\sum_{\tau=p,N_1}\sum_{\mu,\nu,\mu',\nu'\in\{+,-\}} \mathfrak{G}_{\mu,\nu;\mu',\nu'}^{\tau} + \frac{d}{dt}E_{\mathrm{mod}}^{S,\mathrm{low2}}(t)\right| \\
&\quad + \sum_{\mu,\nu,\kappa\in\{+,-\}} \left|\int \overline{\widehat{(SU)_\mu}(\xi)}\widehat{(SU)_\nu}(\xi-\eta)\widehat{U_\kappa}(\eta)p_{\mu,\nu}^{\kappa,\tau}(\xi-\eta,\eta)\psi_{\le 0} \right.\\
&\qquad\qquad \left. \times((1+t)^{5/4+3p_0}|\xi-\eta|)\,d\eta\, d\xi\right| \\
&\lesssim \|SU\|_{H^{N_1,p}}^2\big[(1+t)^{-1/2}\|U\|_{W^3} + \|U\|_{H^{N_0,p}}(1+t)^{-(1+2p_0)}\big] \qquad (3.79)\\
&\quad + (\|SU\|_{H^{N_1,p}} + \|U\|_{H^{N_0,p}})^2\|U\|_{W^3}^2 \\
&\quad + \sum_{2^{k_1}\ge(1+t)^{-5/4-3p_0}}\sum_{2^k\ge(1+t)^{-5/6-5p_0}} 2^{-pk_1-pk}\big(\|\Lambda_{\ge2}[\partial_t SU]\|_{H^{N_1,p}}\|SU\|_{H^{N_1,p}}\|U\|_{W^3}^2 \\
&\qquad\qquad + \|SU\|_{H^{N_1,p}}^2\|U\|_{W^2}\|\Lambda_{\ge2}[\partial_t U]\|_{W^2}\big) \\
&\quad + \sum_{2^{k_1}\ge(1+t)^{-5/4-3p_0}} 2^{-pk_1}\|SU\|_{H^{N_1,p}}^2\|U\|_{W^3}\|\Lambda_{\ge2}[\partial_t U]\|_{W^3} \\
&\lesssim (1+t)^{-1+2p_0}\epsilon_1^3.
\end{aligned}$$

In the above estimate, we used the following simple fact, which is derived from the Sobolev embedding:

$$\|\Lambda_{\ge2}[\partial_t U]\|_{W^3} \lesssim \|U\|_{W^3}\|U\|_{W^4} \lesssim \|U\|_{W^{N_2}}^{3/2+10p_0}\|U\|_{H^{N_0,p}}^{1/2-2p_0}. \qquad \square$$

Combining (3.26), (3.28), (3.53), (3.55), (3.75), and (3.76), we can see that our desired energy estimate (2.31) in Proposition (2.7) holds.

4 Improved Dispersion Estimate

The goal of this section is to derive the sharp decay rate of the nonlinear solution under the bootstrap assumptions (2.29), hence finishing the bootstrap argument. To simplify the notation, we define $\Lambda := |\nabla|^{1/2}$ and $\Lambda(\xi) := |\xi|^{1/2}$ throughout this section. From (2.31) in Proposition (2.7), we have

$$\sup_{t\in[0,T]} (1+t)^{-p_0}\big[\|(U^1,U^2)\|_{H^{N_0,p}} + \|S(U^1,U^2)\|_{H^{N_1,p}}\big] \lesssim \epsilon_0. \tag{4.1}$$

From the improved energy estimate (4.1), we can first rule out the very-low-frequency case and the very-high-frequency case as follows:

$$\begin{aligned} &\sum_{\substack{2^k\leq(1+t)^{-5/3-2p_0}\\ \text{or } 2^k\geq(1+t)^{10/89+p_0}}} \big\|P_k(U^1,U^2)\big\|_{W^{N_2}} \\ &\lesssim \sum_{\substack{2^k\leq(1+t)^{-5/3-2p_0}\\ \text{or } 2^k\geq(1+t)^{10/89+p_0}}} 2^{k/2+N_2k_+}\big\|P_k(U^1,U^2)\big\|_{L^2} \lesssim (1+t)^{-1/2}\epsilon_0. \end{aligned} \tag{4.2}$$

To derive the sharp decay estimate for the case when $(1+t)^{-5/3-2p_0} \leq |\xi| \leq (1+t)^{10/89+p_0}$, we use the following auxiliary Z-normed space:

$$\|h\|_Z := \||\xi|^{\beta}(1+|\xi|^{\gamma})h(\xi)\|_{L^\infty_\xi}, \quad \beta = \frac{3}{4} - p_0, \;\; \gamma = N_2 + 2p_0, \tag{4.3}$$

and the linear decay estimate (4.4) in Lemma 4.2.

Remark 4.1. The Z-normed space in (4.3) is similar but much weaker than the one used in [21]. The β used there is $\frac{1}{100}$. Due to the infinite energy setting, β needs to be greater than $\frac{1}{2}$. Otherwise, the finiteness of the Z-norm implies the finiteness of energy and the neutrality of the physical velocity. The choice of β in (4.3), which is suggested by the linear decay estimate (4.4), guarantees that the Z-norm is weaker than the energy spaces that we used in this paper but strong enough to prove the sharp decay rate of the nonlinear solution.

LEMMA 4.2. *For any* $t, |t|\geq 1$, $k\in\mathbb{Z}$, $\epsilon\in(0,\frac{1}{2}]$, *and* $f\in L^2(\mathbb{R})$, *the following estimate holds:*

$$\begin{aligned} \|e^{it\Lambda}P_k f\|_{L^\infty} \lesssim{}& |t|^{-1/2}2^{3k/4}\|\widehat{f}\|_{L^\infty} \\ &+ |t|^{-(1/2+\epsilon/2)}2^{(1-\epsilon)k/4}\big[2^k\|\partial_\xi \widehat{f}\|_{L^2} + \|f\|_{L^2}\big]. \end{aligned} \tag{4.4}$$

Remark 4.3. From the second term on the right-hand side of the estimate (4.4), we know that p should be strictly less than $1/4$ to guarantee the sharp decay rate (i.e., $t^{-1/2}$) of the nonlinear solution. The main reason for this conclusion is that $2^{pk}[2^k\|\partial_\xi \widehat{f}\|_{L^2} + \|f\|_{L^2}]$, which corresponds to the energy of the nonlinear solution, is expected to grow appropriately.

PROOF. Note that estimate (4.4) is scaling invariant. It is sufficient to prove it for the case when $k=0$. The proof is very similar to the proof of linear decay estimates in [20, 22, 24]. With minor modifications, one can redo the argument in [20] to derive (4.4) without any problems. □

The strategy that we will use to get the sharp decay estimate for U^1 and U^2 is as follows:

(i) Recall (3.1). We will first derive the equation satisfied by the normal form transformation (V_1 and V_2) of U^1 and U^2, which is cubic and higher. Then we show that the decay rates of (U^1, U^2) and (V_1, V_2) are comparable.

(ii) We prove that the Z-norm of the profile of (V_1, V_2) does not grow with respect to time. Hence, V_1 and V_2 decay sharply from Lemma 4.2. As the decay rate of (U^1, U^2) is the same as that for (V_1, V_2), we know that U^1 and U^2 also decay sharply.

Recall (3.1). Define $V = V_1 + iV_2$ and $\mathcal{S} := \{(+,+,+), (+,+,-), (+,-,-), (-,-,-)\}$. As the equation satisfied by V is cubic and higher, we can write it as

$$(\partial_t + i\Lambda)V = \sum_{(\iota_1,\iota_2,\iota_3)\in\mathcal{S}} C_{\iota_1,\iota_2,\iota_3}(V^{\iota_1}, V^{\iota_2}, V^{\iota_3}) + R(t), \tag{4.5}$$

where $R(t)$ represents the quartic and higher-order terms. Define $f(t) := e^{it\Lambda}V$. We use $\widehat{f_k}(t,\xi)$ to abbreviate $\widehat{f}(t,\xi)\psi_k(\xi)$ throughout this section.

LEMMA 4.4. *Under the smallness condition* (2.29) *and the improved energy estimate* (4.1)*, the following estimates hold for* $t \in [0, T]$:

$$\|(V_1 - U^1, V_2 - U^2)(t)\|_{W^{N_2}} \lesssim (1+t)^{-5/8}\epsilon_1^2, \quad \|V(t)\|_{W^{N_2}} \lesssim (1+t)^{-1/2}\epsilon_1, \tag{4.6}$$

$$\sup_{k\in\mathbb{Z}} \big[2^{k_-/5+7k_+}\|P_k f(t)\|_{L^2} + 2^{k_-/5}\|\xi\partial_\xi \widehat{f}(\xi,t)\psi_k(\xi)\|_{L^2}\big] \lesssim (1+t)^{p_0}\epsilon_0. \tag{4.7}$$

PROOF. From the estimate (3.7), the bilinear estimate (2.7) in Lemma 2.3, and Sobolev embedding, our desired estimate (4.6) holds as follows:

$$\begin{aligned}
&\|A_1(U^1,U^1)\|_{W^{N_2}} + \|A_2(U^2,U^2)\|_{W^{N_2}} + \|B(U^1,U^2)\|_{W^{N_2}} \\
&\quad\lesssim \sum_{k\in\mathbb{Z}} 2^{5k/4+N_2k_+}\|(P_{\le k}U^1, P_{\le k}U^2)\|_{W^0}\|(P_kU^1, P_kU^2)\|_{L^\infty}^{1/2}\|(P_kU^1, P_kU^2)\|_{L^2}^{1/2} \\
&\qquad + \sum_{k\in\mathbb{Z}}\sum_{k_1\le k} 2^{k_1/4+k+N_2k_{1,+}}\|P_kU^1\|_{L^4}\|P_kU^2\|_{L^\infty} \\
&\quad\lesssim \|(U^1,U^2)\|_{H^{N_0,p}}^{1/2}\|(U^1,U^2)\|_{W^{N_2}}^{3/2} \lesssim (1+t)^{-5/8}\epsilon_1^2,
\end{aligned}$$

We also have the following L^2- type estimate:

$$\begin{aligned}
&\|A_1(U^1,U^1)(t)\|_{H^{N_0-1,p}} + \|A_2(U^2,U^2)(t)\|_{H^{N_0-1,p}} \\
&\quad + \|B(U^1,U^2)(t)\|_{H^{N_0-1,p}} \\
&\qquad \lesssim \|(U^1,U^2)(t)\|_{H^{N_0,p}}\|(U^1,U^2)(t)\|_{W^0} \lesssim (1+t)^{-(1/2-p_0)}\epsilon_1^2,
\end{aligned}$$

which further gives us the following estimate:

$$\sup_{t\in[0,T]}\sup_{k\in\mathbb{Z}}(1+t)^{-p_0}2^{k_-/5+7k_+}\|P_k f(t)\|_{L^2} \lesssim \epsilon_0. \tag{4.8}$$

Recall that $A(\cdot,\cdot)$ is a symmetric bilinear operator. Hence, the following estimate holds for any fixed $k \in \mathbb{Z}$:

$$\begin{aligned}
&\big\| P_k\big[SA_1(U^1,U^1)\big]\big\|_{L^2} \\
&\quad\lesssim \big\| P_k\big[A_1(SU^1,U^1)\big]\big\|_{L^2} + \big\| P_k\big[A_1(U^1,U^1)\big]\big\|_{L^2} \\
&\quad\lesssim \sum_{k_1\le k-4} 2^k\big[\| P_k U^1\|_{L^2}\| P_{k_1}(SU^1)\|_{L^\infty}\| P_k U^1\|_{L^\infty}\| P_{k_1}(SU^1)\|_{L^2}\big]^{1/2} \\
&\qquad + 2^k\| P_k(SU^1,U^1)\|_{L^2}\|U^1\|_{W^0} \\
&\quad\lesssim \sum_{k_1\le k-4} 2^{k_1/4}2^k\| P_{k_1}(SU^1)\|_{L^2}\| P_k U^1\|_{L^2}^{1/2}\| P_k U^1\|_{L^\infty}^{1/2} + (1+|t|)^{-1/2+p_0}\epsilon_1^2 \\
&\quad\lesssim (1+t)^{-1/4+2p_0}\epsilon_1^2 \lesssim \epsilon_0.
\end{aligned}$$

By using the same argument, we can show that

$$\sup_{k\in\mathbb{Z}}\big\| P_k\big[SA_2(U^2,U^2)\big]\big\|_{L^2} + \big\| P_k[SB(U^1,U^2)]\big\|_{L^2} \lesssim \epsilon_0. \tag{4.9}$$

Note that $\widehat{SV}(\xi) = e^{-it|\xi|^{1/2}}(t\partial_t - 2\xi\partial_\xi - 2)\widehat{f}(t,\xi)$. Hence

$$\sup_{t\in[0,T]}\sup_{k\in\mathbb{Z}}(1+t)^{-p_0}2^{k_-/5}\|(t\partial_t - 2\xi\partial_\xi)\widehat{f}(t,\xi)\psi_k(\xi)\|_{L^2} \lesssim \epsilon_0. \tag{4.10}$$

From the estimate (4.13) in Lemma 4.5, the L^2-L^∞-L^∞-type trilinear estimate (2.8) in Lemma 2.3 and (A.7) in Lemma A.1, the following estimate holds:

$$\begin{aligned}
&\|t\partial_t f\|_{L^2} \\
&\quad\lesssim t\sum_{k_1,k_2,k_3\in\mathbb{Z}} 2^{\mathrm{med}\{k_i\}/2}2^{2\max\{k_i\}}\| P_{\min\{k_i\}}(e^{-it\Lambda}f)\|_{L^\infty} \\
&\qquad\times\| P_{\mathrm{med}\{k_i\}}(e^{-it\Lambda}f)\|_{L^\infty}\| P_{\max\{k_i\}}f\|_{L^2} + t\|R\|_{L^2} \\
&\quad\lesssim t\|V(t)\|_{W^{1/2}}^2\|V\|_{H^{2,p}} \lesssim (1+|t|)^{p_0}\epsilon_1^3 \lesssim (1+|t|)^{p_0}\epsilon_0.
\end{aligned}\tag{4.11}$$

Combining (4.8), (4.10), and (4.11), it is easy to see that the desired estimate (4.7) holds. With minor modifications, we can also derive the following estimate:

$$\sup_{\substack{(1+t)^{-5/3-2p_0}\le 2^k\\ \le(1+t)^{10/89+p_0}}} 2^{k_-/5+N_1k}\|\xi\partial_\xi\widehat{f}(t,\xi)\psi_k(\xi)\|_{L^2} \lesssim (1+t)^{p_0}\epsilon_0. \tag{4.12}$$ □

From (4.7) and (4.12), the following estimate holds:

$$\sum_{\substack{(1+t)^{-5/3-2p_0}\le 2^k\\ \le(1+t)^{10/89+p_0}}} 2^{k/8+N_2k_+}\big[2^k\|\partial_\xi\widehat{f}\psi_k(\xi)\|_{L^2} + \|P_kf\|_{L^2}\big] \lesssim (1+|t|)^{1/4-2p_0}\epsilon_0.$$

Hence, from the above estimate and (4.4) in Lemma 4.2, it is sufficient to derive the improved Z-norm estimate of f to derive the sharp decay estimate of V.

4.1 Setup of the Z-Norm Estimate

After writing the cubic terms on the Fourier side and doing the Littlewood-Paley decomposition for all inputs, we have

$$\begin{aligned}\mathcal{F}(e^{it\Lambda}C_{\iota_1,\iota_2,\iota_3}(V^{\iota_1},V^{\iota_2},V^{\iota_3}))(\xi) &:= iI^{\iota_1,\iota_2,\iota_3}(t,\xi)\\ &= i\sum_{k_1,k_2,k_3\in\mathbb{Z}} I^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}(t,\xi),\end{aligned}$$

$$\begin{aligned}I^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}(t,\xi) = \int_{\mathbb{R}\times\mathbb{R}} & e^{it\,\Phi_{\iota_1,\iota_2,\iota_3}(\xi,\eta,\sigma)}\,c^{\iota_1,\iota_2,\iota_3}(\xi-\eta,\eta-\sigma,\sigma)\widehat{f^{\iota_1}_{k_1}}(t,\xi-\eta)\\ &\times \widehat{f^{\iota_2}_{k_2}}(t,\eta-\sigma)\widehat{f^{\iota_3}_{k_3}}(t,\sigma)d\eta\, d\sigma,\end{aligned}$$

where the phase $\Phi_{\iota_1,\iota_2,\iota_3}(\cdot,\cdot,\cdot)$ is defined as follows:

$$\Phi_{\iota_1,\iota_2,\iota_3}(\xi,\eta,\sigma) = \Lambda(\xi) - \iota_1\Lambda(\xi-\eta) - \iota_2\Lambda(\eta-\sigma) - \iota_3\Lambda(\sigma), \quad (\iota_1,\iota_2,\iota_3)\in\mathcal{S}.$$

The precise formulas of symbols $c^{\iota_1,\iota_2,\iota_3}(\cdot,\cdot,\cdot)$, $(\iota_1,\iota_2,\iota_3)\in\mathcal{S}$ do not play many roles. Knowing the estimate of the $\mathcal{S}^\infty$-norm of symbols $c^{\iota_1,\iota_2,\iota_3}(\cdot,\cdot,\cdot)$ is sufficient. More precisely, we have the following lemma:

LEMMA 4.5. *For any $(\iota_1,\iota_2,\iota_3)\in\mathcal{S}$, $k,k_1,k_2,k_3\in\mathbb{Z}$, the following estimate holds*:

$$\|c^{\iota_1,\iota_2,\iota_3}(\xi-\eta,\eta-\sigma,\sigma)\|_{\mathcal{S}^\infty_{k,k_1,k_2,k_3}} \lesssim 2^{\mathrm{med}\{k_1,k_2,k_3\}/2}2^{2\max\{k_1,k_2,k_3\}}, \tag{4.13}$$

where $\mathrm{med}\{k_1,k_2,k_3\}$ *denotes the median among k_1, k_2, and k_3.*

Remark 4.6. We remark that a detailed analysis of cubic terms of the water waves system (1.3) can be found in [21, app. B]. However, the main purposes in [21] are different from our purposes; the estimate of symbols there is slightly different from (4.13). For completeness, we provide an intuitive proof of (4.13).

PROOF OF (4.13). Recall that $V = V_1 + iV_2$, $V_1 = U^1 + A_1(U^1,U^1) + A_2(U^2,U^2)$, and $V_2 = U^2 + B(U^1,U^2)$. From the Taylor expansion of the Dirichlet-Neumann operator in (2.1) and the definitions of V_i and U_i, $i\in\{1,2\}$, in (2.16) and (3.1), we can calculate explicit formulas of $c^{\iota_1,\iota_2,\iota_3}(\xi-\eta,\eta-\sigma,\sigma)$. Then, we can see that our desired estimate (4.13) holds after applying the estimate (2.9) in Lemma 2.4. Unfortunately, the details of the formulas of $c^{\iota_1,\iota_2,\iota_3}(\xi-\eta,\eta-\sigma,\sigma)$ are very tedious. So, in the following, we will provide a more intuitive proof, which explains why our desired estimate (4.13) should hold.

Note that there are two sources of cubic terms that contribute to the cubic terms inside the equation (4.5) satisfied by V. One of them comes from $\Lambda_3[\partial_t U_1]$ and $\Lambda_3[\partial_t U_2]$. The other one comes from $A_1(\Lambda_2[\partial_t U^1],U^1)$, $A_2(\Lambda_2[\partial_t U^2],U^2)$, $B(\Lambda_2[\partial_t U^1],U^2)$, and $B(U^1,\Lambda_2[\partial_t U^1])$.

Our desired estimate (4.13) follows from the following facts:

(i) The symbols of cubic terms are all homogeneous of degree $\frac{5}{2}$ in terms of U^1 and U^2. Hence, to prove (4.13), we only have to consider the case when $\text{med}\{k_i\} \leq \max\{k_i\} - 10$.

(ii) Cubic terms lose at most two derivatives. From (2.17), we know $\Lambda_3[\partial_t U_1]$ and $\Lambda_3[\partial_t U_2]$ lose at most one derivative. From (3.7) in Lemma 3.1 and (2.24) in Lemma (2.5), we know that cubic terms, $A_1(\Lambda_2[\partial_t U^1], U^1)$, $A_2(\Lambda_2[\partial_t U^2], U^2)$, $B(\Lambda_2[\partial_t U^1], U^2)$, and $B(U^1, \Lambda_2[\partial_t U^1])$, lose at most two derivatives. To sum up, we know that the size of the symbol contributes at least a half degree smallness of the medium frequency.

To see why fact (i) holds, note that the symbols of normal form transformations $A_1(\cdot,\cdot)$, $A_2(\cdot,\cdot)$, and $B(\cdot,\cdot)$ are all homogeneous of degree 1 and the symbols of quadratic terms $\Lambda_2[\partial_t U^1]$ and $\Lambda_2[\partial_t U^2]$ are all homogeneous of degree $\frac{3}{2}$. Therefore, the symbols of all cubic terms that come from the normal form transformation are homogeneous of degree $\frac{5}{2}$. Intuitively speaking, the cubic term of $\partial_t U^1$ is either of type $\partial_x h \partial_x h \partial_x \psi$ or of type $\partial_x |\nabla|\psi|\nabla|\psi|\nabla|\psi$; the cubic term of $\partial_t U^2$ is of type $|\nabla|^{1/2}(\partial_x\psi\partial_x\psi\partial_x h)$ or of type $|\nabla|^{1/2}(h\partial_x h\partial_x h)$. Recall that $\Lambda_1[\psi] = |\nabla|^{-1/2}U^2$ and $\Lambda_1[h] = U^1$. Now it is easy to see that they are all homogeneous of degree $\frac{5}{2}$. □

As in [22], we need to modify the phase of the profile first to successfully close the argument and see the modified scattering property. We define

$$\begin{aligned} c^*(\xi, x, y) &:= c^{+,+,-}(\xi + x, \xi + y, -\xi - x - y), \\ \widetilde{c}(\xi) &:= -8\pi|\xi|^{3/2} c^*(\xi, 0, 0), \end{aligned} \tag{4.14}$$

and the modified phase as follows:

$$L(t,\xi) := \widetilde{c}(\xi)\int_0^t |\widehat{f}(s,\xi)|^2 \frac{ds}{1+s}, \quad g(t,\xi) := e^{iL(t,\xi)}\widehat{f}(t,\xi). \tag{4.15}$$

Hence

$$\begin{aligned} \partial_t g(t,\xi) &= i e^{iL(t,\xi)}\left[I^{+,+,-}(t,\xi) + \widetilde{c}(\xi)\frac{|\widehat{f}(t,\xi)|^2}{1+t}\widehat{f}(t,\xi)\right] \\ &\quad + i e^{iL(t,\xi)}[I^{+,+,+}(t,\xi) + I^{+,-,-}(t,\xi) + I^{-,-,-}(t,\xi)] \\ &\quad + e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi). \end{aligned} \tag{4.16}$$

From (4.16), we can see that the modified phase is only effective for the cubic term $I^{+,+,-}(t,\xi)$, which is the only one that has nonempty space-time resonance set.

Now, our main goal is reduced to proving the following proposition:

PROPOSITION 4.7. *Under the bootstrap assumption* (2.29) *and the assumption*

$$\sup_{t\in[0,T']} \|\widehat{f}(t)\|_Z \lesssim \epsilon_1, \quad T' \in (0,T], \tag{4.17}$$

there exists $p_1 > 0$ *such that the following estimate holds for any* $m \in \mathbb{N}$ *and any* $t_1, t_2 \in [2^{m-1}, 2^{m+1}] \subset [0, T']$:

$$\| |\xi|^{3/4-p_0}(1+|\xi|^{N_2+2p_0})(g(t_2,\xi)-g(t_1,\xi))\|_{L^\infty_\xi} \lesssim 2^{-p_1 m}\epsilon_0. \tag{4.18}$$

Hence, we have $T' = T$ *and*

$$\begin{aligned} \sup_{t\in[0,T]} \|\widehat{f}(t)\|_Z = \sup_{t\in[0,T]} \|g(t)\|_Z \lesssim \epsilon_0, \\ \sup_{t\in[0,T]} (1+t)^{1/2}\|V(t)\|_{W^{N_2}} \lesssim \epsilon_0. \end{aligned} \tag{4.19}$$

Now we restrict ourselves to the time interval $[2^{m-1}, 2^{m+1}] \subset [0, T']$ and reformulate estimates (4.6), (4.7), and (4.17) as follows:

$$\begin{aligned} \sup_{k\in\mathbb{Z}} \|P_k f\|_{L^2} &\lesssim \epsilon_0 2^{p_0 m}2^{-k_-/5-7k_+}, \\ \sup_{k\in\mathbb{Z}} \|\mathcal{F}(P_k f)(\xi)\|_{L^\infty_\xi} &\lesssim \epsilon_1 2^{-\beta k}2^{-\gamma k_+}, \\ \sup_{k\in\mathbb{Z}} \|\xi\partial_\xi \widehat{f}(\xi)\psi_k(\xi)\|_{L^2} &\lesssim \epsilon_0 2^{p_0 m}2^{-k_-/5}, \\ \sup_{k\in\mathbb{Z}} \|e^{-it\Lambda}P_k f\|_{L^\infty_x} &\lesssim \epsilon_1 2^{-m/2}2^{-N_2 k_+}. \end{aligned}$$

LEMMA 4.8. *For any* $k \in \mathbb{Z}$ *and* $f \in L^2(\mathbb{R})$, *we have*

$$\|\widehat{P_k f}\|^2_{L^\infty_\xi} \lesssim 2^{-k}\|\widehat{f}\|_{L^2}\big[2^k\|\partial_\xi \widehat{f}\|_{L^2} + \|\widehat{f}\|_{L^2}\big]. \tag{4.20}$$

PROOF. The proof is standard. Because the desired estimate (4.20) is scale invariant, we only have to prove it for the case when $k = 0$, which follows directly from the Cauchy-Schwarz inequality. Or one can find a detailed proof in [22, 24]. □

From estimate (4.20) in Lemma 4.8, the following estimates hold when $|\xi| \leq 2^{-21p_0 m}$ or $|\xi| \geq 2^{20p_0 m}$,

$$\begin{aligned} &\sup_{\substack{|\xi|\leq 2^{-21p_0 m} \\ \text{or } |\xi|\geq 2^{20p_0 m}}} ||\xi|^{3/4-p_0}(1+|\xi|^{N_2+2p_0})g(t,\xi)| \\ &= \sup_{\substack{|\xi|\leq 2^{-21p_0 m} \\ \text{or } |\xi|\geq 2^{20p_0 m}}} \big||\xi|^{3/4-p_0}(1+|\xi|^{N_2+2p_0})f(t,\xi)\big| \\ &\lesssim \sup_{k\leq -21p_0 m} \epsilon_0 2^{p_0 m}2^{(1/4-p_0)k-k/5} \\ &\quad + \sup_{k\geq 20p_0 m} \epsilon_0 2^{(1/4+N_2+p_0)k-(N_0-1)k/2}2^{p_0 m} \lesssim \epsilon_0 2^{-p_0 m}. \end{aligned}$$

Hence, to prove (4.18), it remains to consider the case when $2^{-21p_0 m} \leq |\xi| \leq 2^{20p_0 m}$. For this case, we need to use the equation satisfied by the modified profile

$g(t,\xi)$. Recalling (4.16), we have the following identity:

$$g(t_2,\xi)-g(t_1,\xi)=\sum_{k_1,k_2,k_3\in\mathbb{Z}}\sum_{(\iota_1,\iota_2,\iota_3)\in\mathcal{S}} J^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3} + \int_{t_1}^{t_2} e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi)\,dt,$$

where

$$J^{+,+,-}_{k_1,k_2,k_3} := \int_{t_1}^{t_2} ie^{iL(t,\xi)}\Big[I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)+\widetilde{c}(\xi)\frac{\widehat{f_{k_1}}(t,\xi)\,\overline{\widehat{f_{k_3}}}(t,-\xi)}{1+t}\widehat{f_{k_2}}(t,\xi)\Big]ds,$$

$$J^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3} := \int_{t_1}^{t_2} ie^{iL(t,\xi)} I^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}(t,\xi)\,dt,$$
$$(\iota_1,\iota_2,\iota_3)\in\{(+,+,+),(+,-,-),(-,-,-)\}.$$

The argument naturally aligns into two parts: a Z-norm estimate for the cubic terms and a Z-norm estimate for the remainder term.

4.2 Z-Norm Estimate for the Cubic Terms

The goal of this subsection is to prove the following proposition:

PROPOSITION 4.9. *For $t_1,t_2\in[2^{m-1},2^{m+1}]\subset[0,T']$, $|\xi|:=2^k\in[2^{-21p_0m}, 2^{20p_0m}]$, the following estimate holds under the bootstrap assumptions* (2.29) *and* (4.17):

$$\sum_{k_1,k_2,k_3\in\mathbb{Z}}\sum_{(\iota_1,\iota_2,\iota_3)\in\mathcal{S}}\big\|J^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}\psi_k(\xi)\big\|_Z\lesssim 2^{-p_0m}\epsilon_1^3. \tag{4.21}$$

LEMMA 4.10. *Under the assumptions in Proposition* 4.9*, the following estimate holds if $k_1,k_2,k_3\in[k-10,k+10]$*:

$$\big\|J^{+,+,-}_{k_1,k_2,k_3}\psi_k(\xi)\big\|_Z\lesssim 2^{-2p_0m}. \tag{4.22}$$

PROOF. The idea of proof is very similar to the proof of [20, lemma 6.4]. Because the setting of function spaces is changed, and also for self-completeness, we still give a detailed proof here.

We first do a change of variables to transform $(\xi-\eta,\eta-\sigma,\sigma)$ into $(\xi+\eta,\xi+\sigma,-\xi-\eta-\sigma)$. Hence, we have $(\eta,\sigma)=(0,0)$ at the critical point $(\xi,\xi,-\xi)$ (corresponding to the space-time resonance set). After the change of variables, we decompose $I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)$ as follows:

$$I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)=\sum_{l_1,l_2=\overline{l}}^{k+10} J_{l_1,l_2}(t,\xi),\quad \overline{l}=-(1-100p_0)m/2+3k/4, \tag{4.23}$$

where

$$J_{l_1,l_2}(t,\xi) = \int_{\mathbb{R}^2} e^{it\Phi(\xi,\eta,\sigma)}\widehat{f_{k_1}}(t,\xi+\eta)\widehat{f_{k_2}}(t,\xi+\sigma)\widehat{\overline{f}_{k_3}}(t,-\xi-\eta-\sigma),$$
$$\times\ c^*(\xi,\eta,\sigma)\psi_{l_1}^{\overline{l}}(\eta)\psi_{l_2}^{\overline{l}}(\sigma)d\eta\, d\sigma,$$

where $c^*(\cdot,\cdot,\cdot)$ is defined in (4.14); the phase $\Phi(\xi,\eta,\sigma)$ is defined as

$$\Phi(\xi,\eta,\sigma) := \Lambda(\xi) - \Lambda(\xi+\eta) - \Lambda(\xi+\sigma) + \Lambda(-\xi-\eta-\sigma),$$

and the cutoff function $\psi_{l_2}^{l_1}(\cdot)$ is defined as

$$\psi_{l_2}^{l_1}(\xi) = \begin{cases} \psi_{l_2}(\xi) & \text{if } l_2 > l_1, \\ \psi_{\le l_1}(\xi) & \text{if } l_2 = l_1. \end{cases}$$

Case 1. We first consider the case when $l_2 \ge \max\{l_1, \overline{l}+1\}$. For this case, we have

$$|\partial_\eta \Phi(\xi,\eta,\sigma)|\psi_{l_2}^{\overline{l}}(\sigma) = |\Lambda'(\xi+\eta+\sigma) - \Lambda'(\xi+\eta)|\psi_{l_2}^{\overline{l}}(\sigma) \gtrsim 2^{l_2}2^{-3k/2}. \tag{4.24}$$

After integrating by parts in η, we can derive the estimate

$$|J_{l_1,l_2}(t,\xi)| \lesssim 2^{-m}\big[\big|J^1_{l_1,l_2}(t,\xi)\big| + \big|J^2_{l_1,l_2}(t,\xi)\big|\big], \tag{4.25}$$

where

$$J^1_{l_1,l_2}(t,\xi) = \int_{\mathbb{R}^2} e^{it\Phi(\xi,\eta,\sigma)}\widehat{f_{k_1}}(t,\xi+\eta)\widehat{f_{k_2}}(t,\xi+\sigma)\widehat{\overline{f}_{k_3}}(t,-\xi-\eta-\sigma)\partial_\eta r_1(\xi,\eta,\sigma)d\eta\, d\sigma,$$
$$J^2_{l_1,l_2}(\xi,s) = \int_{\mathbb{R}^2} e^{it\Phi(\xi,\eta,\sigma)}\partial_\eta\big(\widehat{f_{k_1}}(t,\xi+\eta)\widehat{\overline{f}_{k_3}}(t,-\xi-\eta-\sigma)\big)\widehat{f_{k_2}}(t,\xi+\sigma)r_1(\xi,\eta,\sigma)d\eta\, d\sigma,$$

where

$$r_1(\xi,\eta,\sigma) := \frac{c^*(\xi,\eta,\sigma)\psi_{l_1}^{\overline{l}}(\eta)\psi_{l_2}^{\overline{l}}(\sigma)}{\partial_\eta\Phi(\xi,\eta,\sigma)}.$$

From Lemma 2.4, (4.13) in Lemma 4.5, and (4.24), the following estimate holds:

$$\begin{aligned} &\|r_1(\xi,\eta,\sigma)\psi_k(\xi)\|_{\mathcal{S}^\infty} \lesssim 2^{-l_2+4k}, \\ &\|r_1(\xi,\eta,\sigma)\psi_k(\xi)\|_{\mathcal{S}^\infty} \lesssim 2^{-l_2-l_1+4k} + 2^{-l_2+3k}. \end{aligned} \tag{4.26}$$

Therefore, from (4.26) and the L^2-L^2-L^∞-type trilinear estimate (2.8) in Lemma 2.3, the following estimates hold:

$$\begin{aligned} \big|J^1_{l_1,l_2}(t,\xi)\big| &\lesssim \big(2^{-l_2-l_1}2^{4k} + 2^{-l_2+3k}\big)\big\|\widehat{f_{k_1}}(t,\xi+\eta)\psi_{l_1}^{\overline{l}}(\eta)\big\|_{L^2} \\ &\quad\times \big\|\widehat{f_{k_2}}(t,\xi+\sigma)\psi_{l_2}^{\overline{l}}(\sigma)\big\|_{L^2}\big\|e^{-it\Lambda}P_{k_3}f\big\|_{L^\infty} \\ &\lesssim (2^{-l_2-l_1}2^{4k} + 2^{-l_2+3k})2^{(l_1+l_2)/2}\|\widehat{f}_k(t,\xi)\|^2_{L^\infty}2^{-m/2-N_2k_+\epsilon_1} \\ &\lesssim 2^{-50p_0m}\epsilon_1^3 2^{-6k_+}, \end{aligned}$$

$$\begin{aligned}\big|J^2_{l_1,l_2}(t,\xi)\big| &\lesssim \sum_{\{i,j\}=\{1,3\}} 2^{-l_2+4k}\|\partial_\xi \widehat{f}(\xi,s)\psi_{k_i}(\xi)\|_{L^2}\|P_{k_j}f\|_{L^2}\|e^{-it\Lambda}P_{k_2}f\|_{L^\infty}\\ &\lesssim 2^{-l_2-m/2+2p_0m}2^{(3-2p)k-(N_2+N_0-1)k_+}\epsilon_1^3 \lesssim 2^{-50p_0m}2^{-6k_+}\epsilon_1^3.\end{aligned}$$

To sum up, we have

$$|J_{l_1,l_2}(t,\xi)| \lesssim \epsilon_1^3 2^{-(1+50p_0)m}2^{-6k_+}. \tag{4.27}$$

The symmetric case when $l_1 \geq \max\{l_2,\overline{l}\}$ can be handled similarly.

Case 2. It remains to consider the case when $l_1 = l_2 = \overline{l}$. Note that the following estimate holds for this case:

$$\left|J_{\overline{l},\overline{l}}(\xi,s) + \frac{\widetilde{c}(\xi)\widehat{f_{k_1}}(t,\xi)\widehat{f_{k_2}}(t,\xi)\widehat{\overline{f_{k_3}}}(t,-\xi)}{t+1}\right| \lesssim |\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3|, \tag{4.28}$$

where

$$\begin{aligned}\mathcal{I}_1 &= \int_{\mathbb{R}^2} \big(e^{it\Phi(\xi,\eta,\sigma)} - e^{-it\eta\sigma/(4|\xi|^{3/2})}\big)\widehat{f_{k_1}}(t,\xi+\eta)\widehat{f_{k_2}}(t,\xi+\sigma)\\ &\qquad \times \widehat{\overline{f_{k_3}}}(t,-\xi-\eta-\sigma)c^*(\xi,\eta,\sigma)\psi^{\overline{l}}_{\overline{l}}(\eta)\psi^{\overline{l}}_{\overline{l}}(\sigma)d\eta\, d\sigma,\\ \mathcal{I}_2 &= \int_{\mathbb{R}^2} e^{-it\eta\sigma/(4|\xi|^{3/2})}\Big[\widehat{f_{k_1}}(t,\xi+\eta)\widehat{f_{k_2}}(t,\xi+\sigma)\widehat{\overline{f_{k_3}}}(t,-\xi-\eta-\sigma)c^*(\xi,\eta,\sigma)\\ &\qquad - \widehat{f_{k_1}}(t,\xi)\widehat{f_{k_2}}(t,\xi)\widehat{\overline{f_{k_3}}}(t,-\xi)c^*(\xi,0,0)\Big]\psi^{\overline{l}}_{\overline{l}}(\eta)\psi^{\overline{l}}_{\overline{l}}(\sigma)d\eta\, d\sigma,\\ \mathcal{I}_3 &= \int_{\mathbb{R}^2} e^{-it\eta\sigma/(4|\xi|^{3/2})}\widehat{f_{k_1}}(t,\xi)\widehat{f_{k_2}}(t,\xi)\widehat{\overline{f_{k_3}}}(t,-\xi)c^*(\xi,0,0)\psi^{\overline{l}}_{\overline{l}}(\eta)\psi^{\overline{l}}_{\overline{l}}(\sigma)d\eta\, d\sigma\\ &\qquad + \frac{\widetilde{c}(\xi)\widehat{f_{k_1}}(t,\xi)\widehat{f_{k_2}}(t,\xi)\widehat{\overline{f_{k_3}}}(t,-\xi)}{t+1}.\end{aligned}$$

Note that,

$$\left|\Phi(\xi,\eta,\sigma) + \frac{\eta\sigma}{4|\xi|^{3/2}}\right| \lesssim 2^{-5k/2}(|\eta|+|\sigma|)^3.$$

Hence, after using the estimate (4.13) in Lemma 4.5 and the size of the support of η and σ, the following estimate holds:

$$\begin{aligned}|\mathcal{I}_1| &\lesssim \epsilon_1^3 2^m 2^{5\overline{l}}2^{-9k/4-3N_2k_+}\\ &\lesssim \epsilon_1^3 2^{-6k_+-(3-1000p_0)m/2} \lesssim \epsilon_0 2^{-6k_+}2^{-(1+50p_0)m}.\end{aligned} \tag{4.29}$$

Note that

$$\begin{aligned}&\big|\widehat{f_l}(s,\xi+\rho) - \widehat{f_l}(s,\xi)\big|\\ &\quad\lesssim |\rho|^{1/2}\big\|\partial_\xi \widehat{f_l}(s)\big\|_{L^2}, |c^*(\xi,\eta,\sigma)-c^*(\xi,0,0)|\psi^{\overline{l}}_{\overline{l}}(\eta)\psi^{\overline{l}}_{\overline{l}}(\sigma)\\ &\quad\lesssim 2^{3k/2}2^{\overline{l}}.\end{aligned}$$

From above estimates, the following estimate holds after using the estimate (4.13) in Lemma 4.5 and the size of the support of η and σ:

$$
\begin{aligned}
|\mathcal{I}_2| &\lesssim 2^{-k_-/5}2^{5\bar{l}/2-2N_2k_++k}\epsilon_1^3 + 2^{3k/2-9k/4+3\bar{l}-3N_2k_+}\epsilon_1^3 \\
&\lesssim \epsilon_1^3 2^{-6k_+-(1+50p_0)m}.
\end{aligned}
\tag{4.30}
$$

We proceed to estimate $\mathcal{I}_3$. Note the fact (see also [21]) that

$$
\int_{\mathbb{R}\times\mathbb{R}} e^{-ixy}\widetilde{\psi}(x/N)\widetilde{\psi}(y/N)dx\,dy = 2\pi + \mathcal{O}(N^{-1/2}). \tag{4.31}
$$

Through scaling, the following estimate holds from (4.31),

$$
\left| \int_{\mathbb{R}^2} e^{-is\eta\sigma/(4|\xi|^{3/2})}\psi_{\bar{l}}^{\bar{l}}(\eta)\psi_{\bar{l}}^{\bar{l}}(\sigma)d\eta\,d\sigma - \frac{4|\xi|^{3/2}}{s}(2\pi)\right| \lesssim 2^{-(1+25p_0)m}2^{3k/2},
$$

which leads to the following estimate:

$$
|\mathcal{I}_3| \lesssim \epsilon_1^3 2^{7k/4-3N_2k_+}2^{-(1+25p_0)m} \lesssim \epsilon_1^3 2^{-6k_+}2^{-(1+25p_0)m}. \tag{4.32}
$$

To sum up, from (4.29), (4.30), and (4.32), the following estimate holds:

$$
(4.28) \lesssim \epsilon_1^3 2^{-(1+25p_0)m}2^{-6k_+},
$$

therefore finishing the proof. □

Now, it remains to consider the case when $k_i \notin [k-10, k+10]$ for some $i \in \{1,2,3\}$. Therefore, the term $\widehat{f_{k_1}}(t,\xi)\widehat{\overline{f_{k_3}}}(t,-\xi)\widehat{f_{k_2}}(t,\xi)\psi_k(\xi)$ vanishes. From now on, we can drop this term inside $J^{+,+,-}_{k_1,k_2,k_3}$.

LEMMA 4.11. *For any $(\iota_1,\iota_2,\iota_3) \in \mathcal{S}$, we have the following rough estimates*:

$$
\begin{aligned}
\left|I^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}(t,\xi)\right| &\lesssim 2^{\min\{k_i\}/4}2^{3\,\mathrm{med}\{k_i\}/4-N_2\,\mathrm{med}\{k_i\}_+} \\
&\quad\times 2^{5\max\{k_i\}/4-N_2\max\{k_i\}_+}\epsilon_1^3,
\end{aligned}
\tag{4.33}
$$

$$
\begin{aligned}
|I^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}(t,\xi)| &\lesssim 2^{\mathrm{med}\{k_i\}/2}2^{2\max\{k_i\}}\|e^{-it\Lambda}P_{\min\{k_i\}}f\|_{L^\infty} \\
&\quad\times \|P_{\mathrm{med}\{k_i\}}f\|_{L^2}\|P_{\max\{k_i\}}f\|_{L^2}.
\end{aligned}
\tag{4.34}
$$

PROOF. From (4.13) in Lemma 4.5, the desired estimate (4.33) follows straightforwardly after putting all inputs into the Z-normed space. From (4.13) in Lemma 4.5 and the L^2-L^∞-L^∞-type trilinear estimate (2.8) in Lemma 2.3, (4.34) holds straightforwardly. □

Recall that $2^{-21p_0m} \le |\xi| \le 2^{20p_0m}$. From (4.33) and (4.34), the following estimate holds when $\mathrm{med}\{k_i\} \le -(1+100p_0)m$ or $\min\{k_i\} \le -4(1+100p_0)m$ or $\max\{k_i\} \ge (1+100p_0)m/5$:

$$
\sup_{t\in[2^{m-1},2^{m+1}]} \left||\xi|^{3/4-p_0}(1+|\xi|^{N_2+2p_0})I^{\iota_1,\iota_2,\iota_3}_{k_1,k_2,k_3}(\xi,t)\right| \lesssim \epsilon_1^3 2^{-(1+p_0)m}, \tag{4.35}
$$

which is sufficient for most cases in our desired estimate (4.21). Therefore, it is sufficient to consider fixed k_1, k_2, and k_3 in the following range:

$$\begin{aligned} -(1+100p_0)m \le \mathrm{med}\{k_i\} \le \max\{k_i\} \le (1+100p_0)m/5, \\ \min\{k_i\} \ge -4(1+100p_0)m. \end{aligned} \tag{4.36}$$

LEMMA 4.12. *If k_1, k_2, k_3 satisfies the estimate* (4.36) *and one of the following two conditions is satisfied*:

$$\max\{|k_1-k|, |k_2-k|, |k_3-k|\} \ge 20, \quad \mathrm{med}\{k_i\} - \min\{k_i\} \le 10,$$
$$\max\{|k_1-k|, |k_2-k|, |k_3-k|\} \ge 20, \quad \min\{k_i\} \ge -(5/7-1000p_0)m,$$

then the following estimate holds:

$$\sup_{t\in[2^{m-1},2^{m+1}]} \big||\xi|^{3/4-p_0}(1+|\xi|)^{N_2+2p_0} I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)\big| \lesssim \epsilon_1^3 2^{-(1+2p_0)m}. \tag{4.37}$$

PROOF.

Case 1. We first consider the case when $\max\{|k_1-k_3|, |k_2-k_3|\} \ge 5$ and $\min\{k_i\} \ge -(5/7-1000p_0)m$. Note that

$$\partial_\eta \Phi = \Lambda'(\xi+\eta+\sigma) - \Lambda'(\xi+\eta), \quad \partial_\sigma \Phi(\xi,\eta,\sigma) = \Lambda'(\xi+\eta+\sigma) - \Lambda'(\xi+\sigma).$$

From the symmetry between inputs and without loss of generality, we assume that $|k_1-k_3| \ge 5$. Hence

$$|\partial_\eta \Phi(\xi,\eta,\sigma)| = |\Lambda'(\xi+\eta) - \Lambda'(\xi+\eta+\sigma)| \gtrsim 2^{-\min\{k_1,k_3\}/2}. \tag{4.38}$$

After integration by parts in η, we can derive the following estimate:

$$\big|I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)\big| \lesssim \big|\mathcal{J}^1_{k_1,k_2,k_3}\big| + \big|\mathcal{J}^2_{k_1,k_2,k_3}\big|,$$

$$\begin{aligned} \mathcal{J}^1_{k_1,k_2,k_3} &:= \frac{1}{t}\int_{\mathbb{R}^2} e^{it\Phi(\xi,\eta,\sigma)} \widehat{f^+_{k_1}}(t,\xi+\eta)\widehat{f^+_{k_2}}(t,\xi+\sigma) \\ &\qquad \times \widehat{f^-_{k_3}}(t,-\xi-\eta-\sigma)\partial_\eta r_2(\xi,\eta,\sigma)\,d\eta\,d\sigma, \\ \mathcal{J}^2_{k_1,k_2,k_3} &:= \frac{1}{t}\int_{\mathbb{R}^2} e^{it\Phi(\xi,\eta,\sigma)} \partial_\eta\big(\widehat{f^+_{k_1}}(t,\xi+\eta)\widehat{f^-_{k_3}}(t,-\xi-\eta-\sigma)\big) \\ &\qquad \times \widehat{f^+_{k_2}}(t,\xi+\sigma) r_2(\xi,\eta,\sigma)\,d\eta\,d\sigma, \end{aligned} \tag{4.39}$$

where

$$r_2(\xi,\eta,\sigma) = \frac{c^*(\xi,\eta,\sigma)\psi_{k_1}(\xi+\eta)\psi_{k_2}(\xi+\sigma)\psi_{k_3}(\xi+\eta+\sigma)}{\partial_\eta\Phi(\xi,\eta,\sigma)}.$$

From (4.38), (4.13) in Lemma 4.5, and Lemma 2.4, the following estimates hold:

$$\|r_2(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty} \lesssim 2^{\min\{k_1,k_3\}/2} 2^{\mathrm{med}\{k_1,k_2,k_3\}/2} 2^{2\max\{k_1,k_2,k_3\}}, \tag{4.40}$$

$$\|\partial_\eta r_2(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty} \lesssim 2^{-\min\{k_1,k_3\}/2} 2^{\mathrm{med}\{k_1,k_2,k_3\}/2} 2^{2\max\{k_1,k_2,k_3\}}. \tag{4.41}$$

Therefore, from (4.40), (4.41), and the L^2-L^2-L^∞-type trilinear estimate (2.8) in Lemma 2.3, the following estimates hold:

$$\begin{aligned}
&\big|\mathcal{J}^1_{k_1,k_2,k_3}\big| \\
&\quad\lesssim 2^{-\min\{k_i\}/2+\mathrm{med}\{k_i\}/2+2\max\{k_i\}} \\
&\qquad\times 2^{-m}\|e^{-is\Lambda}P_{\min}f\|_{L^\infty}\|P_{\mathrm{med}}f\|_{L^2}\|P_{\max}f\|_{L^2} \\
&\quad\lesssim \epsilon_1^3 2^{2\max\{k_i\}}2^{-(1+100p_0)m}2^{-4\max\{k_1,k_2,k_3\}_+}\lesssim \epsilon_1^3 2^{-(1+100p_0)m},
\end{aligned}$$

$$\begin{aligned}
&\big|\mathcal{J}^2_{k_1,k_2,k_3}\big| \\
&\quad\lesssim 2^{-m}2^{\min\{k_1,k_3\}/2}2^{\mathrm{med}\{k_1,k_2,k_3\}/2}2^{2\max\{k_1,k_2,k_3\}} \\
&\qquad\times\big(\|\partial_\xi\widehat{f}(\xi,s)\psi_{k_1}(\xi)\|_{L^2}\|P_{\max\{k_2,k_3\}}f\|_{L^2}\|e^{-is\Lambda}P_{\min\{k_2,k_3\}}f\|_{L^\infty} \\
&\qquad+\|\partial_\xi\widehat{f}(\xi,s)\psi_{k_3}(\xi)\|_{L^2}\|e^{-is\Lambda}P_{\min\{k_1,k_2\}}f\|_{L^\infty}\|P_{\max\{k_1,k_2\}}f\|_{L^2}\big) \\
&\quad\lesssim \epsilon_1^3 2^{-7\min\{k_i\}/10}2^{-3m/2}+\epsilon_1^3 2^{-3m/2}2^{3\max\{k_i\}_+/2} \\
&\quad\lesssim \epsilon_1^3 2^{-(1+100p_0)m}.
\end{aligned}\tag{4.42}$$

From the above estimates, (4.39), and the fact that $|\xi|\le 2^{20p_0m}$, the following estimate holds:

$$\big||\xi|^{3/4-p_0}(1+|\xi|)^{N_2+2p_0}I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)\big|\lesssim 2^{-(1+100p_0)m+80p_0m}\epsilon_1^3\lesssim 2^{-(1+2p_0)m}\epsilon_1^3,\tag{4.43}$$

hence finishing the proof. We remark that when $|k_2-k_3|\ge 5$ the argument is very similar as we can do integration by parts with respect to σ.

Case 2. Now we consider the case when $\max\{|k_1-k_3|,|k_2-k_3|\}\le 5$ and $\min\ge -(5/7-1000p_0)m$. Recalling that $\max\{|k_1-k|,|k_2-k|,|k_3-k|\}\ge 20$, we know that $\min\{k_1,k_2,k_3\}\ge k+10$. As a result, the following estimate holds:

$$\begin{aligned}
&|\xi+\eta|\sim|\xi+\sigma|\sim|\xi+\sigma+\eta|\sim 2^{k_1},\ |\xi|\sim 2^k\Longrightarrow|\eta|\sim|\sigma|\sim 2^{k_1}, \\
&|\partial_\eta\Phi(\xi,\eta,\sigma)|=|\Lambda'(\xi+\eta)-\Lambda'(\xi+\eta+\sigma)| \\
&\qquad\gtrsim 2^{-k_1/2}\sim 2^{-\min\{k_1,k_3\}/2}.
\end{aligned}\tag{4.44}$$

Note that estimates (4.44) and (4.38) are the same type. Hence the argument used in Case 1 can be applied to this case. We can use integration by parts in η to derive our desired estimate (4.43).

Case 3. For the case when $\mathrm{med}\{k_i\}-\min\{k_i\}\le 10$, we can use the same method used in Case 1 and Case 2. After integration by parts in η or σ, the loss of $2^{-\min\{k_i\}/2}$ can be covered by $2^{\mathrm{med}\{k_i\}/2}$ from the symbol. As a result, the following estimate holds:

$$\begin{aligned}
\big|I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)\big|&\lesssim \epsilon_1^3 2^{-\mathrm{med}\{k_i\}/5}2^{-3m/2}+\epsilon_1^3 2^{-3m/2}2^{3\max\{k_i\}_+/2} \\
&\lesssim \epsilon_1^3 2^{-(21-1000p_0)m/20}.
\end{aligned}$$

In the above estimate, we have used the fact that $\mathrm{med}\{k_i\} \geq -(1+100p_0)m$ and $\max\{k_i\} \leq (1+100p_0)m/5$; see (4.36). □

LEMMA 4.13. *If* k_1, k_2, k_3 *satisfies the estimate* (4.36), $\max\{|k_1-k|, |k_2-k|, |k_3-k|\} \geq 20$, $\min\{k_i\} < -(5/7-1000p_0)m$, *and* $\mathrm{med}\{k_i\} - \min\{k_i\} \geq 10$, *then the following estimate holds*:

$$\left| |\xi|^{3/4-2p_0}(1+|\xi|^{N_2+2p_0}) \int_{t_1}^{t_2} e^{iL(t,\xi)}\big[I^{+,+,-}_{k_1,k_2,k_3}(t,\xi)\big]dt \right| \lesssim 2^{-2p_0 m}\epsilon_1^3. \tag{4.45}$$

PROOF. Define $\widetilde{k} = \min\{k, \mathrm{med}\{k_i\}\}$. Recall that $-21p_0 m \leq k \leq 20p_0 m$ and $\min\{k_i\} < -(5/7-1000p_0)m$. Hence $\min\{k_1,k_2,k_3\}$ is much smaller than k. As $\mathrm{med}\{k_i\} - \min\{k_i\} \geq 10$, we can see that the smallest number among k, k_1, k_2, k_3, which is $\min\{k_i\}$, is much smaller than the second smallest number, which is $\widetilde{k}$. As mentioned before, a very important observation for this case is that the size of phase $\Phi(\xi,\eta,\sigma)$ is determined by the second smallest number instead of the smallest number. More precisely, the following estimate holds:

$$|\Phi(\xi,\eta,\sigma)| \gtrsim 2^{\widetilde{k}/2}. \tag{4.46}$$

To improve the presentation, we assume (4.46) is true first and postpone the proof of (4.46) to the end of this subsection.

To prove (4.45), we do integration by parts in time. As a result, we have

$$\left| \int_{t_1}^{t_2} e^{iL(\xi,s)}[I^{+,+,-}_{k_1,k_2,k_3;1}(\xi,s)]ds \right| \lesssim |B_1(t_1,\xi)| + |B_2(t_2,\xi)| + |T_1(\xi)| + |T_2(\xi)|,$$

where for $i \in \{1,2\}$,

$$\begin{aligned} B_i(t_i,\xi) = \int_{\mathbb{R}^2} & e^{it_i\Phi(\xi,\eta,\sigma)+iL(\xi,t_i)}\widehat{f_{k_1}}(\xi+\eta,t_i)\widehat{f_{k_2}}(\xi+\sigma,t_i) \\ & \times \widehat{\overline{f_{k_3}}}(-\xi-\eta-\sigma,t_i)r_3(\xi,\eta,\sigma)d\eta\, d\sigma, \end{aligned}$$

$$\begin{aligned} T_1(\xi) = \int_{t_1}^{t_2}\int_{\mathbb{R}^2} & e^{it\Phi(\xi,\eta,\sigma)+iL(t,\xi)}\partial_t L(t,\xi)\widehat{f_{k_1}}(t,\xi+\eta)\widehat{f_{k_2}}(t,\xi+\sigma) \\ & \times \widehat{\overline{f_{k_3}}}(t,-\xi-\eta-\sigma)r_3(\xi,\eta,\sigma)d\eta\, d\sigma\, dt, \end{aligned}$$

$$\begin{aligned} T_2(\xi) = \int_{t_1}^{t_2}\int_{\mathbb{R}^2} & e^{it\Phi(\xi,\eta,\sigma)+iL(t,\xi)} \\ & \times \partial_t\big(\widehat{f_{k_1}}(t,\xi+\eta)\widehat{f_{k_2}}(t,\xi+\sigma)\widehat{\overline{f_{k_3}}}(t,-\xi-\eta-\sigma)\big) \\ & \times r_3(\xi,\eta,\sigma)d\eta\, d\sigma\, dt, \end{aligned}$$

$$r_3(\xi,\eta,\sigma) = \frac{c^*(\xi,\eta,\sigma)\psi_{k_1}(\xi+\eta)\psi_{k_2}(\xi+\sigma)\psi_{k_3}(\xi+\eta+\sigma)}{\Phi(\xi,\eta,\sigma)}.$$

From (4.46), Lemma 2.4, and (4.13) in Lemma 4.5, the following estimate holds:

$$\|r_3(\xi,\eta,\sigma)\|_{\mathcal{S}^\infty} \lesssim 2^{\mathrm{med}\{k_i\}/2+2\max\{k_i\}-\widetilde{k}/2} \lesssim 2^{5\max\{k_i\}/2+11p_0 m}. \tag{4.47}$$

From (4.47), the L^2-L^2-L^∞-type trilinear estimate (2.8) in Lemma 2.3 and the fact that $\mathrm{med}\{k_i\} \geq -(1+100p_0)m$, the following estimate holds:

$$\begin{aligned}|B_i(t_i,\xi)| &\lesssim 2^{5\max\{k_i\}/2+11p_0 m}\left\|e^{-it_i\Lambda}P_{\min\{k_i\}}f\right\|_{L^\infty}\|P_{\mathrm{med}\{k_i\}}f\|_{L^2}\|P_{\max\{k_i\}}f\|_{L^2}\\ &\lesssim 2^{-\mathrm{med}\{k_i\}/5+23\max\{k_i\}/10-4\max\{k_i\}_+-m/2+11p_0 m}\epsilon_1^3\\ &\lesssim \epsilon_1^3 2^{-(3-1000p_0)m/10}.\end{aligned}$$

Recall (4.15). From (4.13) in Lemma 4.5, the following estimate holds:

$$\begin{aligned}|\partial_t L(t,\xi)\psi_k(\xi)| &\lesssim 2^{-m}|\xi|^4\|\widehat{f}(t,\xi)\psi_k(\xi)\|^2_{L^\infty_\xi}\\ &\lesssim 2^{-m+4k_+}\epsilon_1^2 \lesssim 2^{-(1-100p_0)m}\epsilon_0.\end{aligned} \tag{4.48}$$

From (4.48), (4.47), and the L^2-L^2-L^∞-type trilinear estimate (2.8) in Lemma 2.3, the following estimate holds:

$$\begin{aligned}|T_1(\xi)| &\lesssim \epsilon_0 2^{m-(1-111p_0)m+5\max\{k_i\}/2}\|e^{-it\Lambda}P_{\min}f\|_{L^\infty}\|P_{\mathrm{med}}f\|_{L^2}\|P_{\max}f\|_{L^2}\\ &\lesssim \epsilon_1^3 2^{111p_0 m-m/2-\mathrm{med}\{k_i\}/5} \lesssim 2^{-(3-1000p_0)m/10}\epsilon_1^3.\end{aligned}$$

Again, we used the fact that $\mathrm{med}\{k_i\} \geq -(1+100p_0)m$.

To estimate $T_2(\xi)$, we put the input $\partial_t \widehat{f^{\pm}_{k_i}}$ and the input with the smallest frequency in L^2 and the other input in L^∞. More precisely, from (4.47) and the L^2-L^2-L^∞-type trilinear estimate (2.8) in Lemma 2.3, the following estimate holds:

$$\begin{aligned}|T_2(\xi)| &\lesssim \sum_{\{l,m,n\}=\{1,2,3\}} 2^{m+5\max\{k_i\}/2+11p_0 m}\|\partial_t\widehat{f}(t,\xi)\psi_{k_l}(\xi)\|_{L^2}\\ &\qquad\times\|P_{\max\{k_m,k_n\}}f\|_{L^2}\|e^{-it\Lambda}P_{\min\{k_m,k_n\}}f\|_{L^\infty}\\ &\lesssim \epsilon_1^5 2^{-(3-1000p_0)m/10}.\end{aligned}$$

To sum up, our desired estimate (4.45) holds. □

It remains to prove (4.21) for the case when $(\iota_1,\iota_2,\iota_3) \in \{(+,+,+),(+,-,-),(-,-,-)\}$ and k_1, k_2, and k_3 satisfy (4.36). An important observation is that we have the following weak ellipticity estimate for phases $\Phi^{-,-,-}$, $\Phi^{+,-,-}$, $\Phi^{+,+,+}$:

$$\sum_{\substack{(\iota_1,\iota_2,\iota_3)\in\mathcal{S},\\(\iota_1,\iota_2,\iota_3)\neq(+,+,-)}} |\Phi^{\iota_1,\iota_2,\iota_3}(\xi_1,\xi_2,\xi_3)\psi_{k_1}(\xi_1)\psi_{k_2}(\xi_2)\psi_{k_3}(\xi_3)| \gtrsim 2^{\mathrm{med}\{k_i\}/2}. \tag{4.49}$$

Hence, after doing integration by parts in time, the loss of $2^{-\mathrm{med}\{k_i\}/2}$ can be covered by the size of the symbol; see (4.36) in Lemma 4.5. Similar to the argument we used in the proof of Lemma 4.13, with minor modifications one can show that the desired estimate (4.21) holds without any problems.

PROOFS OF (3.70), (4.46), AND (4.49). Without loss of generality, we assume that $|\xi| \geq |\xi-\eta| \geq |\eta-\sigma| \geq |\sigma|$. As a result, we have $|\eta| \leq 2|\xi-\eta|$ and $\xi\eta > 0$, which further gives us $|\eta| \leq 2|\xi|/3$ or $\eta = 2\xi$. Recall that $\Lambda(\xi) := |\xi|^{1/2}$. Note that the following estimate holds:

$$|\Lambda(\xi) - \Lambda(\xi-\eta)| = \left| \frac{|\xi| - |\xi-\eta|}{|\xi|^{1/2} + |\xi-\eta|^{1/2}} \right| \leq \frac{2}{3}|\eta|^{1/2}. \tag{4.50}$$

When $(\iota_1, \iota_2, \iota_3) \in \{(+,+,+), (+,-,-), (-,-,-)\}$, from (4.50) the following estimate holds:

$$\begin{aligned} &\left|\Phi^{\iota_1,\iota_2,\iota_3}(\xi,\eta,\sigma)\psi_{k_1}(\xi-\eta)\psi_{k_2}(\eta-\sigma)\psi_{k_3}(\sigma)\right| \\ &\quad \geq |\Lambda(\eta-\sigma)| + |\Lambda(\sigma)| - |\Lambda(\xi) - \Lambda(\xi-\eta)| \\ &\quad \gtrsim \max\{|\eta-\sigma|^{1/2}, |\sigma|^{1/2}\}. \end{aligned} \tag{4.51}$$

Therefore, our desired estimate (4.49) holds. To prove (4.46), we have another assumption, which is $|\sigma| \leq 2^{-10}|\eta-\sigma|$. As a result, we have $\Lambda(|\eta-\sigma|) - \Lambda(|\sigma|) \geq (1-2^{-5})|\eta-\sigma|^{1/2} \geq (1-2^{-5})^2|\eta|^{1/2}$. Hence, from (4.50), the following estimate holds:

$$\begin{aligned} &|\Phi^{+,+,-}(\xi,\eta,\sigma)\psi_{k_1}(\xi-\eta)\psi_{k_2}(\eta-\sigma)\psi_{k_3}(\sigma)| \\ &\quad \geq |\Lambda(\eta-\sigma)| - |\Lambda(\sigma)| - |\Lambda(\xi) - \Lambda(\xi-\eta)| \\ &\quad \gtrsim |\eta-\sigma|^{1/2} \sim \max\{|\eta-\sigma|^{1/2}, |\sigma|^{1/2}\}. \end{aligned} \tag{4.52}$$

We now finish the proof of (4.46); we prove (3.70). As $|\xi-\eta| \ll \min\{|\xi|, |\eta-\sigma|, |\sigma|\}$, without loss of generality we assume that $|\xi-\eta| \ll |\sigma| \leq |\eta-\sigma| \leq |\xi|$. The desired estimate (3.70) holds straightforwardly when $\mu\mu' = -$ or $\mu\nu' = -$. It remains to consider the case when $\mu\mu' = +$ and $\mu\nu' = +$. If $\mu\nu = +$, then the desired estimate (3.70) holds from (4.51). If $\mu\nu = -$, then the desired estimate (3.70) holds from (4.52), hence finishing the proof. □

4.3 *Z*-Norm Estimate of the Remainder Term

LEMMA 4.14. *Under the bootstrap assumptions* (2.29) *and* (4.17)*, the following estimate holds for* $k \in [-21p_0 m, 20p_0 m]$*,*

$$\sup_{t\in[2^{m-1},2^{m+1}]} \left\| \int_{t_1}^{t_2} e^{iL(t,\xi)+it\Lambda(\xi)} \widehat{R}(t,\xi)\psi_k(\xi)\,dt \right\|_Z \lesssim 2^{-(1+p_0)m}\epsilon_1^3. \tag{4.53}$$

PROOF. From (4.20) in Lemma 4.8, (A.7) in Lemma A.1, and (A.10) in Lemma A.2, we have the following estimate:

$$
\begin{aligned}
&\sup_{t_1,t_2\in[2^{m-1},2^{m+1}]}\Big\|\int_{t_1}^{t_2} e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi)\psi_k(\xi)\,dt\Big\|_{L^\infty_\xi}\\
&\lesssim \sup_{t\in[2^{m-1},2^{m+1}]} 2^{-k/2}\Big\|\int_{t_1}^{t_2} e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi)\psi_k(\xi)\,dt\Big\|_{L^2}^{1/2}\\
&\quad\times\Big[\Big\|\int_{t_1}^{t_2} t\partial_t\big(e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi)\big)\psi_k(\xi)\,dt\Big\|_{L^2}\\
&\qquad + 2^m\big(\|SP_k[R(t)]\|_{L^2}+\|P_kR(t)\|_{L^2}\big)\Big]^{1/2}\\
&\lesssim 2^{-m/2+220p_0m}\epsilon_0^2\\
&\quad + 2^{-m/4+20p_0m}\Big\|\int_{t_1}^{t_2} t\partial_t\big(e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi)\big)\psi_k(\xi)\,dt\Big\|_{L^2}^{1/2}.
\end{aligned}
\tag{4.54}
$$

Note that, in the above estimate, we used the fact that

$$\sup_{t\in[2^{m-1},2^{m+1}]}\|(t\partial_t-2\xi\cdot\partial_\xi)e^{iL(t,\xi)}\|_{L^\infty_\xi}\lesssim 2^{p_0m}\epsilon_1^2,$$

which follows from the definition of $L(t,\xi)$ in (4.15).

After integration by parts in time, from (A.7) in Lemma A.1, it is easy to see that the following estimate holds:

$$
\begin{aligned}
&\Big\|\int_{t_1}^{t_2} t\partial_t\big(e^{iL(t,\xi)+it\Lambda(\xi)}\widehat{R}(t,\xi)\big)\psi_k(\xi)\,dt\Big\|_{L^2}\\
&\lesssim \sup_{t\in[2^{m-1},2^{m+1}]} 2^m\|P_k(R)(t)\|_{L^2}\\
&\lesssim 2^{-m/2+200p_0m}\epsilon_0^2.
\end{aligned}
\tag{4.55}
$$

Combining (4.54) and (4.55), it is easy to see that our desired estimate (4.53) holds, hence finishing the proof. □

Appendix A Remainder Estimates

A.1 Explicit Formulas of Remainder Terms

In this subsection, we give all detailed formulas for all good remainder terms that are postponed in Section 3 and Section 4.

The detailed formulas of $\mathcal{R}_1$ and $\mathcal{R}_2$ in the system (2.17) are given as follows:

$$
\begin{aligned}
\mathcal{R}_1 &= |\nabla|^{1/2}T_\alpha U^2 - T_\alpha|\nabla|^{1/2}U^2 - T_{\sqrt{a}}T_V\partial_x h\\
&\quad + T_V\partial_x T_{\sqrt{a}}h + T_{\sqrt{a}}F(h)\psi,
\end{aligned}
\tag{A.1}
$$

$$\begin{aligned}\mathcal{R}_2 = {} & |\nabla|^{1/2}(T_a h) - T_{\sqrt{a}}|\nabla|^{1/2}T_{\sqrt{a}}h - |\nabla|^{1/2}T_V\partial_x\omega + T_V\partial_x|\nabla|^{1/2}\omega \\ & - T_{V\partial_x h}B + T_V T_{\partial_x h}B - T_V T_{\partial_x B}h + T_{V\partial_x B}h \\ & - \frac{1}{2}R_{\mathcal{B}}(V,V) - R_{\mathcal{B}}(V,\partial_x h) - \frac{1}{2}R_{\mathcal{B}}(B,B),\end{aligned} \tag{A.2}$$

where the bilinear operator $R_{\mathcal{B}}(\cdot,\cdot)$ is defined in (2.3), the Taylor coefficient a is defined in (2.11), $\alpha = \sqrt{a}-1$, and ω is the good unknown variable, which is defined in (2.12).

The detailed formulas of $\widetilde{\mathcal{R}}_1$ and $\widetilde{\mathcal{R}}_2$ inside the system (3.21) are given as follows:

$$\begin{aligned}\widetilde{\mathcal{R}}_1 = {} & \widetilde{Q}_1(\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2),U^2) + \widetilde{Q}_1(U^1,\widetilde{B}(U^1,U^2)) \\ & + 2\widetilde{A}_1(\mathcal{R}_1,U^1) + 2\widetilde{A}_2(\mathcal{R}_2,U^2) - \Lambda_2[\mathcal{R}_1](U^1,\widetilde{B}(U^1,U^2)) \\ & - \Lambda_2[\mathcal{R}_1](\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2),U^2) + \Lambda_{\geq 3}[\mathcal{R}_1],\end{aligned}$$

$$\begin{aligned}\widetilde{\mathcal{R}}_2 = {} & \widetilde{Q}_2\big(\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2),U^1\big) \\ & + \widetilde{Q}_2\big(U^1,\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2)\big) + \widetilde{Q}_3(U^2,\widetilde{B}(U^1,U^2)) \\ & + \widetilde{Q}_3(\widetilde{B}(U^1,U^2),U^2) + \widetilde{B}(\mathcal{R}_1,U^2) + \widetilde{B}(U^1,\mathcal{R}_2) + \Lambda_{\geq 3}[\mathcal{R}_2] \\ & - \Lambda_2[\mathcal{R}_2](\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2),U^1) \\ & - \Lambda_2[\mathcal{R}_2](U^1,\widetilde{A}_1(U^1,U^1) + \widetilde{A}_2(U^2,U^2)) \\ & - \Lambda_2[\mathcal{R}_2](\widetilde{B}(U^1,U^2),U^2) - \Lambda_2[\mathcal{R}_2](U^2,\widetilde{B}(U^1,U^2)).\end{aligned}$$

The detailed formulas of $\widetilde{\mathfrak{R}}_1$ and $\widetilde{\mathfrak{R}}_2$ in (3.49) and (3.50) are given as follows:

$$\begin{aligned}\widetilde{\mathfrak{R}}_1 = {} & \widetilde{C}_{1,\tau}\big(\mathcal{R}_1^S,U_2\big) + \widetilde{C}_{2,\tau}\big(\mathcal{R}_2^S,U^1\big) + \widetilde{C}_{1,\tau}\big(SU^1,\Lambda_{\geq 2}\big[\partial_t U^2\big]\big) \\ & + \widetilde{C}_{2,\tau}\big(SU^2,\Lambda_{\geq 2}\big[\partial_t U^1\big]\big) + \Lambda_{\geq 3}\big[\mathcal{R}_1^S\big] - \widetilde{\mathfrak{Q}}_1^S\big(W_{1,\tau} - SU^1,SU^2\big) \\ & - \widetilde{\mathfrak{Q}}_1^S\big(SU^1,W_{2,\tau} - SU^2\big),\end{aligned} \tag{A.3}$$

$$\begin{aligned}\widetilde{\mathfrak{R}}_2 = {} & \widetilde{C}_{3,\tau}\big(\mathcal{R}_1^S,U_1\big) + \widetilde{C}_{4,\tau}\big(\mathcal{R}_2^S,U^2\big) + \widetilde{C}_{3,\tau}\big(SU^1,\Lambda_{\geq 2}\big[\partial_t U^1\big]\big) \\ & + \widetilde{C}_{4,\tau}\big(SU^2,\Lambda_{\geq 2}\big[\partial_t U^2\big]\big) + \Lambda_{\geq 3}\big[\mathcal{R}_2^S\big] - \widetilde{\mathfrak{Q}}_3^S\big(W_{1,\tau} - SU^1,SU^2\big) \\ & - \widetilde{\mathfrak{Q}}_3^S\big(SU^1,W_{2,\tau} - SU^2\big),\end{aligned} \tag{A.4}$$

where the detailed formulas of $\mathcal{R}_1^S$ and $\mathcal{R}_2^S$ are given in (3.31) and (3.32).

At last, we give the detailed formula for the remainder term $R(t)$ in (4.5). Recall (2.18) and the fact that $R(t)$ is quartic and higher. The following identity holds:

$$\begin{aligned}R(t) = {} & \Lambda_{\geq 4}[C_1 + \mathcal{R}_1] + i\Lambda_{\geq 4}[C_2 + \mathcal{R}_2] + 2A_1(C_1 + \Lambda_{\geq 3}[\mathcal{R}_1],U^1) \\ & + 2A_2(C_2 + \Lambda_{\geq 3}[\mathcal{R}_2],U^2) + iB(C_1 + \Lambda_{\geq 3}[\mathcal{R}_1],U^2) \\ & + iB(U^1,C_2 + \Lambda_{\geq 3}[\mathcal{R}_2]).\end{aligned} \tag{A.5}$$

A.2 L^2-Type Estimate of Remainder Terms

From the detailed formulas of remainder terms in Section A.1, we know that it would be sufficient to estimate the L^2-type norms of remainder terms if we have the necessary ingredients, which are the following:

(i) L^2- and L^∞-type estimates of the nontrivial part of Dirichlet-Neumann operator $B(h)\psi$ and $SB(h)\psi$,

(ii) L^2-type estimate of the remainder term of paralinearization $F(h)\psi$ and $SF(h)\psi$,

(iii) L^2- and L^∞-type estimates of the Taylor coefficient a, Sa, $\partial_t a$, and $S\partial_t a$.

Those necessary ingredients will be discussed in detail in Appendix B and Appendix C.

LEMMA A.1. *Under the bootstrap assumption* (2.29)*, we have the following estimates for the remainder terms*:

$$\begin{aligned}\sup_{t\in[0,T]}(1+t)^{1-p_0}\big(&\|\Lambda_{\geq 3}[\mathcal{R}_1]\|_{H^{N_0,p}}+\|\Lambda_{\geq 3}[\mathcal{R}_2]\|_{H^{N_0,p}}\\&+\|\widetilde{\mathcal{R}}_1\|_{H^{N_0,p}}+\|\widetilde{\mathcal{R}}_2\|_{H^{N_0,p}}\big)\lesssim\epsilon_0^2,\end{aligned}\tag{A.6}$$

$$\sup_{t\in[2^{m-1},2^m]}2^{3m/2-200p_0m}\|R(t)\|_{L^2}\lesssim\epsilon_0^2.\tag{A.7}$$

PROOF. Recall the detailed formulas of $\mathcal{R}_1$ and $\mathcal{R}_2$ in (A.1) and (A.2). From (C.2) in Lemma C.1, the estimates in Lemma C.7, and the estimates in Lemma B.3, the following estimate holds for $i\in\{2,3\}$:

$$\begin{aligned}&\|\Lambda_{\geq i}[\mathcal{R}_1]\|_{H^{N_0,p}}+\|\Lambda_{\geq i}[\mathcal{R}_2]\|_{H^{N_0,p}}\\&\quad\lesssim\|\Lambda_{\geq i}[F(h)\psi]\|_{H^{N_0}}+\|(B,\partial_x h,V)\|_{H^{N_0-1}}\|(V,\partial_x h,B)\|_{\widetilde{W}^1}^{i-1}\\&\qquad+\|h\|_{H^{N_0,p}}(\|\Lambda_{\geq i-1}\partial_t\alpha\|_{\widetilde{W}^0}+\|\alpha\|_{\widetilde{W}^1}\|V\|_{\widetilde{W}^0})\\&\qquad+\|\Lambda_{\geq 2}[(B,V)]\|_{H^{N_0-1}}\|(B,V)\|_{\widetilde{W}^1}+\|h\|_{H^{N_0,p}}\|\alpha\|_{\widetilde{W}^0}^2\\&\quad\lesssim\big(\|h\|_{H^{N_0,p}}+\|\partial_x\psi\|_{H^{N_0-1}}\big)\|(\partial_x h,\partial_x\psi)\|_{W^1}^{i-1}\lesssim(1+t)^{-1+2p_0}\epsilon_0^2.\end{aligned}\tag{A.8}$$

From (3.16), (3.20), Lemma 2.3, (A.8), and the estimates in Lemma B.3, the following estimate holds:

$$\begin{aligned}&\|\Lambda_{\geq 3}[\widetilde{\mathcal{R}}_1]\|_{H^{N_0,p}}+\|\Lambda_{\geq 3}[\widetilde{\mathcal{R}}_2]\|_{H^{N_0,p}}\\&\quad\lesssim\|(U^1,U^2)\|_{H^{N_0,p}}\|(U^1,U^2)\|_{W^3}^2+\|(U^1,U^2)\|_{H^{N_0,p}}\|(\mathcal{R}_1,\mathcal{R}_2)\|_{\widetilde{W}^1}\\&\qquad+\|(U^1,U^2)\|_{W^1}\|(\mathcal{R}_1,\mathcal{R}_2)\|_{H^{N_0,p}}\\&\quad\lesssim\big(\|(U^1,U^2,h)\|_{H^{N_0,p}}+\||\nabla|\psi\|_{H^{N_0-1}}\big)\big(\|(\partial_x h,\partial_x\psi)\|_{W^2}+\|(U^1,U^2)\|_{W^3}\big)\\&\quad\lesssim(1+t)^{-1+2p_0}\epsilon_0^2.\end{aligned}$$

In the above estimate, we used the following rough estimate on the good error terms:

$$\|\mathcal{R}_1\|_{\widetilde{W}^1}+\|\mathcal{R}_2\|_{\widetilde{W}^1}\lesssim\|(\partial_x h,\partial_x\psi)\|_{W^2}^2,$$

which is derived from estimating each term inside (A.1) and (A.2).

To prove (A.7), we only need to estimate the $\dot{H}^p$-norm for each term inside (A.7), as we do not worry about losing derivatives for the desired estimate (A.7). Recall (A.1), (A.2), and (A.5). From (B.18) in Lemma B.4, (C.37), and (C.40) in Lemma C.7, our desired estimate (A.7) follows straightforwardly as it is quartic and higher. □

LEMMA A.2. *Under the bootstrap assumption* (2.29)*, the following estimates hold for the remainder terms*:

$$\begin{aligned}\sup_{t\in[0,T]}(1+t)^{1-p_0}\big(\|\Lambda_{\geq 3}[\mathcal{R}_1^S]\|_{H^{N_1,p}}+\|\Lambda_{\geq 3}[\mathcal{R}_2^S]\|_{H^{N_1,p}}\\ +\|\widetilde{\mathfrak{R}}_1\|_{H^{N_1,p}}+\|\widetilde{\mathfrak{R}}_2\|_{H^{N_1,p}}\big)\lesssim\epsilon_0^2.\end{aligned}\tag{A.9}$$

$$\sup_{t\in[2^{m-1},2^m]}\ \sup_{k\in[2^{-100p_0m},2^{100p_0m}]}2^{3m/2-200p_0m}2^{pk}\|P_k(SR(t))\|_{L^2}\lesssim\epsilon_0^2.\tag{A.10}$$

PROOF. Recall (3.31) and (3.32). From the L^2-L^∞-type bilinear estimate, we can derive the following estimate for $i\in\{2,3\}$:

$$\begin{aligned}&\|\Lambda_{\geq i}[\mathcal{R}_1^S]\|_{H^{N_1,p}}+\|\Lambda_{\geq i}[\mathcal{R}_2^S]\|_{H^{N_1,p}}\\ &\quad\lesssim\|\Lambda_{\geq i}[SF(\eta)\psi]\|_{H^{N_1,p}}+\|S\alpha\|_{H^{-1}}\|(\partial_x h,\partial_x\psi)\|_{W^1}^2\\ &\qquad+\|Sh\|_{H^{N_1,p}}(\|\Lambda_{\geq i-1}[\partial_t\alpha]\|_{\widetilde{W}^0}+\|\partial_x\alpha\|_{\widetilde{W}^0}\|V\|_{\widetilde{W}^1})\\ &\qquad+\|h\|_{W^{2+N_1}}\|\Lambda_{\geq i-1}S\partial_t\alpha\|_{H^{-2}}\\ &\qquad+\sum_{\substack{i_1+i_2=i,\\ i_1,i_2\in\mathbb{N}_+}}\|\Lambda_{\geq i_1}[(SB,SV)]\|_{L^2}\|\Lambda_{\geq i_2}[(B,V)]\|_{\widetilde{W}^2}\\ &\qquad+\|\partial_x h\|_{\widetilde{W}^1}\|(SB,SV,S\partial_x h)\|_{L^2}\|(B,V)\|_{\widetilde{W}^1}\\ &\qquad+\|(\partial_x h,\partial_x\psi)\|_{H^{N_1+3}}\|(\partial_x h,\partial_x\psi)\|_{W^2}^2.\end{aligned}\tag{A.11}$$

From (C.37), (C.38), (C.39), and (C.40) in Lemma C.7, (C.26) in Lemma C.6, and (C.21) in Lemma C.25, we have the estimate

$$\begin{aligned}\text{(A.11)}&\lesssim(\|Sh\|_{H^{N_1,p}}+\||\nabla|S\psi\|_{L^2})\|(\partial_x h,\partial_x\psi)\|_{W^2}^{i-1}\\ &\lesssim(1+t)^{-(i-1)/2+p_0}\epsilon_0^2.\end{aligned}\tag{A.12}$$

Recall (A.3) and (A.4). From (3.48) in Lemma 3.4, (3.37), (3.38), and (A.12), the following estimate holds:

$$\begin{aligned}&\|\widetilde{\mathfrak{R}}_1\|_{H^{N_1,p}}+\|\widetilde{\mathfrak{R}}_2\|_{H^{N_1,p}}\\ &\quad\lesssim(1+t)^{-1+p_0}+(1+t)^{-1/2+p_0}\|(U^1,U^2)\|_{W^3}\\ &\qquad+\|(SU^1,SU^2)\|_{H^{N_1,p}}\big(\|\Lambda_{\geq 2}[(\partial_t U^1,\partial_t U^2)]\|_{\widetilde{2}}+\|(U^1,U^2)\|_{W^3}^2\big)\\ &\qquad+\|(U^1,U^2)\|_{H^{N_0,p}}\|(U^1,U^2)\|_{W^3}^2\lesssim(1+t)^{-1+p_0}\epsilon_0.\end{aligned}$$

Note that losing derivatives is not an issue for the desired estimate (A.10). From (B.18) in Lemma B.4, (C.37) and (C.40) in Lemma C.7, (C.22), and (C.21) in Lemma C.5, our desired estimate (A.10) follows straightforwardly as it is quartic and higher. □

Appendix B Properties of the Dirichlet-Neumann Operator

B.1 A Fixed-Point-Type Formulation of the Dirichlet-Neumann Operator

The idea of analyzing the Dirichlet-Neumann operator is based on the work of Alazard-Delort [3]. A key observation is that there exists a fixed-point-type structure inside the velocity potential, which provides a good way to estimate the Dirichlet-Neumann operator in the small-data regime. From the bootstrap assumption (2.29), the following estimate holds:

$$\sup_{t\in[0,T]} \|h(t)\|_{W^{N_2}} + \|h\|^{1/2}_{H^{N_0,p}} \|h\|^{1/2}_{W^{N_2}} \lesssim \epsilon_1. \tag{B.1}$$

We transform the domain $\Omega(t) := \{(x,y) : y \le h(t,x), x \in \mathbb{R}\}$ to the negative half-space through a change of coordinates as follows: $(x,y) \longmapsto (x,z) := (x, y - h(t,x))$, $z \in (-\infty, 0]$. Define $\varphi(x,z) = \phi(x, z + h(t,x))$. Hence $\phi(x,y) = \varphi(x, y - h(t,x))$. It is easy to verify that the following identities hold:

$$\begin{gathered} \partial_y^2 \phi(x, z + h(t,x)) = \partial_z^2 \varphi(x,z), \\ \partial_x^2 \phi = \partial_x^2 \varphi - 2\partial_x \partial_z \varphi \partial_x h - \partial_z \varphi \partial_x^2 h + \partial_z^2 \varphi (\partial_x h)^2. \end{gathered} \tag{B.2}$$

Recall (1.2). We know that ϕ satisfies a Laplace equation inside $\Omega(t)$. Therefore, from (1.2) and (B.2), the following equalities hold:

$$P\varphi := \big[(1 + (\partial_x h)^2)\partial_z^2 + \partial_x^2 - 2\partial_x h \partial_x \partial_z - \partial_x^2 h \partial_z\big]\varphi = 0, \quad \varphi\big|_{z=0} = \psi. \tag{B.3}$$

Recall (2.10). In the (x,z)-coordinate system, we have

$$B(h)\psi = \partial_z \varphi\big|_{z=0}, \quad G(h)\psi = (1 + (\partial_x h)^2)\partial_z \varphi\big|_{z=0} - \partial_x h \partial_x \psi. \tag{B.4}$$

From the above equalities, one can see that the only nontrivial term inside $G(h)\psi$ is $B(h)\psi$. Therefore, to estimate the Dirichlet-Neumann operator in a X-normed space, it is sufficient to estimate $\partial_z \varphi(z)$ in the $L^\infty_z X$-normed space, where $z \in (-\infty, 0]$.

LEMMA B.1. *Let ψ be in the space $\dot{H}^{1/2+p} \cap L^\infty$ and $h(x) \in W^\gamma(\mathbb{R})$ with $\gamma > 2$: then $\varphi(z,x)$ satisfies the following fixed-point-type formulation*:

$$\begin{aligned} \varphi(z) &= e^{z|\nabla|}\psi \\ &\quad + \frac{1}{2}\int_{-\infty}^{0} e^{(z+z')|\nabla|}\big[\partial_x |\nabla|^{-1}(\partial_x h \partial_z \varphi) - \partial_x h \partial_x \varphi - (\partial_x h)^2 \partial_z \varphi\big] dz' \\ &\quad + \frac{1}{2}\int_{-\infty}^{0} e^{-|z-z'||\nabla|}\big[-\partial_x |\nabla|^{-1}(\partial_x h \partial_z \varphi) + \operatorname{sign}(z - z')(\partial_x h \partial_x \varphi \\ &\qquad\qquad - (\partial_x h)^2 \partial_z \varphi)\big] dz', \end{aligned} \tag{B.5}$$

if $\varphi(\cdot,\cdot)$ satisfies $P\varphi = 0$ and $\varphi|_{z=0} = \psi$.

PROOF. One can refer to [3][lemma 1.1.5] for the detailed proof. To be concise, we only sketch a proof here.

We rewrite (B.3) as follows:

$$(\partial_z + |\nabla|)(\partial_z - |\nabla|)\varphi = g(z) = \partial_z g_1(z) + \partial_x g_2(z), \tag{B.6}$$

where

$$g_1(z) = -(\partial_x h)^2 \partial_z \varphi + \partial_x h \partial_x \varphi, \quad g_2(z) = \partial_x h \partial_z \varphi.$$

By treating the nonlinearity $g(z)$ as given, we can solve φ explicitly from (B.6). Note that $g(z)$ has a term of type $\partial_z^2 \varphi$, which is not in terms of $\nabla_{x,z}\varphi$. Hence, we first decompose it as $\partial_z g_1(z) + \partial_x g_2(z)$, where $g_1(z)$ and $g_2(z)$ linearly depend on $\nabla_{x,z}\varphi$. As a result, after doing integration by parts in z once for $\partial_z g_1(z)$, we can derive (B.5). □

From (B.5), we can derive the following fixed-point formulation for $\nabla_{x,z}\varphi$:

$$\begin{aligned}\nabla_{x,z}\varphi(z) &= e^{z|\nabla|}[\partial_x \psi, |D_x|\psi] \\ &\quad + \int_{-\infty}^{0} K(z,z')M(\partial_x h)\nabla_{x,z}\varphi(z')dz' + [0, \partial_x h \partial_x \varphi - (\partial_x h)^2 \partial_z \varphi],\end{aligned} \tag{B.7}$$

where $K(z,z')$ and $M(\eta')$ are the matrices of operators as follows:

$$\begin{aligned}K(z,z') &= \frac{1}{2} e^{(z+z')|\nabla|} \begin{bmatrix} \partial_x & \partial_x \\ |\nabla| & |\nabla| \end{bmatrix} \\ &\quad + \frac{1}{2} e^{-|z-z'||D_x|} \begin{bmatrix} -\partial_x & -(\text{sign}(z-z'))\partial_x \\ (\text{sign}(z-z'))|\nabla| & |\nabla| \end{bmatrix},\end{aligned} \tag{B.8}$$

$$M(\partial_x h) = \begin{bmatrix} 0 & \partial_x |D_x|^{-1} \circ \partial_x h \\ -\partial_x h & (\partial_x h)^2 \end{bmatrix}. \tag{B.9}$$

Now, it is easy to see that there exists a fixed-point-type structure for $\nabla_{x,z}\varphi$ in (B.7). We remark that (B.7) is the starting point of the L^2-type and L^∞-type estimates of the Dirichlet-Neumann operator.

To carry out a fixed-point-type argument, some basic estimates of the operator $K(z,z')$ are necessary. As a result, we have the following lemma:

LEMMA B.2. *For a well-defined vector function $g : (-\infty, 0] \times \mathbb{R} \longmapsto \mathbb{R}^2$, any $j \in \mathbb{Z}$, and p_1, q_1, p_2, and q_2 such that $2 \le q_1 \le p_1 \le +\infty$ and $1 \le q_2 \le p_2 \le +\infty$, we have*

$$\begin{aligned}\left\| \int_{-\infty}^{0} K(z,z') P_j[g(z')]dz' \right\|_{L_z^{p_2} L_x^{p_1}} &\lesssim \\ &2^{j(1/q_1 + 1/q_2 - 1/p_1 - 1/p_2)} \|P_j g\|_{L_z^{q_2} L_x^{q_1}}.\end{aligned} \tag{B.10}$$

PROOF. From the explicit formula of operator $K(z,z')$ in (B.8), it is easy to find that the kernel of $K(z,z') \circ P_j$ is given by

$$\text{(B.11)} \quad \begin{aligned} &\widetilde{K}_j(z,z',x) := \\ &\quad \int_{\mathbb{R}} \frac{1}{2} e^{ix\xi+(z+z')|\xi|} m_1(\xi)\psi_j(\xi) + \frac{1}{2} e^{ix\xi-|z-z'||\xi|} m_2(\xi)\psi_j(\xi)\, d\xi, \end{aligned}$$

where the matrices $m_1(\xi)$ and $m_2(\xi)$ are

$$m_1(\xi) = \begin{bmatrix} -i\xi & -i\xi \\ |\xi| & |\xi| \end{bmatrix}, \quad m_2(\xi) = \begin{bmatrix} i\xi & i\,\mathrm{sign}(z-z')\xi \\ \mathrm{sign}(z-z')|\xi| & |\xi| \end{bmatrix}.$$

We can do integration by parts in ξ once and gain $(|x| + |z-z'|)^{-1}$ with the price of 2^{-j}. Hence, we have the pointwise estimate for the kernel:

$$\text{(B.12)} \qquad |\widetilde{K}_j(z,z',x)| \lesssim_N 2^{2j}(1 + 2^j|x| + 2^j|z \pm z'|)^{-N}.$$

With the above estimate on the kernel, it is not difficult to see that our desired estimate (B.10) holds after using the Hölder inequality and Bernstein inequality. □

LEMMA B.3. *Under the assumption* (B.1)*, the following estimates hold for* $\gamma \le \gamma'$:

$$\text{(B.13)} \quad \begin{aligned} \|\nabla_{x,z}\varphi\|_{L_z^\infty \widetilde{W^\gamma}} &\lesssim \|(\partial_x\psi)\|_{W^\gamma}, \\ \|\nabla_{x,z}\varphi\|_{L_z^\infty H^k} &\lesssim \||\nabla|\psi\|_{H^k} + \|\partial_x h\|_{H^k}\|\partial_x\psi\|_{W^0}, \\ \|\Lambda_{\ge 2}[\nabla_{x,z}\varphi]\|_{L_z^\infty \widetilde{W^\gamma}} &\lesssim \|\partial_x h\|_{\widetilde{W^\gamma}}\|(\partial_x\psi)\|_{W^\gamma}, \\ \|\Lambda_{\ge 2}[\nabla_{x,z}\varphi]\|_{L_z^\infty H^k} &\lesssim \|\partial_x h\|_{\widetilde{W^0}}\||\nabla|\psi\|_{H^k} + \|\partial_x\psi\|_{W^0}\|\partial_x h\|_{H^k}, \\ \|\Lambda_{\ge 3}[\nabla_{x,z}\varphi]\|_{L_z^\infty H^k} &\lesssim \|\partial_x h\|_{\widetilde{W^0}}^2\||\nabla|\psi\|_{H^k} \\ &\quad + \|\partial_x h\|_{\widetilde{W^0}}\|\partial_x\psi\|_{W^0}\|\partial_x h\|_{H^k}. \end{aligned}$$

PROOF. From (B.7) and (B.10) in Lemma B.2, we have

$$\text{(B.14)} \quad \begin{aligned} \|\nabla_{x,z}\varphi\|_{L_z^\infty H^k} &\lesssim \||\nabla|\psi\|_{H^k} + +\||[M(\partial_x h)\nabla_{x,z}\varphi]\|_{L_z^\infty H^k} \\ &\lesssim \||\nabla|\psi\|_{H^k} + \|\partial_x h\|_{\widetilde{W^0}}\|\nabla_{x,z}\varphi\|_{L_z^\infty H^k} \\ &\quad + \|\nabla_{x,z}\varphi\|_{L_z^\infty \widetilde{W^0}}\|\partial_x h\|_{H^k}. \end{aligned}$$

Again, from (B.7) and (B.10) we have

$$\begin{aligned}\|\nabla_{x,z}\varphi\|_{L^\infty_z \widetilde{W^\gamma}} &\lesssim \|e^{z|\nabla|}[\partial_x\psi, |\nabla|\psi]\|_{L^\infty_z \widetilde{W^\gamma}} \\ &\quad + \left\| \int_{-\infty}^0 K(z,z')M(\partial_x h)(\nabla_{x,z}\varphi(z'))\,dz' \right\|_{L^\infty_z W^\gamma} \\ &\quad + \sum_{l=1,2} \|(\partial_x h)^l \nabla_{x,z}\varphi\|_{L^\infty_z \widetilde{W^\gamma}} \\ &\lesssim \|\partial_x\psi\|_{W^\gamma} + \big[\|\partial_x h\|_{L^2}^{2/q}\|\partial_x h\|_{L^\infty}^{1-2/q} + \|\partial_x h\|_{\widetilde{W^\gamma}}\big] \\ &\quad \times \|\nabla_{x,z}\varphi\|_{L^\infty_z \widetilde{W^\gamma}},\end{aligned}$$

which further gives us the following estimate from the smallness assumption (B.1):

$$\|\nabla_{x,z}\varphi\|_{L^\infty_z \widetilde{W^\gamma}} \lesssim \|\partial_x\psi\|_{W^\gamma}. \tag{B.15}$$

Combining estimates (B.15) and (B.14), we have

$$\begin{aligned}\|\nabla_{x,z}\varphi\|_{L^\infty_z H^k} &\lesssim \||\nabla|\psi\|_{H^k} + \|\partial_x h\|_{\widetilde{W^0}}\|\nabla_{x,z}\varphi\|_{L^\infty_z H^k} \\ &\quad + \|\partial_x h\|_{H^k}\|\partial_x\psi\|_{W^0},\end{aligned} \tag{B.16}$$

which further gives us the following estimate from the smallness assumption (B.1),

$$\|\nabla_{x,z}\varphi\|_{L^\infty_z H^k} \lesssim \||\nabla|\psi\|_{H^k} + \|\partial_x h\|_{H^k}\|\partial_x\psi\|_{W^0}. \tag{B.17}$$

From (B.15) and (B.17), we can see that our desired estimate (B.13) holds. From (B.7), one can derive a fixed-point-type formulation for $\Lambda_{\geq i}[\nabla_{x,z}\varphi]$, $i \in \{2,3\}$. Hence, all other estimates can be derived similarly. We omit details here. □

By the same fixed-point-type argument, we can also derive the following lemma, which is very helpful when the scaling vector field is applied to the Dirichlet-Neumann operator. Note that, unlike the estimates in Lemma B.3, in Lemma B.4 only ψ is put into L^2-type spaces.

LEMMA B.4. *Under the smallness assumption* (B.1)*, the following estimates hold for* $i \in \{1,2,3\}$ *and* $k \leq \gamma' - 1$:

$$\begin{aligned}\|\Lambda_{\geq i}[\nabla_{x,z}\varphi]\|_{L^\infty_z H^k} &\lesssim \|\partial_x h\|_{W^k}^{i-1}\||\nabla|\psi\|_{H^k}, \\ \|\Lambda_{\geq 4}[\nabla_{x,z}\psi]\|_{L^\infty_z L^2} &\lesssim \|\partial_x h\|_{W^0}^3\||\nabla|\psi\|_{L^2}.\end{aligned} \tag{B.18}$$

The following lemma is very helpful if one wants to estimate a term of the type $\psi_2 B(h)\psi_1 - B(h)(\psi_1\psi_2)$. This type of term will appear when the scaling vector field hits the Dirichlet-Neumann operator.

LEMMA B.5. *Given any two well-defined smooth functions* $f(x)$ *and* $g(x)$*, if* $\varphi_1(x,z)$ *satisfies* $P\varphi_1 = 0$ *and* $\varphi_1\big|_{z=0} = fg$*, and meanwhile* $\varphi_2(x,z)$ *satisfies* $P\varphi_2 = 0$ *and* $\varphi_2\big|_{z=0} = g(x)$*, then under the smallness assumption that* $\|h\|_{\widetilde{W^{\gamma+1}}} \leq \delta < 1$*, we have the following estimates for* $k \leq \gamma$*,*

$$\|\nabla_{x,z}\varphi_1(z) - f\nabla_{x,z}\varphi_2(z)\|_{L^\infty_z H^k} \lesssim \|g\|_{\widetilde{W^k}}\|f\|_{H^{k+1,1-\epsilon}},$$

$$\|\Lambda_{\geq 3}[\nabla_{x,z}\varphi_1 - f\nabla_{x,z}\varphi_2]\|_{L_z^\infty H^k} \lesssim \|h\|_{\widetilde{W^{k+1}}}\|g\|_{\widetilde{W^k}}\|f\|_{H^{k+1,1-\epsilon}}, \tag{B.19}$$

$$\|\Lambda_{\geq 4}[\nabla_{x,z}\varphi_1 - f\nabla_{x,z}\varphi_2]\|_{L_z^\infty L^2} \lesssim \|h\|^2_{\widetilde{W^1}}\|g\|_{\widetilde{W^0}}\|f\|_{H^{1,1-\epsilon}}. \tag{B.20}$$

PROOF. Recall (B.7). We have

$$\begin{aligned}\nabla_{x,z}\varphi_1(z) = {} & e^{z|\nabla|}[\partial_x(fg), |\nabla|(fg)] + \int_{-\infty}^0 K(z,z')M(\partial_x h)\nabla_{x,z}\varphi_1(z')dz' \\ & + \big[0, \partial_x h\partial_x\varphi_1 - (\partial_x h)^2\partial_z\varphi_1\big],\end{aligned}$$

$$\begin{aligned}f\nabla_{x,z}\varphi_2(z) = {} & fe^{z|\nabla|}[\partial_x g, |\nabla| g] + \int_{-\infty}^0 fK(z,z')M(\partial_x h)\nabla_{x,z}\varphi_2(z')dz' \\ & + \big[0, \partial_x hf\partial_x\varphi_2 - (\partial_x h)^2 f\partial_z\varphi_2\big].\end{aligned}$$

Note that

$$\begin{aligned}&\mathcal{F}\big[e^{z|\nabla|}[\partial_x(fg), |\nabla|(fg)] - f(\cdot)e^{z|\nabla|}[\partial_x g, |\nabla| g]\big](z,\xi) = \\ &\qquad \int_{\mathbb{R}} e^{z|\xi|}a(\xi)\widehat{f}(\xi-\eta)\widehat{g}(\eta)\,d\eta - \int_{\mathbb{R}}\widehat{f}(\xi-\eta)\widehat{g}(\eta)e^{z|\eta|}a(\eta)d\eta,\end{aligned}$$

where $a(\xi) = [i\xi, |\xi|]$. Note that the following estimate holds for any fixed $z \leq 0$:

$$|e^{z|\xi|}a(\xi) - e^{z|\eta|}a(\eta)| \lesssim \big||\xi| - |\eta|\big|.$$

Hence

$$\begin{aligned}&\|e^{z|\nabla|}[\partial_x(fg), |\nabla|(fg)] - f(\cdot)e^{z|\nabla|}[\partial_x g, |\nabla| g]\|_{L_z^\infty H^k} \lesssim \\ &\qquad\qquad \|f\|_{H^{k+1,1-\epsilon}}\|g\|_{\widetilde{W^k}}.\end{aligned} \tag{B.21}$$

We can write the difference of $\nabla_{x,z}\varphi_1$ and $f\nabla_{x,z}\varphi_2$ as follows:

$$\begin{aligned}&\nabla_{x,z}\varphi_1(z) - f\nabla_{x,z}\varphi_2(z) \\ &\quad = e^{z|\nabla|}[\partial_x(fg), |\nabla|(fg)] - fe^{z|\nabla|}[\partial_x g, |\nabla| g] \\ &\qquad + \big[0, \partial_x h(\partial_x\varphi_1 - f\partial_x\varphi_2) - (\partial_x h)^2(\partial_z\varphi_1 - f\partial_z\varphi_2)\big] \\ &\qquad + \int_{-\infty}^0 K(z,z')M(\partial_x h)[\nabla_{x,z}\varphi_1(z') - f\nabla_{x,z}\varphi_2(z')]dz' \\ &\qquad - \int_{-\infty}^0 fK(z,z')M(\partial_x h)\nabla_{x,z}\varphi_2(z') + K(z,z')M(\partial_x h)[f\nabla_{x,z}\varphi_2(z')]dz'.\end{aligned}$$

Similarly to the estimate of (B.21), we have

$$\begin{aligned}&\|\nabla_{x,z}\varphi_1 - f\nabla_{x,z}\varphi_2\|_{L_z^\infty H^k} \\ &\quad \lesssim \|f\|_{H^{k+1,1-\epsilon}}\|g\|_{\widetilde{W^k}} + \|\partial_x h\|_{W^k}\|\nabla_{x,z}\varphi_1 - f\nabla_{x,z}\varphi_2\|_{L_z^\infty H^k} \\ &\qquad + \|f\|_{H^{k+1,1-\epsilon}}\|\partial_x h\|_{\widetilde{W^k}}\|\nabla_{x,z}\varphi_2\|_{L_z^1\widetilde{W^k}}.\end{aligned} \tag{B.22}$$

From the fixed-point-type formulation in (B.7) for $\nabla_{x,z}\varphi_2$, the following estimate holds:

$$\begin{aligned}\|\nabla_{x,z}\varphi_2\|_{L^1_z\widetilde{W^k}} &\lesssim \|g\|_{\widetilde{W^k}} + \|M(\partial_x h)\nabla_{x,z}\varphi_2\|_{L^1_z\widetilde{W^k}} \\ &\lesssim \|g\|_{\widetilde{W^k}} + \|\partial_x h\|_{\widetilde{W^k}}\|\nabla_{x,z}\varphi_2\|_{L^1_z\widetilde{W^k}},\end{aligned}$$

which gives us the following estimate from the smallness assumption on h:

$$\|\nabla_{x,z}\varphi_2\|_{L^1_z\widetilde{W^k}} \lesssim \|g\|_{\widetilde{W^k}}. \tag{B.23}$$

From (B.22) and (B.23), the following estimate holds:

$$\|\nabla_{x,z}\varphi_1 - f\nabla_{x,z}\varphi_2\|_{L^\infty_z H^k} \lesssim \|g\|_{\widetilde{W^k}}\|f\|_{H^{k+1,1-\epsilon}}. \tag{B.24}$$

Similar to the proof of (B.18) in Lemma B.4, with minor modifications we can show that our desired estimates (B.19) and (B.20) hold. □

Appendix C Paralinearization of the Gravity Water Waves System

LEMMA C.1. *The following paralinearization for the Dirichlet-Neumann operator holds*:

$$G(h)\psi = |\nabla|\omega - T_V\partial_x h + F(h)\psi, \quad \omega := \psi - T_{B(h)\psi}h, \tag{C.1}$$

where $F(h)\psi$ is the good quadratic and higher-order error term that does not lose derivatives. Moreover, under the smallness condition (B.1)*, the following estimate holds for $i \in \{2,3\}$ and $k \geq 0$*:

$$\|\Lambda_{\geq i}[F(h)\psi]\|_{H^k} \lesssim \|(\partial_x h, \partial_x\psi)\|_{H^{k-1}}\|(\partial_x h, \partial_x\psi)\|^{i-1}_{W^1}. \tag{C.2}$$

PROOF. Recall that

$$G(h)\psi = \big((1+|\partial_x h|^2)\partial_z\varphi - \partial_x h\partial_x\varphi\big)\big|_{z=0}.$$

Define $W = \varphi - T_{\partial_z\varphi}h$ and $\underline{V} = \partial_x\varphi - \partial_x h\partial_z\varphi$. By using (2.4), as a result, we have the paralinearization of the above equation as follows:

$$\begin{aligned}&\big((1+|\partial_x h|^2)\partial_z\varphi - \partial_x h\partial_x\varphi\big)\\ &\quad= T_{1+|\partial_x h|^2}\partial_z\varphi + 2T_{\partial_z\varphi}T_{\partial_x h}\partial_x h + T_{\partial_x\varphi}R_{\mathcal{B}}(\partial_x h, \partial_x h)\\ &\qquad - T_{\partial_x h}\partial_x\varphi - T_{\partial_x\varphi}\partial_x h\\ &\quad= T_{1+|\partial_x h|^2}\partial_z(W + T_{\partial_z\varphi}h) + 2T_{\partial_z\varphi}T_{\partial_x h}\partial_x h - T_{\partial_x h}\partial_x(W + T_{\partial_z\varphi}h)\\ &\qquad - T_{\partial_x\varphi}\partial_x h + T_{\partial_x\varphi}R_{\mathcal{B}}(\partial_x h, \partial_x h)\\ &\quad= |\nabla|W + T_{1+|\partial_x h|^2}\partial_z W - T_{\partial_x h}\partial_x W - |\nabla|W - T_{\underline{V}}\partial_x h + \widetilde{\mathcal{R}}_1,\\ &\quad= |\nabla|W - T_{\underline{V}}\partial_x h + T_{1+|\partial_x h|^2}(\partial_z - T_A)W + T_{|\partial_x h|^2}T_{A-|\xi|}W\\ &\qquad - T_{|\partial_x h|^2(A-|\xi|)}W + \widetilde{\mathcal{R}}_1\\ &\quad= |\nabla|W - T_{\underline{V}}\partial_x h + \widetilde{F}(h)\psi,\end{aligned}$$

where

$$A = \frac{1}{1+|\partial_x h|^2}(i\partial_x h\cdot\xi + |\xi|),$$

$$\widetilde{\mathcal{R}}_1 := 2T_{\partial_z\varphi}T_{\partial_x h}\partial_x h - 2T_{\partial_z\varphi\partial_x h}\partial_x h - T_{\partial_x h}T_{\partial_z\varphi}\partial_x h + T_{\partial_z\varphi\partial_x h}\partial_x h - T_{\partial_x h}T_{\partial_x\partial_z\varphi}h + T_{\partial_x\varphi}R_{\mathcal{B}}(\partial_x h,\partial_x h), \tag{C.3}$$

$$\widetilde{F}(h)\psi = \big(T_{1+|\partial_x h|^2}(\partial_z - T_A)W + T_{|\partial_x h|^2}T_{A-|\xi|}W - T_{|\partial_x h|^2(A-|\xi|)}W\big) + \widetilde{\mathcal{R}}_1. \tag{C.4}$$

From the explicit formula for $\widetilde{\mathcal{R}}_1$ in (C.3) and the composition lemma, Lemma 2.2, it is easy to see that all terms inside $\widetilde{F}(h)\psi$ except $T_{1+|\partial_x h|^2}(\partial_z - T_A)W$ are good error terms, which do not lose derivatives at the high frequency or the low frequency.

After evaluation (C.4) at the boundary $z = 0$, we can see that the following estimate holds:

$$\begin{aligned}\|\Lambda_{\geq i}[F(h)\psi]\|_{H^k} &\lesssim \big\|\Lambda_{\geq i}\big[(\partial_z - T_A)W\big|_{z=0}\big]\|_{H^k} + \|\partial_x h\|_{\widetilde{W}^1}^{i-1}\|\omega\|_{H^{k-1}} \\ &\quad + \|(\partial_x h,\partial_x\psi,|\nabla|\psi)\|_{\widetilde{W}^1}^{i-1}\|(\partial_x h,\partial_x\psi)\|_{H^{k-1}} \\ &\lesssim \|(\partial_x h,\partial_x\psi)\|_{H^{k-1}}\|(\partial_x h,\partial_x\psi)\|_{W^1}^{(i-1)}.\end{aligned}$$

Note that we used (C.5) in Lemma C.2 in the above estimate. □

Now, we will show that term $(\partial_z - T_A)W$ in (C.4) also does not lose derivatives. More precisely, the following lemma holds:

LEMMA C.2. *The following estimate holds for* $i\in\{2,3\}$, $k\geq 0$,

$$\big\|\Lambda_{\geq i}\big[(\partial_z - T_A)W\big|_{z=0}\big]\big\|_{H^k} \lesssim \|(\partial_x h,\partial_x\psi)\|_{H^{k-1}}\|(\partial_x h,\partial_x\psi)\|_{W^1}^{(i-1)}. \tag{C.5}$$

PROOF. Note that $\Lambda_1[(\partial_z - T_A)W] = 0$, i.e., $(\partial_z - T_A)W$ itself is quadratic and higher. From (C.10) in Lemma C.3, we have

$$(\partial_z - T_{\widetilde{a}})(\partial_z - T_{\widetilde{A}})W = f, \tag{C.6}$$

where $\widetilde{a}$ and $\widetilde{A}$ are given in (C.11).

From (C.20) in Lemma C.4 and (C.14) in Lemma C.3, the following estimate holds:

$$\begin{aligned}\big\|(\partial_z - T_A)W\big|_{z=0}\big\|_{H^k} &\lesssim \sup_{z\in[-1/4,0]}\|(\partial_z - T_{\widetilde{A}})W(z,x)\|_{H^k} + \big\|T_{A'}W\big|_{z=0}\big\|_{H^k} \\ &\lesssim \sup_{z\in[-1,0]}\|(\partial_z - T_{\widetilde{A}})W(z,x)\|_{H^{k-1}} + \|f(z,\cdot)\|_{H^k} \\ &\quad + \|(\partial_x h,\partial_x\psi)\|_{H^{k-1}}\|(\partial_x h,\partial_x\psi)\|_{W^1} \\ &\lesssim \|(\partial_x h,\partial_x\psi)\|_{H^{k-1}}\|(\partial_x h,\partial_x\psi)\|_{W^1}.\end{aligned}$$

It remains to consider the case when $i = 3$. We define

$$\widetilde{W}(z,x) = (\partial_z - T_{\widetilde{A}})W(z,x) - \widetilde{h}(z,x), \quad z \le 0, \tag{C.7}$$

where the function $\widetilde{h}(z,x)$ is to be determined. Let $\widetilde{h}(0,x)$ to be given as follows:

$$\widetilde{h}(0,x) := \Lambda_2\big[(\partial_z - T_{\widetilde{A}})W(z,x)\big|_{z=0}\big].$$

Hence, from the definition, we know that $\widetilde{W}(0,\cdot)$ are the cubic and higher-order terms of $(\partial_z - T_{\widetilde{A}})W(z,x)|_{z=0}$. Let $\widetilde{h}(z,x)$ to be the solution of the following parabolic equation:

$$\begin{cases} (\partial_z - T_{\widetilde{a}})\widetilde{h} = \Lambda_2[f], \\ \widetilde{h}(0,x) = \Lambda_2\big[(\partial_z - T_{\widetilde{A}})W(z,x)\big|_{z=0}\big]. \end{cases} \tag{C.8}$$

Therefore, we can derive the parabolic equation satisfied by $\widetilde{W}(z,x)$ as follows:

$$(\partial_z - T_{\widetilde{a}})\widetilde{W} = \widetilde{\mathcal{N}} := f - (\partial_z - T_{\widetilde{a}})\widetilde{h} = \Lambda_{\ge 3}[f]. \tag{C.9}$$

From (C.20) in Lemma C.4 and (C.14) in Lemma C.3, the following estimate holds:

$$\begin{aligned}
&\big\|\Lambda_{\ge 3}\big[(\partial_z - T_A)W\big|_{z=0}\big]\big\|_{H^k} \\
&\lesssim \sup_{z\in[-1/4,0]} \|(\partial_z - T_{\widetilde{A}})\widetilde{W}(z,x)\|_{H^k} + \big\|\Lambda_{\ge 3}\big[T_{A'}W\big|_{z=0}\big]\big\|_{H^k} \\
&\lesssim \sup_{z\in[-1,0]} \|(\partial_z - T_A)\widetilde{W}(z,x)\|_{H^{k-1}} + \|\Lambda_{\ge 3}[f(z,\cdot)]\|_{H^k} \\
&\quad + \|(\partial_x h, \partial_x\psi)\|_{H^{k-1}}\|(\partial_x h, \partial_x\psi)\|_{W^1}^2 \\
&\lesssim \|(\partial_x h, \partial_x\psi)\|_{H^{k-1}}\|(\partial_x h, \partial_x\psi)\|_{W^1}^2. \qquad \square
\end{aligned}$$

LEMMA C.3. *The following decomposition holds*:

$$(\partial_z - T_{\widetilde{a}})(\partial_z - T_{\widetilde{A}})W = f, \tag{C.10}$$

where

$$\widetilde{a} = a + a', \quad \widetilde{A} = A + A', \tag{C.11}$$

$$a = \frac{1}{1+|\partial_x h|^2}(i\partial_x h\cdot\xi - |\xi|), \quad A = \frac{1}{1+|\partial_x h|^2}(i\partial_x h\cdot\xi + |\xi|), \tag{C.12}$$

$$\begin{aligned} a' &= \frac{1}{A-a}\left(i\partial_\xi a\partial_x A - \frac{\partial_x^2 h a}{1+|\partial_x h|^2}\right), \\ A' &= \frac{1}{a-A}\left(i\partial_\xi a\partial_x A - \frac{\partial_x^2 h A}{1+|\partial_x h|^2}\right). \end{aligned} \tag{C.13}$$

and f is a good error term. The precise formula of f is given in (C.19). *Moreover, under the smallness condition* (B.1)*, the following estimate holds for $i \in \{2,3\}$*:

$$\sup_{z\le 0}\|\Lambda_{\ge i}[f(z)]\|_{H^k} \lesssim \|(\partial_x h, \partial_x\psi)\|_{H^{k-1}}\|(\partial_x h, \partial_x\psi)\|_{W^1}^{(i-1)}. \tag{C.14}$$

PROOF. Recall that $P\varphi = 0$. After paralinearizing this equation, we have

$$\mathcal{P}\varphi + 2T_{\partial_z^2\varphi\partial_x h}\partial_x h - 2T_{\partial_x\partial_z\varphi}\partial_x h - T_{\partial_z\varphi}\partial_x^2 h + f_1 = 0, \tag{C.15}$$

where

$$\begin{aligned}\mathcal{P} &:= T_{(1+|\partial_x h|^2)}\partial_z^2 + \partial_x^2 - 2T_{\partial_x h}\partial_x\partial_z - T_{\partial_x^2 h}\partial_z,\\ f_1 &= R_{\mathcal{B}}(|\partial_x h|^2, \partial_z^2\varphi) - 2R_{\mathcal{B}}(\partial_x\partial_z\varphi, \partial_x h) - R_{\mathcal{B}}(\partial_x^2 h, \partial_z\varphi)\\ &\quad + T_{\partial_z^2\varphi}R_{\mathcal{B}}(\partial_x h, \partial_x h) + 2\big(T_{\partial_z^2\varphi}T_{\partial_x h} - T_{\partial_z^2\varphi\partial_x h}\big)\partial_x h.\end{aligned}$$

Recall that $W = \varphi - T_{\partial_z\varphi}h$; then from (C.15), we have

$$\mathcal{P}W = f_2, \tag{C.16}$$

$$\begin{aligned} f_2 = &-R_{\mathcal{B}}(|\partial_x h|^2, \partial_z^2\varphi) + 2R_{\mathcal{B}}(\partial_x\partial_z\varphi, \partial_x h) + R_{\mathcal{B}}(\partial_x^2 h, \partial_z\varphi)\\ &- T_{\partial_z^2\varphi}R_{\mathcal{B}}(\partial_x h, \partial_x h) - T_{|\partial_x h|^2}T_{\partial_z^3\varphi}h + T_{\partial_z^3\varphi|\partial_x h|^2}h\\ &+ 2T_{\partial_x h}T_{\partial_x\partial_z^2\varphi}h - 2T_{\partial_x h\partial_x\partial_z^2\varphi}h + T_{\partial_x^2 h}T_{\partial_z^2\varphi}h - T_{\partial_x^2 h\partial_z^2\varphi}h.\end{aligned} \tag{C.17}$$

From the composition lemma, Lemma 2.2, we can see that f_2 does not lose derivatives.

Note that the following identities hold from (C.12) and (C.13),

$$\begin{aligned}\widetilde{a} + \widetilde{A} &= \frac{2i\,\partial_x h\xi}{1+|\partial_x h|^2} + \frac{\partial_x^2 h}{1+|\partial_x h|^2},\\ \widetilde{a} \sharp \widetilde{A} := \widetilde{a}\widetilde{A} + \frac{1}{i}\partial_\xi\widetilde{a}\partial_x\widetilde{A} &= -\frac{|\xi|^2}{1+|\partial_x h|^2} + a'A' - i\partial_\xi\widetilde{a}\partial_x A' - i\partial_\xi a'\partial_x A.\end{aligned}$$

Hence, from (C.16), we have

$$(T_{(1+|\partial_x h|^2)}\partial_z^2 - T_{(\widetilde{a}+\widetilde{A})(1+|\partial_x h|^2)} + T_{\widetilde{a}\sharp\widetilde{A}(1+|\partial_x h|^2)})W = f_2 + R_0, \tag{C.18}$$

where

$$R_0 := T_{(a'A' - i\partial_\xi\widetilde{a}\partial_x A' - i\partial_\xi a'\partial_x A)(1+|\partial_x h|^2)}W.$$

It is easy to verify that R_0 does not lose derivatives. From (C.18), we have

$$\big(T_{(1+|\partial_x h|^2)}(\partial_z - T_{\widetilde{a}})(\partial_z - T_{\widetilde{A}})\big)W = R_1,$$

where

$$\begin{aligned} R_1 = f_2 + R_0 &+ T_{(1+|\partial_x h|^2)}(T_{\widetilde{a}}T_{\widetilde{A}} - T_{\widetilde{a}\sharp\widetilde{A}}) + T_{|\partial_x h|^2}T_{\widetilde{a}\sharp\widetilde{A}}W\\ &- T_{\widetilde{a}\sharp\widetilde{A}|\partial_x h|^2}W - T_{|\partial_x h|^2}T_{\widetilde{a}+\widetilde{A}}W + T_{(\widetilde{a}+\widetilde{A})|\partial_x h|^2}W,\end{aligned}$$

which further implies the following equality:

$$\begin{aligned} (\partial_z - T_{\widetilde{a}})(\partial_z - T_{\widetilde{A}})W &= f,\\ f &= T_{1/(1+|\partial_x h|^2)}R_1\\ &\quad + \big(T_{-|\partial_x h|^4/(1+|\partial_x h|^2)} - T_{-|\partial_x h|^2/(1+|\partial_x h|^2)}T_{|\partial_x h|^2}\big)\\ &\quad \times (\partial_z - T_{\widetilde{a}})(\partial_z - T_{\widetilde{A}})W.\end{aligned} \tag{C.19}$$

Again, from the composition lemma, Lemma 2.2, and above explicit formulas, we can see that R_1 and f do not lose derivatives. After checking terms in (C.19) very carefully, the desired estimate (C.14) follows easily. □

LEMMA C.4. *Let $a \in \Gamma_2^1(\mathbb{R}^2)$ and satisfy the assumption* $\mathrm{Re}[a(x,\xi)] \geq c|\xi|$ *for some positive constant c. If u solves the equation*

$$(\partial_w + T_a)u(w,\cdot) = g(w,\cdot),$$

then the following estimate holds for any fixed, sufficiently small constant τ and arbitrarily small constant $\epsilon > 0$:

$$\sup_{w\in[\tau,0]} \|u(w)\|_{H^k} \lesssim M_2^1(a)\frac{1+|\tau|}{|\tau|}\Big[\sup_{z\in[4\tau,0]}\|u(w)\|_{H^{k-2(1-\epsilon)}} + \sup_{z\in[4\tau,0]}\|g(z)\|_{H^{\mu-(1-\epsilon)}}\Big]. \tag{C.20}$$

PROOF. A detailed proof can be found in [3]; the above result is the combination of [3, lemma 2.1.9] and the proof part of [3, lemma 2.1.10]. □

C.1 Estimate of the Scaling Vector Field Part

LEMMA C.5. *Under the smallness assumption* (2.29)*, the following estimates hold for $i \in \{1,2,3\}$ and any $0 < \epsilon \ll 1$:*

$$\|\Lambda_{\geq i}[SB(h)\psi]\|_{H^k} \lesssim \|(\partial_x h, \partial_x\psi)\|_{W^k}^{i-1} \times [\|(Sh,h)\|_{H^{k+1,1-\epsilon}} + \|(|\nabla|(S\psi), |\nabla|\psi)\|_{H^k}], \tag{C.21}$$

$$\|\Lambda_{\geq 4}[SB(h)\psi]\|_{L^2} \lesssim (\|(Sh,h)\|_{H^{1,1-\epsilon}} + \||\nabla|(S\psi,\psi)\|_{L^2}) \times \|(\partial_x h, \partial_x\psi)\|_{W^2}^3. \tag{C.22}$$

PROOF. From [3, lemma 2.3.3], the following identity holds:

$$\begin{aligned} SB(h)\psi &= B(h)(S\psi - (B(h)\psi)Sh) \\ &\quad + \frac{1}{1+(\partial_x h)^2}\big[-2(V(h)\psi)\partial_x h - 2G(h)\psi + 2G(h)(hB(h)\psi) \\ &\quad - 2\eta G(h)(B(h)\psi) + [\partial_x h\partial_x(B(h)\psi) - \partial_x(V(h)\psi)]Sh\big]. \end{aligned} \tag{C.23}$$

We remark that the above identity is derived from comparing the $S\partial_z\varphi|_{z=0}$ with the $\partial_z\varphi_1\big|_{z=0}$, where $\varphi_1(0) = S\psi$. After studying the two Laplace equations in the (x,z)-coordinate formulation, one can derive (C.23) without any problem.

From the facts that $P\varphi = 0$ and $B(h)\psi = \partial_z\varphi|_{z=0}$, after evaluating $P\varphi(z)$ at the boundary $z = 0$, we can derive the identity

$$-\partial_x[V(h)\psi] = G(h)(B(h)\psi). \tag{C.24}$$

After combining the identities (C.23), (C.24), and (2.10), we have the identity

$$
\begin{aligned}
S[B(h)\psi] = B(h)(S\psi) &- \frac{2\partial_x h V(h)\psi + 2G(h)\psi}{1+(\partial_x h)^2} - \frac{2\partial_x h\partial_x h B(h)\psi}{1+(\partial_x h)^2} \\
&+ 2[B(h)(hB(h)\psi) - hB(h)(B(h)\psi)] \\
&+ [ShB(h)(B(h)\psi) - B(h)(ShB(h)\psi)].
\end{aligned} \tag{C.25}
$$

Since terms inside (C.25) that do not depend on the scaling vector field S can be estimated very easily by using estimates in Lemma B.3, we omit the details for those terms. From Lemma B.5, the following estimate holds for $i \in \{1, 2, 3\}$:

$$
\begin{aligned}
&\|\Lambda_{\geq i}[B(h)(hB(h)\psi) - hB(h)(B(h)\psi)]\|_{H^k} \\
&\quad + \|\Lambda_{\geq i}[ShB(h)(B(h)\psi) - B(h)(ShB(h)\psi)]\|_{H^k} \\
&\quad \lesssim \|(Sh, h)\|_{H^{k+1,1-\epsilon}} \|\Lambda_{\geq i-1}[B(h)\psi]\|_{\widetilde{W}^k} \\
&\quad \lesssim \|(Sh, h)\|_{H^{k+1,1-\epsilon}} \|(\partial_x h, \partial_x \psi, |\nabla|\psi)\|_{\widetilde{W}^k}^{(i-1)}.
\end{aligned}
$$

From Lemma B.4, the following estimate holds:

$$
\|\Lambda_{\geq i}[B(h)S\psi]\|_{H^k} \lesssim \|\partial_x h\|_{W^k}^{i-1} \||\nabla|(S\psi)\|_{H^k}.
$$

To sum up, we can see that our desired estimate (C.21) holds. With the additional estimate (B.18) in Lemma B.4, we can redo the above argument to derive the desired estimate (C.22). We omit the details here. □

LEMMA C.6. *Under the smallness assumption* (B.1)*, the following estimate holds for* $1 \leq k \leq \gamma'$, $p \in (0, 1)$, *and* $i \in \{2, 3\}$:

$$
\begin{aligned}
&\|\Lambda_{\geq i}[SF(h)\psi]\|_{H^{k,p}} \\
&\lesssim (\|Sh\|_{H^{k,p}} + \|h\|_{H^{k+1,1-p}} + \||\nabla|\psi\|_{H^k} + \||\nabla|S\psi\|_{H^{k-1}}) \\
&\quad \times \|(\partial_x h, |\nabla|\psi)\|_{W^{k+1}}^{(i-1)}.
\end{aligned} \tag{C.26}
$$

PROOF. Recall (C.1), we have $F(h)\psi = G(h)\psi - |\nabla|(\psi - T_B h) + \partial_x T_V h$. Hence, from (C.25) in Proposition C.5, we can derive the equality

$$
\begin{aligned}
S(F(h)\psi) = F(h)(S\psi) &+ (-\partial_x(Sh)V + \partial_x T_V Sh) \\
&- (G(h)(ShB) - Sh(G(h)B) - |\nabla|T_B Sh) + \mathcal{G},
\end{aligned} \tag{C.27}
$$

where the good error term $\mathcal{G}$ is

$$
\mathcal{G} = \mathcal{G}_1 + \mathcal{G}_2 + \mathcal{G}_3,
$$

where

$$
\begin{aligned}
\mathcal{G}_1 &= 2(G(h)(hB) - hG(h)(B)), \\
\mathcal{G}_2 &= 2\partial_x hV - 2\Lambda_{\geq 2}[G(h)\psi] - |\nabla|T_{B(\eta)(S\psi)}h - \partial_x T_{V(\eta)(S\psi)}h \\
&\quad + |\nabla|T_{SB}h + \partial_x T_{SV}h, \\
\mathcal{G}_3 &= S[|\nabla|(T_B h) + \partial_x(T_V h)] - |\nabla|T_{SB}h - |\nabla|T_B Sh - \partial_x T_{SV}h - \partial_x T_V Sh.
\end{aligned}
$$

One can also refer to [3] for more detailed computations to see the validity of (C.27).

From Lemma B.5, the following estimate holds:

$$\begin{aligned}\|\Lambda_{\geq i}[\mathcal{G}_1]\|_{H^k} &\lesssim \|h\|_{H^{k+1,1-\epsilon}}\|\Lambda_{\geq i-1}[B(h)\psi]\|_{\widetilde{W}^k}\\ &\lesssim \|h\|_{H^{k+1,1-\epsilon}}\|(\partial_x h,\partial_x\psi)\|_{W^k}.\end{aligned}$$

From (C.21) in Lemma C.5, the following estimate holds:

$$\begin{aligned}\|\Lambda_{\geq i}[\mathcal{G}_2]\|_{H^k} &\lesssim \big(\|\Lambda_{\geq i-1}[(B(h)(S\psi),V(h)(S\psi))]\|_{L^2}\\ &\quad + \|\Lambda_{\geq i-1}[(SB(h)\psi, SV(h)\psi)]\|_{L^2}\big)\|(\partial_x h,\partial_x\psi)\|_{W^k}\\ &\quad + \|(\partial_x h,\partial_x\psi\|_{H^k}\|(\partial_x h,\partial_x\psi)\|^{i-1}_{\widetilde{W}^0}\\ &\lesssim \big(\|Sh\|_{H^{1,1-\epsilon}} + \|(\partial_x h,|\nabla|\psi)\|_{H^k} + \||\nabla|S\psi\|_{L^2}\big)\\ &\quad\times\|(\partial_x h,\partial_x\psi)\|^{i-1}_{W^k}.\end{aligned}$$

Note that the commutator term $\mathcal{G}_3$ does not depend on the scaling vector field.

It is easy to derive the following estimate:

$$\begin{aligned}\|\Lambda_{\geq i}[\mathcal{G}_3]\|_{H^k} &\lesssim \|\Lambda_{\geq i-1}[(B(h)\psi,V(h)\psi)]\|_{H^k}\|(\partial_x h,\partial_x\psi)\|_{W^0}\\ &\quad + \|(\partial_x h,\partial_x\psi)\|_{H^k}\|\Lambda_{\geq i-1}[(B(h)\psi,V(h)\psi)]\|_{\widetilde{W^0}}\\ &\lesssim \|(\partial_x h,\partial_x\psi)\|_{H^k}\|(\partial_x h,\partial_x\psi)\|^{i-1}_{W^0}.\end{aligned}$$

To sum up, we have

$$\begin{aligned}&\|\Lambda_{\geq i}[\mathcal{G}]\|_{H^k}\\ &\lesssim \big(\|Sh\|_{H^{1,1-\epsilon}} + \||\nabla|S\psi\|_{L^2} + \|h\|_{H^{k+1,1-\epsilon}} + \||\nabla|\psi\|_{H^k}\big)\\ &\quad\times\|(\partial_x h,\partial_x\psi)\|^{i-1}_{W^k}.\end{aligned}\tag{C.28}$$

To see the structure inside $(G(h)(ShB)-Sh(G(h)B)-|\nabla|T_B Sh)$, we decompose it as follows:

$$G(h)(ShB) - Sh(G(h)B) - |\nabla|T_B Sh = \sum_{k\in\mathbb{Z}}\mathcal{J}_{1,k} + \mathcal{J}_{2,k},\tag{C.29}$$

where

$$\begin{aligned}\mathcal{J}_{1,k} &:= G(h)(P_k BP_{\leq k-1}(Sh)) - P_{\leq k-1}(Sh)G(h)(P_k B)\\ &= -\partial_x h P_k B\partial_x(P_{\leq k-1}(Sh)) + (1+(\partial_x h)^2)\\ &\quad\times[B(h)(P_k BP_{\leq k-1}(Sh)) - P_{\leq k-1}(Sh)B(h)(P_k B)],\end{aligned}\tag{C.30}$$

$$\begin{aligned}\mathcal{J}_{2,k} &:= G(h)(P_k(Sh)P_{\leq k}B) - P_k(Sh)G(h)(P_{\leq k}B) - |\nabla|T_B Sh\\ &:= \mathcal{J}^1_{2,k} + \mathcal{J}^2_{2,k},\end{aligned}\tag{C.31}$$

$$\begin{aligned}\mathcal{J}^1_{2,k} &:= F(h)(P_k(Sh)P_{\leq k}B) + |\nabla|(P_k(Sh)P_{\leq k-1}B - T_B P_k Sh)\\ &= F(h)(P_k(Sh)P_{\leq k}B) + |\nabla|(P_k(Sh)P_{k-10\leq\cdot\leq k-1}B).\end{aligned}\tag{C.32}$$

$$\mathcal{J}_{2,k}^2 := -\big[|\nabla| T_{B(h)(P_k(Sh)P_{\le k}B)} h + \partial_x T_{V(h)(P_k(Sh)P_{\le k}B)} h + P_k(Sh) G(h)(P_{\le k-1}B)\big]. \tag{C.33}$$

Note that we again used the fact that $F(h)\psi = G(h)\psi - |\nabla|(\psi - T_B h) + \partial_x T_V h$ in the decomposition (C.31).

From Lemma B.5, the following estimate holds for some $\epsilon < (1-p)/100$:

$$\begin{aligned}
&\Big\|\sum_{j\in\mathbb{Z}} \Lambda_{\ge i}[\mathcal{J}_{1,j}]\Big\|_{H^k} \\
&\lesssim \sum_{j\in\mathbb{Z}} \|P_{\le j-1}(Sh)\|_{H^{k+1,1-\epsilon/2}} \|\Lambda_{\ge i-1}[P_j[B(h)\psi]]\|_{\widetilde{W}^k} \\
&\quad + \|Sh\|_{H^{k,1-\epsilon}} \|(\partial_x h, \partial_x \psi)\|_{W^{k+1}}^{(i-1)} \\
&\lesssim \|Sh\|_{H^{k,1-\epsilon}} \|\Lambda_{\ge i-1}[B(h)\psi]\|_{\widetilde{W}^{k+1}} \\
&\quad + \|Sh\|_{H^{k,1-\epsilon}} \|(\partial_x h, \partial_x \psi)\|_{W^{k+1}}^{(i-1)} \\
&\lesssim \|Sh\|_{H^{k,1-\epsilon}} \|(\partial_x h, \partial_x \psi)\|_{W^{k+1}}^{i-1}.
\end{aligned} \tag{C.34}$$

From (C.2) in Lemma C.1, the following estimate holds:

$$\begin{aligned}
&\Big\|\sum_{j\in\mathbb{Z}} \Lambda_{\ge i}[\mathcal{J}_{2,j}^1]\Big\|_{H^k} \\
&\lesssim \sum_{j\in\mathbb{Z}} \|\partial_x(P_j(Sh)P_{\le j}B)\|_{H^{k-1}} \|(\partial_x h, \partial_x \psi)\|_{W^1}^{(i-2)} \\
&\lesssim \|Sh\|_{H^{k,1-\epsilon}} \|(\partial_x h, \partial_x \psi)\|_{W^1}^{i-1}.
\end{aligned} \tag{C.35}$$

$$\begin{aligned}
&\Big\|\sum_{j\in\mathbb{Z}} \Lambda_{\ge i}[\mathcal{J}_{2,j}^2]\|_{H^{k,p}} \\
&\lesssim \|Sh\|_{H^{k,p}} \|\Lambda_{\ge i-1}[B(h)\psi]\Big\|_{\widetilde{W}^0} \\
&\quad + \sum_{j\in\mathbb{Z}} \|\partial_x h\|_{W^k} \|\Lambda_{\ge i-1} \\
&\qquad \times \big[\big(B(h)(P_j(Sh)P_{\le j}B)), V(h)(P_j(Sh)P_{\le j}B))\big)\big]\|_{L^2} \\
&\lesssim \|Sh\|_{H^{k,p}} \|(\partial_x h, \partial_x \psi)\|_{W^k}^{i-1}.
\end{aligned} \tag{C.36}$$

To sum up, from (C.27), (C.28), (C.34), (C.35), and (C.36), our desired esitmate (C.26) holds. □

C.2 Estimate of the Taylor Coefficient

LEMMA C.7. *Under the smallness assumption* (B.1)*, the following estimates hold for k, $\gamma \geq 0$, $\gamma \leq \gamma' - 2$, and $i \in \{1, 2, 3\}$*:

$$\begin{aligned} &\|\Lambda_{\geq i}[\alpha]\|_{\widetilde{W^{\gamma}}} \lesssim \|(\partial_x h, \partial_x \psi)\|_{W^{\gamma+1}}^{i}, \\ &\|\Lambda_{\geq i}[\alpha]\|_{H^k} \lesssim \|(\partial_x h, \partial_x \psi)\|_{W^1}^{i-1} \|(\partial_x h, |\nabla|\psi)\|_{H^{k+1}}, \end{aligned} \tag{C.37}$$

$$\begin{aligned} \|\Lambda_{\geq i}[S\alpha]\|_{H^k} \lesssim{}& \|(\partial_x h, \partial_x \psi)\|_{W^{k+1}}^{i-1} \\ &\times [\|(Sh, h)\|_{H^{k+2,1-\epsilon}} + \|(|\nabla|(S\psi), |\nabla|\psi)\|_{H^{k+1}}], \end{aligned} \tag{C.38}$$

$$\|\Lambda_{\geq i}[\partial_t \alpha]\|_{\widetilde{W^{\gamma}}} \lesssim \|(\partial_x h, \partial_x \psi)\|_{W^{\gamma+2}}^{i}, \tag{C.39}$$

$$\begin{aligned} &\|\Lambda_{\geq i}[\partial_t \alpha]\|_{H^k} \lesssim \|(\partial_x h, \partial_x \psi)\|_{W^2}^{(i-1)} \|(\partial_x h, |\nabla|\psi)\|_{H^{k+2}}, \\ &\|\Lambda_{\geq i}[S\partial_t \alpha]\|_{H^k} \lesssim \|(\partial_x h, \partial_x \psi)\|_{W^{k+2}}^{i-1} \big[\|(Sh, h)\|_{H^{k+3,1-\epsilon}} \\ &\qquad\qquad + \|(|\nabla|(S\psi), |\nabla|\psi)\|_{H^{k+2}}\big]. \end{aligned} \tag{C.40}$$

PROOF. Recall that $\alpha = \sqrt{a} - 1$. It is sufficient to estimate the Taylor coefficient a. We cite the following identity from [3, lemma A.4.2]:

$$a = \frac{1}{2(1 + (\partial_x h)^2)} \big[2 + 2V\partial_x B - 2B\partial_x V - G(h)[V^2 + B^2 + 2h]\big]. \tag{C.41}$$

Recall that $a = 1 + \partial_t B + V\partial_x B$ and $\alpha = \sqrt{a} - 1$. Therefore, we have the identity

$$\begin{aligned} \partial_t B = {}& \frac{1}{2(1 + (\partial_x h)^2)} [2 + 2V\partial_x B - 2B\partial_x V - G(h)(V^2 + B^2 + 2h)] \\ & - 1 - V\partial_x B. \end{aligned} \tag{C.42}$$

As we already have the L^2-type and the L^∞-type estimates for $B(h)\psi$, we can easily derive the L^2-type and the L^∞-type estimates for $\partial_t B$ from (C.42). As a result, from estimates in Lemma B.3 , the following estimate holds for $i \in \{1, 2, 3\}$:

$$\begin{aligned} &\|\Lambda_{\geq i}[\partial_t B]\|_{H^k} \lesssim \|(\partial_x h, \partial_x \psi)\|_{H^{k+1}} \|(\partial_x h, \partial_x \psi)\|_{W^1}^{i-1}, \|\Lambda_{\geq i}[\partial_t B]\|_{\widetilde{W}^{\gamma}} \\ &\qquad \lesssim \|(\partial_x h, \partial_x \psi)\|_{W^{\gamma+1}}^{i}. \end{aligned}$$

From (C.21) in Lemma C.5, the following estimate holds

$$\begin{aligned} &\|\Lambda_{\geq i}[S\partial_t B]\|_{H^k} \lesssim \\ &\qquad \|(\partial_x h, \partial_x \psi)\|_{W^{k+1}}^{i-1} [\|(Sh, h)\|_{H^{k+2,1-\epsilon}} + \|(|\nabla|(S\psi), |\nabla|\psi)\|_{H^{k+1}}]. \end{aligned}$$

After taking another derivative with respect to time t on both sides of (C.42), we have an identity for $\partial_t^2 B$. As a result, from the L^2-type and L^∞-type estimates of

$\partial_t B$ and B, the following estimates hold for $i \in \{1, 2, 3\}$:

$$\begin{aligned}
\|\Lambda_{\geq i}[\partial_t^2 B]\|_{H^k} &\lesssim \|(\partial_x h, \partial_x \psi)\|_{H^{k+2}} \|(\partial_x h, \partial_x \psi)\|_{W^2}^{i-1}, \\
\|\Lambda_{\geq i}[\partial_t^2 B]\|_{\widetilde{W}^\gamma} &\lesssim \|(\partial_x h, \partial_x \psi)\|_{W^{\gamma+2}}^{i}, \\
\|\Lambda_{\geq i}[S\partial_t^2 B]\|_{H^k} &\lesssim \|(\partial_x h, \partial_x \psi)\|_{W^{k+2}}^{i-1} \\
&\quad \times [\|(Sh, h)\|_{H^{k+3,1-\epsilon}} + \|(|\nabla|(S\psi), |\nabla|\psi)\|_{H^{k+2}}].
\end{aligned}$$

Hence, our desired estimates follow from (C.21) in Lemma C.5 and estimates in Lemma B.3. □

Acknowledgments. I would like to thank my Ph.D. advisor Alexandru Ionescu for many helpful discussions. Also, I would like to thank the invitations and hospitality of the Hausdorff Research Institute for Mathematics in Bonn and Fudan University, where a part of this work was done. Lastly, I would like to thank the anonymous referee for many helpful suggestions to improve the presentation of this paper.

Bibliography

[1] Alazard, T.; Burq, N.; Zuily, C. On the water-wave equations with surface tension. *Duke Math. J.* **158** (2011), no. 3, 413–499. doi:10.1215/00127094-1345653

[2] Alazard, T.; Burq, N.; Zuily, C. On the Cauchy problem for gravity water waves. *Invent. Math.* **198** (2014), no. 1, 71–163. doi:10.1007/s00222-014-0498-z

[3] Alazard, T.; Delort, J.-M. Global solutions and asymptotic behavior for two dimensional gravity water waves. *Ann. Sci. Éc. Norm. Supér. (4)* **48** (2015), no. 5, 1149–1238.

[4] Alazard, T.; Métivier, G. Paralinearization of the Dirichlet to Neumann operator, and regularity of three-dimensional water waves. *Comm. Partial Differential Equations* **34** (2009), no. 10–12, 1632–1704. doi:10.1080/03605300903296736

[5] Ambrose, D. M.; Masmoudi, N. The zero surface tension limit of two-dimensional water waves. *Comm. Pure Appl. Math.* **58** (2005), no. 10, 1287–1315. doi:10.1002/cpa.20085

[6] Beyer, K.; Günther, M. On the Cauchy problem for a capillary drop. I. Irrotational motion. *Math. Methods Appl. Sci.* **21** (1998), no. 12, 1149–1183. doi:10.1002/(SICI)1099-1476(199808)21:12<1149::AID-MMA990>3.0.CO;2-C

[7] Castro, A.; Córdoba, D.; Fefferman, C.; Gancedo, F.; Gómez-Serrano, J. Finite time singularities for the free boundary incompressible Euler equations. *Ann. of Math. (2)* **178** (2013), no. 3, 1061–1134. doi:10.4007/annals.2013.178.3.6

[8] Christodoulou, D.; Lindblad, H. On the motion of the free surface of a liquid. *Comm. Pure Appl. Math.* **53** (2000), no. 12, 1536–1602. doi:10.1002/1097-0312(200012)53:12<1536::AID-CPA2>3.0.CO;2-Q

[9] Coutand, D.; Shkoller, S. Well-posedness of the free-surface incompressible Euler equations with or without surface tension. *J. Amer. Math. Soc.* **20** (2007), no. 3, 829–930. doi:10.1090/S0894-0347-07-00556-5

[10] Coutand, D.; Shkoller, S. On the finite-time splash and splat singularities for the 3-D free-surface Euler equations. *Comm. Math. Phys.* **325** (2014), no. 1, 143–183. doi:10.1007/s00220-013-1855-2

[11] Craig, W.; Sulem, C.; Sulem, P.-L. Nonlinear modulation of gravity waves: a rigorous approach. *Nonlinearity* **5** (1992), no. 2, 497–522.

[12] Deift, P.; Zhou, X. Long-time asymptotics for solutions of the NLS equation with initial data in a weighted Sobolev space. *Comm. Pure Appl. Math.* **56** (2003), no. 8, 1029–1077. doi:10.1002/cpa.3034

[13] Germain, P.; Masmoudi, N.; Shatah, J. Global solutions for 3D quadratic Schrödinger equations. *Int. Math. Res. Not. IMRN* **2009** (2009), no. 3, 414–432. doi:10.1093/imrn/rnn135

[14] Germain, P.; Masmoudi, N.; Shatah, J. Global solutions for the gravity water waves equation in dimension 3. *Ann. of Math. (2)* **175** (2012), no. 2, 691–754. doi:10.4007/annals.2012.175.2.6

[15] Germain, P.; Masmoudi, N.; Shatah, J. Global existence for capillary water waves. *Comm. Pure Appl. Math.* **68** (2015), no. 4, 625–687. doi:10.1002/cpa.21535

[16] Hunter, J. K.; Ifrim, M.; Tataru, D. Two dimensional water waves in holomorphic coordinates. *Comm. Math. Phys.* **346** (2016), no. 2, 483–552. doi:10.1007/s00220-016-2708-6

[17] Ifrim, M.; Tataru, D. The lifespan of small data solutions in two dimensional capillary water wave. Preprint, 2014. arXiv:1406.5471 [math.AP]

[18] Ifrim, M.; Tataru, D. Global bounds for the cubic nonlinear Schrödinger equation (NLS) in one space dimension. *Nonlinearity* **28** (2015), no. 8, 2661–2675. doi:10.1088/0951-7715/28/8/2661

[19] Ifrim, M.; Tataru, D. Two dimensional water waves in holomorphic coordinates II: Global solutions. *Bull. Soc. Math. France* **144** (2016), no. 2, 369–394.

[20] Ionescu, A. D.; Pusateri, F. Nonlinear fractional Schrödinger equations in one dimension. *J. Funct. Anal.* **266** (2014), no. 1, 139–176. doi:10.1016/j.jfa.2013.08.027

[21] Ionescu, A. D.; Pusateri, F. Global solutions for the gravity water waves system in 2d. *Invent. Math.* **199** (2015), no. 3, 653–804. doi:10.1007/s00222-014-0521-4

[22] Ionescu, A. D.; Pusateri, F. Global analysis of a model for capillary water waves in two dimensions. *Comm. Pure Appl. Math.* **69** (2016), no. 11, 2015–2071. doi:10.1002/cpa.21654

[23] Ionescu, A.; Pusateri, F. A note on the asymptotic behavior of gravity water waves in two dimensions. Unpublished note. Available at: `https://web.math.princeton.edu/~fabiop/2dWWasym-web.pdf`.

[24] Ionescu, A.; Pusateri, F. Global regularity for 2D water waves with surface tension. *Mem. Amer. Math. Soc.*, forthcoming.

[25] Lannes, D. Well-posedness of the water-waves equations. *J. Amer. Math. Soc.* **18** (2005), no. 3, 605–654. doi:10.1090/S0894-0347-05-00484-4

[26] Lindblad, H. Well-posedness for the motion of an incompressible liquid with free surface boundary. *Ann. of Math. (2)* **162** (2005), no. 1, 109–194. doi:10.4007/annals.2005.162.109

[27] Métivier, G. *Para-differential calculus and applications to the Cauchy problem for nonlinear systems.* Centro di Ricerca Matematica Ennio De Giorgi (CRM) Series, 5. Edizioni della Normale, Pisa, 2008.

[28] Nalimov, V. I. The Cauchy-Poisson problem. *Dinamika Splošn. Sredy Vyp.* **18** Dinamika Židkost. so Svobod. Granicami (1974), 104–210, 254.

[29] Shatah, J. Normal forms and quadratic nonlinear Klein-Gordon equations. *Comm. Pure Appl. Math.* **38** (1985), no. 5, 685–696. doi:10.1002/cpa.3160380516

[30] Shatah, J.; Zeng, C. Geometry and a priori estimates for free boundary problems of the Euler equation. *Comm. Pure Appl. Math.* **61** (2008), no. 5, 698–744. doi:10.1002/cpa.20213

[31] Totz, N.; Wu, S. A rigorous justification of the modulation approximation to the 2D full water wave problem. *Comm. Math. Phys.* **310** (2012), no. 3, 817–883. doi:10.1007/s00220-012-1422-2

[32] Wu, S. Well-posedness in Sobolev spaces of the full water wave problem in 2-D. *Invent. Math.* **130** (1997), no. 1, 39–72. doi:10.1007/s002220050177

[33] Wu, S. Well-posedness in Sobolev spaces of the full water wave problem in 3-D. *J. Amer. Math. Soc.* **12** (1999), no. 2, 445–495. doi:10.1090/S0894-0347-99-00290-8

[34] Wu, S. Almost global wellposedness of the 2-D full water wave problem. *Invent. Math.* **177** (2009), no. 1, 45–135. doi:10.1007/s00222-009-0176-8

[35] Wu, S. Global wellposedness of the 3-D full water wave problem. *Invent. Math.* **184** (2011), no. 1, 125–220. doi:10.1007/s00222-010-0288-1
[36] Yosihara, H. Gravity waves on the free surface of an incompressible perfect fluid of finite depth. *Publ. Res. Inst. Math. Sci.* **18** (1982), no. 1, 49–96. doi:10.2977/prims/1195184016
[37] Zakharov, V. Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J. Appl. Mech. Tech. Phys.* **9** (1968), no. 2, 190–194. doi:10.1007/BF00913182

XUECHENG WANG
Mathematics Department
Princeton University
Princeton, NJ 08544
USA
E-mail: `xuecheng@math.princeton.edu`

Received April 2016.
Revised January 2017.