ANNALES DE LA FACULTÉ DES SCIENCES DE TOULOUSE Mathématiques

Soonsik Kwon, Tadahiro Oh and Haewon Yoon

Normal form approach to unconditional well-posedness of nonlinear dispersive PDEs on the real line

Tome XXIX, nº 3 (2020), p. 649-720.

<http://afst.centre-mersenne.org/item?id=AFST_2020_6_29_3_649_0>

Annales de la faculté des sciences de Toulouse Volume XXIX, n° 3, 2020
pp. 649-720

Normal form approach to unconditional well-posedness of nonlinear dispersive PDEs on the real line [(*)]

Soonsik Kwon [(1)], Tadahiro Oh [(2)] and Haewon Yoon [(3)]

Abstract. — In this paper, we revisit the infinite iteration scheme of normal form reductions, introduced by the first and second authors (with Z. Guo), in constructing solutions to nonlinear dispersive PDEs. Our main goal is to present a simplified approach to this method. More precisely, we study normal form reductions in an abstract form and reduce multilinear estimates of arbitrarily high degrees to successive applications of basic trilinear estimates. As an application, we prove unconditional well-posedness of canonical nonlinear dispersive equations on the real line. In particular, we implement this simplified approach to an infinite iteration of normal form reductions in the context of the cubic nonlinear Schrödinger equation (NLS) and the modified KdV equation (mKdV) on the real line and prove unconditional well-posedness in $H^s(\mathbb{R})$ with (i) $s \geqslant \frac{1}{6}$ for the cubic NLS and (ii) $s > \frac{1}{4}$ for the mKdV. Our normal form approach also allows us to construct weak solutions to the cubic NLS in $H^s(\mathbb{R})$, $0 \leqslant s < \frac{1}{6}$, and distributional solutions to the mKdV in $H^{\frac{1}{4}}(\mathbb{R})$ (with some uniqueness statements).

Résumé. — Dans cet article, nous revisitons le schéma d'itération infinie des réductions de forme normale, introduit par les premier et deuxième auteurs (avec

[(*)] Reçu le 16 mars 2018, accepté le 6 novembre 2018.

Keywords: nonlinear Schrödinger equation, modified KdV equation, normal form reduction, unconditional well-posedness, unconditional uniqueness.

2020 *Mathematics Subject Classification:* 35Q55, 35Q53.

[(1)] Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, 291 Daehak-ro Yuseong-gu, Daejeon 34141 (Republic of Korea) — soonsikk@kaist.edu

[(2)] School of Mathematics, The University of Edinburgh, and The Maxwell Institute for the Mathematical Sciences, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD (United Kingdom) — hiro.oh@ed.ac.uk

[(3)] Department of Mathematical Sciences, Korea Advanced Institute of Science and Technology, 291 Daehak-ro Yuseong-gu, Daejeon 34141 (Republic of Korea) — *Current address:* National Center for Theoretical Sciences, No. 1 Sec. 4 Roosevelt Rd., National Taiwan University, Taipei, 10617 (Taiwan) — hwyoon@ncts.ntu.edu.tw

S.K. was partially supported by National Research Foundation of Korea (grant NRF-2015R1D1A1A01058832). T.O. was supported by the European Research Council (grant no. 637995 "ProbDynDispEq"). H.Y. was partially supported by National Research Foundation of Korea (grant 2012-H1A2A1049375).

Article proposé par Nikolay Tzvetkov.

Z. Guo), dans la construction des solutions des EDP dispersives non linéaires. Notre objectif principal est de présenter une approche simplifiée à cette méthode. Plus précisément, nous étudions les réductions de forme normale dans un cadre abstrait et nous réduisons les estimations multilinéaires de degrés arbitraires aux applications successives des estimations trilinéaires fondamentales. Comme application, nous montrons que des équations dispersives non linéaires canoniques sont inconditionnellement bien-posées sur la droite réelle. En particulier, nous implémentons cette approche simplifiée à l'itération infinie des réductions de forme normale dans le contexte de l'équation de Schrödinger non linéaire cubique (NLS) et de l'équation de KdV modifiée (mKdV) sur la droite réelle et nous prouvons qu'elles sont inconditionnellement bien posées dans $H^s(\mathbb{R})$ avec (i) $s \geqslant \frac{1}{6}$ dans le cas pour NLS cubique et (ii) $s > \frac{1}{4}$ dans le cas pour mKdV. Notre approche de forme normale nous permet également de construire solutions faibles au NLS cubique dans $H^s(\mathbb{R})$, $0 \leqslant s < \frac{1}{6}$, et solutions de distribution au mKdV dans $H^{\frac{1}{4}}(\mathbb{R})$ (avec certaine forme d'unicité).

1. Introduction

1.1. Main results

In this paper, we study the Cauchy problem for some canonical nonlinear dispersive equations on the real line. More specifically, we consider the following cubic nonlinear Schrödinger equation (NLS):

$$\begin{cases} i\partial_t u = \partial_x^2 u \pm |u|^2 u \\ u|_{t=0} = u_0 \in H^s(\mathbb{R}), \end{cases} \qquad (x,t) \in \mathbb{R} \times \mathbb{R}, \tag{1.1}$$

and the modified KdV equation (mKdV):

$$\begin{cases} \partial_t u = \partial_x^3 u \pm \partial_x(u^3) \\ u|_{t=0} = u_0 \in H^s(\mathbb{R}), \end{cases} \qquad (x,t) \in \mathbb{R} \times \mathbb{R}, \tag{1.2}$$

where a solution u is complex-valued in (1.1) and is real-valued in (1.2).

The Cauchy problems (1.1) and (1.2) have been studied extensively by many mathematicians. In particular, multilinear harmonic analysis played an important role in establishing well-posedness of these equations in low regularities. Moreover, these equations are known to be one of the simplest examples of completely integrable equations [31, 32, 43, 44] and such integrable structures play an important role in establishing a priori bounds for these equations in the low regularity setting [29, 23]. In the following, however, we will not focus on such an integrable structure in an explicit manner. Our main goal in this paper is to introduce a new methodology to establish well-posedness of (1.1) and (1.2) (with stronger uniqueness) *without* relying on heavy harmonic analysis or complete integrability.

Let us briefly go over the well-posedness theory of (1.1) and (1.2). We say that the Cauchy problem (1.1) or (1.2) is locally well-posed in $H^s(\mathbb{R})$ if given $u_0 \in H^s(\mathbb{R})$, there exists a unique solution $u \in C([-T,T]; H^s(\mathbb{R}))$ to the equation for some $T = T(u_0) > 0$. Moreover, we impose that the solution map: $u_0 \in H^s(\mathbb{R}) \mapsto u \in C([-T,T]; H^s(\mathbb{R}))$ be continuous. If we can take $T > 0$ to be arbitrarily large, we say that the Cauchy problem is globally well-posed. As we see below, one often needs to employ an auxiliary function space X_T to establish well-posedness. As a result, we have uniqueness of solutions only in $C([-T,T]; H^s(\mathbb{R})) \cap X_T$. In this case, we say that uniqueness holds conditionally. If, instead, uniqueness holds in the entire $C([-T,T]; H^s(\mathbb{R}))$, then we say that the Cauchy problem is *unconditionally* (locally) well-posed in $H^s(\mathbb{R})$. See [19]. Unconditional uniqueness is a notion of uniqueness which does not depend on how solutions are constructed. In the following, we summarize the known analytical results on the well-posedness of (1.1) and (1.2).

A basic strategy for proving local well-posedness of (1.1) or (1.2) is to write the equation in the Duhamel formulation:[(1)]

$$u(t) = e^{-it\partial_x^2} u_0 \mp i \int_0^t e^{-i(t-t')\partial_x^2} |u|^2 u(t')\,\mathrm{d}t'$$

and solve the corresponding fixed point problem. When $s > \frac{1}{2}$, Sobolev's embedding theorem allows us to prove local well-posedness of the cubic NLS (1.1) in $H^s(\mathbb{R})$ via the contraction mapping principle. In [42], Tsutsumi used the Strichartz estimates and proved local well-posedness of (1.1) in $L^2(\mathbb{R})$, which immediately implied global well-posedness in $L^2(\mathbb{R})$ thanks to the L^2-conservation. Note that the uniqueness holds conditionally in [42] due to the use of the Strichartz spaces. By refining the analysis, Kato [19] proved unconditional well-posedness of (1.1) in $H^s(\mathbb{R})$, $s \geqslant \frac{1}{6}$.

Another approach, inherited from quasilinear hyperbolic problems, relies on the energy estimates. In [18], Kato studied the mKdV (1.2) from a viewpoint of a hyperbolic equation and proved its local well-posedness in $H^s(\mathbb{R})$, $s > \frac{3}{2}$. In Kato's proof, the dispersive part ∂_x^3 did not play any role. In [20], Kenig–Ponce–Vega exploited the dispersive nature of the equation in the form of local smoothing and maximal function estimates and proved local well-posedness of (1.2) in $H^s(\mathbb{R})$, $s \geqslant \frac{1}{4}$, via the contraction mapping principle. See also [41] for another proof of the local well-posedness in $H^{\frac{1}{4}}(\mathbb{R})$, utilizing the Fourier restriction norm method (i.e. the $X^{s,b}$-spaces). By combining the $X^{s,b}$-spaces with weights and the I-method, Kishimoto [24] then

(1) Here, we only write the Duhamel formulation of the cubic NLS (1.1).

proved global well-posedness of (1.2) in $H^{\frac{1}{4}}(\mathbb{R})$.(2) Note that uniqueness in [14, 20, 24, 41] holds only conditionally. More recently, by combining the Fourier restriction norm method and the energy method, Molinet–Pilod–Vento [33] proved unconditional well-posedness of (1.2) in $H^s(\mathbb{R})$, $s > \frac{1}{3}$.

In the following, we present a new method for proving well-posedness of (1.1) and (1.2) on the real line. More precisely, we apply normal form reductions to the equation infinitely many times and transform it into an new equation. While this new equation involves nonlinear terms of arbitrarily high degrees, it turns out that these nonlinear terms can be estimated in a rather straightforward manner by successive applications of a basic trilinear estimate (called a localized modulation estimate) *without* using any auxiliary function spaces such as the Strichartz spaces and the $X^{s,b}$-spaces. As a result, we obtain the following unconditional well-posedness of (1.1) and (1.2).

Theorem 1.1. — *Let $s \geqslant \frac{1}{6}$. Then, the cubic NLS* (1.1) *is unconditionally globally well-posed in $H^s(\mathbb{R})$.*

Theorem 1.2. — *Let $s > \frac{1}{4}$. Then, the mKdV* (1.2) *is unconditionally globally well-posed in $H^s(\mathbb{R})$.*

Theorem 1.2 for the mKdV (1.2) extends the previous unconditional uniqueness result in $H^s(\mathbb{R})$, $s > \frac{1}{3}$, by Molinet–Pilod–Vento [33] to $s > \frac{1}{4}$, thus almost matching the local well-posedness result in $H^{\frac{1}{4}}(\mathbb{R})$ [20, 41]. On the other hand, Theorem 1.1 for the cubic NLS (1.1) was already proven in [19]. Let us stress, however, that the main purpose of this paper is to introduce a new method for constructing solutions to nonlinear dispersive PDEs on the real line (and on $\mathbb{R}^d$ in general) via (a simplified approach to) an infinite iteration of normal form reductions, which can be applied to different classes of dispersive equations. In our previous work [15], we introduced an infinite iteration of normal form reductions to construct solutions to nonlinear dispersive PDEs in the periodic setting. In particular, we proved unconditional uniqueness of the cubic NLS on the circle $\mathbb{T}$ in $H^s(\mathbb{T})$, $s \geqslant \frac{1}{6}$. On the one hand, the present work can be viewed as an extension of [15] to the non-periodic case. On the other hand, novelty of this work lies in presenting a simplified approach in treating multilinear estimates appearing in this normal form approach.

(2) In [24], Kishimoto first proved an endpoint local well-posedness of the KdV equation in $H^{-\frac{3}{4}}(\mathbb{R})$ and then combined it with the I-method and the Miura transform to establish global well-posedness of the mKdV (1.2) in $H^{\frac{1}{4}}(\mathbb{R})$. In [14], Guo independently proved local well-posedness of the KdV equation in $H^{-\frac{3}{4}}(\mathbb{R})$. His argument, however, does not seem to lead to the claimed global well-posedness as it is presented in [14] due to the use of the function spaces non-compatible with the I-method.

In order to make sense of the cubic nonlinearity in (1.1) or (1.2) as a distribution, we need to have $u \in L^3_{\text{loc}}(\mathbb{R})$. In view of the embedding: $H^{\frac{1}{6}}(\mathbb{R}) \subset L^3(\mathbb{R})$, we see that $s \geqslant \frac{1}{6}$ is necessary for proving unconditional uniqueness for (1.1) and (1.2) within the framework of the L^2-based Sobolev spaces. Moreover, it is known that the solution map for the mKdV (1.2): $u_0 \in H^s(\mathbb{R}) \mapsto u \in C([-T,T];H^s(\mathbb{R}))$ fails to be locally uniformly continuous for $s < \frac{1}{4}$ [6, 21]. Noting that a Picard iteration yields smoothness of a solution map, we see that the regularity restriction $s \geqslant \frac{1}{4}$ is needed to prove local well-posedness of (1.2) via a Picard iteration (even with conditional uniqueness). Hence, Theorems 1.1 and 1.2 are (almost) sharp in the regime where a Picard iteration is applicable by some other consideration. We also point out that well-posedness of (1.1) for $s < 0$ and (and (1.2) for $s < \frac{1}{4}$, respectively) is a long-standing open problem. See [7, 8, 27, 28] for existence results (without uniqueness) below these threshold regularities.

While we need $s \geqslant \frac{1}{6}$ in order to make sense of the cubic nonlinearity $\mathcal{N}_{\text{NLS}}(u) := |u|^2 u$ as a distribution, our normal form argument allows us to establish an a priori bound on the difference of two (smooth) solutions for the cubic NLS (1.1) in $L^2(\mathbb{R})$. This allows us to establish an existence result of certain weak solutions. Before we state our next result, let us recall the following two notions of weak solutions.

We first recall the notion of *weak solutions in the extended sense*. See [4, 5, 15].

Definition 1.3. — *Let* $0 \leqslant s < \frac{1}{6}$ *and* $T > 0$.

(i) *We define a sequence of Fourier cutoff operators to be a sequence of Fourier multiplier operators* $\{T_N\}_{N\in\mathbb{N}}$ *on* $\mathcal{S}'(\mathbb{R})$ *with multipliers* $m_N : \mathbb{R} \to \mathbb{C}$ *such that*
   - m_N *has a compact support on* $\mathbb{R}$ *for each* $N \in \mathbb{N}$,
   - m_N *is uniformly bounded,*
   - m_N *converges pointwise to* 1, *i.e.* $\lim_{N\to\infty} m_N(\xi) = 1$ *for any* $\xi \in \mathbb{R}$.

(ii) *Let* $u \in C([-T,T];H^s(\mathbb{R}))$. *We say that* $\mathcal{N}_{\text{NLS}}(u)$ *exists and is equal to a distribution* $v \in \mathcal{S}'(\mathbb{R} \times (-T,T))$ *if for every sequence* $\{T_N\}_{N\in\mathbb{N}}$ *of (spatial) Fourier cutoff operators, we have*

$$\lim_{N\to\infty} \mathcal{N}_{\text{NLS}}(T_N u) = v$$

in the sense of distributions on $\mathbb{R} \times (-T,T)$.

(iii) *(weak solutions in the extended sense) We say that* $u \in C([-T,T];H^s(\mathbb{R}))$ *is a weak solution of the cubic NLS* (1.1) *in the extended sense if*
   - $u|_{t=0} = u_0$,

- *the nonlinearity* $\mathcal{N}_{\mathrm{NLS}}(u)$ *exists in the sense of* (ii) *above,*
- u *satisfies* (1.1) *in the distributional sense on* $\mathbb{R} \times (-T, T)$*, where the nonlinearity* $\mathcal{N}_{\mathrm{NLS}}(u)$ *is interpreted as above.*

See also [13] for a similar notion of weak solutions, where the nonlinearity is defined as a distributional limit of smoothed nonlinearities.

Next, we introduce the following notion of *sensible weak solutions.* See [12, 37, 39].

Definition 1.4 (sensible weak solutions). — *Let* $0 \leqslant s < \frac{1}{6}$ *and* $T > 0$*. Given* $u_0 \in H^s(\mathbb{R})$*, we say that* $u \in C([-T, T]; H^s(\mathbb{R}))$ *is a sensible weak solution to the cubic NLS* (1.1) *on* $[-T, T]$ *if, for any sequence* $\{u_{0,m}\}_{m\in\mathbb{N}}$ *of Schwartz functions tending to* u_0 *in* $H^s(\mathbb{R})$*, the corresponding Schwartz solutions* u_m *with* $u_m|_{t=0} = u_{0,m}$ *converge to* u *in* $C([-T, T]; H^s(\mathbb{R}))$*. Moreover, we impose that there exists a distribution* v *such that* $\mathcal{N}_{\mathrm{NLS}}(u_m)$ *converges to* v *in the space-time distributional sense, independent of the choice of the approximating sequence.*

By using the equation, the convergence of u_m to u in $C([-T, T]; H^s(\mathbb{R}))$ implies that $\mathcal{N}_{\mathrm{NLS}}(u_m)$ converges to some v in the space-time distributional sense. Hence, the last part of Definition 1.4 is not quite necessary. We, however, keep it for clarity.

Note that sensible weak solutions are unique by definition. See [17, 22] for analogous notions of solutions (with uniqueness embedded in the definition). On the other hand, weak solutions in the extended sense are not unique in general. In fact, Christ [5] proved non-uniqueness of weak solutions in the extended sense for the renormalized cubic NLS on $\mathbb{T}$ in negative Sobolev spaces. These notions of weak solutions in Definitions 1.3 and 1.4 are rather weak and we need to interpret the cubic nonlinearity $\mathcal{N}_{\mathrm{NLS}}(u)$ as a (unique) limit of smoothed nonlinearities $\mathcal{N}_{\mathrm{NLS}}(T_N u)$ or the nonlinearities $\mathcal{N}_{\mathrm{NLS}}(u_m)$ of smooth approximating solutions u_m. This in particular implies that weak solutions in the sense of Definitions 1.3 or 1.4 do not have to satisfy the equation even in the distributional sense.

Our normal form approach yields the following result without relying on any auxiliary function spaces.

Theorem 1.5. — *Let* $0 \leqslant s < \frac{1}{6}$*. Then, the cubic NLS* (1.1) *is globally well-posed in* $H^s(\mathbb{R})$

- *in the sense of weak solutions in the extended sense and*
- *in the sense of sensible weak solutions.*

As for the mKdV (1.2), our normal form argument provides an a priori bound in $H^{\frac{1}{4}}(\mathbb{R})$. In this regularity, the cubic nonlinearity $\partial_x(u^3)$ makes sense as a distribution and thus we do not need the notion of weak solutions in the extended sense in Definition 1.3. On the other hand, we can define sensible weak solutions to the mKdV (1.2) as in Definition 1.4.

THEOREM 1.6. — *The mKdV* (1.2) *is globally well-posed in $H^{\frac{1}{4}}(\mathbb{R})$ in the sense of sensible weak solutions. These solutions are indeed distributional solutions to* (1.2).

Note that solutions constructed in Theorems 1.5 and 1.6 agree with those from the previous well-posedness results in [20, 41, 42]. This easily follows from the unconditional uniqueness in higher regularities (for example, in Theorems 1.1 and 1.2) and the conditional well-posedness results in low regularities [20, 41, 42], which provides uniqueness as a limit of classical solutions. We point out, however, that the importance of Theorems 1.5 and 1.6 does not lie in their statements but in the method of the construction of solutions. Our normal form approach transforms the equations (1.1) and (1.2) to the normal form equations (see (1.8) and (3.38)), at least for smooth solutions belonging to $H^s(\mathbb{R})$ with $s \geqslant \frac{1}{6}$ for the cubic NLS and $s > \frac{1}{4}$ for the mKdV. We then prove unconditional global well-posedness of the normal form equations in $H^s(\mathbb{R})$ with the regularities specified in Theorems 1.5 and 1.6, i.e. $s \geqslant 0$ for the cubic NLS and $s \geqslant \frac{1}{4}$ for the mKdV. See Theorem 3.18 below. Then, Theorems 1.5 and 1.6 follow as corollaries to this unconditional well-posedness on the normal form equation.

Lastly, note that while Theorems 1.1, 1.2, 1.5, and 1.6 claim global-in-time results, it suffices to prove these theorems only locally in time thanks to the (conditional) global well-posedness [10, 24, 42]. More precisely, in the following, we perform local-in-time construction of solutions on a time interval of length $T = T(\|u_0\|_{H^s}) > 0$ with $s \geqslant 0$ for the cubic NLS (1.1) and $s \geqslant \frac{1}{4}$ for the mKdV (1.2). Noting that the global well-posedness results in [10, 24, 42] provide an a priori estimate of the form: $\sup_{t\in[-T,T]} \|u(t)\|_{H^s} \lesssim C(\|u_0\|_{H^s}, T)$ for any $T > 0$, we can simply iterate the local-in-time argument to prove Theorems 1.1, 1.2, 1.5, and 1.6. Since our analysis is of local-in-time nature, the focusing/defocusing nature of the equations does not play any role. Hence, we assume that the equations are defocusing, i.e. with the $-$ signs in (1.1) and (1.2).

Remark 1.7. — In [38], Y. Wang and the second author introduced the notion of enhanced uniqueness, which is uniqueness among all solutions (with the same initial data) equipped with *some* smooth approximating solutions. They used an infinite iteration of normal form reductions for the fourth order cubic NLS (4NLS) in negative Sobolev spaces and proved such enhanced uniqueness. This notion of enhanced uniqueness allows us to compare solutions belonging to various auxiliary functions spaces (so that the cubic nonlinearity makes sense in some appropriate manner). On the one hand, this notion was useful in [38] since there was no known (conditional) well-posedness for 4NLS in negative Sobolev spaces at that time. We point out, however, that such notion becomes useless once we have (i) conditional well-posedness in relevant low regularity and (ii) unconditional well-posedness in high regularities. In such a case, this notion of enhanced uniqueness coincides with uniqueness as a limit of classical solutions. This is precisely the situation for the cubic NLS and the mKdV under consideration.

1.2. Normal form approach

In this subsection, we briefly explain our strategy for proving Theorems 1.1, 1.2, 1.5, and 1.6. As mentioned above, our main tool is the normal form method. In particular, we apply normal form reductions to (1.1) and (1.2) infinitely many times to transform them into new equations. These new equations involve infinite series of nonlinearities of arbitrarily high degrees and thus are more complicated algebraically than the original equations. As we see later, however they are easier to handle analytically. Namely, we renormalize the equations into analytically simpler equations at the expense of algebraic and notational complexity.

In the following, we consider the cubic NLS (1.1) as an example. Letting $v(t) = e^{it\partial_x^2}u(t)$ denote the interaction representation of u, we can rewrite the equation (1.1) as[(3)]

$$\partial_t v = \mathcal{N}(v) := \mathcal{F}^{-1}\left\{ i \int_{\xi=\xi_1-\xi_2+\xi_3} e^{-i\Phi(\bar{\xi})t} \prod_{j=1}^{3} \widehat{v}(\xi_j, t) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \right\}, \tag{1.3}$$

where the modulation function[(4)] $\Phi(\bar{\xi})$ is defined by

$$\begin{aligned}\Phi(\bar{\xi}) = \Phi(\xi, \xi_1, \xi_2, \xi_3) &= \xi^2 - \xi_1^2 + \xi_2^2 - \xi_3^2 \\ &= 2(\xi_2 - \xi_1)(\xi_2 - \xi_3) = 2(\xi - \xi_1)(\xi - \xi_3). \end{aligned} \tag{1.4}$$

Note that the last two equalities hold under the condition $\xi = \xi_1 - \xi_2 + \xi_3$. We point out that it is natural to consider the equation in terms of the interaction representation if we want to exploit the oscillatory factor $e^{-i\Phi(\bar{\xi})t}$ in (1.3).

(3) For simplicity of the exposition, we drop the complex conjugate sign on $\widehat{v}(\xi_2)$.

(4) In [41], this phase function is referred to as a resonance function. For our analysis, resonance does not play any important role. Instead, modulation (as in the Fourier restriction norm method) plays an important role. For this reason, we refer to $\Phi(\bar{\xi})$ as a modulation function.

Such a formulation in terms of the interaction representation is classical and already appears in the work of Kato [18] in the context of the (generalized) KdV equation. By integrating (1.3) in time, we obtain

$$v(t) = u_0 + \int_0^t \mathcal{N}(v)(t')\mathrm{d}t'. \tag{1.5}$$

On the one hand, when $s > \frac{1}{2}$, we can easily estimate (1.5) by the algebra property of $H^s(\mathbb{R})$. On the other hand, when $s \leqslant \frac{1}{2}$, we must exploit the dispersion, namely, the oscillation coming from the oscillatory factor $e^{-i\Phi(\bar{\xi})t}$ in (1.3). This is often manifested in the form of the Strichartz estimates and/or the Fourier restriction norm method. In the following, we simply rely on integration by parts. By taking the spatial Fourier transform of (1.5) and (formally) integrating by parts,[5] we have

$$\begin{aligned} \widehat{v}(\xi) = \widehat{u}_0(\xi) - \int_{\xi=\xi_1-\xi_2+\xi_3} \frac{e^{-i\Phi(\bar{\xi})t'}}{\Phi(\bar{\xi})} \prod_{j=1}^3 \widehat{v}(\xi_j,t')\mathrm{d}\xi_1\mathrm{d}\xi_2\bigg|_{t'=0}^{t} \\ + \int_0^t \int_{\xi=\xi_1-\xi_2+\xi_3} \frac{e^{-i\Phi(\bar{\xi})t'}}{\Phi(\bar{\xi})} \partial_t \left(\prod_{j=1}^3 \widehat{v}(\xi_j,t') \right) \mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}t'. \end{aligned} \tag{1.6}$$

Note that we have gained a full power of the modulation thanks to $\Phi(\bar{\xi})$ in the denominator. Compare this with the usual application of the Fourier restriction norm method where one only gains $\sim \frac{1}{2}$-power of the modulation.

At this point, there are several issues in (1.6). First, note that the modulation function $\Phi(\bar{\xi})$ appearing in the denominator may be 0. This corresponds to the so-called *resonance*. Even if $\Phi(\bar{\xi}) \neq 0$, integration by parts does not seem to help if $\Phi(\bar{\xi})$ is small, corresponding to the *nearly resonant* case. In order to resolve this issue, we separately estimate the contributions from (i) nearly resonant case: $|\Phi(\bar{\xi})| \leqslant N$ and (ii) (highly) non-resonant case: $|\Phi(\bar{\xi})| > N$ for some parameter $N = N(\|u_0\|_{H^s}) > 1$. In particular, we perform integration by parts only in the non-resonant case (ii). Thanks to the restriction on the modulation, we can estimate the contribution from the nearly resonant case (i) in $C_tH_x^s$, $s \geqslant 0$ (and $s \geqslant \frac{1}{4}$ for the mKdV), in a straightforward manner. See Lemmas 2.3 and 2.6.

The second issue is that we have increased the degree of the nonlinearity in (1.6). In view of (1.3), the last term in (1.6) is now quintic. Indeed, by

(5) In fact, this integration by parts basically corresponds to the (Poincaré–Dulac) normal form reduction. See the introduction of [15] by the first two authors (with Z. Guo), relating the integration-by-parts (or differentiation-by-parts) procedure with the normal form reductions. See Arnold [1] for a general discussion of the Poincaré–Dulac normal form reductions in the finite dimensional setting.

assuming that the time derivative falls on the first factor, we can write the last term in (1.6) as

$$
\begin{aligned}
&\sim \int_0^t \int_{\xi=\xi_1-\xi_2+\xi_3} \frac{e^{-i\Phi(\bar{\xi})t'}}{\Phi(\bar{\xi})} \widehat{\mathcal{N}}(v)(\xi_1, t') \prod_{j=2}^{3} \widehat{v}(\xi_j, t') \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}t' \\
&\sim \int_0^t \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ \xi_1=\zeta_1-\zeta_2+\zeta_3}} \frac{e^{-i(\Phi(\bar{\xi})+\Phi(\bar{\zeta}))t'}}{\Phi(\bar{\xi})} \prod_{k=1}^{3} \widehat{v}(\zeta_k, t') \prod_{j=2}^{3} \widehat{v}(\xi_j, t') \mathrm{d}\zeta_1 \mathrm{d}\zeta_2 \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}t',
\end{aligned}
\tag{1.7}
$$

where $\Phi(\bar{\zeta}) := \Phi(\xi_1, \zeta_1, \zeta_2, \zeta_3)$. The main idea is to perform integration by parts once again. In order to exploit the oscillation of $e^{-i(\Phi(\bar{\xi})+\Phi(\bar{\zeta}))t'}$, we separately estimate the contributions from (i) nearly resonant case: $|\Phi(\bar{\xi}) + \Phi(\bar{\zeta})| \leqslant N_1$ and (ii) non-resonant case: $|\Phi(\bar{\xi}) + \Phi(\bar{\zeta})| > N_1$ for some suitable threshold $N_1 > 1$.[(6)] Then, we integrate (1.7) by parts only in the non-resonant case (ii), thus introducing a septic nonlinearity.

By formally iterating this procedure indefinitely, we arrive at the following *normal form equation*:

$$
v(t) = u_0 + \sum_{j=2}^{\infty} \mathcal{N}_0^{(j)}(v(t'))\bigg|_{t'=0}^{t} + \int_0^t \sum_{j=1}^{\infty} \mathcal{N}_1^{(j)}(v(t'))\, \mathrm{d}t', \tag{1.8}
$$

where $\mathcal{N}_0^{(j)}(v)$ and $\mathcal{N}_1^{(j)}(v)$ are $(2j-1)$- and $(2j+1)$-multilinear terms, respectively. See (3.38) below. These multilinear terms $\mathcal{N}_0^{(j)}(v)$ and $\mathcal{N}_1^{(j)}(v)$ appear as a result of $(j-1)$-many iterations of the normal form reductions. Then, the main task is to estimate each term of the infinite series in (1.8) in the $C_t H_x^s$-norm in a summable manner. There are, however, three potential difficulties:

(1) The degrees of the nonlinearities can be arbitrarily high.
(2) In performing integration by parts in the Jth step, the number of factors on which the time derivative falls is $2J+1$. Thus, the constants grow like $3 \cdot 5 \cdot 7 \cdot \ldots \cdot (2J+1)$.
(3) Our multilinear estimates need to provide small constants on the terms *without* time integration, i.e. on the boundary terms, such as the second term on the right-hand side of (1.6) and $\mathcal{N}_0^{(j)}(v)$ in (1.8). (We can introduce small constants for the terms inside time integration by making the time interval of integration sufficiently short.)

In Section 3, we will treat these issues and prove that the normal form equation is unconditionally well-posed (Theorem 3.18). Theorems 1.1, 1.2,

(6) As we see later, we choose $N_1 \sim |\Phi(\bar{\xi})|^{1-\delta}$ for some $\delta \in (0,1)$. See (3.7).

1.5, and 1.6 then follow as corollaries to this unconditional well-posedness of the normal form equation.

In [15], we implemented an infinite iteration of normal form reductions sketched above in the context of the cubic NLS on the circle $\mathbb{T}$. In particular, we introduced the notion of ordered trees (see Definition 3.3) and indexed all the multilinear terms by such ordered trees, handling the issues (1), (2), and (3). Moreover, in handling the multilinear estimates, we exploited the discrete structure of the spatial frequency space $\mathbb{Z} = (\mathbb{T})^*$ in the form of the divisor counting argument. In the non-periodic setting, such number theoretic tools are no longer available. In this paper, we change our viewpoint and view these multilinear terms as iterative compositions of trilinear operators (see Definition 3.13 and (3.26)) with modulation restrictions. We first establish trilinear localized modulation estimates in Section 2 as a fundamental building block. Then, by applying such trilinear localized modulation estimates in an iterative manner, we estimate the multilinear terms of arbitrarily high degrees, appearing in (1.8). This provides a simplified framework for implementing an infinite iteration of normal form reductions.[7]

Lastly, let us mention the role of two different topologies for this normal form argument. Roughly speaking, we

(i) establish a priori estimates in a stronger topology (in $H^s(\mathbb{R})$ with $s \geqslant 0$ for the cubic NLS and $s \geqslant \frac{1}{4}$ for the mKdV) but
(ii) justify all the formal computations in a weaker topology (in the Fourier–Lebesgue space $\mathcal{F}L^\infty(\mathbb{R})$ defined in (1.9) below) for *smoother* solutions ($s \geqslant \frac{1}{6}$ for the cubic NLS and $s > \frac{1}{4}$ for the mKdV), thus making sense of the identity (1.8) in the distributional sense.

By formally performing an infinite iteration of normal form reductions, we derive the normal form equation (1.8) in Section 3. In establishing a priori estimates in $H^s(\mathbb{R})$, we estimate each multilinear term in the H^s-norm with $s \geqslant 0$ for the cubic NLS and $s \geqslant \frac{1}{4}$ for the mKdV. In Section 4, we justify all the formal computations performed in Section 3, in particular the integration-by-parts steps, where we switch time derivatives and integrations over spatial frequencies. See (1.6) for example. For this purpose, we work in a weaker topology. Indeed, we justify all the steps of the normal form reductions for each *fixed* frequency $\xi \in \mathbb{R}$ of the interaction representation $\widehat{v}(\xi)$ (and hence of each multilinear term in (1.8)). It is in this step where we need

[7] During the preparation of this manuscript, we learned that Kishimoto [25, 26] independently used a similar abstraction of a basic multilinear estimate as a fundamental building block in the application of an infinite iteration of normal form reductions to prove unconditional well-posedness for various dispersive PDEs in the periodic setting.

to assume a higher regularity: $s \geqslant \frac{1}{6}$ for the cubic NLS and $s > \frac{1}{4}$ for the mKdV. In the case of the mKdV, we also need to handle the derivative loss in the equation. In particular, in each step of the normal form reductions (i.e. integration by parts), we use the equation (1.2) to replace $\partial_t \widehat{v}$ by the cubic nonlinearity (see (2.14)), which introduces a derivative loss at each step. Since we work for each fixed $\xi \in \mathbb{R}$, the derivative loss in the first "generation" (i.e. in the original equation) does not cause any problem. We then *shift part of the derivative loss up by one generation* to reduce the derivative loss in the last generation. See Subsection 4.2 for a further discussion.

1.3. Remarks and comments

A precursor to this normal form approach appeared in the work of Babin–Ilyin–Titi [2] for the KdV on $\mathbb{T}$, establishing unconditional well-posedness of the KdV in $L^2(\mathbb{T})$. See also [30] for an analogous unconditional well-posedness result for the periodic mKdV in $H^{\frac{1}{2}}(\mathbb{T})$. Note that two iterations were sufficient in [2, 30]. In [15], we further developed this normal form approach and introduced an infinite iteration scheme of normal form reductions in the context of the cubic NLS on the circle. By performing normal form reductions infinitely many times, we proved unconditional well-posedness of the periodic cubic NLS in $H^{\frac{1}{6}}(\mathbb{T})$. In this series of work, the viewpoint of unconditional well-posedness was first introduced in [30], while the viewpoint of the (Poincaré–Dulac) normal form reductions was first introduced in [15].

More recently, by combining an infinite iteration of normal form reductions and the Cole–Hopf transform, we proved unconditional global well-posedness for the quadratic derivative NLS on $\mathbb{T}$ for small mean-zero initial data [9]. Moreover, this method allowed us to construct an infinite sequence of invariant quantities under the dynamics. Kishimoto [25] adapted our infinite iteration approach and proved unconditional well-posedness for higher dimensional NLS, the Zakharov system on $\mathbb{T}^d$, $d = 1, 2$, the derivative cubic NLS on $\mathbb{T}$, the Benjamin–Ono and modified Benjamin–Ono equations in the periodic setting.

One may naturally expect that an infinite iteration of normal form reductions is needed to prove Theorem 1.1 for the cubic NLS on the real line just as in the periodic case [15].[(8)] It is, however, to our surprise to see that

[(8)] We point out recent works [40, 3] on the construction of solutions to the cubic NLS on the real line via an infinite iteration of normal form reductions. Their implementation of normal form reductions follows closely the original argument in [15] and unconditional uniqueness in modulation spaces (including Theorem 1.1 above) is established. In [12], this construction was extended to almost critical Fourier–Lebesgue and modulation spaces.

we also need to perform normal form reductions infinitely many times in proving Theorem 1.2 for the mKdV on the real line. This is in sharp contrast with the mKdV on the circle, where two iterations were sufficient [30]. In this paper, we chose to study the cubic NLS (1.1) and the mKdV (1.2) as canonical examples. As in the periodic case [25], our method of an infinite iteration scheme of normal form reductions is fairly general that it can be applied to study a wide variety of equations in the Euclidean space $\mathbb{R}^d$ of general dimensions.

This normal form approach has various applications beyond establishing unconditional uniqueness. It has been used to exhibit nonlinear smoothing [11], to prove a good approximation property in proving symplectic non-squeezing [16], and establishing effective energy estimates with smoothing in proving quasi-invariance of Gaussian measures on periodic functions under dispersive PDEs [36].(9) More recently, the second author introduced a way to perform normal form reductions infinitely many times in establishing energy estimates [35, 38]. In particular, the notion of ordered trees was extended to that of ordered bi-trees to accommodate normal form reductions on energy quantities. Note that such an infinite iteration of normal form reductions on an energy quantity basically amounts to adding infinitely many correction terms in the I-method terminology, going far beyond the known application of the I-method [10], where only finitely many correction terms were considered.

The main novelty of this paper is to reduce multilinear estimates to successive applications of a basic trilinear localized modulation estimate and in fact to reduce the entire problem of proving unconditional well-posedness to simply proving two basic trilinear estimates (i.e. localized modulation estimates in the strong norm and in the weak norm: Lemmas 2.3 and 4.1 for the cubic NLS and Lemma 2.6 and 4.9 for the mKdV). Such reduction can easily be implemented in the context of our previous work [15, 35, 38], except for [9] where the algebraic property of the equation played an important role inducing cancellation of resonant terms via symmetrization at each step of the normal form reductions. See also the concluding remark at the end of this paper.

This paper is organized as follows. In Section 2 we establish crucial trilinear estimates (localized modulation estimates) for the cubic NLS (1.1) and the mKdV (1.2). In Section 3, we perform an infinite iteration of normal form reductions and derive the normal form equation. We carry out a computation in Section 3 at a formal level. In Subsection 3.4, we prove unconditional local well-posedness of the normal form equation in $H^s(\mathbb{R})$ with $s \geqslant 0$ for the

(9) Such an application of normal form reductions in energy estimates is more classical and precedes the work of [2].

cubic NLS and $s \geqslant \frac{1}{4}$ for the mKdV (Theorem 3.18) and discuss the proofs of Theorems 1.1, 1.2, 1.5, and 1.6, assuming that smooth solutions satisfy the normal form equation. In Section 4, we justify the formal computation in Section 3 and then conclude the proofs of the main theorems.

Notations

We use $A \lesssim B$ to denote an estimate of the form $A \leqslant CB$ for some constant $C > 0$, which may vary from line to line and depend on various parameters. We also use $A \sim B$ to denote $A \lesssim B \lesssim A$, while we use $A \ll B$ to denote $A \leqslant \varepsilon B$ for some small absolute constant $\varepsilon > 0$. We use $a+$ to denote $a+\varepsilon$ for arbitrarily small $\varepsilon \ll 1$, where an implicit constant is allowed to depend on $\varepsilon > 0$ (and it usually diverges as $\varepsilon \to 0$).

Given a function f on $\mathbb{R}$, we define its Fourier transform by

$$\mathcal{F}(f)(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e^{-2\pi i x\xi} \mathrm{d}x.$$

We drop the harmless factor of 2π in the following. We define the Fourier–Lebesgue space $\mathcal{F}L^p(\mathbb{R})$, $p \geqslant 1$, by the norm:

$$\|f\|_{\mathcal{F}L^p} = \|\widehat{f}\|_{L^p}. \tag{1.9}$$

Any summation over capitalized variables such as $N_1, N_2, \ldots$ are presumed to be dyadic, i.e. these variables range over dyadic numbers of the form 2^k, $k \in \mathbb{Z}_{\geqslant 0}$. We also use the following shorthand notations: ξ_{ij} and ξ_{i-j} for $\xi_i + \xi_j$ and $\xi_i - \xi_j$, respectively.

Given dyadic $N \geqslant 1$, we use $\mathbf{P}_N$ to denote the Littlewood–Paley projector onto the spatial frequencies $\{|\xi| \sim N\}$. Given $k \in \mathbb{Z}$, we use Π_k to denote the (spatial) frequency projector onto the interval $[k, k+1)$:

$$\Pi_k v(\xi) = \mathbb{1}_{[k,k+1)}(\xi) \cdot v(\xi). \tag{1.10}$$

We use $S(t)$ to denote the linear propagator for the linear Schrödinger equation: $i\partial_t u = \partial_x^2 u$ and the Airy equation: $\partial_t u = \partial_x^3 u$, depending on the context. Namely, $S(t) = e^{-it\partial_x^2}$ for the linear Schrödinger equation and $S(t) = e^{t\partial_x^3}$ for the Airy equation. Then, given a function u on $\mathbb{R} \times \mathbb{R}$, we define its interaction representation v by

$$v(t) = S(-t)u(t). \tag{1.11}$$

We mainly perform our analysis in terms of the interaction representation.

In the following, we only consider positive times for the simplicity of the presentation.

2. Localized modulation estimates

In this section, we establish crucial trilinear estimates (called localized modulation estimates) for the cubic NLS (1.1) and the mKdV (1.2). See Lemmas 2.3 and 2.6. While their proofs are very elementary, these trilinear estimates constitute a fundamental building block for multilinear estimates on the nonlinear terms (of arbitrarily high degrees) appearing in the normal form reductions in Section 3.

2.1. Localized modulation estimates for the cubic NLS

We first consider the cubic NLS (1.1). On the Fourier side, we write (1.1) as

$$i\partial_t \widehat{u}(\xi) = -\xi^2 \widehat{u}(\xi) - \int_{\xi=\xi_1-\xi_2+\xi_3} \widehat{u}(\xi_1)\overline{\widehat{u}(\xi_2)}\widehat{u}(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2. \tag{2.1}$$

Let $v(t) = S(-t)u(t)$ be the interaction representation defined in (1.11). Then, we have $\widehat{v}(\xi,t) = e^{-i\xi^2 t}\widehat{u}(\xi,t)$.

Define a trilinear operator $\mathcal{N}(v_1,v_2,v_3)$ by

$$\widehat{\mathcal{N}}(v_1,v_2,v_3)(\xi,t) := i\int_{\xi=\xi_1-\xi_2+\xi_3} e^{-i\Phi(\bar{\xi})t}\widehat{v}_1(\xi_1)\overline{\widehat{v}_2(\xi_2)}\widehat{v}_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2, \tag{2.2}$$

where the modulation function $\Phi(\bar{\xi})$ is as in (1.4). With this notation, we can write (2.1) as

$$\partial_t v = \mathcal{N}(v,v,v). \tag{2.3}$$

Remark 2.1.

(i) When there is no confusion, we simply denote $\widehat{v}(\xi,t)$ and $\widehat{\mathcal{N}}(v_1,v_2,v_3)$ by $v(\xi,t)$ and $\mathcal{N}(v_1,v_2,v_3)$ in the following. For example, we write (2.2) as

$$\mathcal{N}(v_1,v_2,v_3)(\xi,t) = i\int_{\xi=\xi_1-\xi_2+\xi_3} e^{-i\Phi(\bar{\xi})t}v_1(\xi_1)\overline{v_2(\xi_2)}v_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2$$

under this convention. Note that while the equation (2.3) can be interpreted as an equation on the physical side or on the Fourier side under this convention, this does not cause any confusion in terms of its meaning. A similar comment applies to other multilinear operators.

(ii) Due to the presence of the time-dependent phase factor $e^{-i\Phi(\bar{\xi})t}$, the trilinear expression $\mathcal{N}(v_1, v_2, v_3)$ is non-autonomous and in fact depends on t. For simplicity of notations, however, we suppress such t-dependence when there is no confusion. We also set $\mathcal{N}(v) = \mathcal{N}(v, v, v)$, when all the three arguments are identical. We apply this convention to all the multilinear operators appearing in this paper.

For $M \geqslant 1$ and $\alpha \in \mathbb{R}$, we also define trilinear operators $\mathcal{N}^\alpha_{\leqslant M}$, $\mathcal{N}^\alpha_{>M}$, and $\mathcal{N}^\alpha_M$ with modulation restrictions:

$$\mathcal{N}^\alpha_{\leqslant\,(>)\,M}(v_1, v_2, v_3)(\xi, t) := i \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ |\Phi(\bar{\xi})-\alpha|\leqslant\,(>)\,M}} e^{-i\Phi(\bar{\xi})t} v_1(\xi_1)\overline{v_2(\xi_2)}v_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2, \tag{2.4}$$

$$\begin{aligned}\mathcal{N}^\alpha_M(v_1, v_2, v_3)(\xi, t) &:= \mathcal{N}^\alpha_{\leqslant 2M}(v_1, v_2, v_3)(\xi, t) - \mathcal{N}^\alpha_{\leqslant M}(v_1, v_2, v_3)(\xi, t) \\ &= i \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ |\Phi(\bar{\xi})-\alpha|\sim M}} e^{-i\Phi(\bar{\xi})t} v_1(\xi_1)\overline{v_2(\xi_2)}v_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2,\end{aligned}$$

where $|\Phi(\bar{\xi}) - \alpha| \sim M$ is a shorthand notation for $M < |\Phi(\bar{\xi}) - \alpha| \leqslant 2M$. The following trilinear operator also plays an important role in our analysis:

$$\mathcal{I}^\alpha_M(v_1, v_2, v_3)(\xi, t) := i \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ |\Phi(\bar{\xi})-\alpha|\sim M}} \frac{e^{-i\Phi(\bar{\xi})t}}{\Phi(\bar{\xi}) - \alpha} v_1(\xi_1)\overline{v_2(\xi_2)}v_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2.$$

We also define $\mathcal{I}^\alpha_{>M}$ in an obvious manner. In the subsequent part of this paper, we use the following conventions:

- When $\alpha = 0$, we drop the superscript and simply write $\mathcal{N}_M, \mathcal{N}_{\leqslant M}$, $\dots$, for $\mathcal{N}^0_M, \mathcal{N}^0_{\leqslant M}, \dots$.
- In Section 3, these multilinear operators appear in an iterative manner. For clarity, we often write $\mathcal{N}^\alpha_{|\Phi(\bar{\xi})-\alpha|\sim M}$ for $\mathcal{N}^\alpha_M$, thus explicitly showing the variable of restriction.

Remark 2.2. — Recall that the (time) resonance corresponds to $\Phi(\bar{\xi}) = 0$. Thus, the term $\mathcal{N}^0_{\leqslant M}$ corresponds to the *nearly resonant* contribution to the nonlinearity $\mathcal{N}$ (with the cutoff size M).

We now state the localized modulation estimates for the cubic NLS. These trilinear estimates play a key role in our analysis in Section 3.

Lemma 2.3 (Localized modulation estimates for the cubic NLS). — *Let $s \geqslant 0$. Then, we have*

$$\|\mathcal{N}^{\alpha}_{\leqslant M}(v_1, v_2, v_3)\|_{H^s} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^{3} \|v_j\|_{H^s}, \tag{2.5}$$

$$\|\mathcal{N}^{\alpha}_{\leqslant M}(v) - \mathcal{N}^{\alpha}_{\leqslant M}(w)\|_{H^s} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+} \big(\|v\|^2_{H^s} + \|w\|^2_{H^s}\big) \|v - w\|_{H^s}, \tag{2.6}$$

for any $M \geqslant 1$ and $\alpha \in \mathbb{R}$.

Remark 2.4. — Recall that the trilinear operator $\mathcal{N}^{\alpha}_{\leqslant M}(v_1, v_2, v_3)$ depends on $t \in \mathbb{R}$ in an non-autonomous manner. Hence, strictly speaking, we should have written the first estimate (2.5) as

$$\sup_{t \in \mathbb{R}} \|\mathcal{N}^{\alpha}_{\leqslant M}(v_1, v_2, v_3)\|_{H^s} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^{3} \|v_j\|_{H^s}.$$

Note that, in the definition (2.4), the non-autonomous parameter $t \in \mathbb{R}$ appears only in the oscillatory factor $e^{-i\Phi(\bar{\xi})t}$. We, however, do not make use of this oscillatory factor in the proof of (2.5). See (2.7) below. In particular, (2.5) holds uniformly in $t \in \mathbb{R}$. In view of this observation, we simply write (2.5) with the understanding that the estimate holds uniformly in the non-autonomous parameter $t \in \mathbb{R}$. We use this convention for all the multilinear estimates appearing in this paper.

Let us also note that the "spatial" estimate (2.5) immediately implies the following space-time estimate:

$$\|\mathcal{N}^{\alpha}_{\leqslant M}(v_1, v_2, v_3)\|_{L^\infty_T H^s_x} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^{3} \|v_j\|_{L^\infty_T H^s_x}$$

for all $v_j \in L^\infty([-T, T]; H^s(\mathbb{R}))$. The same remark also applies to the other multilinear estimates.

Proof. — In the following, we only present the proof of (2.5), since the second estimate (2.6) on the difference follows from (2.5) and the multilinearity of $\mathcal{N}^{\alpha}_{\leqslant M}$. By the triangle inequality with $s \geqslant 0$, we have $\langle \xi \rangle^s \lesssim \langle \xi_1 \rangle^s \langle \xi_2 \rangle^s \langle \xi_3 \rangle^s$ under $\xi = \xi_1 - \xi_2 + \xi_3$. Hence, it suffices to prove (2.5) for $s = 0$.

By duality, the desired estimate (2.5) follows once we prove

$$\left| \int_{\xi = \xi_1 - \xi_2 + \xi_3} \mathbb{1}_{|\Phi(\bar{\xi}) - \alpha| \leqslant M} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) v_4(\xi) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}\xi \right| \\ \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^{4} \|v_j\|_{L^2} \tag{2.7}$$

for all non-negative functions $v_1, \dots, v_4 \in L^2_\xi(\mathbb{R})$.

Case 1: $\min(|\xi_{2-1}|, |\xi_{2-3}|) \leqslant 1$. — Let $\zeta = \xi_2 - \xi_1 = \xi_3 - \xi$. Without loss of generality, we assume that $|\zeta| \leqslant 1$. Then, it follows from Hölder's inequality that

$$\begin{aligned}
\text{LHS of (2.7)} &= \bigg| \int_{|\zeta|\leqslant 1} \int_{\xi_1} v_1(\xi_1) v_2(\xi_1+\zeta) \mathrm{d}\xi_1 \int_{\xi_3} v_3(\xi_3) v_4(\xi_3 - \zeta) \mathrm{d}\xi_3 \mathrm{d}\zeta \bigg| \\
&\leqslant \bigg\| \int_{\xi_1} v_1(\xi_1) v_2(\xi_1+\zeta) \mathrm{d}\xi_1 \bigg\|_{L^\infty_\zeta} \bigg\| \int_{\xi_3} v_3(\xi_3) v_4(\xi_3 - \zeta) \mathrm{d}\xi_3 \bigg\|_{L^\infty_\zeta} \\
&\lesssim \prod_{j=1}^4 \|v_j\|_{L^2}.
\end{aligned}$$

This proves (2.7).

Case 2: $\min(|\xi_{2-3}|, |\xi_{2-1}|) > 1$. — Without loss of generality, assume that $\xi - \xi_3 > 1$. Under $|\Phi(\bar{\xi}) - \alpha| \leqslant M$, it follows from (1.4) that

$$\frac{\alpha - M}{2(\xi - \xi_3)} \leqslant \xi - \xi_1 \leqslant \frac{\alpha + M}{2(\xi - \xi_3)}. \tag{2.8}$$

Then, by the standard Cauchy–Schwarz argument with (2.8), we have

$$\begin{aligned}
\text{LHS of (2.7)} &\leqslant \bigg\| \int_{\xi = \xi_1 - \xi_2 + \xi_3} \mathbb{1}_{|\Phi(\bar{\xi}) - \alpha| \leqslant M} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_3 \bigg\|_{L^2_\xi} \|v_4\|_{L^2} \\
&\leqslant \sup_\xi \bigg(\int_{\xi = \xi_1 - \xi_2 + \xi_3} \mathbb{1}_{|\Phi(\bar{\xi}) - \alpha| \leqslant M} \mathrm{d}\xi_1 \mathrm{d}\xi_3 \bigg)^{\frac{1}{2}} \prod_{j=1}^4 \|v_j\|_{L^2} \\
&\leqslant \sup_\xi \bigg(\int_{1 < \xi - \xi_3 \lesssim |\alpha| + M} \frac{M}{\xi - \xi_3} \mathrm{d}\xi_3 \bigg)^{\frac{1}{2}} \prod_{j=1}^4 \|v_j\|_{L^2} \\
&\lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^4 \|v_j\|_{L^2},
\end{aligned} \tag{2.9}$$

where we used the assumption that $|\xi - \xi_1| > 1$ and $|\Phi(\bar{\xi})| \leqslant |\alpha| + M$ in the third inequality. This completes the proof of Lemma 2.3. $\square$

Next, we estimate the trilinear operators $\mathcal{I}^\alpha_M$ and $\mathcal{I}^\alpha_{>M}$.

Lemma 2.5. — *Let $s \geqslant 0$. Then, we have*

$$\|\mathcal{I}^\alpha_M(v)\|_{H^s} \lesssim \langle \alpha \rangle^{0+} M^{-\frac{1}{2}+} \|v\|^3_{H^s}, \tag{2.10}$$

$$\|\mathcal{I}^\alpha_M(v) - \mathcal{I}^\alpha_M(w)\|_{H^s} \lesssim \langle \alpha \rangle^{0+} M^{-\frac{1}{2}+} \big(\|v\|^2_{H^s} + \|w\|^2_{H^s} \big) \|v - w\|_{H^s}, \tag{2.11}$$

and

$$\|\mathcal{I}^\alpha_{>M}(v)\|_{H^s} \lesssim \langle\alpha\rangle^{0+} M^{-\frac{1}{2}+} \|v\|^3_{H^s}, \tag{2.12}$$

$$\|\mathcal{I}^\alpha_{>M}(v) - \mathcal{I}^\alpha_{>M}(w)\|_{H^s} \lesssim \langle\alpha\rangle^{0+} M^{-\frac{1}{2}+} \big(\|v\|^2_{H^s} + \|w\|^2_{H^s}\big) \|v-w\|_{H^s}, \tag{2.13}$$

for any $M \geqslant 1$ *and* $\alpha \in \mathbb{R}$.

Proof. — In the following, we only prove (2.10) and (2.12) since (2.11) and (2.13) follow in a similar manner.

Note that we did not exploit the oscillatory nature of the exponential factor $e^{-i\Phi(\bar{\xi})t}$ in the proof of Lemma 2.3. See (2.7). Hence, by Lemma 2.3, we have

$$\begin{aligned}
\|\mathcal{I}^\alpha_M(v)\|_{H^s} &= \bigg\| \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ |\Phi(\bar{\xi})-\alpha|\sim M}} \frac{e^{-i\Phi(\bar{\xi})t}}{\Phi(\bar{\xi})-\alpha} v(\xi_1)\overline{v(\xi_2)}v(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2 \bigg\|_{H^s} \\
&\lesssim \frac{1}{M} \bigg\| \int_{\xi=\xi_1-\xi_2+\xi_3} \mathbb{1}_{|\Phi(\bar{\xi})-\alpha|\sim M} \prod_{j=1}^{3} |v(\xi_j)| \mathrm{d}\xi_1\mathrm{d}\xi_2 \bigg\|_{H^s} \\
&\lesssim \langle\alpha\rangle^{0+} M^{-\frac{1}{2}+} \|v\|^3_{H^s}.
\end{aligned}$$

This proves (2.10). Similarly, we have

$$\begin{aligned}
\|\mathcal{I}^\alpha_{>M}(v)\|_{H^s} &= \bigg\| \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ |\Phi(\bar{\xi})-\alpha|>M}} \frac{e^{-i\Phi(\bar{\xi})t}}{\Phi(\bar{\xi})-\alpha} v(\xi_1)\overline{v(\xi_2)}v(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2 \bigg\|_{H^s} \\
&\leqslant \sum_{\substack{N\geqslant M \\ \text{dyadic}}} \bigg\| \int_{\substack{\xi=\xi_1-\xi_2+\xi_3 \\ N<|\Phi(\bar{\xi})-\alpha|\leqslant 2N}} \frac{e^{-i\Phi(\bar{\xi})t}}{\Phi(\bar{\xi})-\alpha} v(\xi_1)\overline{v(\xi_2)}v(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2 \bigg\|_{H^s} \\
&\lesssim \langle\alpha\rangle^{0+} \sum_{\substack{N\geqslant M \\ \text{dyadic}}} N^{-\frac{1}{2}+} \|v\|^3_{H^s} \\
&\lesssim \langle\alpha\rangle^{0+} M^{-\frac{1}{2}+} \|v\|^3_{H^s}.
\end{aligned}$$

This proves (2.12). □

2.2. Localized modulation estimates for the mKdV

In this subsection, we perform similar analysis on the mKdV (1.2) and establish localized modulation estimates on the relevant trilinear operators. Let $v(t) = S(-t)u(t)$ be the interaction representation defined in (1.11). Then, we have $\widehat{v}(\xi,t) = e^{i\xi^3 t}\widehat{u}(\xi,t)$.

Define a trilinear operator $\mathcal{N}(v_1, v_2, v_3)$ by[10]

$$\mathcal{N}(v_1, v_2, v_3)(\xi, t) := -i \int_{\xi=\xi_1+\xi_2+\xi_3} \xi e^{i\Psi(\bar{\xi})t} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2, \quad (2.14)$$

where the modulation function $\Psi(\bar{\xi})$ is given by

$$\begin{aligned}\Psi(\bar{\xi}) = \Psi(\xi, \xi_1, \xi_2, \xi_3) &= \xi^3 - \xi_1^3 - \xi_2^3 - \xi_3^3 \\ &= 3(\xi_1 + \xi_2)(\xi_2 + \xi_3)(\xi_3 + \xi_1). \quad (2.15)\end{aligned}$$

Here, the last equality holds under the condition $\xi = \xi_1 + \xi_2 + \xi_3$. With this notation, we can write the mKdV (1.2) as

$$\partial_t v = \mathcal{N}(v). \quad (2.16)$$

As before, we define several trilinear operators. Given $M \geqslant 1$ and $\alpha \in \mathbb{R}$, we let

$$\mathcal{N}^{\alpha}_{\leqslant(>)M}(v_1, v_2, v_3)(\xi, t) := -i \int_{\substack{\xi=\xi_1+\xi_2+\xi_3 \\ |\Psi(\bar{\xi})-\alpha| \leqslant (>) M}} \xi e^{i\Psi(\bar{\xi})t} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2,$$

$$\mathcal{N}^{\alpha}_{M}(v_1, v_2, v_3)(\xi, t) := -i \int_{\substack{\xi=\xi_1+\xi_2+\xi_3 \\ |\Psi(\bar{\xi})-\alpha| \sim M}} \xi e^{i\Psi(\bar{\xi})t} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2.$$

We also define the following trilinear operator:

$$\mathcal{I}^{\alpha}_{M}(v_1, v_2, v_3)(\xi, t) := -i \int_{\substack{\xi=\xi_1+\xi_2+\xi_3 \\ |\Psi(\bar{\xi})-\alpha| \sim M}} \frac{\xi e^{i\Psi(\bar{\xi})t}}{\Psi(\bar{\xi}) - \alpha} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2$$

and define $\mathcal{I}^{\alpha}_{>M}$ in an obvious manner.

We now present the localized modulation estimates for the mKdV. While the proof does not employ any sophisticated analytical tools, it is more involved than the proof of Lemma 2.3.

Lemma 2.6 (Localized modulation estimates for the mKdV). — *Let $s \geqslant \frac{1}{4}$. Then, we have*

$$\|\mathcal{N}^{\alpha}_{\leqslant M}(v_1, v_2, v_3)\|_{H^s} \lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^{3} \|v_j\|_{H^s}, \quad (2.17)$$

$$\|\mathcal{N}^{\alpha}_{\leqslant M}(v) - \mathcal{N}^{\alpha}_{\leqslant M}(w)\|_{H^s} \lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+} \big(\|v\|^2_{H^s} + \|w\|^2_{H^s}\big) \|v - w\|_{H^s}, \quad (2.18)$$

for any $M \geqslant 1$ and $\alpha \in \mathbb{R}$.

[10] We follow the conventions introduced in Remark 2.1.

Proof. — In the following, we only present the proof of (2.17), since the second estimate (2.18) on the difference follows from (2.17) and the multilinearity of $\mathcal{N}^{\alpha}_{\leqslant M}$. By the triangle inequality: $\langle\xi\rangle^\sigma \lesssim \langle\xi_1\rangle^\sigma\langle\xi_2\rangle^\sigma\langle\xi_3\rangle^\sigma$ for $\sigma \geqslant 0$ under $\xi = \xi_1 + \xi_2 + \xi_3$, it suffices to prove (2.17) for $s = \frac{1}{4}$.

By duality, the desired estimate (2.17) follows once we prove

$$\left| \int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m(\bar{\xi}) \prod_{j=1}^{3} v_j(\xi_j) v_4(\xi) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \mathrm{d}\xi \right| \lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+} \prod_{j=1}^{4} \|v_j\|_{L^2} \quad (2.19)$$

for all non-negative functions $v_1, \ldots, v_4 \in L^2_\xi(\mathbb{R})$, where the multiplier $m(\bar{\xi})$ is given by

$$m(\bar{\xi}) = m(\xi, \xi_1, \xi_2, \xi_3) = \frac{|\xi|\langle\xi\rangle^{\frac{1}{4}}}{\langle\xi_1\rangle^{\frac{1}{4}}\langle\xi_2\rangle^{\frac{1}{4}}\langle\xi_3\rangle^{\frac{1}{4}}}. \quad (2.20)$$

By the standard Cauchy–Schwarz argument, we have

$$\begin{aligned} &\text{LHS of (2.19)} \\ &\quad \leqslant \left\| \int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m(\bar{\xi}) \prod_{j=1}^{3} v_j(\xi_j) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \right\|_{L^2_\xi} \|v_4\|_{L^2} \\ &\quad \leqslant \sup_\xi \left(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2(\bar{\xi}) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \right)^{\frac{1}{2}} \prod_{j=1}^{4} \|v_j\|_{L^2}. \quad (2.21) \end{aligned}$$

Hence, it suffices to show that

$$\sup_\xi \left(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2(\bar{\xi}) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \right)^{\frac{1}{2}} \lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+}. \quad (2.22)$$

In the following, we either prove (2.22) or directly establish (2.19).

Case 1: $|\xi| \lesssim 1$. — By Cauchy–Schwarz, Hölder's, and Young's inequalities followed by Hölder's inequality, we have

$$\begin{aligned}
\text{LHS of (2.19)} &\lesssim \bigg\| \int_{\xi=\xi_1+\xi_2+\xi_3} \prod_{j=1}^{3} \langle \xi_j \rangle^{-\frac{1}{4}} v_j(\xi_j) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \bigg\|_{L^2_{|\xi|\lesssim 1}} \|v_4\|_{L^2_\xi} \\
&\lesssim \bigg\| \int_{\xi=\xi_1+\xi_2+\xi_3} \prod_{j=1}^{3} \langle \xi_j \rangle^{-\frac{1}{4}} v_j(\xi_j) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \bigg\|_{L^\infty_{|\xi|\lesssim 1}} \|v_4\|_{L^2} \\
&\lesssim \prod_{j=1}^{3} \| \langle \xi_j \rangle^{-\frac{1}{4}} v_j(\xi_j) \|_{L^{3/2}_{\xi_j}} \|v_4\|_{L^2} \lesssim \prod_{j=1}^{4} \|v_j\|_{L^2}.
\end{aligned}$$

In the following, we consider the case $|\xi| \gg 1$. Without loss of generality, we assume that $|\xi_{12}| \geqslant |\xi_{23}| \geqslant |\xi_{31}|$.

Case 2: $|\xi| \gg 1$ *and* $|\xi_{31}| \leqslant |\xi_{23}| \leqslant 1$. — In this case, we have $|\xi + \xi_3| = |\xi_{31} + \xi_{23}| \lesssim 1$. Since $|\xi| \gg 1$, this yields

$$|\xi_{12}| = |\xi - \xi_3| \sim |\xi| \gg 1. \tag{2.23}$$

Moreover, we have $|\xi_1| \sim |\xi_2| \sim |\xi_3| \sim |\xi| \gg 1$. Thus, we have

$$m(\bar{\xi}) \sim |\xi|^{\frac{1}{2}}$$

in this case. Let $\zeta_1 = \xi_{23}$, $\zeta_2 = \xi_{31}$, and $\zeta_3 = \xi_{12}$. Then, it follows from (2.15) that

$$\Psi(\bar{\xi}) = 3\zeta_1\zeta_2\zeta_3. \tag{2.24}$$

In the following, we freely use (partial) changes of variables between ξ_1, ξ_2, ξ_3, ξ and $\zeta_1, \zeta_2, \zeta_3$. Note that we have $|\zeta_2| \leqslant |\zeta_1| \leqslant 1$.

Subcase 2.a: $|\alpha| \lesssim M$. — For fixed $|\xi| \gg 1$, the condition $|\Psi(\bar{\xi}) - \alpha| \leqslant M$ with (2.23) and (2.24) implies that

$$|\zeta_2| \leqslant |\zeta_1|^{\frac{1}{2}} |\zeta_2|^{\frac{1}{2}} \lesssim \frac{(|\alpha| + M)^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}} \lesssim \frac{M^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}}.$$

Then, by a change of variables and Cauchy–Schwarz inequality, we have

$$\begin{aligned}
\text{LHS of (2.19)} \lesssim \sum_{\substack{N \gg 1 \\ \text{dyadic}}} N^{\frac{1}{2}} \int_{|\zeta_2| \lesssim \frac{M^{1/2}}{N^{1/2}}} &\bigg(\int_{|\xi_1|\sim N} v_1(\xi_1) v_3(-\xi_1 + \zeta_2) \mathrm{d}\xi_1 \bigg) \\
&\times \bigg(\int_{|\xi|\sim N} v_2(\xi - \zeta_2) v_4(\xi) \mathrm{d}\xi \bigg) \mathrm{d}\zeta_2
\end{aligned}$$

$$\lesssim M^{\frac{1}{2}}\|v_2\|_{L^2}\|v_3\|_{L^2}\sum_{\substack{N\gg 1\\ \text{dyadic}}}\|\mathbf{P}_N v_1\|_{L^2}\|\mathbf{P}_N v_4\|_{L^2}$$
$$\lesssim M^{\frac{1}{2}}\prod_{j=1}^{4}\|v_j\|_{L^2},$$

yielding (2.19). Here, $\mathbf{P}_N$ denotes the Littlewood–Paley projector onto the spatial frequencies $\{|\xi|\sim N\}$.

Subcase 2.b: $|\alpha|\gg M$. — For fixed $M\geqslant 1$, write $|\alpha|\sim 2^K M$ for some $K\in\mathbb{N}$. Note that we have

$$K\sim\log\left(\frac{|\alpha|}{M}\right). \tag{2.25}$$

If $|\zeta_2|\lesssim\frac{M^{1/2}}{|\xi|^{1/2}}$, then we can proceed as in Subcase 2.a. Hence, we assume that

$$|\zeta_1|\geqslant|\zeta_2|\gg\frac{M^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}}$$

in the following.

If $|\zeta_1|\gtrsim\frac{|\alpha|+M}{M^{1/2}|\xi|^{1/2}}\sim\frac{|\alpha|}{M^{1/2}|\xi|^{1/2}}$, then the condition $|\Psi(\bar{\xi})-\alpha|\leqslant M$ implies that

$$|\zeta_2|\lesssim\frac{|\alpha|+M}{|\zeta_1||\xi|}\lesssim\frac{M^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}},$$

thus reducing to the previous case. Therefore, it remains to consider the case

$$\frac{M^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}}\ll|\zeta_1|\ll\frac{|\alpha|}{M^{\frac{1}{2}}|\xi|^{\frac{1}{2}}}\sim\frac{2^K M^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}}, \tag{2.26}$$

where K satisfies (2.25).

Now, suppose that $|\zeta_1|\sim\frac{2^k M^{1/2}}{|\xi|^{1/2}}$ for some $1\leqslant k\leqslant K$. Then, for fixed ξ and ζ_1, the condition $|\Psi(\bar{\xi})-\alpha|\leqslant M$ implies that

$$\frac{\alpha-M}{3|\zeta_1|}\leqslant|F(\zeta_2)|\leqslant\frac{\alpha+M}{3|\zeta_1|}, \tag{2.27}$$

where $F(\zeta_2)$ is defined by

$$F(\zeta_2)=\zeta_2^2-(2\xi-\zeta_1)\zeta_2. \tag{2.28}$$

Note that the graph of $F(\zeta_2)$ is a parabola with a vertex $\sim(\xi,-\xi^2)$ in view of $|\zeta_1|\leqslant 1\ll|\xi|$. In particular, the slope of this parabola when $|\zeta_2|\leqslant 1$ is

$-2\xi + O(1)$. Hence, it follows from (2.27) and the assumption on the size of $|\zeta_1|$ that ζ_2 belongs to an interval $I_k = I_k(\zeta_1, \xi)$ of length

$$|I_k(\zeta_1, \xi)| \sim \frac{M}{|\zeta_1||\xi|} \sim \frac{M^{\frac{1}{2}}}{2^k |\xi|^{\frac{1}{2}}}. \tag{2.29}$$

Then, from (2.26) and (2.29), we obtain

$$\begin{aligned}
\text{LHS of (2.22)} &\lesssim \sup_\xi |\xi|^{\frac{1}{2}} \bigg(\sum_{k=1}^{K} \int_{|\zeta_1| \sim \frac{2^k M^{1/2}}{|\xi|^{1/2}}} \int_{\zeta_2 \in I_k(\zeta_1, \xi)} 1 \, \mathrm{d}\zeta_2 \mathrm{d}\zeta_1 \bigg)^{\frac{1}{2}} \\
&\lesssim \sup_\xi |\xi|^{\frac{1}{4}} M^{\frac{1}{4}} \bigg(\sum_{k=1}^{K} \int_{|\zeta_1| \sim \frac{2^k M^{1/2}}{|\xi|^{1/2}}} 2^{-k} \mathrm{d}\zeta_1 \bigg)^{\frac{1}{2}} \\
&\lesssim K^{\frac{1}{2}} M^{\frac{1}{2}} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}},
\end{aligned}$$

where the last inequality follows from (2.25).

Case 3: $|\xi| \gg 1$ and $|\xi_{31}| \leqslant 1 < |\xi_{23}| \leqslant |\xi_{12}|$. — In this case, we have $|\xi_2| \sim |\xi| \gg 1$ and $\langle \xi_1 \rangle \sim \langle \xi_3 \rangle$. Thus, we have

$$m(\bar{\xi}) \sim \frac{|\xi|}{\langle \xi_1 \rangle^{\frac{1}{2}}}. \tag{2.30}$$

Subcase 3.a: $|\xi_1| \gtrsim |\xi|$. — Since $|\xi| \gg 1 \geqslant |\xi_{31}| = |\xi - \xi_2|$, we have $|\xi_{23} + \xi_{12}| = |\xi + \xi_2| \sim |\xi|$. By the triangle inequality with $|\xi_{23}| \leqslant |\xi_{12}|$, we have $|\xi_{12}| \gtrsim |\xi| \gg 1$. Let $F(\zeta_2)$ be as in (2.28). Then, noting that

$$F'(\zeta_2) = 2\zeta_2 - 2\xi + \zeta_1 = -\xi_{12} + \zeta_2 = -\xi_{12} + O(1),$$

it follows from (2.27) that ζ_2 belongs to an interval $I = I(\zeta_1, \xi)$ of length

$$|I(\zeta_1, \xi)| \lesssim \frac{M}{|\zeta_1||\xi|} \leqslant \frac{M}{|\xi|} \tag{2.31}$$

for each fixed ξ and ζ_1 and hence for each fixed ξ and $\xi_1 = \xi - \zeta_1$. Given $k \in \mathbb{Z}$, let Π_k be the frequency projector onto the interval $[k, k+1)$ defined in (1.10). Then, using a variant of the Cauchy–Schwarz argument (2.21) with (2.30) and (2.31), we have

$$
\begin{aligned}
&\text{LHS of (2.19)}\\
&\leqslant \bigg\| \sum_{|k|\gg 1} \int_{|\xi_1|\in[k,k+1)} \int_{|\zeta_2|\leqslant 1} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m(\bar{\xi}) \\
&\qquad\qquad \times v_1(\xi_1) v_2(\xi-\zeta_2) v_3(-\xi_1+\zeta_2)\mathrm{d}\zeta_2\mathrm{d}\xi_1 \bigg\|_{L^2_\xi} \|v_4\|_{L^2}\\
&\leqslant \sup_{|k|\gg 1}\sup_{\xi}\bigg(\int_{|\xi_1|\in[k,k+1)} \int_{\zeta_2\in I(\zeta_1,\xi)} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2(\bar{\xi})\mathrm{d}\zeta_2\mathrm{d}\xi_1\bigg)^{\frac{1}{2}}\\
&\qquad\qquad \times \sum_{|k|\gg 1}\sum_{\ell=0}^{2} \|\Pi_k v_1\|_{L^2}\|\Pi_{-k-\ell}v_3\|_{L^2}\|v_2\|_{L^2}\|v_4\|_{L^2}\\
&\lesssim M^{\frac{1}{2}} \prod_{j=1}^{4}\|v_j\|_{L^2}.
\end{aligned}
\tag{2.32}
$$

Subcase 3.b: $|\xi_1| \ll |\xi|$. — In this case, we have $|\zeta_1| \sim |\xi|$. Then, arguing as in Subcase 3.a, we conclude that ζ_2 belongs to an interval $I = I(\zeta_1, \xi)$ of length

$$
|I(\zeta_1,\xi)| \lesssim \frac{M}{|\zeta_1||\xi|} \sim \frac{M}{|\xi|^2}
$$

for each fixed ξ and $\zeta_1 = \xi - \xi_1$. In particular, we have

$$
\sup_{|k|\gg 1}\sup_{\xi}\bigg(\int_{|\xi_1|\in[k,k+1)} \int_{\zeta_2\in I(\zeta_1,\xi)} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2(\bar{\xi})\mathrm{d}\zeta_2\mathrm{d}\xi_1\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.
$$

The rest follows as in (2.32).

Case 4: $|\xi| \gg 1$ *and* $|\xi_{12}|, |\xi_{23}|, |\xi_{31}| > 1$. — Noting that $\max(|\xi_{12}|, |\xi_{23}|, |\xi_{31}|) \gtrsim |\xi| \gg 1$, the condition $|\Psi(\bar{\xi}) - \alpha| \leqslant M$ with (2.15) implies that

$$
|\alpha| + M \gtrsim \max(|\xi|, |\xi_{12}|, |\xi_{23}|, |\xi_{31}|). \tag{2.33}
$$

In the following, the size relation of $|\xi_{12}|, |\xi_{23}|, |\xi_{31}|$ does not play any role. Without loss of generality, assume that $|\xi_1| \geqslant |\xi_2| \geqslant |\xi_3|$.

Subcase 4.a: $|\xi_1| \sim |\xi| \gg |\xi_2| \geqslant |\xi_3|$. — In this case, by viewing Ψ as a function of ξ_2 for fixed ξ and ξ_3, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |(\xi-\xi_3)(\xi-2\xi_2-\xi_3)| =$

$|\xi_{12}||\xi_{1-2}| \gtrsim |\xi|^2 \gg 1$. Hence, with (2.33), we have

$$\begin{aligned}\text{LHS of (2.22)} &\lesssim \sup_{\xi}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \frac{|\xi|^2}{\langle\xi_2\rangle^{\frac{1}{2}}\langle\xi_3\rangle^{\frac{1}{2}}} d\xi_2 d\xi_3\bigg)^{\frac{1}{2}} \\ &\lesssim M^{\frac{1}{2}}\bigg(\int_{|\xi_3|\ll|\xi|} \frac{1}{\langle\xi_3\rangle} d\xi_3\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}(\log|\xi|)^{\frac{1}{2}} \lesssim \langle\alpha\rangle^{0+}M^{\frac{1}{2}+},\end{aligned}$$

yielding (2.22).

Subcase 4.b: $|\xi_1|, |\xi_2| \gtrsim |\xi| \gg |\xi_3|$. — Note that, in the first step of (2.21), we can perform Cauchy–Schwarz inequality in ξ_2 instead of ξ. Then, (2.19) follows once we prove

$$\sup_{\xi_2}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2(\bar{\xi}) d\xi_1 d\xi_3\bigg)^{\frac{1}{2}} \lesssim \langle\alpha\rangle^{0+}M^{\frac{1}{2}+}. \tag{2.34}$$

If $|\xi_1| \sim |\xi_2| \gg |\xi| \gg |\xi_3|$, then $|\xi+\xi_1| \sim |\xi_2|$, and $|\xi_{23}| \sim |\xi_2|$. Then, by viewing Ψ as a function of ξ_1 for fixed ξ_2 and ξ_3, we have

$$|\partial_{\xi_1}\Psi(\bar{\xi})| = |\xi_{23}(\xi+\xi_1)| \gtrsim |\xi_2|^2 \tag{2.35}$$

and thus

$$\begin{aligned}\text{LHS of (2.34)} &\lesssim \sup_{\xi_2}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \frac{|\xi|^{\frac{3}{2}}}{\langle\xi_3\rangle^{\frac{1}{2}}} d\xi_1 d\xi_3\bigg)^{\frac{1}{2}} \\ &\lesssim M^{\frac{1}{2}} \sup_{\xi_2} \frac{1}{\langle\xi_2\rangle^{\frac{1}{4}}}\bigg(\int_{|\xi_3|\ll|\xi_2|} \frac{1}{\langle\xi_3\rangle^{\frac{1}{2}}} d\xi_3\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.\end{aligned} \tag{2.36}$$

If $|\xi_1| \sim |\xi_2| \sim |\xi| \gg |\xi_3|$, we have

$$\max(|\xi+\xi_1|, |\xi+\xi_2|) \gtrsim |2\xi+\xi_{12}| = |3\xi-\xi_3| \sim |\xi|.$$

Without loss of generality, assume that $|\xi+\xi_1| \gtrsim |\xi|$. (Otherwise, we switch the role of ξ_1 and ξ_2 in (2.34).) Then, (2.35) and hence (2.36) hold in this case as well.

Subcase 4.c: $|\xi_1|, |\xi_2|, |\xi_3| \gtrsim |\xi|$. — In this case, we have

$$\max(|\xi+\xi_1|, |\xi+\xi_2|, |\xi+\xi_3|) \gtrsim |3\xi+\xi_{123}| = 4|\xi|. \tag{2.37}$$

Without loss of generality, assume that $|\xi+\xi_1| \gtrsim |\xi|$. Then, by viewing Ψ as a function of ξ_1 for fixed ξ_2 and ξ_3, we have

$$|\partial_{\xi_1}\Psi(\bar{\xi})| \sim |\xi_{23}(\xi+\xi_1)| \gtrsim |\xi||\xi_{23}|. \tag{2.38}$$

Note that, By performing Cauchy–Schwarz inequality in ξ_3 instead of ξ in the first step of (2.21), it suffices to prove

$$\sup_{\xi_3}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar\xi)-\alpha|\leqslant M}\cdot m^2(\bar\xi)\mathrm{d}\xi_1\mathrm{d}\xi_2\bigg)^{\frac12} \lesssim \langle\alpha\rangle^{0+}M^{\frac12+}. \tag{2.39}$$

From (2.38) and (2.33), we have

$$\begin{aligned}\text{LHS of (2.39)} &\lesssim \sup_{\xi_3}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar\xi)-\alpha|\leqslant M}|\xi|\mathrm{d}\xi_1\mathrm{d}\xi_2\bigg)^{\frac12}\\ &\lesssim M^{\frac12}\sup_{\xi_3}\bigg(\int_{1\leqslant|\xi_{23}|\lesssim|\alpha|+M}\frac{1}{|\xi_{23}|}\mathrm{d}\xi_2\bigg)^{\frac12} \lesssim \langle\alpha\rangle^{0+}M^{\frac12+}.\end{aligned}$$

This completes the proof of Lemma 2.6. ☐

Remark 2.7. — While the simple Cauchy–Schwarz argument (2.21) works for most of the cases in the proof of Lemma 2.6, it does not seem to work for Case 2 in the endpoint case: $s=\frac14$. We point out that the Cauchy–Schwarz argument suffices for Case 2 in the non-endpoint case: $s>\frac14$.

As an immediate corollary to Lemma 2.6, we obtain the following lemma. The proof is analogous to that of Lemma 2.5.

Lemma 2.8. — *Let $s\geqslant\frac14$. Then, we have*

$$\begin{aligned}\|\mathcal{I}^\alpha_M(v)\|_{H^s} &\lesssim \langle\alpha\rangle^{0+}M^{-\frac12+}\|v\|^3_{H^s},\\ \|\mathcal{I}^\alpha_M(v)-\mathcal{I}^\alpha_M(w)\|_{H^s} &\lesssim \langle\alpha\rangle^{0+}M^{-\frac12+}\big(\|v\|^2_{H^s}+\|w\|^2_{H^s}\big)\|v-w\|_{H^s},\end{aligned}$$

and

$$\begin{aligned}\|\mathcal{I}^\alpha_{>M}(v)\|_{H^s} &\lesssim \langle\alpha\rangle^{0+}M^{-\frac12+}\|v\|^3_{H^s},\\ \|\mathcal{I}^\alpha_{>M}(v)-\mathcal{I}^\alpha_{>M}(w)\|_{H^s} &\lesssim \langle\alpha\rangle^{0+}M^{-\frac12+}\big(\|v\|^2_{H^s}+\|w\|^2_{H^s}\big)\|v-w\|_{H^s},\end{aligned}$$

for any $M\geqslant 1$ and $\alpha\in\mathbb{R}$.

3. Normal form reductions

In this section, we implement an infinite iteration scheme of normal form reductions at a formal level. We perform normal form reductions in an iterative manner, transforming part of the nonlinearity into nonlinearities of higher and higher degrees. In the end, we formally arrive at an equation involving infinite series of nonlinearities of arbitrarily high degrees (Subsection 3.4).

Such an infinite iteration of normal form reductions was first introduced in Guo–Kwon–Oh [15] in proving unconditional well-posedness of the cubic NLS on $\mathbb{T}$. While the implementation of normal form reductions in [15] was systematic, the multilinear estimates heavily depended on the structure of the equation as well as some elementary number theory (the divisor counting argument). In the following, we perform normal form reductions in a rather abstract manner. This allows us to handle the cubic NLS (1.1) and the mKdV (1.2) in an identical manner by applying the localized modulation estimates obtained in Section 2.

Before proceeding further, we need to set up some notations. In the following, we simply denote the Fourier coefficient $v(\xi) = \widehat{v}(\xi)$ by v_ξ. When the complex conjugate sign on v_ξ does not play any significant role, we drop the complex conjugate sign. We often drop the complex number i and simply use 1 for ± 1 and $\pm i$.

In the following presentation of normal form reductions, we restrict our attention to the cubic NLS (1.1). In view of the localized modulation estimates (Lemmas 2.6 and 2.8), one can easily modify the argument to handle the mKdV (1.2). All the computations in this section (such as switching summations and integrals) are formal, assuming that u (and hence v) is a smooth solution. In Section 4, we justify our formal computations when $u \in C_t H_x^s$ with (i) $s \geqslant \frac{1}{6}$ for the cubic NLS and (ii) $s > \frac{1}{4}$ for the mKdV, respectively.

3.1. Notation: index by trees

When we apply a normal form reduction, i.e. integration by parts as in (1.6),(11) a time derivative can fall on any of the factors v_{ξ_j}, transforming the nonlinearity into that of a higher degree. In each step of normal form reductions, we need to keep track of where a time derivative falls. which may be a cumbersome task in general. In [15], we introduced the notion of *ordered trees* for indexing such terms arising in the general steps of normal form reductions. In order to carry out our analysis, we will need to supplement more notations related to ordered trees in the following.

Definition 3.1. — *Given a partially ordered set $\mathcal{T}$ with partial order $\leqslant$, we say that $b \in \mathcal{T}$ with $b \leqslant a$ and $b \neq a$ is a child of $a \in \mathcal{T}$, if $b \leqslant c \leqslant a$ implies either $c = a$ or $c = b$. If the latter condition holds, we also say that a is the parent of b.*

(11) In fact, we proceed without an integration symbol in the following. Namely, we perform differentiation by parts.

As in [4, 34], our trees refer to a particular subclass of ternary trees.

Definition 3.2. — *A tree $\mathcal{T}$ is a finite partially ordered set satisfying the following properties:*

- *Let $a_1, a_2, a_3, a_4 \in \mathcal{T}$. If $a_4 \leqslant a_2 \leqslant a_1$ and $a_4 \leqslant a_3 \leqslant a_1$, then we have $a_2 \leqslant a_3$ or $a_3 \leqslant a_2$.*
- *A node $a \in \mathcal{T}$ is called terminal, if it has no child. A non-terminal node $a \in \mathcal{T}$ is a node with exactly three children denoted by a_1, a_2 and a_3.*[12]
- *There exists a maximal element $r \in \mathcal{T}$ (called the root node) such that $a \leqslant r$ for all $a \in \mathcal{T}$. We assume that the root node is non-terminal.*
- *$\mathcal{T}$ consists of the disjoint union of $\mathcal{T}^0$ and $\mathcal{T}^\infty$, where $\mathcal{T}^0$ and $\mathcal{T}^\infty$ denote the collection of non-terminal nodes and terminal nodes, respectively.*

Note that the number $|\mathcal{T}|$ of nodes in a tree $\mathcal{T}$ is $3j+1$ for some $j \in \mathbb{N}$, where $|\mathcal{T}_0| = j$ and $|\mathcal{T}^\infty| = 2j+1$. We use $T(j)$ to denote the collection of trees of the jth generation, namely, with j parental nodes.

Next, we recall the notion of ordered trees introduced in [15]. Roughly speaking, an ordered tree "remembers how it grew".

Definition 3.3. — *We say that a sequence $\{\mathcal{T}_j\}_{j=1}^J$ is a chronicle of J generations, if*

- *$\mathcal{T}_j \in T(j)$ for each $j = 1, \dots, J$,*
- *$\mathcal{T}_{j+1}$ is obtained by changing one of the terminal nodes in $\mathcal{T}_j$, denoted by $p^{(j)}$, into a non-terminal node (with three children), $j = 1, \dots, J-1$.*

Given a chronicle $\{\mathcal{T}_j\}_{j=1}^J$ of J generations, we refer to $\mathcal{T}_J$ as an ordered tree of the Jth generation. *We use $\mathfrak{T}(J)$ to denote the collection of the ordered trees of the Jth generation. Note that the cardinality of $\mathfrak{T}(J)$ is given by*

$$|\mathfrak{T}(J)| = 1 \cdot 3 \cdot 5 \cdot \dots \cdot (2J-1) =: c_J \tag{3.1}$$

Remark 3.4. — Given two ordered trees $\mathcal{T}_J$ and $\widetilde{\mathcal{T}}_J$ of the Jth generation, it may happen that $\mathcal{T}_J = \widetilde{\mathcal{T}}_J$ as trees (namely as graphs) while $\mathcal{T}_J \neq \widetilde{\mathcal{T}}_J$ as ordered trees according to Definition 3.3. Henceforth, when we refer to an ordered tree $\mathcal{T}_J$ of the Jth generation, it is understood that there is an underlying chronicle $\{\mathcal{T}_j\}_{j=1}^J$.

[12] Note that the order of children plays an important role in our discussion. We refer to a_j as the jth child of a non-terminal node $a \in \mathcal{T}$. In terms of the planar graphical representation of a tree, we set the jth node from the left as the jth child a_j of $a \in \mathcal{T}$.

Definition 3.5.

(i) *Given an ordered tree* $\mathcal{T}_J \in \mathfrak{T}(J)$ *with a chronicle* $\{\mathcal{T}_j\}_{j=1}^J$*, we define a "projection"* π_j*,* $j = 1, \dots, J$*, from* $\mathcal{T}_J$ *to subtrees in* $\mathcal{T}_J$ *of one generation by setting*
- $\pi_1(\mathcal{T}_J) = \mathcal{T}_1$*,*
- $\pi_j(\mathcal{T}_J)$ *to be the tree formed by the three terminal nodes in* $\mathcal{T}_j \setminus \mathcal{T}_{j-1}$ *and its parent,* $j = 2, \dots, J$*. Intuitively speaking,* $\pi_j(\mathcal{T}_J)$ *is the tree added in transforming* $\mathcal{T}_{j-1}$ *into* $\mathcal{T}_j$*.*

We use $r^{(j)}$ *to denote the root node of* $\pi_j(\mathcal{T}_J)$ *and refer to it as the* jth root node*. By definition, we have*

$$r^{(j)} = p^{(j-1)}. \tag{3.2}$$

Note that $p^{(j-1)}$ *is not necessarily a node in* $\pi_{j-1}(\mathcal{T}_J)$*.*

(ii) *Given* $j \in \{1, \dots, J-1\}$*,* $p^{(j)}$ *appears as a terminal node of* $\pi_k(\mathcal{T})$ *for exactly one* $k \in \{1, 2 \dots, j-1\}$*. In particular,* $p^{(j)}$ *is the* ℓ*th child of the* k*th root note* $r^{(k)}$ *for some* $\ell \in \{1, 2, 3\}$*. We define* the order of $p^{(j)}$*, denoted by* $\#p^{(j)}$*, to be this number* $\ell \in \{1, 2, 3\}$*.*

(iii) *We define the* essential terminal nodes $\pi_j^\infty(\mathcal{T}_J)$ *of the* j*th generation by setting*

$$\pi_j^\infty(\mathcal{T}_J) := \pi_j(\mathcal{T}_J)^\infty \cap \mathcal{T}_J^\infty = (\mathcal{T}_j \setminus \mathcal{T}_{j-1}) \cap \mathcal{T}_J^\infty.$$

By definition, $\pi_j^\infty(\mathcal{T}_J)$ *may be empty. Note that* $\{\pi_j^\infty(\mathcal{T}_J)\}_{j=1}^J$ *forms a partition of* $\mathcal{T}_J^\infty$*.*

We record the following simple observation. This will be useful in Subsections 3.3 and 4.3.

Remark 3.6. — Let $\mathcal{T} \in \mathfrak{T}(J)$ be an ordered tree. Then, for each fixed $j = 2, \dots, J$, there exists a path[13] $a_1, a_2, \dots, a_K$, starting at the root node $r = r^{(1)}$ and ending at the jth root node $r^{(j)}$ such that $a_k \neq r^{(\ell)}$ for any $k = 1, \dots, K$ and $\ell \geqslant j+1$. Namely, we can move from $r^{(1)}$ to $r^{(j)}$ without hitting a root node of a higher generation.

More concretely, given $r^{(j)}$, we know that it appears as a terminal node of $\pi_{j_1}(\mathcal{T})$ for exactly one $j_1 \in \{1, 2 \dots, j-1\}$. Similarly, $r^{(j_1)}$ appears as a terminal node of $\pi_{j_2}(\mathcal{T})$ for exactly one $j_2 \in \{1, 2 \dots, j_1 - 1\}$. We can iterate this process, which must terminate in a finite number of steps with $j_k = 1$. This generates the shortest path $r^{(j_k)}, r^{(j_{k-1})}, \dots, r^{(j_1)}, r^{(j)}$ from $r^{(1)}$ to $r^{(j)}$ and we denote it by $P(r^{(1)}, r^{(j)})$. Similarly, given $a \in \mathcal{T} \setminus \{r^{(1)}\}$, one can easily construct the shortest path from $r^{(1)}$ to a since a is a terminal node of $\pi_k(\mathcal{T})$ for some k. We denote this shortest path by $P(r^{(1)}, a)$.

[13] A path is a sequence of nodes $a_1, a_2, \dots, a_K$ such that a_k and a_{k+1} are adjacent.

Given an ordered tree, we need to consider all possible frequency assignments to nodes that are "consistent".

Definition 3.7. — *Given an ordered tree* $\mathcal{T} \in \mathfrak{T}(J)$, *we define an* index function $\boldsymbol{\xi} : \mathcal{T} \to \mathbb{R}$ *such that*

$$\xi_a = \xi_{a_1} - \xi_{a_2} + \xi_{a_3} \tag{3.3}$$

for $a \in \mathcal{T}^0$, *where* a_1, a_2, *and* a_3 *denote the children of* a. *Here, we identified* $\boldsymbol{\xi} : \mathcal{T} \to \mathbb{R}$ *with* $\{\xi_a\}_{a\in\mathcal{T}} \in \mathbb{R}^{\mathcal{T}}$. *We use* $\Xi(\mathcal{T}) \subset \mathbb{R}^{\mathcal{T}}$ *to denote the collection of such index functions* $\boldsymbol{\xi}$.

Remark 3.8.

(i) If we associate functions $v_a = v_a(\xi_a)$ to each node $a \in \mathcal{T}$, then the relation (3.3) implies that $v_a = v_{a_1} * \overline{v_{a_2}} * v_{a_3}$.
(ii) For the mKdV, we need to replace (3.3) by $\xi_a = \xi_{a_1} + \xi_{a_2} + \xi_{a_3}$.

Given an ordered tree $\mathcal{T}_J \in \mathfrak{T}(J)$ with a chronicle $\{\mathcal{T}_j\}_{j=1}^J$ and associated index functions $\boldsymbol{\xi} \in \Xi(\mathcal{T}_J)$, we use superscripts to keep track of "generations" of frequencies.

Consider $\mathcal{T}_1$ of the first generation. We define the first generation of frequencies by

$$\big(\xi^{(1)}, \xi^{(1)}_1, \xi^{(1)}_2, \xi^{(1)}_3\big) := (\xi_r, \xi_{r_1}, \xi_{r_2}, \xi_{r_3}),$$

where r_j denotes the three children of the root node r.

In general, the ordered tree $\mathcal{T}_j$ of the jth generation is obtained from $\mathcal{T}_{j-1}$ by changing one of its terminal nodes $a \in \mathcal{T}^\infty_{j-1}$ into a non-terminal node. Then, we define the jth generation of frequencies by

$$\big(\xi^{(j)}, \xi^{(j)}_1, \xi^{(j)}_2, \xi^{(j)}_3\big) := (\xi_a, \xi_{a_1}, \xi_{a_2}, \xi_{a_3}),$$

where a_j denotes the three children of the node $a \in \mathcal{T}^\infty_{j-1}$. Note that the parent node a is nothing but the jth root node $r^{(j)}$ defined in Definition 3.5.

Our main analytical tool is the localized modulation estimates from Section 2. Hence, it is important to keep track of the modulation for frequencies in each generation. We use μ_j to denote the corresponding modulation function introduced at the jth generation. Namely, we set[14]

$$\begin{aligned}\mu_j = \mu_j\big(\xi^{(j)}, \xi^{(j)}_1, \xi^{(j)}_2, \xi^{(j)}_3\big) &:= \big(\xi^{(j)}\big)^2 - \big(\xi^{(j)}_1\big)^2 + \big(\xi^{(j)}_2\big)^2 - \big(\xi^{(j)}_3\big)^2\\ &= 2\big(\xi^{(j)}_2 - \xi^{(j)}_1\big)\big(\xi^{(j)}_2 - \xi^{(j)}_3\big) = 2\big(\xi^{(j)} - \xi^{(j)}_1\big)\big(\xi^{(j)} - \xi^{(j)}_3\big),\end{aligned}$$

(14) For the mKdV, the modulation function μ_j is given by

$$\mu_j := \big(\xi^{(j)}\big)^3 - \big(\xi^{(j)}_1\big)^3 - \big(\xi^{(j)}_2\big)^3 - \big(\xi^{(j)}_3\big)^3.$$

where the last two equalities hold in view of (3.3). We also use the following shorthand notation:

$$\widetilde{\mu}_j := \sum_{k=1}^{j} \mu_k.$$

3.2. Normal form reductions: second and third generations

We are now ready to perform normal form reductions. As we mentioned earlier, we only consider the cubic NLS (1.1) in $H^s(\mathbb{R})$, $s \geqslant 0$, in the following. Since our implementation is carried out at an abstract level, a minor modification suffices for the mKdV in $H^s(\mathbb{R})$, $s \geqslant \frac{1}{4}$.

Fix dyadic $N > 1$ (to be determined later). We first write (2.3) as

$$\begin{aligned} \partial_t v = \mathcal{N}(v) &= \mathcal{N}_{\leqslant N}(v) + \mathcal{N}_{>N}(v) \\ &=: \mathcal{N}_1^{(1)}(v) + \mathcal{N}_2^{(1)}(v). \end{aligned}$$

By Lemma 2.3, we can estimate the low modulation part:

$$\|\mathcal{N}_1^{(1)}(v)\|_{H^s} = \|\mathcal{N}_{\leqslant N}(v)\|_{H^s} \lesssim N^{\frac{1}{2}+}\|v\|_{H^s}^3 \tag{3.4}$$

for $s \geqslant 0$. The main point is that the restriction $|\Phi(\bar{\xi})| \leqslant N$ provides a restriction on the possible range of frequencies.

The high modulation part $\mathcal{N}_2^{(1)}(v) = \mathcal{N}_{>N}(v)$ with $|\Phi(\bar{\xi})| > N$ can not benefit such a frequency restriction. In this case, we exploit a rapid oscillation due to the high modulation, introducing cancellation under a time integration. For this purpose, we iteratively apply differentiation by parts and transform $\mathcal{N}_2^{(1)}(v)$ into infinite series of multilinear terms.

Let C_0 denote the domain of $\mathcal{N}_2^{(1)}(v) = \mathcal{N}_{>N}(v)$:

$$C_0 := \big\{|\mu_1| > N\big\}. \tag{3.5}$$

By taking differentiation by parts[15] with (2.3), we have

$$\begin{aligned} \mathcal{N}_2^{(1)}(v)(\xi,t) &= \mathcal{N}_{>N}(v)(\xi,t) \\ &= \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0} e^{-i\mu_1 t} \prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a} \end{aligned}$$

[15] When we apply differentiation by parts, we keep the minus sign on the second term for emphasis.

$$
\begin{aligned}
&= \partial_t\Bigg[\int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg] \\
&\qquad\qquad - \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\partial_t\Bigg(\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg) \\
&= \partial_t\Bigg[\int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg] \\
&\qquad\qquad - \sum_{\mathcal{T}_2\in\mathfrak{T}(2)}\int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_2)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i(\mu_1+\mu_2)t}}{\mu_1}\prod_{a\in\mathcal{T}_2^\infty} v_{\xi_a} \\
&=: \partial_t\mathcal{N}_0^{(2)}(v)(\xi,t) + \mathcal{N}^{(2)}(v)(\xi,t).
\end{aligned}
\tag{3.6}
$$

From Lemma 2.5, we have the following estimate on the boundary term $\mathcal{N}_0^{(2)}(v)$.

Lemma 3.9. — *Let $s \geqslant 0$. Then, we have*

$$
\begin{aligned}
\|\mathcal{N}_0^{(2)}(v)\|_{H^s} &\lesssim N^{-\frac{1}{2}+}\|v\|_{H^s}^3, \\
\|\mathcal{N}_0^{(2)}(v) - \mathcal{N}_0^{(2)}(w)\|_{H^s} &\lesssim N^{-\frac{1}{2}+}\big(\|v\|_{H^s}^2 + \|w\|_{H^s}^2\big)\|v-w\|_{H^s}.
\end{aligned}
$$

Next, we decompose the frequency space into

$$
C_1 := \big\{|\mu_1+\mu_2| \leqslant 5^3|\mu_1|^{1-\delta}\big\} \tag{3.7}
$$

and its complement C_1^c,[16] where $\delta > 0$ is a small constant. Then, we decompose $\mathcal{N}^{(2)}$ as

$$
\mathcal{N}^{(2)} = \mathcal{N}_1^{(2)} + \mathcal{N}_2^{(2)}, \tag{3.8}
$$

where $\mathcal{N}_1^{(2)} := \mathcal{N}^{(2)}|_{C_1}$ is defined as the restriction of $\mathcal{N}^{(2)}$ on C_1 and $\mathcal{N}_2^{(2)} := \mathcal{N}^{(2)} - \mathcal{N}_1^{(2)}$, namely $\mathcal{N}_2^{(2)}$ is the restriction of $\mathcal{N}^{(2)}$ on C_1^c. Note that we have

$$
\mathcal{N}_{>N} = \partial_t\mathcal{N}_0^{(2)} + \mathcal{N}_1^{(2)} + \mathcal{N}_2^{(2)}
$$

at this point. Thanks to the restriction (3.7) on the modulation, we can estimate the first term $\mathcal{N}_1^{(2)}$. However, we do not have a direct control of $\mathcal{N}_2^{(2)}$. In the following, we apply another normal form reduction to $\mathcal{N}_2^{(2)}$.

[16] Clearly, the number 5^3 in (3.7) does not play any role at this point. However, we insert it to match with (3.23). See also (3.15) and (3.21).

Lemma 3.10. — *Let $s \geqslant 0$. Then, we have*

$$\|\mathcal{N}_1^{(2)}(v)\|_{H^s} \lesssim N^{-\frac{\delta}{2}+}\|v\|_{H^s}^5, \tag{3.9}$$

$$\|\mathcal{N}_1^{(2)}(v) - \mathcal{N}_1^{(2)}(w)\|_{H^s} \lesssim N^{-\frac{\delta}{2}+}\big(\|v\|_{H^s}^4 + \|w\|_{H^s}^4\big)\|v-w\|_{H^s}, \tag{3.10}$$

for $0 < \delta < 1$.

Proof. — We only present the proof of (3.9) since (3.10) follows in a similar manner in view of the multilinearity of $\mathcal{N}_1^{(2)}$. Moreover, by the triangle inequality, it suffices to prove (3.9) for $s = 0$. From (3.6) and (3.8) with (3.7), we have

$$\begin{aligned}
&\mathcal{N}_1^{(2)}(v)(\xi,t)\\
&= \sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_2)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a_1\in\pi_1^\infty(\mathcal{T}_2)} v_{\xi_{a_1}} \cdot \mathbb{1}_{C_1}e^{-i\mu_2 t}\prod_{a_2\in\pi_2^\infty(\mathcal{T}_2)} v_{\xi_{a_2}}\\
&= \sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_2)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a_1\in\pi_1^\infty(\mathcal{T}_2)} v_{\xi_{a_1}} \int_{\substack{\boldsymbol{\xi}^{(2)}\in\Xi(\pi_2(\mathcal{T}_2))\\ \boldsymbol{\xi}^{(2)}_{r^{(2)}}=\xi^{(2)}}} \mathbb{1}_{C_1}e^{-i\mu_2 t}\prod_{a_2\in\pi_2^\infty(\mathcal{T}_2)} v_{\xi_{a_2}}\\
&= \sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_2)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a_1\in\pi_1^\infty(\mathcal{T}_2)} v_{\xi_{a_1}} \cdot \mathcal{N}^{\mu_1}_{\leqslant 5^3|\mu_1|^{1-\delta}}(v)(\xi^{(2)},t).
\end{aligned}$$

In the second line, we slightly abused notations in the domain of the second integration for clarity since, strictly speaking, it is already included in the domain of the first integral. Note that the second integral is over three variables $\{\xi_{a_2}\}_{a_2\in\pi_2^\infty(\mathcal{T}_2)}$, while the first integral is over two variables $\{\xi_{a_1}\}_{a_1\in\pi_1^\infty(\mathcal{T}_2)}$, with one constraint $\boldsymbol{\xi}_r = \xi$.

Then, from Lemmas 2.3 and 2.5 with (3.1) and (3.5), we have

$$\begin{aligned}
\|\mathcal{N}_1^{(2)}(v)\|_{L^2} &\lesssim \sum_{T_2\in\mathfrak{T}(2)}\sum_{\substack{M\geqslant N\\ \text{dyadic}}} \|\mathcal{I}_M(v,v,\mathcal{N}^{\mu_1}_{\leqslant 5^3|\mu_1|^{1-\delta}}(v))\|_{L^2}\\
&\lesssim \sum_{\substack{M\geqslant N\\ \text{dyadic}}} M^{-\frac{1}{2}+}\|v\|_{L^2}^2\|\mathcal{N}^{M}_{\lesssim 5^3 M^{1-\delta}}(v)\|_{L^2}\\
&\lesssim N^{-\frac{\delta}{2}+}\|v\|_{L^2}^5.
\end{aligned}$$

This proves (3.9). □

Next, we apply a normal form reduction to $\mathcal{N}_2^{(2)}$. On the support of $\mathcal{N}_2^{(2)}$, namely, on $C_0 \cap C_1^c$, we have

$$|\mu_1 + \mu_2| > 5^3|\mu_1|^{1-\delta} > N^{1-\delta}. \tag{3.11}$$

By applying differentiation by parts once again, we have

$$
\begin{aligned}
\mathcal{N}_2^{(2)}(v)(\xi) &= \partial_t \Bigg[\sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_2),\, \boldsymbol{\xi}_r = \xi}}^{C_0 \cap C_1^c} \frac{e^{-i(\mu_1+\mu_2)t}}{\mu_1(\mu_1+\mu_2)} \prod_{a \in \mathcal{T}_2^\infty} v_{\xi_a} \Bigg] \\
&\qquad - \sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_2),\, \boldsymbol{\xi}_r = \xi}}^{C_0 \cap C_1^c} \frac{e^{-i(\mu_1+\mu_2)t}}{\mu_1(\mu_1+\mu_2)} \partial_t \Bigg(\prod_{a \in \mathcal{T}_2^\infty} v_{\xi_a} \Bigg) \\
&= \partial_t \Bigg[\sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_2),\, \boldsymbol{\xi}_r = \xi}}^{C_0 \cap C_1^c} \frac{e^{-i(\mu_1+\mu_2)t}}{\mu_1(\mu_1+\mu_2)} \prod_{a \in \mathcal{T}_2^\infty} v_{\xi_a} \Bigg] \\
&\qquad - \sum_{\mathcal{T}_3 \in \mathfrak{T}(3)} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_3),\, \boldsymbol{\xi}_r = \xi}}^{C_0 \cap C_1^c} \frac{e^{-i(\mu_1+\mu_2+\mu_3)t}}{\mu_1(\mu_1+\mu_2)} \prod_{a \in \mathcal{T}_3^\infty} v_{\xi_a} \\
&=: \partial_t \mathcal{N}_0^{(3)}(v)(\xi) + \mathcal{N}^{(3)}(v)(\xi).
\end{aligned} \tag{3.12}
$$

We can easily estimate the boundary term $\mathcal{N}_0^{(3)}(v)$ as follows.

Lemma 3.11. — *Let $s \geqslant 0$. Then, we have*

$$
\|\mathcal{N}_0^{(3)}(v)\|_{H^s} \lesssim N^{-1+\frac{\delta}{2}+} \|v\|_{H^s}^5, \tag{3.13}
$$

$$
\|\mathcal{N}_0^{(3)}(v) - \mathcal{N}_0^{(3)}(w)\|_{H^s} \lesssim N^{-1+\frac{\delta}{2}+} \big(\|v\|_{H^s}^4 + \|w\|_{H^s}^4\big) \|v-w\|_{H^s}, \tag{3.14}
$$

for $0 < \delta < 1$.

Proof. — We only present the proof of (3.13) since (3.14) follows in a similar manner. Moreover, by the triangle inequality, it suffices to prove (3.13) for $s = 0$. We proceed as in the proof of Lemma 3.10. By an iterative application of Lemma 2.5 with (3.11), we have

$$
\begin{aligned}
\|\mathcal{N}_0^{(3)}(v)\|_{L^2} &\leqslant \Bigg\| \sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_2),\, \boldsymbol{\xi}_r = \xi}}^{C_0 \cap C_1^c} \frac{e^{i(\mu_1+\mu_2)t}}{\mu_1(\mu_1+\mu_2)} \prod_{a \in \mathcal{T}_2^\infty} v_{\xi_a} \Bigg\|_{L^2} \\
&\lesssim \sum_{T_2 \in \mathfrak{T}(2)} \sum_{\substack{M \geqslant N \\ \text{dyadic}}} \|\mathcal{I}_M(v, v, \mathcal{I}^{\mu_1}_{>5^3|\mu_1|^{1-\delta}}(v))\|_{L^2} \\
&\lesssim \sum_{\substack{M \geqslant N \\ \text{dyadic}}} M^{-\frac{1}{2}+} \|v\|_{L^2}^2 \|\mathcal{I}^M_{\gtrsim 5^3 M^{1-\delta}}(v)\|_{L^2} \\
&\lesssim N^{-1+\frac{\delta}{2}+} \|v\|_{L^2}^5,
\end{aligned}
$$

yielding the desired estimate (3.13). ☐

As in the first step, we decompose $\mathcal{N}^{(3)}$ as

$$
\mathcal{N}^{(3)} = \mathcal{N}_1^{(3)} + \mathcal{N}_2^{(3)},
$$

where $\mathcal{N}_1^{(3)}$ is the restriction of $\mathcal{N}^{(3)}$ onto

$$C_2 := \big\{|\widetilde{\mu}_3| \leqslant 7^3|\widetilde{\mu}_2|^{1-\delta}\big\} \cup \big\{|\widetilde{\mu}_3| \leqslant 7^3|\mu_1|^{1-\delta}\big\} \tag{3.15}$$

and $\mathcal{N}_2^{(3)} := \mathcal{N}^{(3)} - \mathcal{N}_1^{(3)}$. At this point, we have

$$\mathcal{N}_{>N} = \sum_{j=2}^{3} \partial_t \mathcal{N}_0^{(j)} + \sum_{j=2}^{3} \mathcal{N}_1^{(j)} + \mathcal{N}_2^{(3)}.$$

As before, the modulation restriction (3.15) allows us to estimate the first term $\mathcal{N}_1^{(3)}$.

Lemma 3.12. — *Let $s \geqslant 0$. Then, we have*

$$\|\mathcal{N}_1^{(3)}(v)\|_{H^s} \lesssim N^{-\frac{1}{2}+}\|v\|_{H^s}^5, \tag{3.16}$$

$$\|\mathcal{N}_1^{(3)}(v) - \mathcal{N}_1^{(3)}(w)\|_{H^s} \lesssim N^{-\frac{1}{2}+}\big(\|v\|_{H^s}^4 + \|w\|_{H^s}^4\big)\|v-w\|_{H^s}, \tag{3.17}$$

for $0 < \delta < 1$.

Proof. — We only present the proof of (3.16) since (3.17) follows in a similar manner. Moreover, by the triangle inequality, it suffices to prove (3.16) for $s = 0$. As in the proof of Lemma 3.10, with a slight abuse of notations, we have

$$\begin{aligned}\mathcal{N}_1^{(3)}(v)(\xi,t) = \sum_{\mathcal{T}_3 \in \mathfrak{T}(3)} &\int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_3)\\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \frac{e^{-i\mu_1 t}}{\mu_1} \prod_{a_1 \in \pi_1^\infty(\mathcal{T}_3)} v_{\xi_{a_1}} \\ &\times \int_{\substack{\boldsymbol{\xi}^{(2)} \in \Xi(\pi_2(\mathcal{T}_3))\\ \boldsymbol{\xi}^{(2)}_{r^{(2)}} = \xi^{(2)}}} \mathbb{1}_{C_1^c} \frac{e^{-i\mu_2 t}}{\widetilde{\mu}_2} \prod_{a_2 \in \pi_2^\infty(\mathcal{T}_3)} v_{\xi_{a_2}} \\ &\times \int_{\substack{\boldsymbol{\xi}^{(3)} \in \Xi(\pi_3(\mathcal{T}_3))\\ \boldsymbol{\xi}^{(3)}_{r^{(3)}} = \xi^{(3)}}} \mathbb{1}_{C_2} e^{-i\mu_3 t} \prod_{a_3 \in \pi_3^\infty(\mathcal{T}_3)} v_{\xi_{a_3}}.\end{aligned} \tag{3.18}$$

Note that the last integral is over three variables $\{\xi_{a_3}\}_{a_3 \in \pi_3^\infty(\mathcal{T}_3)}$, while the first and second integrals are over two and two variables (or one and three variables) $\{\xi_{a_1}\}_{a_1 \in \pi_1^\infty(\mathcal{T}_3)}$ and $\{\xi_{a_2}\}_{a_2 \in \pi_2^\infty(\mathcal{T}_3)}$, with one constraint $\boldsymbol{\xi}_r = \xi$.

We first consider the case $|\widetilde{\mu}_3| \leqslant 7^3|\widetilde{\mu}_2|^{1-\delta}$. For each fixed ordered tree $\mathcal{T}_3 \in \mathfrak{T}(3)$, each septilinear term in (3.18) can be written as

$$\begin{aligned}&\mathcal{N}_1^{(3)}\big|_{\mathcal{T}_3} \\ &= \mathcal{I}_{|\mu_1|>N}\Big(v, v, \mathcal{I}^{\mu_1}_{|\mu_2+\mu_1|>5^3|\mu_1|^{1-\delta}}\big(v, v, \mathcal{N}^{\widetilde{\mu}_2}_{|\mu_3+\widetilde{\mu}_2|\leqslant 7^3|\widetilde{\mu}_2|^{1-\delta}}(v,v,v)\big)\Big)\end{aligned} \tag{3.19}$$

or

$$\mathcal{N}_1^{(3)}\big|_{\mathcal{T}_3} = \mathcal{I}_{|\mu_1|>N}\Big(\mathcal{I}^{\mu_1}_{|\mu_2+\mu_1|>5^3|\mu_1|^{1-\delta}}(v,v,v), v, \mathcal{N}^{\tilde{\mu}_2}_{|\mu_3+\tilde{\mu}_2|\leqslant 7^3|\tilde{\mu}_2|^{1-\delta}}(v,v,v)\Big) \quad (3.20)$$

up to permutations of terminal nodes within a subtree of one generation. In the following, we only consider (3.19) since (3.20) can be estimated in a similar manner. By dyadically decomposing μ_1 and $\widetilde{\mu}_2$, we have

$$(3.19) \sim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\gtrsim N_1^{1-\delta}\\ \text{dyadic}}} \mathcal{I}_{|\mu_1|\sim N_1}\Big(v,v,\mathcal{I}^{\mu_1}_{|\tilde{\mu}_2|\sim N_2}\big(v,v,\mathcal{N}^{\tilde{\mu}_2}_{|\mu_3+\tilde{\mu}_2|\leqslant 7^3|\tilde{\mu}_2|^{1-\delta}}(v,v,v)\big)\Big).$$

Then, by Lemmas 2.3 and 2.5, we can estimate (3.19) as

$$\begin{aligned}
\|(3.19)\|_{L^2} &\lesssim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\gtrsim N_1^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac12+}\|v\|_{L^2}^2 \\
&\qquad\qquad \times \Big\|\mathcal{I}^{\mu_1}_{|\tilde{\mu}_2|\sim N_2}\big(v,v,\mathcal{N}^{\tilde{\mu}_2}_{|\mu_3+\tilde{\mu}_2|\leqslant 7^3|\tilde{\mu}_2|^{1-\delta}}(v,v,v)\big)\Big\|_{L^2} \\
&\lesssim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\gtrsim N_1^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac12+}N_2^{-\frac12+}\|v\|_{L^2}^4 \\
&\qquad\qquad \times \Big\|\mathbb{1}_{|\tilde{\mu}_2|\sim N_2}\cdot\mathcal{N}^{\tilde{\mu}_2}_{|\mu_3+\tilde{\mu}_2|\leqslant 7^3|\tilde{\mu}_2|^{1-\delta}}(v,v,v)\Big\|_{L^2} \\
&\lesssim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\gtrsim N_1^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac12+}N_2^{-\frac{\delta}{2}+}\|v\|_{L^2}^7 \\
&\lesssim N^{-\frac12-\frac{\delta}{2}+\frac{\delta^2}{2}+}\|v\|_{L^2}^7.
\end{aligned}$$

Next, we consider the case $|\widetilde{\mu}_3| \leqslant 7^3|\mu_1|^{1-\delta}$. In this case, we need to estimate the terms of the form (3.19) and (3.20) with $|\mu_3+\widetilde{\mu}_2| \leqslant 7^3|\widetilde{\mu}_2|^{1-\delta}$ replaced by $|\mu_3+\widetilde{\mu}_2| \leqslant 7^3|\mu_1|^{1-\delta}$. Proceeding as before with Lemmas 2.3 and 2.5, we have

$$\begin{aligned}
\|(3.19)\|_{L^2} &\lesssim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\gtrsim N_1^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac12+}N_2^{-\frac12+}\|v\|_{L^2}^4 \\
&\qquad\qquad \times \Big\|\mathbb{1}_{|\tilde{\mu}_2|\sim N_2}\cdot\mathcal{N}^{\tilde{\mu}_2}_{|\mu_3+\tilde{\mu}_2|\leqslant 7^3|\mu_1|^{1-\delta}}(v,v,v)\Big\|_{L^2} \\
&\lesssim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\gtrsim N_1^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac12+}N_2^{-\frac12+}N_1^{\frac12-\frac{\delta}{2}+}\|v\|_{L^2}^7 \\
&\lesssim N^{-\frac12+}\|v\|_{L^2}^7.
\end{aligned}$$

This completes the proof of Lemma 3.12. □

As in the previous step, we can not estimate $\mathcal{N}_2^{(3)}$ in a direct manner. Hence, we perform the third step of normal form reductions:

$$\begin{aligned}\mathcal{N}_2^{(3)}(v)(\xi) = \partial_t\Bigg[\sum_{\mathcal{T}_3\in\mathfrak{T}(3)} \int_{\substack{C_0\cap C_1^c\cap C_2^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_3),\,\boldsymbol{\xi}_r=\xi}} \frac{e^{-i\tilde{\mu}_3 t}}{\prod_{j=1}^3 \widetilde{\mu}_j} \prod_{a\in\mathcal{T}_3^\infty} v_{\xi_a}\Bigg] \\ - \sum_{\mathcal{T}_4\in\mathfrak{T}(4)} \int_{\substack{C_0\cap C_1^c\cap C_2^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_4),\,\boldsymbol{\xi}_r=\xi}} \frac{e^{-i\tilde{\mu}_4 t}}{\prod_{j=1}^3 \widetilde{\mu}_j} \prod_{a\in\mathcal{T}_4^\infty} v_{\xi_a} \\ =: \partial_t\mathcal{N}_0^{(4)}(v)(\xi) + \mathcal{N}^{(4)}(v)(\xi).\end{aligned}$$

The boundary term $\mathcal{N}_0^{(4)}(v)$ can be estimated as in Lemmas 3.9 and 3.11. As for $\mathcal{N}^{(4)}(v)$, we decompose it as $\mathcal{N}^{(4)} = \mathcal{N}_1^{(4)} + \mathcal{N}_2^{(4)}$ corresponding to the restrictions onto

$$C_3 = \left\{|\widetilde{\mu}_4| \leqslant 9^3|\widetilde{\mu}_3|^{1-\delta}\right\} \cup \left\{|\widetilde{\mu}_4| \leqslant 9^3|\mu_1|^{1-\delta}\right\} \tag{3.21}$$

and its complement C_3^c, respectively. On the one hand, the modulation restriction (3.21) allows us to estimate $\mathcal{N}_1^{(4)}$ as in Lemmas 3.10 and 3.12. On the other hand, we apply the fourth step of normal form reductions to $\mathcal{N}_2^{(4)}$. In this way, we continue normal form reductions in an indefinite manner. In the next subsection, we describe this procedure in the general Jth step.

3.3. General Jth step

In this subsection, we discuss the general Jth step in this normal form procedure. Given an ordered tree $\mathcal{T} \in \mathfrak{T}(J)$, we introduce the following multilinear operators $\mathfrak{S}_0(\mathcal{T};\,\cdot\,)$ and $\mathfrak{S}_1(\mathcal{T};\,\cdot\,)$, which allow us to estimate the multilinear terms (associated with the ordered tree $\mathcal{T}$) in an efficient manner. For simplicity of notations, we set M_j by

$$M_j := \max(|\widetilde{\mu}_j|, |\mu_1|).$$

Definition 3.13. — *Let $k = 0, 1$. Then, we define $\mathfrak{S}_0$ and $\mathfrak{S}_1$ as mappings:*

$$\mathcal{T} \in \bigcup_{j=1}^\infty \mathfrak{T}(j) \longmapsto a\ (2j+1)\text{-linear map } \mathfrak{S}_k(\mathcal{T};\,\cdot\,) \text{ on } \mathcal{S}(\mathbb{R})^{\otimes 2j+1},\ k=0,1,$$

by the following rules. Let $v \in \mathcal{S}(\mathbb{R})$.

Definition of $\mathfrak{S}_0(\mathcal{T}; v)$:

(i) *Replace a terminal node (denoted as "•") by v.*

(ii) *Replace the* J*th root node* $r^{(J)}$ *(denoted as "$\diamond$") by the trilinear operator* $\mathcal{I}^{\tilde{\mu}_{J-1}}_{|\mu_J+\tilde{\mu}_{J-1}|>(2J+1)^3 M_{J-1}^{1-\delta}}$ *whose arguments are given by the functions associated with its three children (namely* v *in this case).*

(iii) *Let* $j = J-1$*. Replace the* j*th root node (denoted as "$\circ$") by the trilinear operator* $\mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j+\tilde{\mu}_{j-1}|>(2j+1)^3 M_{j-1}^{1-\delta}}$ *whose arguments are given by the functions associated with its three children. Repeat this process for* $j = J-2, J-3, \dots, 2$.

(iv) *Replace the root node* $r = r^{(1)}$ *(denoted as "$\star$") by the trilinear operator* $\mathcal{I}_{|\mu_1|>N}$ *whose arguments are given by the functions associated with its three children.*

Definition of $\mathfrak{S}_1(\mathcal{T}; v)$*:*

(i) *Replace a terminal node (denoted as "$\bullet$") by* v.

(ii) *Replace the* J*th root node* $r^{(J)}$ *(denoted as "$\diamond$") by the trilinear operator* $\mathcal{N}^{\tilde{\mu}_{J-1}}_{|\mu_J+\tilde{\mu}_{J-1}|\leqslant(2J+1)^3 M_{J-1}^{1-\delta}}$ *whose arguments are given by the functions associated with its three children (namely* v *in this case).*

(iii) *Let* $j = J-1$*. Replace the* j*th root node (denoted as "$\circ$") by the trilinear operator* $\mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j+\tilde{\mu}_{j-1}|>(2j+1)^3 M_{j-1}^{1-\delta}}$ *whose arguments are given by the functions associated with its three children. Repeat this process for* $j = J-2, J-3, \dots, 2$.

(iv) *Replace the root node* $r = r^{(1)}$ *(denoted as "$\star$") by the trilinear operator* $\mathcal{I}_{|\mu_1|>N}$ *whose arguments are given by the functions associated with its three children.*

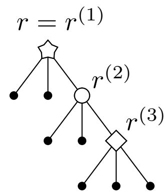
Note that the only difference between the two definitions appears in Step (ii). The operators $\mathfrak{S}_0(\mathcal{T};\,\cdot\,)$ and $\mathfrak{S}_1(\mathcal{T};\,\cdot\,)$ are a priori defined from $\mathcal{S}(\mathbb{R})^{\otimes 2j+1}$ to $\mathcal{S}'(\mathbb{R})$. In the following, we show that they are bounded on $L^2(\mathbb{R})$.

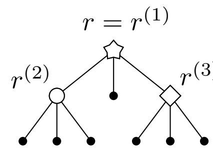
Remark 3.14. — In the above definition, we only defined $\mathfrak{S}_0(\mathcal{T}; v)$ and $\mathfrak{S}_1(\mathcal{T}; v)$, namely, when all the $2j+1$ arguments are identical. Let us now describe how to define $\mathfrak{S}_k(\mathcal{T}; v_1, \dots, v_{2j+1})$, $k = 0, 1$, in general.

Given a tree $\mathcal{T} \in \mathfrak{T}(j)$, label its terminal nodes by $a_1, \dots, a_{2j+1}$ (say, by moving from left to right in the planar graphical representation of the tree). Given functions $v_1, \dots, v_{2j+1} \in \mathcal{S}(\mathbb{R})$, we only need to modify Step (i) in Definition 3.13 as follows:

(i′) Replace terminal nodes $a_\ell \in \mathcal{T}^\infty$ by v_ℓ.

Before proceeding further, let us consider the following examples of ordered trees of the third generation:

$\mathcal{T} =$ 

$\mathcal{T}' =$ 

It is easy to see that $\mathfrak{S}_1(\mathcal{T}; v)$ and $\mathfrak{S}_1(\mathcal{T}'; v)$ correspond to the septilinear terms (3.19) and (3.20), respectively.

Next, let $\mathbf{T}$ be the collection of formal sums of elements in $\bigcup_{j=1}^\infty \mathfrak{T}(j)$. Then, we extend the definitions of $\mathfrak{S}_0$ and $\mathfrak{S}_1$ to elements in $\mathbf{T}$ by imposing the "additivity":

$$\mathfrak{S}_k\bigg(\sum_{\alpha\in\mathcal{A}} \mathcal{T}^\alpha;\, \cdot\,\bigg) := \sum_{\alpha\in\mathcal{A}} \mathfrak{S}_k(\mathcal{T}^\alpha;\, \cdot\,) \tag{3.22}$$

for a finite index set $\mathcal{A}$. With this definition, we can write $\mathcal{N}_0^{(3)}(v)$ and $\mathcal{N}_1^{(3)}(v)$ from the previous subsection as

$$\mathcal{N}_0^{(3)}(v) = \mathfrak{S}_0\bigg(\sum_{\mathcal{T}\in\mathfrak{T}(2)} \mathcal{T}; v\bigg) \qquad \text{and} \qquad \mathcal{N}_1^{(3)}(v) = \mathfrak{S}_1\bigg(\sum_{\mathcal{T}\in\mathfrak{T}(3)} \mathcal{T}; v\bigg).$$

Now, we are ready to discuss the general Jth step of the normal form reductions. Define C_j by

$$\begin{aligned} C_j &= \big\{|\widetilde{\mu}_{j+1}| \leqslant (2j+3)^3 M_j^{1-\delta}\big\} \\ &= \big\{|\widetilde{\mu}_{j+1}| \leqslant (2j+3)^3 |\widetilde{\mu}_j|^{1-\delta}\big\} \cup \big\{|\widetilde{\mu}_{j+1}| \leqslant (2j+3)^3 |\mu_1|^{1-\delta}\big\} \end{aligned} \tag{3.23}$$

for $j \in \mathbb{N}$. Then, after J steps, we have

$$\begin{aligned} \mathcal{N}_2^{(J)}(v)(\xi) &= \sum_{\mathcal{T}_J\in\mathfrak{T}(J)} \int_{\substack{C_0\cap\bigcap_{j=1}^{J-1} C_j^c \\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_J),\, \boldsymbol{\xi}_r=\xi}} \frac{e^{-i\tilde{\mu}_J t}}{\prod_{j=1}^{J-1}\widetilde{\mu}_j} \prod_{a\in\mathcal{T}_J^\infty} v_{\xi_a} \\ &= \partial_t\Bigg[\sum_{\mathcal{T}_J\in\mathfrak{T}(J)} \int_{\substack{C_0\cap\bigcap_{j=1}^{J-1} C_j^c \\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_J),\, \boldsymbol{\xi}_r=\xi}} \frac{e^{-i\tilde{\mu}_J t}}{\prod_{j=1}^{J}\widetilde{\mu}_j} \prod_{a\in\mathcal{T}_J^\infty} v_{\xi_a}\Bigg] \\ &\qquad\qquad - \sum_{\mathcal{T}_{J+1}\in\mathfrak{T}(J+1)} \int_{\substack{C_0\cap\bigcap_{j=1}^{J-1} C_j^c\cap C_J \\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_{J+1}),\, \boldsymbol{\xi}_r=\xi}} \frac{e^{-i\tilde{\mu}_{J+1} t}}{\prod_{j=1}^{J}\widetilde{\mu}_j} \prod_{a\in\mathcal{T}_{J+1}^\infty} v_{\xi_a} \\ &\qquad\qquad - \sum_{\mathcal{T}_{J+1}\in\mathfrak{T}(J+1)} \int_{\substack{C_0\cap\bigcap_{j=1}^{J} C_j^c \\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_{J+1}),\, \boldsymbol{\xi}_r=\xi}} \frac{e^{-i\tilde{\mu}_{J+1} t}}{\prod_{j=1}^{J}\widetilde{\mu}_j} \prod_{a\in\mathcal{T}_{J+1}^\infty} v_{\xi_a} \end{aligned}$$

$$=: \partial_t \mathcal{N}_0^{(J+1)}(v)(\xi) + \mathcal{N}_1^{(J+1)}(v)(\xi) + \mathcal{N}_2^{(J+1)}(v)(\xi). \tag{3.24}$$

As in the previous subsection, let

$$\mathcal{N}^{(J+1)} := \mathcal{N}_1^{(J+1)} + \mathcal{N}_2^{(J+1)}. \tag{3.25}$$

In view of Definition 3.13, we have

$$\mathcal{N}_0^{(J+1)}(v) = \mathfrak{S}_0\Bigg(\sum_{\mathcal{T}\in\mathfrak{T}(J)} \mathcal{T}; v\Bigg)$$

$$\text{and} \quad \mathcal{N}_1^{(J+1)}(v) = \mathfrak{S}_1\Bigg(\sum_{\mathcal{T}\in\mathfrak{T}(J+1)} \mathcal{T}; v\Bigg). \tag{3.26}$$

In the following, we estimate $\mathcal{N}_0^{(J+1)}$ and $\mathcal{N}_1^{(J+1)}$ for general $J \in \mathbb{N}$. As for the last term $\mathcal{N}_2^{(J+1)}$ in (3.24), we perform a normal form reduction once again and obtain (3.24) with J replaced by $J+1$. In Section 4, we show that the remainder term $\mathcal{N}_2^{(J+1)}$ tends to 0 in an appropriate sense as $J \to \infty$.

Lemma 3.15. — *Let $s \geqslant 0$. Then, we have*

$$\|\mathcal{N}_0^{(J+1)}(v)\|_{H^s} \lesssim N^{-\frac{J}{2}+\frac{J-1}{2}\delta+} \|v\|_{H^s}^{2J+1}, \tag{3.27}$$

$$\|\mathcal{N}_0^{(J+1)}(v) - \mathcal{N}_0^{(J+1)}(w)\|_{H^s} \lesssim N^{-\frac{J}{2}+\frac{J-1}{2}\delta+} \big(\|v\|_{H^s}^{2J} + \|w\|_{H^s}^{2J}\big)\|v-w\|_{H^s}, \tag{3.28}$$

for $0<\delta<1$.

Proof. — We only present the proof of (3.27) since (3.28) follows in a similar manner. Note that there is an extra factor $\sim J$ when we estimate the difference in (3.28) since $|a^{2J+1} - b^{2J+1}| \lesssim \big(\sum_{j=1}^{2J+1} a^{2J+1-j} b^{j-1}\big)|a-b|$ has $O(J)$ many terms. This, however, does not cause a problem since the constant we obtain decays like a power of a factorial in J (as we see below in (3.32)). The same comment applies to Lemma 3.16 below.

Moreover, we claim that it suffices to prove (3.27) for $s=0$. When $s>0$, we argue as follows. Fix an ordered tree $\mathcal{T} \in \mathfrak{T}(J)$ and an index function $\boldsymbol{\xi} \in \Xi(\mathcal{T})$ with $\boldsymbol{\xi}_r = \xi$. By the triangle inequality, we have $\max_{k=1,2,3}\langle \xi_k^{(j)}\rangle \geqslant \frac{1}{3}\langle \xi^{(j)}\rangle$, since we have $\xi^{(j)} = \xi_1^{(j)} - \xi_2^{(j)} + \xi_3^{(j)}$. Hence, there exists at least one terminal node $a \in \mathcal{T}^\infty$ such that

$$\langle \xi\rangle^s \leqslant 3^{Js}\langle \xi_a\rangle^s.$$

Note that the constant grows exponentially in J. However, this exponential growth does not cause a problem thanks to the factorial decay in the denominator in (3.32) below.

From (3.22) and (3.26), we have

$$\|\mathcal{N}_0^{(J+1)}(v)\|_{L^2} \leqslant c_J \sup_{\mathcal{T}\in\mathfrak{T}(J)} \|\mathfrak{S}_0(\mathcal{T};v)\|_{L^2}, \tag{3.29}$$

where $c_J = |\mathfrak{T}(J)|$ as in (3.1). We now decompose $\mathfrak{S}_0(\mathcal{T};v)$ into dyadic pieces in terms of modulations $\widetilde{\mu}_j$. Given dyadic N_j, $j=1,\dots,J$, define $\widetilde{M}_j$ by

$$\widetilde{M}_j := \max(N_j, N_1). \tag{3.30}$$

With $\bar{N} = (N_1,\dots,N_J)$, we define $\mathfrak{S}_{0,\bar{N}}(\mathcal{T};v)$ by making the following modifications in Steps (ii), (iii), and (iv) of the definition of $\mathfrak{S}_0(\mathcal{T};v)$:

$$\begin{array}{cccc}
\text{(ii)} & \mathcal{I}^{\tilde{\mu}_{J-1}}_{|\mu_J+\tilde{\mu}_{J-1}|>(2J+1)^3 M_{J-1}^{1-\delta}} & \Longrightarrow & \mathcal{I}^{\tilde{\mu}_{J-1}}_{|\mu_J+\tilde{\mu}_{J-1}|\sim N_J}, \\
\text{(iii)} & \mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j+\tilde{\mu}_{j-1}|>(2j+1)^3 M_{j-1}^{1-\delta}} & \Longrightarrow & \mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j+\tilde{\mu}_{j-1}|\sim N_j}, \\
\text{(iv)} & \mathcal{I}_{|\mu_1|>N} & \Longrightarrow & \mathcal{I}_{|\mu_1|\sim N_1}.
\end{array}$$

Then, we have

$$\mathfrak{S}_0(\mathcal{T};v) \sim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\geqslant 5^3\widetilde{M}_1^{1-\delta}\\ \text{dyadic}}} \cdots \sum_{\substack{N_J\geqslant (2J+1)^3\widetilde{M}_{J-1}^{1-\delta}\\ \text{dyadic}}} \mathfrak{S}_{0,\bar{N}}(\mathcal{T};v). \tag{3.31}$$

Fix an ordered tree $\mathcal{T}\in\mathfrak{T}(J)$. In view of Remark 3.6, we can estimate $\mathfrak{S}_{0,\bar{N}}(\mathcal{T};v)$ by applying Lemma 2.5 in a successive manner in the following order:

$$\mathcal{I}_{|\mu_1|\sim N_1},\ \mathcal{I}^{\mu_1}_{|\mu_2+\mu_1|\sim N_2},\ \mathcal{I}^{\tilde{\mu}_2}_{|\mu_3+\tilde{\mu}_2|\sim N_3},\dots,\ \mathcal{I}^{\tilde{\mu}_{J-1}}_{|\mu_J+\tilde{\mu}_{J-1}|\sim N_J}.$$

Then, it follows from Lemma 2.5 with (3.29), (3.31), and (3.30), that

$$\begin{aligned}
\|\mathcal{N}_0^{(J+1)}(v)\|_{L^2} &\leqslant c_J \sup_{\mathcal{T}\in\mathfrak{T}(J)} \|\mathfrak{S}_0(\mathcal{T};v)\|_{L^2} \\
&\lesssim c_J \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\geqslant 5^3\widetilde{M}_1^{1-\delta}\\ \text{dyadic}}} \cdots \\
&\qquad\qquad \sum_{\substack{N_J\geqslant (2J+1)^3\widetilde{M}_{J-1}^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac{1}{2}+} \prod_{j=2}^J N_j^{-\frac{1}{2}+} \|v\|_{L^2}^{2J+1} \\
&\lesssim \frac{c_J}{\prod_{j=2}^J (2j+1)^{\frac{3}{2}-}} \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} N_1^{-\frac{J}{2}+\frac{J-1}{2}\delta+} \|v\|_{L^2}^{2J+1} \qquad (3.32)\\
&\lesssim N^{-\frac{J}{2}+\frac{J-1}{2}\delta+} \|v\|_{L^2}^{2J+1}.
\end{aligned}$$

This completes the proof of Lemma 3.15. $\square$

A similar argument yields the following bounds on $\mathcal{N}_1^{(J+1)}(v)$.

LEMMA 3.16. — *Let $s \geqslant 0$. Then, we have*

$$\|\mathcal{N}_1^{(J+1)}(v)\|_{H^s} \lesssim N^{-\frac{J-1}{2}+\frac{J-2}{2}\delta+}\|v\|_{H^s}^{2J+3}, \tag{3.33}$$
$$\|\mathcal{N}_1^{(J+1)}(v)-\mathcal{N}_1^{(J+1)}(w)\|_{H^s} \lesssim N^{-\frac{J-1}{2}+\frac{J-2}{2}\delta+}\big(\|v\|_{H^s}^{2J+2}+\|w\|_{H^s}^{2J+2}\big)\|v-w\|_{H^s},$$

for $0<\delta<1$.

Proof. — Arguing as in the proof of Lemma 3.15, it suffices to prove (3.33) for $s=0$. From (3.22) and (3.26), we have

$$\|\mathcal{N}_1^{(J+1)}(v)\|_{L^2} \leqslant c_{J+1} \sup_{\mathcal{T}\in\mathfrak{T}(J+1)} \|\mathfrak{S}_1(\mathcal{T};v)\|_{L^2}. \tag{3.34}$$

As in the proof of Lemma 3.15, we decompose $\mathfrak{S}_1(\mathcal{T};v)$ into dyadic pieces in terms of modulations $\widetilde{\mu}_j$. With $\bar{N}=(N_1,\dots,N_{J+1})$, we define $\mathfrak{S}_{1,\bar{N}}(\mathcal{T};v)$ by making the following modifications in Steps (ii), (iii), and (iv) of the definition of $\mathfrak{S}_1(\mathcal{T};v)$ (with J replaced by $J+1$):

$$\begin{array}{lccc}
\text{(ii)} & \mathcal{N}^{\tilde{\mu}_J}_{|\mu_{J+1}+\tilde{\mu}_J|\leqslant(2J+3)^3M_J^{1-\delta}} & \Longrightarrow & \mathcal{N}^{\tilde{\mu}_J}_{|\mu_{J+1}+\tilde{\mu}_J|\sim N_{J+1}}, \\
\text{(iii)} & \mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j+\tilde{\mu}_{j-1}|>(2j+1)^3M_{j-1}^{1-\delta}} & \Longrightarrow & \mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j+\tilde{\mu}_{j-1}|\sim N_j}, \\
\text{(iv)} & \mathcal{I}_{|\mu_1|>N} & \Longrightarrow & \mathcal{I}_{|\mu_1|\sim N_1},
\end{array}$$

where $\widetilde{M}_j$ is as in (3.30). Then, we have

$$\mathfrak{S}_1(\mathcal{T};v)\sim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}}\sum_{\substack{N_2\geqslant 5^3\widetilde{M}_1^{1-\delta}\\ \text{dyadic}}}\cdots \sum_{\substack{N_J\geqslant(2J+1)^3\widetilde{M}_{J-1}^{1-\delta}\\ \text{dyadic}}}\ \sum_{\substack{N_{J+1}\leqslant 2^{-1}\cdot(2J+3)^3\widetilde{M}_J^{1-\delta}\\ \text{dyadic}}} \mathfrak{S}_{1,\bar{N}}(\mathcal{T};v). \tag{3.35}$$

Fix an ordered tree $\mathcal{T}\in\mathfrak{T}(J+1)$. Proceeding as before, we can estimate $\mathfrak{S}_{1,\bar{N}}(\mathcal{T};v)$ by applying Lemmas 2.3 and 2.5 in a successive manner in the following order:

$$\mathcal{I}_{|\mu_1|\sim N_1},\ \mathcal{I}^{\mu_1}_{|\mu_2+\mu_1|\sim N_2},\dots,\ \mathcal{I}^{\tilde{\mu}_{J-1}}_{|\mu_J+\tilde{\mu}_{J-1}|\sim N_J},\ \mathcal{N}^{\tilde{\mu}_J}_{|\mu_{J+1}+\tilde{\mu}_J|\sim N_{J+1}}.$$

We first consider the contribution from the case $\widetilde{M}_J \sim N_J$. It follows from Lemmas 2.3 and 2.5 with (3.34), (3.35), and (3.30) that

$$\begin{aligned}
\|\mathcal{N}_1^{(J+1)}(v)\|_{L^2} &\leqslant c_{J+1} \sup_{\mathcal{T}\in\mathfrak{T}(J)} \|\mathfrak{S}_1(\mathcal{T};v)\|_{L^2} \\
&\lesssim c_{J+1} \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\geqslant 5^3\widetilde{M}_1^{1-\delta}\\ \text{dyadic}}} \cdots \sum_{\substack{N_J\geqslant (2J+1)^3\widetilde{M}_{J-1}^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac{1}{2}+} \prod_{j=2}^{J} N_j^{-\frac{1}{2}+} \\
&\qquad \times \sum_{\substack{N_{J+1}\leqslant 2^{-1}\cdot(2J+3)^3 N_J^{1-\delta}\\ \text{dyadic}}} N_{J+1}^{\frac{1}{2}+} \|v\|_{L^2}^{2J+3} \\
&\lesssim c_{J+1}(2J+3)^{\frac{3}{2}+} \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\geqslant 5^3\widetilde{M}_1^{1-\delta}\\ \text{dyadic}}} \cdots \sum_{\substack{N_J\geqslant (2J+1)^3\widetilde{M}_{J-1}^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac{1}{2}+} N_J^{-\frac{\delta}{2}+} \\
&\qquad \times \prod_{j=2}^{J-1} N_j^{-\frac{1}{2}+} \|v\|_{L^2}^{2J+3} \\
&\lesssim \frac{c_{J+1}(2J+3)^{\frac{3}{2}+}}{\prod_{j=2}^{J-1}(2j+1)^{\frac{3}{2}-}} \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} N_1^{-\frac{1}{2}+} N_1^{-\frac{\delta}{2}+\frac{\delta^2}{2}+} N_1^{-\frac{J-2}{2}+\frac{J-2}{2}\delta+} \|v\|_{L^2}^{2J+3} \\
&\lesssim \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} N_1^{-\frac{J-1}{2}+\frac{J-3}{2}\delta+\frac{\delta^2}{2}+} \|v\|_{L^2}^{2J+3} \lesssim N^{-\frac{J-1}{2}+\frac{J-3}{2}\delta+\frac{\delta^2}{2}+} \|v\|_{L^2}^{2J+3}.
\end{aligned}$$

Next, we consider the contribution from the case $\widetilde{M}_J \sim N_1$. Proceeding as above, we have

$$\begin{aligned}
&\|\mathcal{N}_1^{(J+1)}(v)\|_{L^2} \\
&\qquad \leqslant c_{J+1} \sup_{\mathcal{T}\in\mathfrak{T}(J)} \|\mathfrak{S}_1(\mathcal{T};v)\|_{L^2} \\
&\qquad \lesssim c_{J+1} \sum_{\substack{N_1\geqslant N\\ \text{dyadic}}} \sum_{\substack{N_2\geqslant 5^3\widetilde{M}_1^{1-\delta}\\ \text{dyadic}}} \cdots \sum_{\substack{N_J\geqslant (2J+1)^3\widetilde{M}_{J-1}^{1-\delta}\\ \text{dyadic}}} N_1^{-\frac{1}{2}+} \prod_{j=2}^{J} N_j^{-\frac{1}{2}+} \\
&\qquad\qquad \times \sum_{\substack{N_{J+1}\leqslant 2^{-1}\cdot(2J+3)^3 N_1^{1-\delta}\\ \text{dyadic}}} N_{J+1}^{\frac{1}{2}+} \|v\|_{L^2}^{2J+3}
\end{aligned}$$

$$
\begin{aligned}
&\lesssim c_{J+1}(2J+3)^{\frac{3}{2}+} \sum_{\substack{N_1 \geqslant N \\ \text{dyadic}}} \sum_{\substack{N_2 \geqslant 5^3 \widetilde{M}_1^{1-\delta} \\ \text{dyadic}}} \cdots \\
&\qquad\qquad \sum_{\substack{N_J \geqslant (2J+1)^3 \widetilde{M}_{J-1}^{1-\delta} \\ \text{dyadic}}} N_1^{-\frac{\delta}{2}+} \prod_{j=2}^{J} N_j^{-\frac{1}{2}+} \|v\|_{L^2}^{2J+3} \\
&\lesssim \frac{c_{J+1}(2J+3)^{\frac{3}{2}+}}{\prod_{j=2}^{J}(2j+1)^{\frac{3}{2}-}} \sum_{\substack{N_1 \geqslant N \\ \text{dyadic}}} N_1^{-\frac{J-1}{2}+\frac{J-2}{2}\delta+} \|v\|_{L^2}^{2J+3} \\
&\lesssim N^{-\frac{J-1}{2}+\frac{J-2}{2}\delta+} \|v\|_{L^2}^{2J+3}.
\end{aligned}
$$

This completes the proof of Lemma 3.16. □

Remark 3.17. — As mentioned at the beginning of this section, we can perform an analogous analysis for the mKdV (1.2). In this case, it follows from Lemmas 2.6 and 2.8 that Lemmas 3.15 and 3.16 hold for $s \geqslant \frac{1}{4}$.

3.4. Normal form equation

After the Jth step of the normal form reductions, we transformed the original equation (2.3) to

$$
\begin{aligned}
\partial_t v(\xi) &= \mathcal{N}_{\leqslant N}(v)(\xi) + \mathcal{N}_{>N}(v)(\xi) \\
&= \sum_{j=2}^{J+1} \partial_t \mathcal{N}_0^{(j)}(v)(\xi) + \sum_{j=1}^{J+1} \mathcal{N}_1^{(j)}(v)(\xi) + \mathcal{N}_2^{(J+1)}(v)(\xi).
\end{aligned} \tag{3.36}
$$

By iterating this procedure indefinitely, we *formally*[17] arrive at the following limit equation:

$$
\partial_t v(\xi) = \partial_t \left(\sum_{j=2}^{\infty} \mathcal{N}_0^{(j)}(v)(\xi) \right) + \sum_{j=1}^{\infty} \mathcal{N}_1^{(j)}(v)(\xi). \tag{3.37}
$$

[17] This means that the derivation can be easily justified for smooth solutions but not for rough solutions. Here, we assume that the remainder term $\mathcal{N}_2^{(J+1)}(v)(\xi)$ tends to 0 as $J \to \infty$. In Section 4, we justify all the computations for rough functions, namely, $u \in C_t H^s$ with $s \geqslant \frac{1}{6}$ for the cubic NLS and $s > \frac{1}{4}$ for the mKdV.

Integrating (3.37) in time and applying the Fourier inversion formula, we obtain the following *normal form equation*:

$$v(t) = \Gamma_{u_0}(v)$$
$$:= u_0 + \left[\sum_{j=2}^{\infty} \mathcal{N}_0^{(j)}(v(t)) - \sum_{j=2}^{\infty} \mathcal{N}_0^{(j)}(u_0)\right] + \int_0^t \sum_{j=1}^{\infty} \mathcal{N}_1^{(j)}(v(t'))\mathrm{d}t'. \tag{3.38}$$

Theorem 3.18. — *The normal form equation* (3.38) *is unconditionally locally well-posed in* $H^s(\mathbb{R})$ *with*

$$\text{(i) } s \geqslant 0 \textit{ for the cubic NLS } (1.1) \quad \textit{and} \quad \text{(ii) } s \geqslant \tfrac{1}{4} \textit{ for the mKdV } (1.2). \tag{3.39}$$

Proof. — Given $u_0 \in H^s(\mathbb{R})$, let $R = 1 + \|u_0\|_{H^s}$. Given $T > 0$, we use B_{2R} to denote the closed ball of radius $2R$ in $C_T H^s := C([0,T]; H^s(\mathbb{R}))$ centered at the origin. By (3.4), Lemmas 3.9–3.12, 3.15, and 3.16, we have

$$\begin{aligned} \|\Gamma_{u_0}(v)\|_{C_T H^s} \leqslant \|u_0\|_{H^s} + C \sum_{j=2}^{J} N^{-\frac{j-1}{2}+\frac{j-2}{2}\delta+} \big(\|v\|_{C_T H^s}^{2j-1} + \|u_0\|_{H^s}^{2j-1}\big) \\ + CT\left\{N^{\frac{1}{2}+}\|v\|_{C_T H^s}^3 + \sum_{j=2}^{J} N^{-\frac{j-2}{2}+\frac{j-3}{2}\delta+}\|v\|_{C_T H^s}^{2j+1}\right\} \end{aligned} \tag{3.40}$$

for s satisfying (3.39). Note that the estimate (3.4) on $\mathcal{N}_1^{(1)}$ is the only estimate with a positive power of N. However, it appears inside the time integral in (3.38). This allows us to (i) choose $N = N(R) \gtrsim 1$, guaranteeing the convergence of the geometric series in (3.40) for $v \in B_{2R}$, and then (ii) choose $T = T(N) = T(R) > 0$ sufficiently small to control $TN^{\frac{1}{2}+}\|v\|_{C_T H^s}^3$. A similar estimate also holds on the difference $\|\Gamma_{u_0}(v) - \Gamma_{u_0}(w)\|_{C_T H^s}$ for $v, w \in B_{2R}$. Then, by a standard fixed point argument, we can show that the normal form equation (3.38) is locally well-posed in $C([0,T]; H^s(\mathbb{R}))$, provided that s satisfies (3.39). While the fixed point argument yields this uniqueness only in the ball $B_{2R} \subset C([0,T]; H^s(\mathbb{R}))$, we can apply a standard continuity argument to upgrade uniqueness to that in the entire $C([0,T]; H^s(\mathbb{R}))$ (by possibly shrinking the local existence time). See Remark 2.9 in [9] for example. Lastly, by considering the difference of two solutions $v_1, v_2 \in C([0,T]; H^s(\mathbb{R}))$ to (3.38), we also obtain

$$\|v_1 - v_2\|_{C_T H^s} \lesssim \|v_1(0) - v_2(0)\|_{H^s}$$

for small $T = T(\|v_1(0)\|_{H^s}, \|v_2(0)\|_{H^s}) > 0$. This proves Theorem 3.18. □

In the following, we sketch the proofs of Theorems 1.5 and 1.6, assuming that smooth solutions to the cubic NLS (1.1) (or the mKdV (1.2)) satisfy the normal form equation (3.38) (which we will show in the next section). By starting with two smooth solutions $u_1, u_2 \in C([0,T]; H^\infty(\mathbb{R}))$ to the cubic NLS (1.1) (or the mKdV (1.2)), the above analysis yields

$$\|u_1 - u_2\|_{C_T H^s} \lesssim \|u_1(0) - u_2(0)\|_{H^s} \tag{3.41}$$

for s satisfying (3.39). The difference estimate (3.41) in particular implies the convergence of approximating solutions (to a unique limiting function), yielding the local well-posedness in the sense of sensible weak solutions claimed in Theorems 1.5 and 1.6. See [37] for details. Furthermore, arguing as in [15], we can also show that solutions to the normal form equation (3.38) are indeed weak solutions in the extended sense to the original equation. Since the argument is straightforward, we omit details.

If we justify that solutions to the cubic NLS in $C([0,T]; H^s(\mathbb{R}))$, $s \geqslant \frac{1}{6}$ (and $s > \frac{1}{4}$ for the mKdV), indeed satisfy the normal form equation (3.38), then the difference estimate (3.41) yields uniqueness. Since our analysis does not involve any auxiliary function spaces, such uniqueness is unconditional, thus establishing Theorems 1.1 and 1.2. In the next section, we justify all the steps of the normal form reductions and thus the derivation of the normal form equation (3.38) under the regularity assumption above.

4. Unconditional well-posedness

In this section, we present the proof of Theorems 1.1 and 1.2. In view of the (conditional) well-posedness results in $H^s(\mathbb{R})$: $s \geqslant 0$ for the cubic NLS [42] and $s \geqslant \frac{1}{4}$ for the mKdV [10, 20, 24], we focus on proving unconditional uniqueness, locally in time. As mentioned above, the main task is to make the formal computations in Section 3 rigorous. Once this is achieved, the difference estimate (3.41) yields unconditional uniqueness. In the following, we justify

(i) the application of the product rule and
(ii) switching time derivatives and integrals in spatial frequencies (for each fixed $\xi \in \mathbb{R}$)

in the normal form reductions (3.6), (3.12), and (3.24). Moreover, we show that

(iii) the remainder term $\mathcal{N}_2^{(J+1)}(v)(\xi)$ in (3.36) tends to 0 as $J \to \infty$ (for each fixed $\xi \in \mathbb{R}$).[18]

[18] This part is not explicitly written in [15]. It is, however, easy to see that the computation in [15, (5.3)] and its generalization for the Jth generation (which follows as

It is crucial to note that we verify (i)–(iii) for each *fixed* $\xi \in \mathbb{R}$, namely, in a weaker topology than the H^s-topology used in Section 3. Moreover, while all the multilinear estimates (Lemmas 3.9–3.12, 3.15, and 3.16) for the cubic NLS in Section 3 hold for $s \geqslant 0$, we need an extra regularity $s \geqslant \frac{1}{6}$ in justifying (i), (ii), and (iii). As for the mKdV, the regularity $s \geqslant \frac{1}{4}$ suffices for the multilinear estimates in Section 3, while a slightly higher regularity $s > \frac{1}{4}$ is needed for justifying the normal form reductions.

4.1. Unconditional well-posedness for the cubic NLS

Let u be a solution to (1.1) in $C(\mathbb{R}; H^s(\mathbb{R}))$ for some $s \geqslant \frac{1}{6}$ and let v be the corresponding interaction representation defined by (1.11). On the one hand, by Sobolev's inequality, we have $|u|^2 u \in C(\mathbb{R}; L^1(\mathbb{R}))$. On the other hand, it follows from (2.2) and (2.3) that $\widehat{v}(\xi)$ satisfies

$$\partial_t \widehat{v}(\xi, t) = i e^{-it\xi^2} \mathcal{F}(|u|^2 u)(\xi, t)$$

for each $\xi \in \mathbb{R}$. Hence, by Riemann–Lebesgue lemma, we conclude that $\widehat{v}(\xi)$ is a C^1-function in t for each fixed $\xi \in \mathbb{R}$. This justifies (i) the application of the product rule in Section 3, provided that $s \geqslant \frac{1}{6}$.

Next, we justify the exchange of time derivatives and integrals in spatial frequencies. Before proceeding further, we first need to present several multilinear estimates. From (2.3) with Hausdorff–Young's inequality, Sobolev's inequality, and the unitarity of the linear propagator $S(t) = e^{-it\partial_x^2}$, we have

$$\|\partial_t v\|_{\mathcal{F}L^\infty} = \|\mathcal{N}(v)\|_{\mathcal{F}L^\infty} \leqslant \|u\|_{L^3}^3 \lesssim \|u\|_{H^{1/6}}^3 = \|v\|_{H^{1/6}}^3, \tag{4.1}$$

where the $\mathcal{F}L^\infty$-norm is defined in (1.9). Note that the same estimate holds for $\mathcal{N}_M^\alpha$ and $\mathcal{N}_{\leqslant M}^\alpha$.[19]

We also need the following $\mathcal{F}L^\infty$-estimates, i.e. uniform estimates in spatial frequencies. Our main goal is to prove Lemma 4.3 below, controlling the $\mathcal{F}L^\infty$-norms of the multilinear terms $\mathcal{N}^{(J+1)}(v)$ and $\mathcal{N}_2^{(J+1)}(v)$ in terms of the $H^{\frac{1}{6}}$-norm of v.

---

a minor modification of [15, Lemma 3.11] with (4.1) below) would imply (iii) for the cubic NLS on $\mathbb{T}$.

(19) In this case, we simply take the absolute values of the Fourier coefficients of each argument and drop a modulation restriction. For example, we have

$$\|\mathcal{N}_M^\alpha(v)\|_{\mathcal{F}L^\infty} \leqslant \left\|\mathcal{F}^{-1}(|\widehat{v}|)\right\|_{L^3}^3 \lesssim \left\|\mathcal{F}^{-1}(|\widehat{v}|)\right\|_{H^{1/6}}^3 = \|v\|_{H^{1/6}}^3,$$

where we used the fact that the $H^{\frac{1}{6}}(\mathbb{R})$ is a Fourier lattice in the last step.

LEMMA 4.1 (Localized modulation estimates for the cubic NLS in the weak norm). — *Given $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that*

$$\|\mathcal{N}^\alpha_{\leqslant M}(v_1, v_2, v_3)\|_{\mathcal{F}L^\infty} \leqslant C_\varepsilon M^{\frac{1}{2}} \min_{j=1,2,3}\left(\|v_j\|_{\mathcal{F}L^\infty} \prod_{\substack{k=1\\k\neq j}}^{3} \|v_k\|_{H^\varepsilon}\right) \tag{4.2}$$

for any $M \geqslant 1$ and $\alpha \in \mathbb{R}$.

Proof. — By duality, (4.2) follows once we prove[20]

$$\left|\int_{\xi=\xi_1-\xi_2+\xi_3} \mathbb{1}_{|\Phi(\bar{\xi})-\alpha|\leqslant M} v_1(\xi_1)v_2(\xi_2)v_3(\xi_3)v_4(\xi)\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}\xi\right| \\ \lesssim M^{\frac{1}{2}} \min_{j=1,2,3}\left(\|v_j\|_{L^\infty_\xi} \prod_{\substack{k=1\\k\neq j}}^{3} \|\langle\xi\rangle^\varepsilon v_k\|_{L^2_\xi}\right)\|v_4\|_{L^1_\xi} \tag{4.3}$$

for all non-negative functions $v_1, \ldots, v_4 \in L^2_\xi(\mathbb{R})$.

Case 1: $\max(|\xi_{2-1}|, |\xi_{2-3}|) \leqslant 1$. — Let $\zeta = \xi_2 - \xi_1$ and $\widetilde{\zeta} = \xi_2 - \xi_3$. Then, thanks to the restriction $|\zeta|, |\widetilde{\zeta}| \leqslant 1$, we have

$$\begin{aligned} &\text{LHS of (4.3)} \\ &\quad\leqslant \sup_\xi \left|\int_{|\zeta|\leqslant 1}\int_{|\tilde{\zeta}|\leqslant 1} v_1(\xi+\widetilde{\zeta})v_2(\xi+\zeta+\widetilde{\zeta})v_3(\xi+\zeta)\mathrm{d}\widetilde{\zeta}\mathrm{d}\zeta\right| \cdot \|v_4\|_{L^1_\xi} \\ &\quad\lesssim \min_{j=1,2,3}\left(\|v_j\|_{L^\infty_\xi} \prod_{\substack{k=1\\k\neq j}}^{3} \|v_k\|_{L^2_\xi}\right)\|v_4\|_{L^1_\xi}. \end{aligned}$$

Case 2: $\max(|\xi_{2-3}|, |\xi_{2-1}|) > 1$. — Without loss of generality, assume that $\xi - \xi_3 > 1$. Proceeding as in (2.9) with (2.8), we have

$$\begin{aligned} &\text{LHS of (4.3)} \\ &\leqslant \left\|\int_{\xi=\xi_1-\xi_2+\xi_3} \mathbb{1}_{|\Phi(\bar{\xi})-\alpha|\leqslant M} v_1(\xi_1)v_2(\xi_2)v_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_3\right\|_{L^\infty_\xi} \|v_4\|_{L^1_\xi} \\ &\leqslant \sup_\xi\left(\int_{\xi=\xi_1-\xi_2+\xi_3} \mathbb{1}_{|\Phi(\bar{\xi})-\alpha|\leqslant M} \langle\xi_3\rangle^{-\varepsilon}\mathrm{d}\xi_1\mathrm{d}\xi_3\right)^{\frac{1}{2}} \\ &\qquad\qquad \times \|v_1\|_{L^2_\xi}\|v_2\|_{L^\infty_\xi}\|\langle\xi\rangle^\varepsilon v_3\|_{L^2_\xi}\|v_4\|_{L^1_\xi} \end{aligned}$$

[20] Recall our convention of denoting $\widehat{v}(\xi)$ by $v(\xi)$ when there is no confusion. See Remark 2.1.

$$\leqslant \sup_{\xi}\bigg(\int_{\xi-\xi_3>1}\frac{M}{(\xi-\xi_3)\langle\xi_3\rangle^{\varepsilon}}\mathrm{d}\xi_3\bigg)^{\frac{1}{2}}\|v_1\|_{L^2_\xi}\|v_2\|_{L^\infty_\xi}\|\langle\xi\rangle^{\varepsilon}v_3\|_{L^2_\xi}\|v_4\|_{L^1_\xi}$$
$$\lesssim M^{\frac{1}{2}}\|v_1\|_{L^2_\xi}\|v_2\|_{L^\infty_\xi}\|\langle\xi\rangle^{\varepsilon}v_3\|_{L^2_\xi}\|v_4\|_{L^1_\xi}. \tag{4.4}$$

An analogous computation yields

$$\text{LHS of (4.3)} \lesssim M^{\frac{1}{2}}\|v_1\|_{L^\infty_\xi}\|v_2\|_{L^2_\xi}\|\langle\xi\rangle^{\varepsilon}v_3\|_{L^2_\xi}\|v_4\|_{L^1_\xi}. \tag{4.5}$$

Lastly, with $\langle\xi-\xi_3\rangle^{\varepsilon}=\langle\xi_1-\xi_2\rangle^{\varepsilon}\lesssim\langle\xi_1\rangle^{\varepsilon}\langle\xi_2\rangle^{\varepsilon}$, we have

$$\begin{aligned}
&\text{LHS of (4.3)}\\
&\leqslant \bigg\|\int_{\xi=\xi_1-\xi_2+\xi_3}\mathbb{1}_{|\Phi(\bar{\xi})-\alpha|\leqslant M}v_1(\xi_1)v_2(\xi_2)v_3(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_3\bigg\|_{L^\infty_\xi}\|v_4\|_{L^1_\xi}\\
&\leqslant \sup_{\xi}\bigg(\int_{\xi=\xi_1-\xi_2+\xi_3}\mathbb{1}_{|\Phi(\bar{\xi})-\alpha|\leqslant M}\langle\xi-\xi_3\rangle^{-\varepsilon}\mathrm{d}\xi_1\mathrm{d}\xi_3\bigg)^{\frac{1}{2}}\bigg(\prod_{j=1}^{2}\|\langle\xi\rangle^{\varepsilon}v_j\|_{L^2_\xi}\bigg)\\
&\qquad\qquad\times\|v_3\|_{L^\infty_\xi}\|v_4\|_{L^1_\xi}\\
&\leqslant \sup_{\xi}\bigg(\int_{\xi-\xi_3>1}\frac{M}{(\xi-\xi_3)\langle\xi-\xi_3\rangle^{\varepsilon}}\mathrm{d}\xi_3\bigg)^{\frac{1}{2}}\bigg(\prod_{j=1}^{2}\|\langle\xi\rangle^{\varepsilon}v_j\|_{L^2_\xi}\bigg)\|v_3\|_{L^\infty_\xi}\|v_4\|_{L^1_\xi}\\
&\lesssim M^{\frac{1}{2}}\bigg(\prod_{j=1}^{2}\|\langle\xi\rangle^{\varepsilon}v_j\|_{L^2_\xi}\bigg)\|v_3\|_{L^\infty_\xi}\|v_4\|_{L^1_\xi}.
\end{aligned}\tag{4.6}$$

Putting (4.4), (4.5), and (4.6), we obtain (4.3) in this case. □

As a corollary to Lemma 4.1, we have the following estimates on $\mathcal{I}^{\alpha}_{M}$ and $\mathcal{I}^{\alpha}_{>M}$.

Lemma 4.2. — *Given $\varepsilon>0$, there exists $C_\varepsilon>0$ such that*

$$\|\mathcal{I}^{\alpha}_{M}(v)\|_{\mathcal{F}L^\infty}\leqslant C_\varepsilon M^{-\frac{1}{2}}\|v\|^2_{H^\varepsilon}\|v\|_{\mathcal{F}L^\infty},$$
$$\|\mathcal{I}^{\alpha}_{>M}(v)\|_{\mathcal{F}L^\infty}\leqslant C_\varepsilon M^{-\frac{1}{2}}\|v\|^2_{H^\varepsilon}\|v\|_{\mathcal{F}L^\infty},$$

for any $M\geqslant 1$ and $\alpha\in\mathbb{R}$.

Let $\mathcal{N}^{(J+1)}$ be as in (3.25). Then, from (3.24), we have

$$\begin{aligned}
&\mathcal{N}^{(J+1)}(v)(\xi)\\
&\qquad=\sum_{\mathcal{T}_{J+1}\in\mathfrak{T}(J+1)}\int_{\substack{C_0\cap\bigcap_{j=1}^{J-1}C_j^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_{J+1}),\,\boldsymbol{\xi}_r=\xi}}\frac{e^{-i\tilde{\mu}_{J+1}t}}{\prod_{j=1}^{J}\widetilde{\mu}_j}\prod_{a\in\mathcal{T}^\infty_{J+1}}v_{\xi_a}\\
&\qquad=\sum_{\mathcal{T}_J\in\mathfrak{T}(J)}\sum_{b\in\mathcal{T}^\infty_J}\int_{\substack{C_0\cap\bigcap_{j=1}^{J-1}C_j^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_J),\,\boldsymbol{\xi}_r=\xi}}\frac{e^{-i\tilde{\mu}_J t}}{\prod_{j=1}^{J}\widetilde{\mu}_j}\,\partial_t v_{\xi_b}\prod_{a\in\mathcal{T}^\infty_J\setminus\{b\}}v_{\xi_a}.
\end{aligned}\tag{4.7}$$

Now, given $\mathcal{T}_J \in \mathfrak{T}(J)$, we label its terminal nodes by $a_1, \dots, a_{2J+1}$. Then, it follows from Definition 3.13 and (4.7) with (3.22) and Remark 3.14 that

$$\mathcal{N}^{(J+1)}(v) = \sum_{\mathcal{T}_J \in \mathfrak{T}(J)} \sum_{k=1}^{2J+1} \mathfrak{S}_0\Big(\mathcal{T}_J; \mathbf{v}_k\Big), \tag{4.8}$$

where $\mathbf{v}_k = (v, \dots, v, \underbrace{\partial_t v}_{k\text{th spot}}, v, \dots, v)$. Compare this with $\mathcal{N}_0^{(J+1)}$ in (3.24) and (3.26).

Lemma 4.3. — *Let $\mathcal{N}^{(J+1)}(v)$ be as in* (3.25)*. Then, we have*

$$\|\mathcal{N}^{(J+1)}(v)\|_{\mathcal{F}L^\infty} \lesssim N^{-\frac{J}{2}+\frac{J-1}{2}\delta}\|v\|_{H^{1/6}}^{2J+3}, \tag{4.9}$$

$$\|\mathcal{N}_2^{(J+1)}(v)\|_{\mathcal{F}L^\infty} \lesssim N^{-\frac{J}{2}+\frac{J-1}{2}\delta}\|v\|_{H^{1/6}}^{2J+3}, \tag{4.10}$$

for $0 < \delta < 1$.

Proof. — We use the representation (4.8) for $\mathcal{N}^{(J+1)}(v)$. Proceeding as in the proof of Lemma 3.15[21] with Lemma 4.2, we have

$$\|\mathcal{N}^{(J+1)}(v)\|_{\mathcal{F}L^\infty} \lesssim N^{-\frac{J}{2}+\frac{J-1}{2}\delta}\|v\|_{H^\varepsilon}^{2J}\|\partial_t v\|_{\mathcal{F}L^\infty}. \tag{4.11}$$

Then, (4.9) follows from (4.11) and (4.1). The second estimate (4.10) differs from (4.9) only in the modulation restriction C_J^c for $\partial_t v$ (viewed as a cubic term). Noting that the product estimate (4.1) holds even with the modulation restriction, we see that the second estimate (4.10) follows in an analogous manner. □

Remark 4.4. — Note that we do not make use of the oscillatory factor in establishing the estimates (4.1), (4.9), and (4.10). In particular, the integrals in spatial frequencies in (4.1), (4.9), and (4.10) converge absolutely.

Let us now consider the first step of the normal form reductions. By rearranging (3.6), we have

$$\begin{aligned}
\partial_t\Bigg[\int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg] &= \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \partial_t\Bigg(\mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg)\\
&= \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0} e^{-i\mu_1 t}\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a} + \sum_{\mathcal{T}_1\in\mathfrak{T}(1)}\int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0}\frac{e^{-i\mu_1 t}}{\mu_1}\partial_t\Bigg(\prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg)\\
&= \mathcal{N}_2^{(1)}(v)(\xi) - \mathcal{N}^{(2)}(v)(\xi).
\end{aligned}$$

[21] Note that we have an $O(J)$ loss due to the summation in k. This, however, does not cause any trouble thanks to the fast decay in (3.32).

Then, in view of Lemma 4.3 with Remark 4.4, we can justify the switching of the time derivative and the integration in the first equality above by applying Fubini's theorem to the integrated (in time) formulation.

Similarly, by rearranging (3.24), we have

$$\begin{aligned}\partial_t\Bigg[\sum_{\mathcal{T}_J\in\mathfrak{T}(J)}\int_{\substack{C_0\cap\bigcap_{j=1}^{J-1}C_j^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_J),\,\boldsymbol{\xi}_r=\xi}}\frac{e^{-i\tilde{\mu}_J t}}{\prod_{j=1}^J\widetilde{\mu}_j}\prod_{a\in\mathcal{T}_J^\infty}v_{\xi_a}\Bigg]\\
&=\sum_{\mathcal{T}_J\in\mathfrak{T}(J)}\int_{\substack{C_0\cap\bigcap_{j=1}^{J-1}C_j^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_J),\,\boldsymbol{\xi}_r=\xi}}\partial_t\Bigg(\frac{e^{-i\tilde{\mu}_J t}}{\prod_{j=1}^J\widetilde{\mu}_j}\prod_{a\in\mathcal{T}_J^\infty}v_{\xi_a}\Bigg)\\
&=\mathcal{N}_2^{(J)}(\xi)-\mathcal{N}^{(J+1)}(\xi)\end{aligned}$$

in the general case. Once again, in view of Lemma 4.3 with Remark 4.4, we can justify the switching of the time derivative and the integration in the first equality above by applying Fubini's theorem to the integrated (in time) formulation

Lastly, it follows from (4.10) in Lemma 4.3 that, for each fixed $\xi\in\mathbb{R}$, the remainder term $\mathcal{N}_2^{(J+1)}(v)(\xi)$ tends to 0 as $J\to\infty$, provided that $u\in C(\mathbb{R};H^{\frac{1}{6}}(\mathbb{R}))$. This justifies the derivation of the normal form equation[(22)] (3.38) and hence the difference estimates (3.4) and (3.41) for the cubic NLS. By iterating the local-in-time argument, this yields unconditional uniqueness in the class $C(\mathbb{R};H^{\frac{1}{6}}(\mathbb{R}))$. This completes the proof of Theorem 1.1.

Remark 4.5. — As in [15], it is also possible to justify the exchange of time derivatives and integrals in spatial frequencies in the distributional sense under a milder regularity assumption that $v\in C(\mathbb{R};L^2(\mathbb{R}))$. Given a family $\{f_\xi\}_{\xi\in\mathbb{R}}$ of temporal distributions in $\mathcal{D}_t'$. we *define* $\int f_\xi \mathrm{d}\xi\in\mathcal{D}_t'$ by

$$\Big\langle\int f_\xi\mathrm{d}\xi,\phi\Big\rangle:=\int\langle f_\xi,\phi\rangle\mathrm{d}\xi \tag{4.12}$$

for $\phi\in\mathcal{D}_t$, provided that the integral on the right-hand side is well defined (in the Lebesgue sense) for each $\phi\in\mathcal{D}_t$. Then, as a distributional derivative,

(22) At this point, the normal form equation (3.38) is justified only for each fixed $\xi\in\mathbb{R}$ on the Fourier side. In view of Lemmas 3.15, and 3.16, we can a posteori show that the normal form equation (3.38) indeed holds in $C([0,T];H^{\frac{1}{6}}(\mathbb{R}))$. See the proof of Proposition 2.1 in [37]. A similar comment applies to the mKdV.

$\partial_t \int f_\xi \mathrm{d}\xi \in \mathcal{D}_t'$ is given by

$$\left\langle \partial_t \int f_\xi \mathrm{d}\xi, \phi \right\rangle = -\left\langle \int f_\xi \mathrm{d}\xi, \partial_t \phi \right\rangle \overset{(4.12)}{=} -\int \langle f_\xi, \partial_t \phi \rangle \mathrm{d}\xi = \int \langle \partial_t f_\xi, \phi \rangle \mathrm{d}\xi \\ \overset{(4.12)}{=} \left\langle \int \partial_t f_\xi \mathrm{d}\xi, \phi \right\rangle,$$

provided that $\int f_\xi \mathrm{d}\xi$ is well defined in the sense of (4.12). Namely, we have

$$\partial_t \int f_\xi \mathrm{d}\xi = \int \partial_t f_\xi \mathrm{d}\xi \tag{4.13}$$

as elements in $\mathcal{D}_t'$, as long as $\int f_\xi \mathrm{d}\xi$ exists. Compare this with Lemma 5.1 in [15]. As usual, we have

$$\left\langle \int f_\xi \mathrm{d}\xi, \phi \right\rangle = \int \langle f_\xi, \phi \rangle \mathrm{d}\xi = \iint f_\xi(t)\phi(t)\mathrm{d}t\mathrm{d}\xi \tag{4.14}$$

for locally integrable functions $f_\xi(t)$.

Now, let us consider the exchange of the time differentiation and the integration in spatial frequencies in (3.6). Lemma 3.9 with $s = 0$ states that, for almost every $\xi \in \mathbb{R}$, the integral

$$\mathcal{N}_0^{(2)}(v)(\xi) = \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \frac{e^{-i\mu_1 t}}{\mu_1} \prod_{a \in \mathcal{T}_1^\infty} v_{\xi_a}(t) =: \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} X(\boldsymbol{\xi}, t)$$

converges absolutely and uniformly on compact time intervals, if $v \in C(\mathbb{R}; L^2(\mathbb{R}))$. Then, for almost every $\xi \in \mathbb{R}$, we have

$$\begin{aligned} \langle \mathcal{N}_0^{(2)}(v)(\xi), \phi \rangle &= \left\langle \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} X(\boldsymbol{\xi}), \phi \right\rangle \overset{(4.12)}{=} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \langle X(\boldsymbol{\xi}), \phi \rangle \\ &\overset{(4.14)}{=} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \int X(\boldsymbol{\xi}, t)\phi(t)\mathrm{d}t \\ &= \int \mathcal{N}_0^{(2)}(v)(\xi, t)\phi(t)\mathrm{d}t, \end{aligned}$$

where the last equality follows from Lemma 3.9 and Fubini's theorem, since the right-hand side is absolutely integrable for $v \in C(\mathbb{R}; L^2(\mathbb{R}))$ and $\phi \in \mathcal{D}_t$. This in particular show that

$$\int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} X(\boldsymbol{\xi})$$

is well defined as an integral of temporal distributions in the sense described above. Therefore, from (4.13), we conclude that, for almost every $\xi \in \mathbb{R}$,

$$\partial_t \Bigg[\int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \frac{e^{-i\mu_1 t}}{\mu_1} \prod_{a \in \mathcal{T}_1^\infty} v_{\xi_a} \Bigg] = \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \partial_t \Bigg(\frac{e^{-i\mu_1 t}}{\mu_1} \prod_{a \in \mathcal{T}_1^\infty} v_{\xi_a} \Bigg).$$

in the (temporal) distributional sense. A similar argument can be used to justify the exchange of the time differentiation and the integration in the Jth step of the normal form reductions in this mild sense, provided that $v \in C(\mathbb{R}; L^2(\mathbb{R}))$. We, however, point out that the justification of (i) and (iii) requires a higher regularity of $s \geqslant \frac{1}{6}$, which is sufficient for switching time derivatives and integrals in the usual sense.

4.2. Unconditional well-posedness for the mKdV

In this subsection, we discuss the proof of Theorem 1.2. As in Subsection 4.1, our goal is to justify (i), (ii), and (iii) in the normal form reductions for the mKdV (1.2). While the structure of the argument follows closely that of the proof of Theorem 1.1, we need to reformulate the problem in order to handle the derivative in the nonlinearity.

Given a solution u to (1.2), let v be the corresponding interaction representation defined by (1.11). It follows from (2.14) and (2.16) that $\widehat{v}(\xi)$ satisfies

$$\partial_t \widehat{v}(\xi, t) = -i\xi e^{it\xi^3} \mathcal{F}(u^3)(\xi, t)$$

for each $\xi \in \mathbb{R}$. Arguing as in Subsection 4.1, we see that $\widehat{v}(\xi)$ is a C^1-function in t for each *fixed* $\xi \in \mathbb{R}$, provided that $u \in C(\mathbb{R}; H^{\frac{1}{6}}(\mathbb{R})) \subset C(\mathbb{R}; L^3(\mathbb{R}))$. This justifies (i) the application of the product rule in the normal form reductions, provided that $s \geqslant \frac{1}{6}$.

Next, we discuss the issues (ii) and (iii). For this purpose, we first need to review the normal form reductions in Section 3. By writing out the first step (3.6) of the normal form reductions for the mKdV, we have

$$\begin{aligned}
\mathcal{N}_2^{(1)}(v)(\xi, t) &= \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \xi e^{i\mu_1 t} \prod_{a \in \mathcal{T}_1^\infty} v_{\xi_a} \\
&= \partial_t \Bigg[\int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_1) \\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \xi^{(1)} \frac{e^{i\mu_1 t}}{\mu_1} \prod_{a \in \mathcal{T}_1^\infty} v_{\xi_a} \Bigg] \\
&\qquad - \sum_{\mathcal{T}_2 \in \mathfrak{T}(2)} \int_{\substack{\boldsymbol{\xi} \in \Xi(\mathcal{T}_2) \\ \boldsymbol{\xi}_r = \xi}} \mathbb{1}_{C_0} \Bigg(\prod_{j=1}^{2} \xi^{(j)} \Bigg) \frac{e^{i(\mu_1 + \mu_2)t}}{\mu_1} \prod_{a \in \mathcal{T}_2^\infty} v_{\xi_a}
\end{aligned}$$

$$= \partial_t \mathcal{N}_0^{(2)}(v)(\xi,t) + \mathcal{N}^{(2)}(v)(\xi,t). \tag{4.15}$$

The main issue here is the derivative loss in the last generation. More precisely, an analogue of the $\mathcal{F}L^\infty$-estimate (4.1) on $\partial_t v$ does not hold for the mKdV, even if we use the $H^{\frac{1}{4}}$-norm on the right-hand side. We instead have the following lemma.

Lemma 4.6. — *The following estimate holds:*

$$\left| \int_{\xi=\xi_1+\xi_2+\xi_3} |\xi|^{\frac{1}{4}} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \right| \lesssim \prod_{j=1}^{3} \|v_j\|_{H^{1/4}} \tag{4.16}$$

for any $\xi \in \mathbb{R}$.

Proof. — Without loss of generality, assume that $|\xi_1| \gtrsim |\xi|$ and set $w_1(\xi_1) = \langle \xi_1 \rangle^{\frac{1}{4}} v_1(\xi_1)$. Then, by Hausdorff–Young's inequality followed by Sobolev's inequality, we have

$$\begin{aligned} \text{LHS of (4.16)} &\lesssim \left\| \mathcal{F}^{-1}(|w_1| * |v_2| * |v_3|) \right\|_{L^1_x} \\ &\leqslant \|\mathcal{F}^{-1}(|w_1|)\|_{L^2} \|\mathcal{F}^{-1}(|v_2|)\|_{L^4} \|\mathcal{F}^{-1}(|v_3|)\|_{L^4} \\ &\leqslant \prod_{j=1}^{3} \|v_j\|_{H^{1/4}}. \end{aligned}$$

This proves Lemma 4.6. □

Remark 4.7. — Let $\sigma_0 \geqslant 0$. Then, proceeding as in the proof of Lemma 4.6, we have

$$\sup_\xi \left| \int_{\xi=\xi_1+\xi_2+\xi_3} |\xi|^{\sigma_0} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2 \right| \lesssim \prod_{j=1}^{3} \|v_j\|_{H^\sigma}$$

for $\sigma \geqslant \max(\sigma_0, \frac{1}{4})$. Note that the regularity restriction on σ is sharp by considering the case $|\xi_1| \sim |\xi| \gg |\xi_2|, |\xi_3|$ and its permutations. In particular, when $\sigma = \frac{1}{4}$, we can absorb precisely $\frac{1}{4}$-power of $|\xi|$ in this trilinear estimate.

On the one hand, Lemma 4.6 shows that we can absorb $\frac{1}{4}$-derivative in the second generation. On the other hand, we still need to handle the remaining $\frac{3}{4}$-derivative. We can resolve this issue by reformulating the normal form reductions as follows. By the construction, we have $\xi^{(2)} = \xi_k^{(1)}$ for some $k \in \{1,2,3\}$. See (3.2) and Definitions 3.3 and 3.5. Hence, we can rewrite (4.15) as

$$\begin{aligned}\mathcal{N}_2^{(1)}(v)(\xi,t) &= |\xi|^{\frac{3}{4}} \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0} \operatorname{sgn}(\xi)|\xi|^{\frac{1}{4}} e^{i\mu_1 t} \prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a} \\
&= |\xi|^{\frac{3}{4}} \cdot \partial_t \Bigg[\int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0} \operatorname{sgn}(\xi)|\xi|^{\frac{1}{4}} \frac{e^{i\mu_1 t}}{\mu_1} \prod_{a\in\mathcal{T}_1^\infty} v_{\xi_a}\Bigg] \\
&\quad - |\xi|^{\frac{3}{4}} \sum_{\mathcal{T}_1\in\mathfrak{T}(1)} \sum_{p^{(1)}\in\mathcal{T}_1^\infty} \int_{\substack{\boldsymbol{\xi}\in\Xi(\mathcal{T}_1)\\ \boldsymbol{\xi}_r=\xi}} \mathbb{1}_{C_0} \operatorname{sgn}(\xi)|\xi|^{\frac{1}{4}} |\xi_{p^{(1)}}|^{\frac{3}{4}} \\
&\qquad\qquad \times \frac{e^{i\mu_1 t}}{\mu_1} \mathcal{M}(v)(\xi_{p^{(1)}}) \prod_{a\in\mathcal{T}_1^\infty\setminus\{p^{(1)}\}} v_{\xi_a} \\
&= \partial_t \mathcal{N}_0^{(2)}(v)(\xi,t) + \mathcal{N}^{(2)}(v)(\xi,t),\end{aligned}$$

where $\operatorname{sgn}(\xi) = \pm 1$ denotes the sign[23] of ξ and $\mathcal{M}(v) = \mathcal{M}(v,v,v)$ is defined by

$$\begin{aligned}&\mathcal{M}(v_1,v_2,v_3)(\xi,t) \\
&\qquad := -i \int_{\xi=\xi_1+\xi_2+\xi_3} \operatorname{sgn}(\xi)|\xi|^{\frac{1}{4}} e^{i\Psi(\bar{\xi})t} v(\xi_1)v(\xi_2)v(\xi_3)\mathrm{d}\xi_1\mathrm{d}\xi_2. \qquad (4.17)\end{aligned}$$

In particular, we have *shifted $\frac{3}{4}$-derivative up by one generation* so that there is only $\frac{1}{4}$-derivative in the second generation, for which Lemma 4.6 is applicable. Similarly, with (3.2) and Remark 3.6, we can express $\mathcal{N}^{(J+1)}$ appearing in the Jth step as

$$\begin{aligned}&\mathcal{N}^{(J+1)}(v)(\xi) \\
&= \sum_{\mathcal{T}_{J+1}\in\mathfrak{T}(J+1)} \int_{\substack{C_0\cap\bigcap_{j=1}^{J-1} C_j^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_{J+1}),\,\boldsymbol{\xi}_r=\xi}} \Bigg(\prod_{j=1}^{J+1}\xi^{(j)}\Bigg) \frac{e^{i\tilde{\mu}_{J+1}t}}{\prod_{j=1}^J \widetilde{\mu}_j} \prod_{a\in\mathcal{T}_{J+1}^\infty} v_{\xi_a} \\
&= |\xi|^{\frac{3}{4}} \sum_{\mathcal{T}_J\in\mathfrak{T}(J)} \sum_{p^{(J)}\in\mathcal{T}_J^\infty} \int_{\substack{C_0\cap\bigcap_{j=1}^{J-1} C_j^c\\ \boldsymbol{\xi}\in\Xi(\mathcal{T}_J),\,\boldsymbol{\xi}_r=\xi}} \operatorname{sgn}(\xi)|\xi|^{\frac{1}{4}} \\
&\quad \times \Bigg(\prod_{j\in\#P(r^{(1)},p^{(J)})\setminus\{1\}} |\xi_{p^{(j-1)}}|^{\frac{3}{4}} \cdot \operatorname{sgn}(\xi_{p^{(j-1)}})|\xi_{p^{(j-1)}}|^{\frac{1}{4}}\Bigg) \\
&\quad \times \Bigg(\prod_{j\notin\#P(r^{(1)},p^{(J)})} \xi_{r^{(j)}}\Bigg)|\xi_{p^{(J)}}|^{\frac{3}{4}} \frac{e^{i\tilde{\mu}_J t}}{\prod_{j=1}^J \widetilde{\mu}_j} \mathcal{M}(v)(\xi_{p^{(J)}}) \prod_{a\in\mathcal{T}_J^\infty\setminus\{p^{(J)}\}} v_{\xi_a}.\end{aligned}$$
(4.18)

[23] When $\xi = 0$, we have $\mathcal{N}_2^{(1)}(v)(\xi,t) = 0$ and hence we may assume $\xi \neq 0$.

Here, $P(r^{(1)}, p^{(J)})$ is the shortest path from $r^{(1)}$ to $p^{(J)}$ defined in Remark 3.6 and $\#P(r^{(1)}, p^{(J)})$ is defined by

$$\#P(r^{(1)}, p^{(J)}) = \big\{ j \in \{1, \dots, J\} : r^{(j)} \in P(r^{(1)}, p^{(J)}) \big\}. \tag{4.19}$$

Note that $1 \in \#P(r^{(1)}, p^{(J)})$.

In view of (3.24), we see that $\mathcal{N}_1^{(J+1)}$ and $\mathcal{N}_2^{(J+1)}$ are given by (4.18) after modifying the frequency restriction to $C_0 \cap \bigcap_{j=1}^{J-1} C_j^c \cap C_J$ and $C_0 \cap \bigcap_{j=1}^{J} C_j^c$, respectively. Then, we have the following ξ-dependent estimates, replacing the $\mathcal{F}L^\infty$-estimates in Lemma 4.3 for the cubic NLS.

Lemma 4.8. — *Let $s > \frac{1}{4}$. Then, we have*

$$|\mathcal{N}^{(J+1)}(v)(\xi)| \lesssim |\xi|^{\frac{3}{4}} N^{-\frac{J}{3} + \frac{J-1}{3}\delta+} \|v\|_{H^s}^{2J} \|v\|_{H^{1/4}}^3, \tag{4.20}$$

$$|\mathcal{N}_2^{(J+1)}(v)(\xi)| \lesssim |\xi|^{\frac{3}{4}} N^{-\frac{J}{3} + \frac{J-1}{3}\delta+} \|v\|_{H^s}^{2J} \|v\|_{H^{1/4}}^3, \tag{4.21}$$

for $0 < \delta < 1$ and $\xi \in \mathbb{R}$.

We present the proof of this lemma in the next subsection. On the one hand, the estimates in Lemma 4.8 depend on $\xi \in \mathbb{R}$. On the other hand, we only need to justify the normal form reductions for each *fixed* $\xi \in \mathbb{R}$. Hence, this ξ-dependence does not cause any trouble. In fact, once we have Lemma 4.8, we can proceed as in Subsection 4.1 and justify

(ii) switching time derivatives and integrals in spatial frequencies and
(iii) the remainder term $\mathcal{N}_2^{(J+1)}(v)(\xi)$ tends to 0 as $J \to \infty$ (for each fixed $\xi \in \mathbb{R}$)

in the normal form reductions.

4.3. Proof of Lemma 4.8

We conclude this paper by presenting the proof of Lemma 4.8. We first need to introduce new trilinear operators. For $j \in \{1, 2, 3\}$, $M \geqslant 1$, and $\alpha \in \mathbb{R}$, define trilinear operators $\mathcal{N}_{j,\leqslant M}^{\alpha}$ and $\mathcal{I}_{j,M}^{\alpha}$ by

$$\begin{aligned}
&\mathcal{N}_{j,\leqslant M}^{\alpha}(v_1, v_2, v_3)(\xi, t) \\
&\qquad := \int_{\substack{\xi = \xi_1 + \xi_2 + \xi_3 \\ |\Psi(\bar{\xi}) - \alpha| \leqslant M}} |\xi|^{\frac{1}{4}} |\xi_j|^{\frac{3}{4}} e^{i\Psi(\bar{\xi})t} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2, \\
&\mathcal{I}_{j,M}^{\alpha}(v_1, v_2, v_3)(\xi, t) \\
&\qquad := \int_{\substack{\xi = \xi_1 + \xi_2 + \xi_3 \\ |\Psi(\bar{\xi}) - \alpha| \sim M}} |\xi|^{\frac{1}{4}} |\xi_j|^{\frac{3}{4}} \frac{e^{i\Psi(\bar{\xi})t}}{\Psi(\bar{\xi}) - \alpha} v_1(\xi_1) v_2(\xi_2) v_3(\xi_3) \mathrm{d}\xi_1 \mathrm{d}\xi_2.
\end{aligned} \tag{4.22}$$

As in Section 2, we also define $\mathcal{N}^{\alpha}_{j,M}$, $\mathcal{N}^{\alpha}_{j,>M}$, and $\mathcal{I}^{\alpha}_{j,>M}$ in an analogous manner.

Lemma 4.9 (Localized modulation estimates for the mKdV in the weak norm). — *Let $s > \frac{1}{4}$. Then, we have*

$$\|\mathcal{N}^{\alpha}_{j,\leqslant M}(v_1, v_2, v_3)\|_{\mathcal{F}L^\infty} \lesssim \max\{|\alpha|, M\}^{\frac{1}{12}} M^{\frac{1}{2}} \|v_j\|_{\mathcal{F}L^\infty} \prod_{\substack{k=1\\k\neq j}}^{3} \|v_k\|_{H^s} \tag{4.23}$$

for any $j \in \{1, 2, 3\}$, $M \geqslant 1$, and $\alpha \in \mathbb{R}$, where the implicit constant is independent of $j \in \{1, 2, 3\}$.

We postpone the proof of Lemma 4.9 to the end of this section. As an immediate corollary to Lemma 4.9, we have the following estimates on $\mathcal{I}^{\alpha}_{j,M}$ and $\mathcal{I}^{\alpha}_{j,>M}$.

Lemma 4.10. — *Let $s > \frac{1}{4}$. Then, we have*

$$\|\mathcal{I}^{\alpha}_{j,M}(v_1, v_2, v_3)\|_{\mathcal{F}L^\infty} \lesssim \max\{|\alpha|, M\}^{\frac{1}{12}} M^{-\frac{1}{2}} \|v_j\|_{\mathcal{F}L^\infty} \prod_{\substack{k=1\\k\neq j}}^{3} \|v_k\|_{H^s},$$

$$\|\mathcal{I}^{\alpha}_{j,>M}(v_1, v_2, v_3)\|_{\mathcal{F}L^\infty} \lesssim \max\{|\alpha|, M\}^{\frac{1}{12}} M^{-\frac{1}{2}} \|v_j\|_{\mathcal{F}L^\infty} \prod_{\substack{k=1\\k\neq j}}^{3} \|v_k\|_{H^s},$$

for any $j \in \{1, 2, 3\}$, $M \geqslant 1$, and $\alpha \in \mathbb{R}$, where the implicit constant is independent of $j \in \{1, 2, 3\}$.

Now, we are ready to prove Lemma 4.8 (and hence Theorem 1.2), assuming Lemma 4.9. Given $\mathcal{T} \in \mathfrak{T}(J)$, we first define $\widetilde{\mathfrak{S}}_0(\mathcal{T}, v)$ by making the following modifications in Steps (ii), (iii), and (iv) of the definition of $\mathfrak{S}_0(\mathcal{T}; v)$ in Definition 3.13:

(ii) and (iii): Let $j = 2, \dots, J$. Recall the definitions of $\#P(r^{(1)}, p^{(J)})$ and the order $\#p^{(j)}$ of $p^{(j)}$ from (4.19) and Definition 3.5.

- If $j \in \#P(r^{(1)}, p^{(J)})$, then we make the following change: $\mathcal{I}^{\tilde{\mu}_{j-1}}_{|\mu_j + \tilde{\mu}_{j-1}| > (2j+1)^3 M_{j-1}^{1-\delta}} \implies \mathcal{I}^{\tilde{\mu}_{j-1}}_{\#p^{(k)}, |\mu_j + \tilde{\mu}_{j-1}| > (2j+1)^3 M_{j-1}^{1-\delta}}$, where $\mathcal{I}^{\alpha}_{j,M}$ is as in (4.22) and $p^{(k)}$ is the unique node such that $p^{(k)} \in \pi_j(\mathcal{T})^\infty \cap P(r^{(1)}, p^{(J)})$.
- If $j \notin \#P(r^{(1)}, p^{(J)})$, we do not make any modification.

(iv): $\mathcal{I}_{|\mu_1| > N} \implies |\xi|^{\frac{3}{4}} \mathcal{I}_{\#p^{(k)}, |\mu_1| > N}$, where $p^{(k)}$ is the unique node such that $p^{(k)} \in \pi_1(\mathcal{T})^\infty \cap P(r^{(1)}, p^{(J)})$.

Given $\mathcal{T} \in \mathfrak{T}(J)$, we label its terminal nodes by $a_1, \dots, a_{2J+1}$. Then, it follows from the definition above for $\widetilde{\mathfrak{S}}_0(\mathcal{T};\,\cdot\,)$ and (4.18) with (3.22) and Remark 3.14 that

$$\mathcal{N}^{(J+1)}(v) = \sum_{\mathcal{T}\in\mathfrak{T}(J)} \sum_{k=1}^{2J+1} \widetilde{\mathfrak{S}}_0\left(\mathcal{T};\mathbf{v}_k\right), \tag{4.24}$$

where $\mathbf{v}_k = (v, \dots, v, \underbrace{\mathcal{M}(v)}_{k\text{th spot}}, v, \dots, v)$. Compare this with (4.8).

Lemma 4.6 with (4.17) yields

$$\|\mathcal{M}(v)\|_{\mathcal{F}L^\infty} \lesssim \|v\|_{H^{1/4}}^3. \tag{4.25}$$

Then, a slight modification[24] of the proof of Lemma 3.15 with Lemmas 2.8 and 4.10 and (4.25) yields the first estimate (4.20) in Lemma 4.8. The only difference between $\mathcal{N}^{(J+1)}(v)$ and $\mathcal{N}_2^{(J+1)}(v)$ is the modulation restriction C_J^c in the last generation. In particular, $\mathcal{N}_2^{(J+1)}(v)$ can be expressed as (4.24) with an extra modulation restriction on $\mathcal{M}(v)$. Since the proof of Lemma 4.6 remains true with such a modulation restriction, the second estimate (4.21) in Lemma 4.8 follows in an analogous manner. This completes the proof of Lemma 4.8 and hence the proof of Theorem 1.2, assuming Lemma 4.9.

It remains to prove Lemma 4.9. The remaining part of this paper is devoted to the proof of Lemma 4.9.

Proof of Lemma 4.9. — For convenience, let $A_j = \{1,2,3\}\setminus\{j\}$. Then, by duality, (4.23) follows once we prove

$$\begin{aligned}\left|\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar\xi)-\alpha|\leqslant M}\cdot|\xi|^{\frac14}|\xi_j|^{\frac34}\prod_{k=1}^{3} v_k(\xi_k)v_4(\xi)\mathrm{d}\xi_1\mathrm{d}\xi_2\mathrm{d}\xi\right| \\ \lesssim \max\{|\alpha|,M\}^{\frac1{12}}M^{\frac12}, \|v_j\|_{\mathcal{F}L^\infty}\left(\prod_{k\in A_j}\|v_k\|_{H^s}\right)\|v_4\|_{\mathcal{F}L^1}\end{aligned} \tag{4.26}$$

(24) In particular, in (3.32), we replace $N_1^{-\frac12+}\prod_{j=2}^J N_j^{-\frac12+}$ by

$$N_1^{-\frac12+\frac1{12}+}\prod_{j=2}^{J}\left(N_j^{-\frac12+}\max(N_{j-1}^{\frac1{12}},N_j^{\frac1{12}})\right)\leqslant\prod_{j=1}^{J}N_j^{-\frac13+}.$$

for all non-negative functions $v_1, \dots, v_4$ and $j \in \{1,2,3\}$. Given $s \geqslant \frac{1}{4}$ and $j \in \{1,2,3\}$, define $m_{s,j}(\bar{\xi})$ by

$$m_{s,j}(\bar{\xi}) = \frac{|\xi|^{\frac{1}{4}}|\xi_j|^{\frac{3}{4}}}{\prod_{k\in A_j}\langle \xi_k\rangle^s}. \tag{4.27}$$

When $s = \frac{1}{4}$, we simply denote $m_{\frac{1}{4},j}$ by m_j. By a variant of the Cauchy–Schwarz argument, we have

$$\begin{aligned}\text{LHS of (4.26)} &\leqslant \bigg\| \int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot |\xi|^{\frac{1}{4}}|\xi_j|^{\frac{3}{4}} \prod_{k=1}^{3} v_k(\xi_k)\bigg\|_{L^\infty_\xi} \|v_4\|_{L^1_\xi} \\ &\leqslant \sup_\xi \bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2_{s,j}(\bar{\xi}) \mathrm{d}\xi_1\mathrm{d}\xi_2\bigg)^{\frac{1}{2}} \\ &\qquad\qquad \times \|v_j\|_{\mathcal{F}L^\infty} \prod_{k\in A_j} \|v_k\|_{H^s}\|v_4\|_{L^1_\xi}. \end{aligned} \tag{4.28}$$

Hence, it suffices to show that

$$\sup_\xi \bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m^2_{s,j}(\bar{\xi}) \mathrm{d}\xi_1\mathrm{d}\xi_2\bigg)^{\frac{1}{2}} \lesssim \max\{|\alpha|, M\}^{\frac{1}{12}} M^{\frac{1}{2}}, \tag{4.29}$$

uniformly in $j \in \{1,2,3\}$ for $s > \frac{1}{4}$.

While we do not explicitly state so, it is understood that all the estimates and statements in the following hold uniformly in $j \in \{1,2,3\}$. Also, we will see that the estimate (4.23) in fact holds at the endpoint regularity $s = \frac{1}{4}$ for many of the following cases. For those cases, by monotonicity of $\langle \xi\rangle^s$ in s, it suffices to prove (4.23) for $s = \frac{1}{4}$.

Case 1: $|\xi| \lesssim 1$, $|\xi_j| \lesssim 1$. — In this case, we prove (4.29) with $s = \frac{1}{4}$. Without loss of generality, we assume $j = 1$.

Subcase 1.a: $|\xi_{23}|$ and $|\xi_{2-3}| \gtrsim 1$. — By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_1, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\xi_{23}||\xi_{2-3}| \gtrsim 1$. Then, with $|m_1(\bar{\xi})| \lesssim 1$, we have

$$\text{LHS of (4.29)} \lesssim \sup_\xi \bigg(\int_{|\xi_1|\lesssim 1} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \mathrm{d}\xi_2\mathrm{d}\xi_1\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.$$

Subcase 1.b: $|\xi_{23}|$ or $|\xi_{2-3}| \ll 1$. — In this case, we have $\langle\xi_2\rangle \sim \langle\xi_3\rangle$. When $|\xi_2|, |\xi_3| \lesssim 1$, it is easy to see that the left-hand side of (4.29) is $O(1)$ with $|m_1(\bar{\xi})| \lesssim 1$ and integration in ξ_2 and ξ_3. Now, suppose that $|\xi_2| \sim |\xi_3| \gg 1$. By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_3, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\xi_{12}||\xi_{1-2}| \sim |\xi_2|^2 \gg 1$ since $|\xi_1| \lesssim 1 \ll |\xi_2|$. Then, with $|m_1(\bar{\xi})| \lesssim \langle \xi_2 \rangle^{-\frac{1}{2}}$, we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi}\Bigg(\sum_{\substack{N\gg 1\\ \text{dyadic}}} \frac{1}{N}\int_{|\xi_2|\sim|\xi_3|\sim N} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\mathrm{d}\xi_2\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}$$
$$\lesssim \Bigg(\sum_{\substack{N\gg 1\\ \text{dyadic}}} \frac{M}{N^3}N\Bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.$$

Case 2: $|\xi| \gg 1$, $|\xi_j| \lesssim 1$. — In this case, we prove (4.29) with $s = \frac{1}{4}$. We denote $A_j = \{1,2,3\} \setminus \{j\} = \{k_1, k_2\}$.

Subcase 2.a: $|\xi_{k_1}| \sim |\xi_{k_2}| \gtrsim |\xi| \gg 1 \gtrsim |\xi_j|$, *where* k_1, $k_2 \in A_j$. — By viewing Ψ as a function of ξ_{k_1} for fixed ξ and ξ_{k_2}, we have $|\partial_{\xi_{k_1}}\Psi(\bar{\xi})| \sim |\xi_{jk_1}||\xi_{j-k_1}| \sim |\xi_{k_2}|^2$. Then, with $|m_j(\bar{\xi})| \lesssim |\xi_{k_2}|^{-\frac{1}{4}}$, we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi}\Bigg(\int_{|\xi_{k_1}|\sim|\xi_{k_2}|\gg 1} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot\frac{1}{|\xi_{k_2}|^{\frac{1}{2}}}\mathrm{d}\xi_{k_1}\mathrm{d}\xi_{k_2}\Bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.$$

Subcase 2.b: $|\xi_{k_1}| \sim |\xi| \gg \max(|\xi_{k_2}|, |\xi_j|, 1)$, *where* k_1, $k_2 \in A_j$. — In this case, $|\xi_{k_1k_2}| \sim |\xi_{k_1}|$ and $|\xi_{k_1-k_2}| \sim |\xi_{k_1}|$. By viewing Ψ as a function of ξ_{k_1} for fixed ξ and ξ_j, we have $|\partial_{\xi_{k_1}}\Psi(\bar{\xi})| \sim |\xi_{k_1k_2}||\xi_{k_1-k_2}| \sim |\xi|^2$. Then, with $|m_j(\bar{\xi})| \lesssim 1$, we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi}\Bigg(\int_{|\xi_j|\lesssim 1} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\mathrm{d}\xi_{k_1}\mathrm{d}\xi_j\Bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.$$

Case 3: $|\xi| \lesssim 1$, $|\xi_j| \gg 1$. — We prove (4.29) with $s = \frac{1}{4}$ in Subcases 3.a and 3.b, while Subcase 3.c requires $s > \frac{1}{4}$. In Subcases 3.a and 3.b, we only need the condition $|\xi| \ll |\xi_j|$ and their relative sizes with respect to 1 is not important.

Subcase 3.a: $|\xi_{k_1}| \sim |\xi_{k_2}| \gg |\xi_j| \gg |\xi|$, *where* k_1, $k_2 \in A_j$. — Let $s = \frac{1}{4}$. We have $|\xi_{jk_2}| \sim |\xi_{j-k_2}| \sim |\xi_{k_1}|$. By viewing Ψ as a function of ξ_{k_2} for fixed ξ and ξ_{k_1}, we have $|\partial_{\xi_{k_2}}\Psi(\bar{\xi})| \sim |\xi_{jk_2}||\xi_{j-k_2}| \sim |\xi_{k_1}|^2$. Then, we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi}\Bigg(\int_{|\xi_j|\lesssim 1} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot\frac{|\xi|^{\frac{1}{2}}|\xi_j|^{\frac{3}{2}}}{|\xi_{k_1}|}\mathrm{d}\xi_{k_2}\mathrm{d}\xi_{k_1}\Bigg)^{\frac{1}{2}}$$
$$\lesssim M^{\frac{1}{2}}\sup_{\xi}\Bigg(\int_{|\xi_{k_1}|\gg|\xi|} \frac{|\xi|^{\frac{1}{2}}}{|\xi_{k_1}|^{\frac{3}{2}}}\mathrm{d}\xi_{k_1}\Bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.$$

Subcase 3.b: $|\xi_1| \sim |\xi_2| \sim |\xi_3| \gg |\xi|$. — Let $s = \frac{1}{4}$. We have $m_j(\bar{\xi}) \lesssim |\xi|^{\frac{1}{4}}|\xi_1|^{\frac{1}{4}}$ and $|\xi_{12}| \sim |\xi_{23}| \sim |\xi_{31}| \sim |\xi_1|$. We claim that $\max\{|\xi_{1-2}|, |\xi_{2-3}|,$

$|\xi_{3-1}|\} \gtrsim |\xi_1|$. Otherwise, i.e. if $\max\{|\xi_{1-2}|, |\xi_{2-3}|, |\xi_{3-1}|\} \ll |\xi_1|$, then ξ_1, ξ_2, and ξ_3 must have the same sign and thus $|\xi| = |\xi_1+\xi_2+\xi_3| \sim |\xi_1|$, leading to a contradiction. Without loss of generality, we assume $|\xi_{2-3}| \sim |\xi_1|$. By viewing Ψ as a function of ξ_3 for fixed ξ and ξ_1, we have $|\partial_{\xi_3}\Psi(\bar{\xi})| \sim |\xi_{23}||\xi_{2-3}| \sim |\xi_1|^2$. Hence, we have

$$\begin{aligned}\text{LHS of (4.29)} &\lesssim \sup_{\xi}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot|\xi|^{\frac12}|\xi_1|^{\frac12}\mathrm{d}\xi_3\mathrm{d}\xi_1\bigg)^{\frac12}\\ &\lesssim M^{\frac12}\sup_{\xi}\bigg(\int_{|\xi_1|\gg|\xi|}\frac{|\xi|^{\frac12}}{|\xi_1|^{\frac32}}\mathrm{d}\xi_1\bigg)^{\frac12} \lesssim M^{\frac12}.\end{aligned}$$

Subcase 3.c: $|\xi_{k_1}| \sim |\xi_j| \gg |\xi_{k_2}|$, $|\xi|$, where k_1, $k_2 \in A_j$. — In this case, we have $|\xi_{k_1k_2}| \sim |\xi_{k_1-k_2}| \sim |\xi_j|$. By viewing Ψ as a function of ξ_{k_1} for fixed ξ and ξ_j, we have $|\partial_{\xi_{k_1}}\Psi(\bar{\xi})| \sim |\xi_{k_1k_2}||\xi_{k_1-k_2}| \sim |\xi_j|^2$. Hence, with $|m_{j,s}(\bar{\xi})| \lesssim |\xi|^{\frac14}|\xi_j|^{\frac34-s}$, we have

$$\begin{aligned}\text{LHS of (4.29)} &\lesssim \sup_{|\xi|\lesssim1}\bigg(\int_{|\xi_j|\gg1}\mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}|\xi|^{\frac12}|\xi_j|^{\frac32-2s}\mathrm{d}\xi_{k_1}\mathrm{d}\xi_j\bigg)^{\frac12}\\ &\lesssim M^{\frac12}\sup_{|\xi|\lesssim1}\bigg(\int_{|\xi_j|\gg1}\frac{1}{|\xi_j|^{\frac12+2s}}\mathrm{d}\xi_j\bigg)^{\frac12}\lesssim M^{\frac12},\end{aligned}$$

provided that $s > \frac14$.

In the remaining part of the proof, we split the case $|\xi|$, $|\xi_j| \gg 1$ into three subcases, depending on the sizes of $|\xi_{12}|$, $|\xi_{23}|$, and $|\xi_{31}|$. Without loss of generality, we assume $|\xi_{12}| \geqslant |\xi_{23}| \geqslant |\xi_{31}|$. Then, by the triangle inequality, we have

$$|\xi_{12}| \gtrsim |\xi|. \tag{4.30}$$

Case 4: $|\xi|, |\xi_j| \gg 1$ and $|\xi_{31}| \leqslant |\xi_{23}| \leqslant 1$. — Arguing as in Case 2 of the proof of Lemma 2.6, we have

$$|\xi_1| \sim |\xi_2| \sim |\xi_3| \sim |\xi| \gg 1 \tag{4.31}$$

in this case. In particular, we have

$$m_j(\bar{\xi}) \sim |\xi|^{\frac12} \sim m(\bar{\xi}) \tag{4.32}$$

for any $j \in \{1,2,3\}$, where $m(\bar{\xi})$ is as in (2.20). Let $\zeta_1 = \xi_{23}$, $\zeta_2 = \xi_{31}$, and $\zeta_3 = \xi_{12}$ as before.

We prove (4.26) with $s = \frac14$ in Subcases 4.a and 4.b, while Subcase 4.c requires $s > \frac14$.

Subcase 4.a: $|\xi| \lesssim M$. — Let $s = \frac{1}{4}$. By Hölder's inequality with $|\zeta_2| \leqslant |\zeta_1| \leqslant 1$, we have

$$\begin{aligned}\text{LHS of (4.26)} &\lesssim \|v_4\|_{L^1_\xi} \sum_{\substack{1 \ll N \lesssim M \\ \text{dyadic}}} N^{\frac{1}{2}} \sup_{|\xi| \sim N} \int_{|\zeta_2| \leqslant |\zeta_1| \leqslant 1} \prod_{k \in A_j} \langle \xi_k \rangle^{\frac{1}{4}} \\ &\qquad\qquad \times v_1(\xi - \zeta_1) v_2(\xi - \zeta_2) v_3(\zeta_1 + \zeta_2 - \xi) \mathrm{d}\zeta_1 \mathrm{d}\zeta_2 \\ &\lesssim M^{\frac{1}{2}} \|v_j\|_{\mathcal{F}L^\infty} \Bigg(\prod_{k \in A_j} \|v_k\|_{H^{1/4}} \Bigg) \|v_4\|_{\mathcal{F}L^1}.\end{aligned}$$

In the following, we assume that $|\xi| \gg M$.

Subcase 4.b: $|\zeta_2| \leqslant |\zeta_1| \lesssim \frac{M^{1/2}}{|\xi|^{1/2}}$. — Let $s = \frac{1}{4}$. By Hölder's inequality with (4.31) and (4.32), we have

$$\begin{aligned}\text{LHS of (4.26)} &\lesssim \|v_4\|_{L^1_\xi} \sum_{\substack{N \gg M \\ \text{dyadic}}} N^{\frac{1}{2}} \sup_{|\xi| \sim N} \int_{|\zeta_2| \leqslant |\zeta_1| \lesssim \frac{M^{1/2}}{N^{1/2}}} \prod_{k \in A_j} \langle \xi_k \rangle^{\frac{1}{4}} \\ &\qquad\qquad \times v_1(\xi - \zeta_1) v_2(\xi - \zeta_2) v_3(\zeta_1 + \zeta_2 - \xi) \mathrm{d}\zeta_1 \mathrm{d}\zeta_2 \\ &\lesssim M^{\frac{1}{2}} \|v_j\|_{L^\infty_\xi} \|v_4\|_{L^1_\xi} \sum_{\substack{N \gg M \\ \text{dyadic}}} \prod_{k \in A_j} \|\mathbf{P}_N v_k\|_{H^{1/4}} \\ &\lesssim M^{\frac{1}{2}} \|v_j\|_{\mathcal{F}L^\infty} \Bigg(\prod_{k \in A_j} \|v_k\|_{H^{1/4}} \Bigg) \|v_4\|_{\mathcal{F}L^1},\end{aligned}$$

where we used Cauchy–Schwarz inequality (in N) in the last step.

Subcase 4.c: $\frac{M^{1/2}}{|\xi|^{1/2}} \ll |\zeta_1|$. — In this case, we use $m_{s,j}(\bar{\xi}) \sim |\xi|^{1-2s}$. By viewing Ψ as a function of ξ_3 for fixed ξ and ζ_1, we have $|\partial_{\xi_3} \Psi(\bar{\xi})| \sim |\xi_{23}||\xi_{2-3}| \sim |\zeta_1||\xi|$, since $|\xi_{2-3}| = |2\xi_3 - \zeta_1| \sim |\xi|$. Hence, we have

$$\begin{aligned}\text{LHS of (4.29)} &\lesssim \sup_\xi \Bigg(\int_{\frac{M^{1/2}}{|\xi|^{1/2}} \ll |\zeta_1| \leqslant 1} \mathbb{1}_{|\Psi(\bar{\xi}) - \alpha| \leqslant M} \cdot |\xi|^{2-4s} \mathrm{d}\xi_3 \mathrm{d}\zeta_1 \Bigg)^{\frac{1}{2}} \\ &\lesssim M^{\frac{1}{2}} \sup_\xi |\xi|^{\frac{1}{2} - 2s} \Bigg(\int_{\frac{M^{1/2}}{|\xi|^{1/2}} \ll |\zeta_1| \leqslant 1} \frac{1}{|\zeta_1|} \mathrm{d}\zeta_1 \Bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}},\end{aligned}$$

provided that $s > \frac{1}{4}$.

Remark 4.11. — When $|\zeta_1| \geqslant |\zeta_2| \gg \frac{M^{1/2}}{|\xi|^{1/2}}$, it follows from (2.15) and (4.30) that

$$|\zeta_1| \lesssim \frac{|\alpha| + M}{|\xi||\zeta_2|} \ll \frac{|\alpha| + M}{M} \cdot \frac{M^{\frac{1}{2}}}{|\xi|^{\frac{1}{2}}} \tag{4.33}$$

under the condition $|\Psi(\bar{\xi}) - \alpha| \leqslant M$. In particular, (4.33) with $|\zeta_1| \gg \frac{M^{1/2}}{|\xi|^{1/2}}$ implies that $|\alpha| \gg M$. Then, in view of (4.32), the desired estimate (4.29) with $s = \frac{1}{4}$ follows from Subcase 2.b in the proof of Lemma 2.6.

Case 5: $|\xi|, |\xi_j| \gg 1$ *and* $|\xi_{31}| \leqslant 1 < |\xi_{23}| \leqslant |\xi_{12}|$. — In this case, we have

$$\langle \xi_1 \rangle \sim \langle \xi_3 \rangle \qquad \text{and} \qquad \langle \xi_2 \rangle \sim \langle \xi \rangle. \tag{4.34}$$

First, we consider the case $j = 1$. We have

$$m_{s,1}(\bar{\xi}) \sim \frac{|\xi|^{\frac{1}{4}} |\xi_1|^{\frac{3}{4}}}{\langle \xi_2 \rangle^s \langle \xi_3 \rangle^s} \lesssim |\xi|^{\frac{1}{4} - s} |\xi_1|^{\frac{3}{4} - s}. \tag{4.35}$$

In the following, we take $s > \frac{1}{4}$.

Subcase 5.a: $|\xi| \gtrsim |\xi_1| \gg 1$. — Arguing as in Subcase 3.a in the proof of Lemma 2.6, we see that ζ_2 belongs to an interval $I = I(\zeta_1, \xi)$ of length

$$|I(\zeta_1, \xi)| \lesssim \frac{M}{|\zeta_1||\xi|} \tag{4.36}$$

for each fixed ξ and ζ_1 and hence for each fixed ξ and $\xi_1 = \xi - \zeta_1$. Then, using a variant of (4.28) with (4.35), (4.36), and $|\zeta_1| = |\xi - \xi_1| \lesssim |\xi|$, we have

$$\begin{aligned}
&\text{LHS of (4.26)} \\
&\leqslant \bigg\| \int_{1 \leqslant |\zeta_1| \lesssim |\xi|} \int_{|\zeta_2| \leqslant 1} \mathbb{1}_{|\Psi(\bar{\xi}) - \alpha| \leqslant M} \cdot m_{s,1}(\bar{\xi}) \prod_{k \in A_1} \langle \xi_k \rangle^s \\
&\qquad\qquad \times v_1(\xi - \zeta_1) v_2(\xi - \zeta_2) v_3(-\xi + \zeta_1 + \zeta_2) \mathrm{d}\zeta_2 \mathrm{d}\zeta_1 \bigg\|_{L^\infty_\xi} \|v_4\|_{L^1_\xi} \\
&\lesssim \sup_\xi \bigg(\int_{1 \leqslant |\zeta_1| \lesssim |\xi|} \int_{\zeta_2 \in I(\zeta_1, \xi)} \mathbb{1}_{|\Psi(\bar{\xi}) - \alpha| \leqslant M} \cdot |\xi|^{\frac{1}{2} - 2s} |\xi_1|^{\frac{3}{2} - 2s} \mathrm{d}\zeta_2 \mathrm{d}\zeta_1 \bigg)^{\frac{1}{2}} \\
&\qquad\qquad \times \|v_1\|_{L^\infty_\xi} \bigg(\prod_{k \in A_1} \|v_k\|_{H^s} \bigg) \|v_4\|_{L^1_\xi} \\
&\lesssim M^{\frac{1}{2}} \sup_\xi \bigg(\int_{1 \leqslant |\zeta_1| \lesssim |\xi|} \frac{|\xi|^{1-4s}}{|\zeta_1|} \mathrm{d}\zeta_1 \bigg)^{\frac{1}{2}} \|v_1\|_{L^\infty_\xi} \bigg(\prod_{k \in A_1} \|v_k\|_{H^s} \bigg) \|v_4\|_{L^1_\xi} \\
&\lesssim M^{\frac{1}{2}} \|v_1\|_{\mathcal{F}L^\infty} \bigg(\prod_{k \in A_1} \|v_k\|_{H^s} \bigg) \|v_4\|_{\mathcal{F}L^1},
\end{aligned} \tag{4.37}$$

provided that $\frac{1}{4} < s \leqslant \frac{3}{4}$. For $s > \frac{3}{4}$, we simply use $|\xi_1|^{\frac{3}{2}-2s} \lesssim 1$ and repeat the computation in (4.37).

Subcase 5.b: $|\xi_1| \gg |\xi|$. — By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_1, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\xi_{23}||\xi_{2-3}| \sim |\xi_1|^2$ thanks to (4.34). Hence, with (4.35), we have

$$\begin{aligned}
\text{LHS of (4.29)} &\lesssim \sup_\xi \bigg(\int_{|\xi_1|\gg|\xi|} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot |\xi_1|^{\frac{3}{2}-2s} \mathrm{d}\xi_2 \mathrm{d}\xi_1\bigg)^{\frac{1}{2}} \\
&\lesssim M^{\frac{1}{2}} \sup_\xi \bigg(\int_{|\xi_1|\gg 1} |\xi_1|^{-\frac{1}{2}-2s} \mathrm{d}\xi_1\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}},
\end{aligned}$$

provided that $s > \frac{1}{4}$.

Next, we consider the case $j = 2$. It follows from (4.27) and (4.34) that

$$m_{s,2}(\bar{\xi}) \sim \frac{|\xi|^{\frac{1}{4}}|\xi_2|^{\frac{3}{4}}}{\langle\xi_1\rangle^s\langle\xi_3\rangle^s} \sim \frac{|\xi|}{\langle\xi_1\rangle^{2s}}$$

Subcase 5.c: $|\xi_2| \lesssim |\xi_1|$. — Proceeding as in Subcase 5.a with $|\zeta_1| = |\xi - \xi_1| \lesssim |\xi_1|$, we obtain

$$\begin{aligned}
&\text{LHS of (4.26)} \\
&\leqslant \bigg\| \int_{1\leqslant|\zeta_1|\lesssim|\xi_1|} \int_{|\zeta_2|\leqslant 1} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot m_{s,2}(\bar{\xi}) \prod_{k\in A_2} \langle\xi_k\rangle^s \\
&\qquad\qquad \times v_1(\xi-\zeta_1)v_2(\xi-\zeta_2)v_3(-\xi+\zeta_1+\zeta_2)\mathrm{d}\zeta_2\mathrm{d}\zeta_1 \bigg\|_{L^\infty_\xi} \|v_4\|_{L^1_\xi} \\
&\lesssim \sup_\xi \bigg(\int_{1\leqslant|\zeta_1|\lesssim|\xi_1|} \int_{\zeta_2\in I(\zeta_1,\xi)} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot |\xi|^2|\xi_1|^{-4s} \mathrm{d}\zeta_2\mathrm{d}\zeta_1\bigg)^{\frac{1}{2}} \\
&\qquad\qquad \times \|v_2\|_{L^\infty_\xi} \bigg(\prod_{k\in A_2} \|v_k\|_{H^s}\bigg) \|v_4\|_{L^1_\xi} \\
&\lesssim M^{\frac{1}{2}} \sup_\xi \bigg(\int \frac{1}{\langle\zeta_1\rangle\langle\xi-\zeta_1\rangle^{4s-1}} \mathrm{d}\zeta_1\bigg)^{\frac{1}{2}} \|v_2\|_{L^\infty_\xi} \bigg(\prod_{k\in A_2} \|v_k\|_{H^s}\bigg) \|v_4\|_{L^1_\xi} \\
&\lesssim M^{\frac{1}{2}} \|v_2\|_{\mathcal{F}L^\infty} \bigg(\prod_{k\in A_2} \|v_k\|_{H^s}\bigg) \|v_4\|_{\mathcal{F}L^1},
\end{aligned}$$

provided that $s > \frac{1}{4}$.

Subcase 5.d: $|\xi_2| \gg |\xi_1|$. — By viewing Ψ as a function of ξ_1 for fixed ξ and ξ_3, we have $|\partial_{\xi_1}\Psi(\bar{\xi})| \sim |\xi_{12}||\xi_{1-2}| \sim |\xi|^2$. Hence, we have

$$\begin{aligned}\text{LHS of (4.29)} &\lesssim \sup_{\xi}\bigg(\int_{|\xi_3|\ll|\xi|} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\frac{|\xi|^2}{\langle\xi_3\rangle^{4s}}\mathrm{d}\xi_1\mathrm{d}\xi_3\bigg)^{\frac{1}{2}}\\ &\lesssim M^{\frac{1}{2}}\sup_{\xi}\bigg(\int\frac{1}{\langle\xi_3\rangle^{4s}}\mathrm{d}\xi_3\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}},\end{aligned}$$

provided that $s > \frac{1}{4}$.

Lastly, we consider the case $j = 3$.

Subcase 5.e: $|\xi_3| \lesssim |\xi|$. — If $|\xi_{23}| \gtrsim |\xi|$, then this case follows from Subcase 5.a by switching $1 \leftrightarrow 3$. Now, suppose that $|\xi_{23}| \ll |\xi|$. Then, it follows from (4.34) that $\langle\xi_1\rangle \sim \langle\xi_2\rangle \sim \langle\xi_3\rangle \sim \langle\xi\rangle$. In particular, we have $m_{s,3}(\bar{\xi}) \sim |\xi|^{1-2s}$. Hence, we can apply Subcase 5.a by replacing $m_1(\bar{\xi})$ with $m_3(\bar{\xi})$ (without switching $1 \leftrightarrow 3$).

Subcase 5.f: $|\xi_3| \gg |\xi|$. — This case follows from Subcase 5.b by switching $1 \leftrightarrow 3$.

Case 6: $|\xi|, |\xi_j| \gg 1$ *and* $|\xi_{12}|, |\xi_{23}|, |\xi_{31}| > 1$. — From (2.33) with $\xi = \xi_1 + \xi_2 + \xi_3$, we have

$$|\alpha| + M \gtrsim \max(|\xi|, |\xi_1|, |\xi_2|, |\xi_3|), \tag{4.38}$$

which allows us to prove (4.29) with $s = \frac{1}{4}$ in this case. In the following, the size relation of $|\xi_{12}|, |\xi_{23}|, |\xi_{31}|$ does not play any role. Hence, without loss of generality, assume that $|\xi_1| \geqslant |\xi_2| \geqslant |\xi_3|$. Recall that $A_j = \{1, 2, 3\} \setminus \{j\} = \{k_1, k_2\}$.

Subcase 6.a: $|\xi_1| \sim |\xi| \gg |\xi_2| \geqslant |\xi_3|$. — Let $s = \frac{1}{4}$. By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_3, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\xi_{12}||\xi_{1-2}| \sim |\xi|^2$. Hence, with (4.38), we have

$$\begin{aligned}\text{LHS of (4.29)} &\lesssim \sup_{\xi}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot\frac{|\xi|^{\frac{1}{2}}|\xi_j|^{\frac{3}{2}}}{\langle\xi_{k_1}\rangle^{\frac{1}{2}}\langle\xi_{k_2}\rangle^{\frac{1}{2}}}\mathrm{d}\xi_2\mathrm{d}\xi_3\bigg)^{\frac{1}{2}}\\ &\lesssim \sup_{\xi}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot\frac{|\xi|^2}{\langle\xi_3\rangle}\mathrm{d}\xi_2\mathrm{d}\xi_3\bigg)^{\frac{1}{2}}\\ &\lesssim M^{\frac{1}{2}}\bigg(\int_{|\xi_3|\ll|\xi|}\frac{1}{\langle\xi_3\rangle}\mathrm{d}\xi_3\bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}(\log|\xi|)^{\frac{1}{2}} \lesssim \langle\alpha\rangle^{0+}M^{\frac{1}{2}+}.\end{aligned}$$

Subcase 6.b: $|\xi_1| \sim |\xi_2| \gg |\xi| \gg |\xi_3|$. — Let $s = \frac{1}{4}$. We have $|\xi_{12}| = |\xi - \xi_3| \sim |\xi|$ and $|\xi_{1-2}| \sim |\xi_1|$, since ξ_1 and ξ_2 have opposite signs in this case. By viewing Ψ as a function of ξ_1 for fixed ξ and ξ_3, we have $|\partial_{\xi_1}\Psi(\bar{\xi})| \sim |\xi_{12}||\xi_{1-2}| \sim |\xi||\xi_1|$. Then, noting that

$$|m_{s,j}(\bar{\xi})| \lesssim \frac{|\xi|^{\frac{1}{4}}|\xi_1|^{\frac{1}{2}}}{\langle\xi_3\rangle^s},$$

it follows from (4.38) that

$$\begin{aligned}\text{LHS of (4.29)} &\lesssim \sup_{\xi}\Bigg(\sum_{\substack{|\xi|\ll N\lesssim|\alpha|+M\\ \text{dyadic}}}\int_{\substack{|\xi_1|\sim N\\ |\xi_3|\ll|\xi|}}\mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot\frac{|\xi|^{\frac{1}{2}}N}{\langle\xi_3\rangle^{\frac{1}{2}}}\mathrm{d}\xi_1\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}\\ &\lesssim M^{\frac{1}{2}}\sup_{\xi}\Bigg(\sum_{\substack{|\xi|\ll N\lesssim|\alpha|+M\\ \text{dyadic}}}\int_{|\xi_3|\ll|\xi|}\frac{1}{|\xi|^{\frac{1}{2}}\langle\xi_3\rangle^{\frac{1}{2}}}\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}\lesssim\langle\alpha\rangle^{0+}M^{\frac{1}{2}+}.\end{aligned}$$

Subcase 6.c: $|\xi_1| \sim |\xi_2| \sim |\xi| \gg |\xi_3|$. — In this case, the desired estimate holds for $s = \frac{1}{4}$ but it requires an extra factor of $\max\{|\alpha|, M\}^{\frac{1}{12}}$. We have $|\xi_{12}| = |\xi - \xi_3| \sim |\xi|$, $|\xi_{23}| \sim |\xi|$, and $|\xi_{13}| \sim |\xi|$. Hence, the condition $|\Psi(\bar{\xi})| \leqslant |\alpha| + M$ with (2.15) implies $|\xi|^{\frac{1}{4}} \lesssim \max\{|\alpha|, M\}^{\frac{1}{12}}$. In particular, we have

$$|m_j(\bar{\xi})| \lesssim \frac{|\xi|^{\frac{3}{4}}}{\langle\xi_3\rangle^{\frac{1}{4}}} \lesssim \max\{|\alpha|, M\}^{\frac{1}{12}}|\xi|^{\frac{1}{2}}$$

for any $j \in \{1, 2, 3\}$. By viewing Ψ as a function of ξ_1 for fixed ξ and ξ_2, we have $|\partial_{\xi_1}\Psi(\bar{\xi})| \sim |\xi_{13}||\xi_{1-3}| \sim |\xi|^2$. Hence, we have

$$\begin{aligned}&\text{LHS of (4.29)}\\ &\quad\lesssim \max\{|\alpha|, M\}^{\frac{1}{12}}\sup_{\xi}\bigg(\int_{\xi=\xi_1+\xi_2+\xi_3}\mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot|\xi|\mathrm{d}\xi_1\mathrm{d}\xi_2\bigg)^{\frac{1}{2}}\\ &\quad\lesssim \max\{|\alpha|, M\}^{\frac{1}{12}}M^{\frac{1}{2}}\sup_{\xi}\bigg(\int_{|\xi_2|\sim|\xi|}\frac{1}{|\xi|}\mathrm{d}\xi_2\bigg)^{\frac{1}{2}}\lesssim\max\{|\alpha|, M\}^{\frac{1}{12}}M^{\frac{1}{2}}.\end{aligned}$$

In the following three subcases, we deal with the case $|\xi_1|, |\xi_2|, |\xi_3| \gtrsim |\xi|$.

Subcase 6.d: $|\xi_1| \sim |\xi_2| \gg |\xi_3| \gtrsim |\xi|$. — Let $s = \frac{1}{4}$. By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_1, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\xi_{23}||\xi_{2-3}| \sim |\xi_2|^2$.

Hence, with $m_{s,j}(\bar{\xi}) \lesssim |\xi_1|^{\frac{1}{2}}$ and (4.38), we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi} \Bigg(\sum_{\substack{|\xi| \ll N \lesssim |\alpha|+M \\ \text{dyadic}}} \int_{|\xi_1|\sim|\xi_2|\sim N} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot N \mathrm{d}\xi_2 \mathrm{d}\xi_1 \Bigg)^{\frac{1}{2}}$$

$$\lesssim M^{\frac{1}{2}} \sup_{\xi} \Bigg(\sum_{\substack{|\xi| \ll N |\alpha|+M \\ \text{dyadic}}} \int_{|\xi_1|\sim N} \frac{1}{N} \mathrm{d}\xi_1 \Bigg)^{\frac{1}{2}} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+}.$$

Subcase 6.e: $|\xi_1| \sim |\xi_2| \sim |\xi_3| \gg |\xi|$. — This case (with $s = \frac{1}{4}$) follows from Subcase 3.b.

Subcase 6.f: $|\xi_1| \sim |\xi_2| \sim |\xi_3| \sim |\xi|$. — In this case, we have $m_{s,j}(\bar{\xi}) \sim |\xi|^{1-2s}$. In the following, we do not use the size relation between $|\xi_1|$, $|\xi_2|$, and $|\xi_3|$.

Subsubcase 6.f.i: $\min\{|\zeta_1|, |\zeta_2|, |\zeta_3|\} \ll |\xi|$. — Let $s = \frac{1}{4}$. Without loss of generality, assume $|\zeta_3| = |\xi_{12}| \ll |\xi|$. Then, we have $|\xi_{1-2}| = |2\xi_1 - \xi_{12}| \sim |\xi|$. By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_3 (or $\zeta_3 = \xi - \xi_3$), we have $|\partial_{\xi_2} \Psi(\bar{\xi})| \sim |\zeta_3||\xi_{1-2}| \sim |\zeta_3||\xi|$. Hence, with (2.33), we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi} \bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot |\xi| \mathrm{d}\xi_2 \mathrm{d}\zeta_3 \bigg)^{\frac{1}{2}}$$

$$\lesssim M^{\frac{1}{2}} \sup_{\xi} \bigg(\int_{1 \leqslant |\zeta_3| \lesssim |\alpha|+M} \frac{1}{|\zeta_3|} \mathrm{d}\zeta_3 \bigg)^{\frac{1}{2}} \lesssim \langle \alpha \rangle^{0+} M^{\frac{1}{2}+}.$$

In the following, we consider the case: $\min\{|\zeta_1|, |\zeta_2|, |\zeta_3|\} \sim |\xi|$. By the triangle inequality, we have

$$\max(|\xi - \xi_1|, |\xi - \xi_2|, |\xi - \xi_3|) \geqslant |3\xi - \xi_{123}| = 2|\xi|.$$

Hence, we have $|\zeta_1| \sim |\zeta_2| \sim |\zeta_3| \sim |\xi|$.

Subsubcase 6.f.ii: $|\zeta_1| \sim |\zeta_2| \sim |\zeta_3| \sim |\xi|$ *and* $\max\{|\xi_{1-2}|, |\xi_{2-3}|, |\xi_{3-1}|\} \sim |\xi|$. — Let $s = \frac{1}{4}$. Without loss of generality, assume that $|\xi_{1-2}| \sim |\xi|$. By viewing Ψ as a function of ξ_2 for fixed ξ and ξ_3, we have $|\partial_{\xi_2} \Psi(\bar{\xi})| \sim |\zeta_3||\xi_{1-2}| \sim |\xi|^2$. Hence we have

$$\text{LHS of (4.29)} \lesssim \sup_{\xi} \bigg(\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M} \cdot |\xi| \mathrm{d}\xi_2 \mathrm{d}\xi_3 \bigg)^{\frac{1}{2}}$$

$$\lesssim M^{\frac{1}{2}} \sup_{\xi} \bigg(\int_{|\xi_3|\sim|\xi|} \frac{1}{|\xi|} \mathrm{d}\xi_3 \bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.$$

In the three subsubcases, we assume that $|\xi_{1-2}| \geqslant |\xi_{2-3}| \geqslant |\xi_{3-1}|$ without loss of generality.

Subsubcase 6.f.iii: $|\zeta_1| \sim |\zeta_2| \sim |\zeta_3| \sim |\xi|$ and $1 \lesssim |\xi_{1-2}| \ll |\xi|$. — Let $s = \frac{1}{4}$. For fixed ξ and a dyadic number $1 \lesssim N \ll |\xi|$, suppose that $|\xi_{1-2}| \sim N$. Then, by writing $\xi = 3\xi_3 + \xi_{1-2} + 2\xi_{2-3}$, we see that ξ_3 is contained in an interval $I_3(\xi, N)$ of length $\lesssim N$. Moreover, for fixed ξ and ξ_3, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\zeta_3||\xi_{1-2}| \sim |\xi|N$. Hence, we obtain

$$\begin{aligned}
\text{LHS of (4.29)} &\lesssim \sup_{\xi}\Bigg(\sum_{1\lesssim N \ll |\xi|}\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot \mathbb{1}_{|\xi_{1-2}|\sim N}|\xi|\mathrm{d}\xi_2\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}\\
&\lesssim M^{\frac{1}{2}}\sup_{\xi}\Bigg(\sum_{1\lesssim N \ll |\xi|}\int_{\xi_3\in I_3(\xi,N)}\frac{1}{N}\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}\\
&\lesssim \langle\alpha\rangle^{0+}M^{\frac{1}{2}+}\sup_{\xi}\Bigg(\sum_{1\lesssim N \ll |\xi|}\frac{1}{|\xi|^{0+}}\Bigg)^{\frac{1}{2}} \lesssim \langle\alpha\rangle^{0+}M^{\frac{1}{2}},
\end{aligned}$$

where we used (2.33) in the penultimate step.

Subsubcase 6.f.iv: $|\zeta_1| \sim |\zeta_2| \sim |\zeta_3| \sim |\xi|$, $|\xi_{1-2}| \ll 1$, and $|\xi_{3-1}| \lesssim \frac{M}{|\xi|}$. Let $s = \frac{1}{4}$. Arguing as in Subsubcase 6.f.iii, we see that for fixed ξ, ξ_3 is contained in an interval $I(\xi)$ of length $O(1)$. Hence we have

$$\begin{aligned}
\text{LHS of (4.29)} &\lesssim \sup_{\xi}\Bigg(\int_{\substack{\xi_3\in I(\xi)\\ |\xi_1-\xi_3|\lesssim \frac{M}{|\xi|}}} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot|\xi|\mathrm{d}\xi_1\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}\\
&\lesssim M^{\frac{1}{2}}\sup_{\xi}\Bigg(\int_{\xi_3\in I(\xi)}1\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}} \lesssim M^{\frac{1}{2}}.
\end{aligned}$$

Subsubcase 6.f.v: $|\zeta_1| \sim |\zeta_2| \sim |\zeta_3| \sim |\xi|$, and $\frac{M}{|\xi|} \ll |\xi_{3-1}| \leqslant |\xi_{1-2}| \ll 1$. — Let $s = \frac{1}{4}$. For fixed ξ and a dyadic number $1 \ll N \ll \frac{|\xi|}{M}$, suppose that $|\xi_{1-2}| \sim N^{-1}$. Then, arguing as in Subsubcase 6.f.iii, we see that ξ_3 is contained in an interval $\widetilde{I}_3(\xi, N)$ of length $\lesssim N^{-1}$. Moreover, for fixed ξ and ξ_3, we have $|\partial_{\xi_2}\Psi(\bar{\xi})| \sim |\zeta_3||\xi_{1-2}| \sim |\xi|N^{-1}$. Hence, with (2.33), we obtain

LHS of (4.29)

$$\lesssim \sup_{\xi}\Bigg(\sum_{1\ll N \ll \frac{|\xi|}{M}}\int_{\xi=\xi_1+\xi_2+\xi_3} \mathbb{1}_{|\Psi(\bar{\xi})-\alpha|\leqslant M}\cdot \mathbb{1}_{|\xi_{1-2}|\sim N^{-1}}|\xi|\mathrm{d}\xi_2\mathrm{d}\xi_3\Bigg)^{\frac{1}{2}}$$

$$\lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+} \sup_{\xi} \Bigg(\sum_{1 \ll N \ll \frac{|\xi|}{M}} \int_{\xi_3 \in \widetilde{I}_3(\xi, N)} \frac{N}{|\xi|^{0+}} \mathrm{d}\xi_3 \Bigg)^{\frac{1}{2}}$$

$$\lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+} \sup_{\xi} \Bigg(\sum_{N \gg 1} \frac{1}{M^{0+} N^{0+}} \Bigg)^{\frac{1}{2}} \lesssim \langle\alpha\rangle^{0+} M^{\frac{1}{2}+}.$$

This completes the proof of Lemma 4.9. □

Concluding remark

In Section 3, we presented the full details of the normal form reductions since this is the first paper, where we handle multilinear estimates by successive applications of the trilinear localized modulation estimate. The essential part for establishing an a priori estimate in $L^2(\mathbb{R})$ for the cubic NLS and in $H^{\frac{1}{4}}(\mathbb{R})$ for the mKdV appears in Subsection 3.3, where we applied the localized modulation estimates from Section 2. In Section 4, we also needed to prove another localized modulation estimate (in the weak norm: Lemmas 4.1 and 4.9) for justifying the formal computations in Section 3, where an extra complication was introduced for the mKdV problem due to the derivative nonlinearity. In essence, our method allows one to reduce the entire problem of proving unconditional well-posedness to simply proving two basic trilinear estimates (i.e. localized modulation estimates in the strong norm and in the weak norm: Lemmas 2.3 and 4.1 for the cubic NLS and Lemma 2.6 and 4.9 for the mKdV). This reduction is the main novelty of the paper and such a reduction provides a significant simplification in studying unconditional well-posedness for various dispersive PDEs on $\mathbb{R}^d$ and $\mathbb{T}^d$.

Acknowledgements

S.K. would like to express his gratitude to the School of Mathematics at the University of Edinburgh for its hospitality during his visit in the academic year 2015–2016. The authors would like to thank Nobu Kishimoto for a helpful discussion and Didier Pilod for a helpful comment. They are also grateful to Razvan Mosincat for careful proofreading.

Bibliography

[1] V. I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, second ed., Grundlehren der Mathematischen Wissenschaften, vol. 250, Springer, 1988, xiv+351 pages.

[2] A. V. Babin, A. A. Ilyin & E. S. Titi, "On the regularization mechanism for the periodic Korteweg-de Vries equation", *Commun. Pure Appl. Math.* **64** (2011), no. 5, p. 591-648.

[3] L. Chaichenets, D. Hundertmark, P. Kunstmann & N. Pattakos, “Nonlinear Schrödinger equation, differentiation by parts and modulation spaces”, *J. Evol. Equ.* **19** (2019), no. 3, p. 803-843.

[4] M. Christ, “Power series solution of a nonlinear Schrödinger equation”, in *Mathematical aspects of nonlinear dispersive equations*, Annals of Mathematics Studies, vol. 163, Princeton University Press, 2007, p. 131-155.

[5] ———, “Nonuniqueness of weak solutions of the nonlinear Schrödinger equation”, `https://arxiv.org/abs/math/0503366`, 2018.

[6] M. Christ, J. Colliander & T. Tao, “Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations”, *Am. J. Math.* **125** (2003), no. 6, p. 1235-1293.

[7] ———, “A priori bounds and weak solutions for the nonlinear Schrödinger equation in Sobolev spaces of negative order”, *J. Funct. Anal.* **254** (2008), no. 2, p. 368-395.

[8] M. Christ, J. Holmer & D. Tataru, “Low regularity a priori bounds for the modified Korteweg–de Vries equation”, *Lib. Math. (N.S.)* **32** (2012), no. 1, p. 51-75.

[9] J. Chung, Z. Guo, S. Kwon & T. Oh, “Normal form approach to global well-posedness of the quadratic derivative nonlinear Schrödinger equation on the circle”, *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **34** (2017), no. 5, p. 1273-1297.

[10] J. Colliander, M. Keel, G. Staffilani, H. Takaoka & T. Tao, “Sharp global well-posedness for KdV and modified KdV on ℝ and 𝕋”, *J. Am. Math. Soc.* **16** (2003), no. 3, p. 705-749.

[11] M. B. Erdoğan & N. Tzirakis, “Global smoothing for the periodic KdV evolution”, *Int. Math. Res. Not.* **20** (2013), p. 4589-4614.

[12] J. Forlano & T. Oh, “Normal form approach to the one-dimensional cubic nonlinear Schrödinger equation in Fourier-amalgam spaces”, preprint.

[13] M. Gubinelli, “Rough solutions for the periodic Korteweg de Vries equation”, *Commun. Pure Appl. Anal.* **11** (2012), no. 2, p. 709-733.

[14] Z. Guo, “Global well posedness of Korteweg de Vries equation in $H^{-3/4}(\mathbb{R})$”, *J. Math. Pures Appl.* **91** (2009), no. 6, p. 583-597.

[15] Z. Guo, S. Kwon & T. Oh, “Poincaré Dulac normal form reduction for unconditional well-posedness of the periodic cubic NLS”, *Commun. Math. Phys.* **322** (2013), no. 1, p. 19-48.

[16] S. Hong & S. Kwon, “Nonsqueezing property of the coupled KdV type system without Miura transform”, `https://arxiv.org/abs/1509.08114`, 2015.

[17] T. Kappeler & P. Topalov, “Global wellposedness of KdV in $H^{-1}(\mathbb{T},\mathbb{R})$”, *Duke Math. J.* **135** (2006), no. 2, p. 327-360.

[18] T. Kato, “On the Korteweg-de Vries equation”, *Manuscr. Math.* **28** (1979), no. 1-3, p. 89-99.

[19] ———, “On nonlinear Schrödinger equations. II. H^s-solutions and unconditional well-posedness”, *J. Anal. Math.* **67** (1995), p. 281-306.

[20] C. E. Kenig, G. Ponce & L. Vega, “Well posedness and scattering results for the generalized Korteweg de Vries equation via the contraction principle”, *Commun. Pure Appl. Math.* **46** (1993), no. 4, p. 527-620.

[21] ———, “On the ill-posedness of some canonical dispersive equations”, *Duke Math. J.* **106** (2001), no. 3, p. 617-633.

[22] R. Killip & M. Vişan, “KdV is wellposed in H^{-1}”, *Ann. Math.* **190** (2019), no. 1, p. 249-305.

[23] R. Killip, M. Vişan & X. Zhang, “Low regularity conservation laws for integrable PDE”, *Geom. Funct. Anal.* **28** (2018), no. 4, p. 1062-1090.

[24] N. Kishimoto, "Well posedness of the Cauchy problem for the Korteweg de Vries equation at the critical regularity", *Differ. Integral Equ.* **22** (2009), no. 5-6, p. 447-464.

[25] ———, "Unconditional uniqueness of solutions for nonlinear dispersive equations", 2015, Proceedings of the 40th Sapporo Symposium on Partial Differential Equations, available at `http://eprints3.math.sci.hokudai.ac.jp/2375/`, 78-82 pages.

[26] ———, "private communication", 2016.

[27] H. Koch & D. Tataru, "A priori bounds for the 1D cubic NLS in negative Sobolev spaces", *Int. Math. Res. Not.* **16** (2007), article ID rnm053 (36 pages).

[28] ———, "Energy and local energy bounds for the 1-d cubic NLS equation in $H^{-\frac{1}{4}}$", *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **29** (2012), no. 6, p. 955-988.

[29] ———, "Conserved energies for the cubic nonlinear Schrödinger equation in one dimension", *Duke Math. J.* **167** (2018), no. 17, p. 3207-3313.

[30] S. Kwon & T. Oh, "On unconditional well posedness of modified KdV", *Int. Math. Res. Not.* **15** (2012), p. 3509-3534.

[31] R. M. Miura, "Korteweg de Vries equation and generalizations. I. A remarkable explicit nonlinear transformation", *J. Math. Phys.* **9** (1968), p. 1202-1204.

[32] R. M. Miura, C. S. Gardner & M. D. Kruskal, "Korteweg de Vries equation and generalizations. II. Existence of conservation laws and constants of motion", *J. Math. Phys.* **9** (1968), p. 1204-1209.

[33] L. Molinet, D. Pilod & S. Vento, "Unconditional uniqueness for the modified Korteweg-de Vries equation on the line", *Rev. Mat. Iberoam.* **34** (2018), no. 4, p. 1563-1608.

[34] T. Oh, "A remark on norm inflation with general initial data for the cubic nonlinear Schrödinger equations in negative Sobolev spaces", *Funkc. Ekvacioj* **60** (2017), no. 2, p. 259-277.

[35] T. Oh, P. Sosoe & N. Tzvetkov, "An optimal regularity result on the quasi-invariant Gaussian measures for the cubic fourth order nonlinear Schrödinger equation", *J. Éc. Polytech., Math.* **5** (2018), p. 793-841.

[36] T. Oh & N. Tzvetkov, "Quasi-invariant Gaussian measures for the cubic fourth order nonlinear Schrödinger equation", *Probab. Theory Relat. Fields* **169** (2017), no. 3-4, p. 1121-1168.

[37] T. Oh & Y. Wang, "Normal form approach to the one dimensional periodic cubic nonlinear Schrödinger equation in almost critical Fourier–Lebesgue spaces", `https://arxiv.org/abs/1811.04868`, to appear in *J. Anal. Math.*

[38] ———, "Global well posedness of the periodic cubic fourth order NLS in negative Sobolev spaces", *Forum Math. Sigma* **6** (2018), article ID e5 (80 pages).

[39] ———, "Global well-posedness of the one-dimensional cubic nonlinear Schrödinger equation in almost critical spaces", *J. Differ. Equations* **269** (2020), no. 1, p. 612-640.

[40] N. Pattakos, "NLS in the modulation space $M_{2,q}(\mathbb{R})$", *J. Fourier Anal. Appl.* **25** (2019), no. 4, p. 1447-1486.

[41] T. Tao, "Multilinear weighted convolution of L^2-functions, and applications to nonlinear dispersive equations", *Am. J. Math.* **123** (2001), no. 5, p. 839-908.

[42] Y. Tsutsumi, "L^2-solutions for nonlinear Schrödinger equations and nonlinear groups", *Funkc. Ekvacioj* **30** (1987), no. 1, p. 115-125.

[43] M. Wadati, "The modified Korteweg–de Vries equation", *J. Phys. Soc. Japan* **34** (1973), p. 1289-1296.

[44] V. E. Zakharov & A. B. Shabat, "Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media", *Sov. Phys., JETP* **34** (1972), no. 1, p. 62-69, translated from Ž. Èksper. Teoret. Fiz. **61** (1971), no. 1, p. 118-134.