



Optimal Local Well-Posedness for the Periodic Derivative Nonlinear Schrödinger Equation

Yu Deng¹, Andrea R. Nahmod² , Haitian Yue¹

¹ Department of Mathematics, University of Southern California, Los Angeles, CA 90089, USA

² Department of Mathematics, University of Massachusetts, 710 N. Pleasant Street, Amherst, MA 01003, USA

Received: 24 June 2020 / Accepted: 1 September 2020

Published online: 14 November 2020 – © Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract: We prove local well-posedness for the periodic derivative nonlinear Schrödinger equation, which is L^2 critical, in Fourier-Lebesgue spaces which scale like $H^s(\mathbb{T})$ for $s > 0$. Our result is optimal in the sense that it covers the full subcritical regime. In particular we close the existing gap in the subcritical theory by improving the result of Grünrock and Herr (SIAM J Math Anal 39(6):1890–1920, 2008), which established local well-posedness in Fourier-Lebesgue spaces which scale like $H^s(\mathbb{T})$ for $s > \frac{1}{4}$. We achieve this result by a delicate analysis of the structure of the solution and the construction of an adapted nonlinear submanifold of a suitable function space. Together these allow us to construct the unique solution to the given subcritical data. This constructive procedure is inspired by the theory of para-controlled distributions developed by Gubinelli et al. (Forum Math Pi, 3:75, 2015) and Catellier and Chouk (Ann Probab 46(5):2621–2679, 2018) in the context of stochastic PDE. Our proof and results however, are purely deterministic.

1. Introduction

The derivative nonlinear Schrödinger equation

$$iu_t + \partial_x^2 u = i\partial_x(|u|^2 u), \quad (1.1)$$

where $(t, x) \in (-T, T) \times \mathbb{T}$ (periodic) or $(-T, T) \times \mathbb{R}$ (non-periodic), is a Hamiltonian PDE obtained in a long-wave, weakly nonlinear scaling regime from the nonlinear propagation of compressible magneto-hydrodynamic wave trains parallel to the magnetic field in an ideal plasma with dispersion due to Hall currents in the Ohm law [12, 47, 51, 52]. Equation (1.1) is well-known to be a completely integrable system after work by Kaup and Newell, who showed in [48] the general scheme for solving (1.1) in terms of the inverse scattering method (see also [44–47] and references therein). In particular (1.1)

Andrea R. Nahmod is partially supported by NSF-DMS-1463714 and NSF-DMS-1800852.

conserves mass and energy. However, unlike the usual one dimensional cubic NLS, equation, (1.1) is not Galilean invariant. Global in time existence of solutions in the Schwartz class on \mathbb{R} was shown by Lee in [49].

The Cauchy initial value problem problem for (1.1) is scale invariant for data in L^2 , that is, if $u(t, x)$ is a solution then so is $u_\lambda(t, x) = \lambda^{\frac{1}{2}}u(\lambda^2t, \lambda x)$ with the same L^2 norm. Thus *a priori* one expects local well-posedness for (1.1) with initial data in H^s for $s \geq 0$. However, while local well-posedness in H^s for (1.1) is known for $s \geq \frac{1}{2}$ [40, 63], one has *ill-posedness* in H^s for $s < \frac{1}{2}$ [3, 40, 63]. One way to close the gap between the scaling heuristics and actual local well-posedness results is by considering data in the Fourier-Lebesgue spaces H_p^σ , where $p \geq 2$. These spaces are defined as

$$\|u_0\|_{H_p^\sigma} := \|\langle k \rangle^\sigma \widehat{u}_0(k)\|_{L_k^p} \quad (1.2)$$

(with L_k^p replaced by ℓ_k^p in the periodic case). These spaces have naturally arisen in the literature and we refer the reader to e.g. [15, 24–27, 42, 68] for some instances. Note that in one dimension the Fourier-Lebesgue space H_p^σ has the same scaling¹ as the Sobolev space H^s for

$$s = \sigma + \frac{1}{p} - \frac{1}{2}; \quad (1.3)$$

in particular, $H_\infty^{\frac{1}{2}}$ has the scaling of L^2 , and $H_2^{\frac{1}{2}} = H^{\frac{1}{2}}$.

In the non-periodic case, Grünrock [25] proved optimal local well-posedness for (1.1) in $H_p^\sigma(\mathbb{R})$ for $\sigma \geq \frac{1}{2}$ and $p < \infty$, which allows the corresponding Sobolev regularity s to be arbitrarily close to 0, thus covering the full subcritical range. The proof combines the gauge transformation introduced in [37] (used also in [38, 39, 61, 63]) and new bilinear and trilinear estimates for the gauged equation in an appropriate variant of Bourgain’s Fourier restriction norm spaces [8] (see Sect. 3 below for details) which follow from the dispersion and the smoothing properties of the Schrödinger propagator on \mathbb{R} .

In the periodic case, however, local well-posedness for (1.1) in $H_p^\sigma(\mathbb{T})$ is only known for $\sigma \geq \frac{1}{2}$ and $2 \leq p < 4$, which scaling-wise correspond to Sobolev regularity $s > \frac{1}{4}$. This is the work of Grünrock and Herr [27]. Their proof is based on the adapted periodic gauge transformation in [40] and new multilinear estimates for the gauged equation in adapted variants of the Fourier restriction norm spaces. Moreover, it is proved in [27] that the crucial trilinear estimates² become false when $p \geq 4$, so this result as well as the existing gap in the local well-posedness theory between $s > 1/4$ and the scaling prediction $s > 0$, cannot be improved *within the framework* of [27].

In this paper we close this existing gap in the periodic case. More precisely, we prove local well-posedness for (1.1) in $H_p^\sigma(\mathbb{T})$ for $\sigma \geq \frac{1}{2}$ and $p < \infty$ which covers the entire subcritical regime, hence yielding optimal local well-posedness. Our main theorem is stated as follows:

Theorem 1.1. *Fix $\sigma \geq \frac{1}{2}$ and $p_0 < \infty$. For any $A > 0$, there exists $T = T(p_0, A) > 0$, such that if $\|u_0\|_{H_{p_0}^\sigma} \leq A$, then there exists a unique solution $u \in \mathcal{Z} \subset C_t^0 H_{p_0}^\sigma(J)$ to (1.1) with initial data $u(0) = u_0$, where $J = [-T, T]$. Here \mathcal{Z} is an explicitly defined*

¹ Here and henceforth we mean that the homogenous part of the Fourier-Lebesgue norm scales like the corresponding homogeneous Sobolev norm.

² More precisely the trilinear estimates containing as one of their inputs the derivative term.

sub-manifold of $C_t^0 H_{p_0}^\sigma(J)$, see Definition 4.3 below. The map $u_0 \mapsto u$ is continuous with respect to the $C_t^0 H_{p_0}^\sigma(J)$ metric.

Remark 1.2. The solution we construct solves (1.1) in the sense that it solves the integral equation

$$u(t) = e^{it\partial_x^2} u_0 + \int_0^t e^{i(t-s)\Delta} \partial_x (|u(s)|^2 u(s)) \, ds. \quad (1.4)$$

It is also the unique limit of smooth solutions: given $A > 0$, and any smooth initial data u_0 in the A -ball of $H_{p_0}^\sigma$, the classical solution exists for time $T = T(p_0, A) > 0$, and the data-to-solution map extends continuously to all of this ball. Moreover, if $p_0 < 4$, our solution coincides with the solution constructed in [27], for as long as the latter exists.

Remark 1.3. We will only prove Theorem 1.1 with $\sigma = \frac{1}{2}$ and $p_0 \geq 4$. The extension to $\sigma > \frac{1}{2}$ is standard (see Proposition 7.1 for a sketch), and when $2 \leq p_0 < 4$ the result follows directly from [40].

Remark 1.4. In the periodic case, we suspect that local well-posedness for (1.1) in the scaling critical spaces (e.g. $L^2(\mathbb{T})$ and $H_\infty^{\frac{1}{2}}(\mathbb{T})$) is false. For a dyadic number N , consider some carefully chosen initial data u_0 whose Fourier transform $\widehat{u}_0(k)$ is supported in $\{k \in \mathbb{Z} : k \sim N\}$ and is of size $N^{-\frac{1}{2}}$. Then it directly follows that $u_0 \in L^2 \cap H_\infty^{\frac{1}{2}}$ uniformly in N . Denote by u_{lin} its linear evolution $e^{it\partial_x^2} u_0$, and consider the second order iteration³ $u_{(2)} := (i\partial_t + \partial_x^2)^{-1} \partial_x (|u_{\text{lin}}|^2 u_{\text{lin}})$ and the third order iteration $u_{(3)} := (i\partial_t + \partial_x^2)^{-1} \partial_x (|u_{(2)}|^2 u_{(2)})$. Then, the spacetime Fourier transform $\widehat{u}_{(2)}(k, \xi)$ of the second iterate $u_{(2)}$ is supported in $\{(k, \xi) \in \mathbb{Z} \times \mathbb{R} : |\xi + k^2| \lesssim N^2\}$ and has size at least $N^{-\frac{1}{2}} |\xi + k^2|^{-1}$ by some simple calculations from the linear evolution and the definition of $u_{(2)}$. Similarly, the spacetime Fourier transform $\widehat{u}_{(3)}(k, \xi)$, even restricted to $|\xi + k^2| \sim 1$, will have size at least $N^{-\frac{1}{2}} \log(N)$. This log divergence implies that the third iteration $u_{(3)}$ has already left the data space $C_t^0 L^2$ or $C_t^0 H_\infty^{\frac{1}{2}}$. In fact, similar log divergences also appears in the Cauchy problem for the one-dimensional quintic⁴ and two-dimensional cubic periodic nonlinear Schrödinger equations with data in L^2 where Bourgain showed in particular the failure of the $L^6(\mathbb{T})$ and $L^4(\mathbb{T}^2)$ Strichartz estimates for data in L^2 (see [8]). In fact in both cases the local well posedness in L^2 remain open to date.

Remark 1.5. Global well-posedness for the Cauchy problem for (1.1) is known to hold for data in H^s , $s \geq \frac{1}{2}$ both on \mathbb{R} [16, 17, 50] and on \mathbb{T} [40, 53, 69]. Furthermore, one has almost sure global well posedness for data in Fourier Lebesgue spaces $H_p^\sigma(\mathbb{T})$ that have the scaling of $H^{\frac{1}{2}-\varepsilon}(\mathbb{T})$, $\varepsilon > 0$ [58, 59]. In this paper our primary goal is to close the gap in the local well-posedness Cauchy theory. One may then study the question of deterministic global well-posedness below $H^{\frac{1}{2}}(\mathbb{T})$ which requires quite different techniques such as for example exploiting the integrability of the equation and seeking suitable new conservation laws below $H^{\frac{1}{2}}$.

³ The operator $(i\partial_t + \partial_x^2)^{-1}$ is just the Duhamel operator I as in (2.14). For technical reasons, here we in fact use the truncated version of the Duhamel operator \mathcal{I} as in (2.16).

⁴ Note that for data with finite mass, Eq. (1.1) is in fact gauged equivalent to (1.5).

1.1. The standard approach, and difficulties. Generally speaking the difficulty one faces in solving (1.1) is a derivative loss arising from the term $i|u|^2 u_x$ in the nonlinearity of (1.1), and hence for low regularity data the key is to somehow make up for this loss. The first step towards this goal is a gauge transformation [37–40, 63] which removes this bad resonant term in the nonlinearity that loses derivatives and makes the estimates uncontrollable. Matters are then reduced to studying the gauged derivative nonlinear Schrödinger equation which we schematically write as

$$(\partial_t - i\partial_x^2)v = \mathcal{C}(v, v, v) + \mathcal{Q}(v, \dots, v), \quad (1.5)$$

where the nonlinearity $\partial_x(|u|^2 u)$ has been transformed into the sum of the ‘better’ cubic term

$$\mathcal{C}(v, v, v) \sim i\bar{v}_x \cdot v^2,$$

plus a quintic term which contains no derivative terms in it and which we momentarily neglect in this discussion as being ‘lower order’. Once the bad nonlinear term is gauged away from (1.1), the solution v to (1.5) is constructed by a fixed point argument, which follows from proving multilinear estimates in suitable Fourier restriction norm function spaces adapted to the data space. In the non-periodic case [25] these spaces in conjunction with the dispersion and smoothing effects available on \mathbb{R} suffice, as we mentioned above, to prove optimal local well-posedness⁵ for (1.5) and hence for (1.1) in H_p^σ , where $\sigma \geq \frac{1}{2}$ and $p < \infty$. In the periodic case however, the authors need to introduce in [27] a fourth parameter q in the Fourier restriction norm function spaces, namely they define⁶

$$\|u\|_{X_{p,q}^{\sigma,b}} = \|\langle k \rangle^\sigma \langle \xi + k^2 \rangle^b \widehat{u}(k, \xi)\|_{\ell_k^p L_\xi^q}$$

and prove that, if $\sigma = \frac{1}{2}$ and $p < 4$ the trilinear estimate

$$\left\| \int_0^t e^{i(t-s)\partial_x^2} (\partial_x \bar{v}_1 \cdot v_2 v_3) ds \right\|_{X_{p,q}^{\frac{1}{2},b}} \lesssim \prod_{j=1}^3 \|v_j\|_{X_{p,q}^{\frac{1}{2},b}} \quad (1.6)$$

holds true for $(b, q) = (\frac{1}{2}+, 2)$. Furthermore, they construct explicit counterexamples showing that for $\sigma = \frac{1}{2}$ and $p \geq 4$, the trilinear estimate (1.6) fails for *any* choice of (b, q) [27]. In other words, when $\sigma = \frac{1}{2}$ and $p \geq 4$, which in the Sobolev scale corresponds to regularity $0 < s \leq \frac{1}{4}$, the local solution to (1.5) cannot be constructed directly by a contraction mapping argument.⁷

To prove our Theorem 1.1 we must and will take a different approach. After performing the gauge transformation, our point of departure is the following observation: let $\sigma = \frac{1}{2}$ and $p \geq 4$. If one compares the profiles of the counterexamples constructed in [27] with the profiles of the terms occurring in the formal Picard iterations of (1.5), then they will never coincide, although they belong to exactly the same adapted Fourier

⁵ Local well-posedness for the gauged equation (1.5) implies local existence, uniqueness and continuity of the flow map for (1.1) [27, 40].

⁶ When $p = q = 2$ these spaces coincide with Bourgain’s Fourier restriction norm spaces associated to the Schrödinger equation, and are simply denoted by $X^{s,b}$.

⁷ In principle, it might be possible that a trilinear estimate holds in some exotic Banach space not of form $X_{p,q}^{s,b}$ but, if not unlikely, this would at least require a rather sophisticated construction.

restriction norm $X_{p,q}^{\sigma,b}$ spaces. In particular, let us consider two functions: (1) the second iteration u^* in a formal expansion corresponding to a ‘high–low–low’ interaction (high in the derivative term) where the high input is u_{lin} (as in Remark 1.4) and the two low inputs are just spacetime Schwartz functions; (2) the ‘intruding’ term u^{**} whose spacetime Fourier transform is supported in $|k| \sim N$, $|(\xi + k^2) - 2k| \sim 1$ and has size $\sim N^{-1}$. We can check that spacetime Fourier transform of u^* is actually supported in $\{|k| \sim N, |(\xi + k^2) - 2k^2| \sim 1\}$ and has size $\sim N^{-\frac{3}{2}}$. Note that both u^* and u^{**} belong to the same $X_{\infty,q}^{\frac{1}{2},b}$ space, namely when $b \leq \frac{1}{2}$; however, u^* is an iteration term that actually appears in the solution, while u^{**} is precisely the ‘intruding’ term that leads to the counterexample of (1.6) in [27]. Namely, if one inputs in the cubic derivative nonlinearity an ‘intruding’ term with derivative, a linear evolution u_{lin} and a spacetime Schwartz function, then the output function, even restricted to frequencies $k = 0$ and $|\xi| \sim 1$, will have size $\sim N^{\frac{1}{2}}$ which yields a contradiction to (1.6). Therefore it is reasonable to imagine that, the solution v to (1.5) still exists in one of these spaces—say $b = \frac{1}{2}+$ and $q = 2$ for definiteness—but will have some *specific structure* such that it precisely avoids the counterexamples constructed in [27]. To that effect we will construct v in a *nonlinear submanifold* \mathcal{W} of the Banach space $X_{p,2}^{\sigma,\frac{1}{2}+}$ containing functions of a specific structure whence the trilinear estimate (1.6) will actually hold true when $\sigma = \frac{1}{2}$ and $p \geq 4$. The heart of this paper will be to identify this precise structure.

In order to motivate our approach we take a step back and review some of the methods developed in the probabilistic (random data, or stochastically forced) context. We note in passing that an immediate corollary of our main Theorem 1.1 is that for random initial data of form

$$u^\omega(0) := \sum_{k \in \mathbb{Z}} \frac{g_k(\omega)}{\langle k \rangle^{\frac{1}{2}+\theta}} e^{ikx} \quad (1.7)$$

where g_k are i.i.d Gaussian random variables, $\langle k \rangle := \sqrt{1 + |k|^2}$, and $\theta > 0$ is fixed but arbitrary, the solution to (1.1) or equivalently (1.5) almost surely exists for a positive time.

1.2. Ideas from the probabilistic setting. In the probabilistic PDE context (i.e. random data theory for dispersive and wave equations or parabolic stochastic PDEs) where one deals with randomized initial data or a random forcing term, the idea of exploiting the structure of the solution has been used for a long time, see for example Bourgain [6, 7] in the context of the defocusing (Wick ordered) cubic nonlinear Schrödinger equation,⁸ and Da Prato-Debussche [19, 20] in the context of the stochastic Navier–Stokes and the stochastic quantization equations. More recently this idea has been exploited in a large body of work by many authors. See for example [4–7, 10, 18, 22, 23, 41, 60, 66, 70] and references therein for some works on the random data local Cauchy theory in the context of nonlinear Schrödinger equations. The key point is that, if one considers the linear evolution of random data (or random forcing), then almost surely, it enjoys much better estimates than arbitrary functions of the same regularity. In turn this allows one to *re-center* the solution around the linear evolution of random data (or around higher order iterates), and conclude that the difference between the two belongs to a Banach

⁸ See also, later work by Burq and Tzvetkov in the context of nonlinear wave equations [9].

space of higher regularity than the one dictated by the (weaker) regularity of the random initial data.

For example, in Bourgain [6], which deals with the cubic nonlinear Schrödinger equation on \mathbb{T}^2 , the initial data ϕ^ω belongs to Sobolev $H^{-\varepsilon}$ almost surely, for any $\varepsilon > 0$, whence its linear evolution only belongs to the Fourier restriction norm spaces $X^{-\varepsilon, \frac{1}{2}+}$ almost surely. On the other hand, the equation is L^2 critical, so if one were to try to prove local well posedness via a fixed point argument, the needed trilinear estimates would fail for arbitrary functions in $X^{-\varepsilon, \frac{1}{2}+}$. Instead, Bourgain [6] constructed solutions u centered around the random linear evolution $\Psi^\omega := e^{it\bar{\Delta}}\phi^\omega$. That is of the form:

$$u = \Psi^\omega + R, \quad \text{where} \quad (i\partial_t + \Delta)R = \mathcal{N}(\Psi^\omega + R), \quad (1.8)$$

and where we have denoted by \mathcal{N} the Wick ordered cubic nonlinearity. Then, almost surely, the needed trilinear estimates for $\mathcal{N}(\Psi^\omega + R)$ hold true and the solution R to the difference equation in (1.8) can be constructed in a smoother space $X^{\varepsilon, \frac{1}{2}+}$ by a contraction mapping argument. Heuristically, one should view (1.8) as a ‘hybrid equation’ which on the one hand behaves subcritically in R , thus locally well-posed in H^ε ; while on the other hand the random linear evolutions Ψ^ω behave better than an arbitrary function in $X^{-\varepsilon, \frac{1}{2}+}$ when they are entries in $\mathcal{N}(\Psi^\omega + R)$ thanks to large deviation estimates. A similar phenomenon happens in Da Prato–Debussche’s argument for-for example-the stochastic Navier–Stokes equation on \mathbb{T}^2 with spacetime white noise forcing ζ [19] where the role of Ψ^ω is replaced by Z , the linear evolution of white noise, $Z_t - \Delta Z = \zeta$.

In both cases, the method can be understood as constructing solutions in a (random affine) submanifold \mathcal{W} consisting of functions belonging to a ball in a smoother space, centered at the random linear evolution.

In the past few years, Gubinelli, Imkeller and Perkowski [28, 29] (see also [11] and [30]) developed a far-reaching generalization of this re-centering method based on the idea of *para-controlled distributions*. This is an analytic counterpart to the theory of regularity structures developed by Hairer [33–36]. Roughly speaking, in addition to the linear evolution and possibly (suitably renormalized) higher order expressions of the linear evolution, one moves to the new ‘center’ terms that are ‘para-controlled’ by such expressions. Here a function f is said to be para-controlled by a function g if, up to some smoother ‘remainder’ terms, f can be written as the Bony para-product between high frequencies of g and low frequencies of some auxiliary function h , namely that

$$f = \Pi_>(g, h) + R := \sum_N P_N g \cdot P_{\ll N} h + R, \quad (1.9)$$

where for dyadic frequencies N , P_N and $P_{\ll N}$ are the standard Littlewood-Paley operators projecting onto frequencies $\sim N$ and $\ll N$ respectively, and R is smoother than f . An example is the (parabolic) Φ_3^4 model, which is the cubic heat equation on \mathbb{T}^3 with white noise forcing [11, 28], where one constructs solutions of the form

$$u = Z + I(P_3(Z)) + I\Pi_>(P_2(Z), u - Z) + R, \quad (1.10)$$

where Z is the linear evolution of the white noise, $P_3(Z)$ consists of other structured components which themselves are given in term of (suitably renormalized cubic) powers of Z , I is the Duhamel operator $(\partial_t - \Delta)^{-1}$, and R is a remainder, that has higher regularity. Here then the solution consists of the linear evolution Z , a higher order

expression $I(P_3(Z))$, a para-controlled part term $I \Pi_>(Z^2, u - Z)$ and a remainder and thus belongs to a random submanifold which is much more nonlinear.

These ideas have been extensively used in various stochastic contexts in recent years by many authors. We refer the reader for example to work by Mourrat and Weber [56] and to Mourrat, Weber and Xu [57] and to references therein for further work in the context of the Φ_3^4 model and to Chandra and Weber [13] and references therein for a nice survey of these ideas. See also [2]. We also refer to recent work by Gubinelli, Koch and Oh [31, 32] where these ideas were applied to the stochastic nonlinear wave equation with quadratic nonlinearity in \mathbb{T}^2 and in \mathbb{T}^3 respectively.

1.3. The deterministic context of DNLS. Inspired by the ideas in the probabilistic setting described above, in this paper we develop a new *deterministic* method to describe the structure of solutions v to the Cauchy initial value problem for (1.1) with data at almost critical regularity. A review of all of the above examples suggests that, if we were in the probabilistic setting (i.e. (1.7)), we should look for solutions essentially of form

$$u = w + (\text{terms para-controlled by } w) + (\text{smooth remainders}),$$

where w is the combination of the random linear evolution, and multilinear expressions dictated by the random linear evolution. The choice of such w is forced upon us (one can at most choose the order of expansion) by the fact that one needs to (and indeed can) gain from the exact Gaussian structure.

In the deterministic setting, there is no gain from randomness. One could try to mimic the probabilistic construction of para-controlled terms in previous works, and arrive at the ansatz

$$v = w + I \Pi_>^{(1)}(\partial_x \bar{w}, v, v) + I \Pi_>^{(2)}(\partial_x \bar{w}, w, v) + (\text{smooth remainders}), \quad (1.11)$$

where I is the Duhamel operator

$$I F(t) = \int_0^t e^{i(t-s)\partial_x^2} F(s) \, ds$$

and the cubic para-products are defined by

$$\begin{aligned} \Pi_>^{(1)}(\partial_x \bar{w}, v, v) &= \sum_N P_N \partial_x \bar{w} \cdot (P_{\ll N} v)^2, \\ \Pi_>^{(2)}(\partial_x \bar{w}, w, v) &= \sum_N P_N \partial_x \bar{w} \cdot (P_N w P_{\ll N} v). \end{aligned} \quad (1.12)$$

However unlike the probabilistic setting, we are no longer guided by the Gaussians and need to find the right w ourselves. The naive choice of linear evolution for w is doomed to fail, and even if one includes multilinear expressions of the linear evolution, calculations show that in the absence of randomness, one would need to expand to a very high (if not infinite) order before unearthing ‘smooth remainders’ that have enough regularity (namely $H_{4-}^{\frac{1}{2}}$ due to [27]) to close the estimates. With a high order of expansion, the terms involved then quickly become too complex to control in our setting. Work in this direction was considered by the second author together with Chanillo, Czubak, Mendelson and Staffilani in the context of the nonlinear wave equation with quadratic derivative nonlinearities, see [14] for details.

To get out of this maze, in this paper we will give up the idea of fixing w to be some explicit multilinear expression dictated by the linear evolution. Instead we will construct this w , which ‘para-controls’⁹ the solution v , *dynamically*. That is, we take all the linear and higher order terms in the above-mentioned expansion, as well as the presumed smooth error terms, and put them into a single ‘center’ w . This leads to the new ansatz

$$v = w + I \Pi_{>}^{(1)}(\partial_x \bar{w}, v, v) + I \Pi_{>}^{(2)}(\partial_x \bar{w}, w, v), \quad (1.13)$$

where w is the ‘center’ which itself moves together with v and belongs to some subspace of $X_{p,2}^{\sigma, \frac{1}{2}+}$, $p \geq 4$. As it turns out, uncovering the final structure of v is slightly more complex but (1.13) conveys the main philosophy¹⁰ (see Sect. 4 for details). Since w does not have a specific multilinear structure, one difficulty is identifying the right space where w will lie. By carefully analyzing the terms that are expected to appear in w , we can specify this space¹¹ to be $X_{p,\infty-}^{\frac{1}{2}, 1-}$.

A final complication comes from the fact that unlike the parabolic setting where the Duhamel operator I automatically gains two derivatives, such gain is not automatic for the Schrödinger equation. Rather, it has to be manually *induced by performing a frequency cut-off also in the Fourier variable of time*, so as to restrict to the region where the parabolic weight in frequency (which is the one that appears in the $X_{p,q}^{s,b}$ norms) is large. In principle this would require that we replace in our ansatz (1.13) the Duhamel operator I by a frequency cut-off version of it, which would introduce non-locality in time which could be incompatible with local in time solutions. Fortunately the frequency cut-off can be substituted by a suitable time convolution

$$\tilde{I}F(t) := \int_0^t \chi(k^2(t-s)) e^{i(t-s)\partial_x^2} F(s) ds \quad (1.14)$$

which has the same effect for some carefully chosen χ . See Sect. 4.1 for details.

With the above discussion, we can now fix the submanifold \mathcal{W} , in which the solution v to (1.5) is uniquely constructed, to be

$$\begin{aligned} \mathcal{W} = \{v \in X_{p,2}^{\sigma, \frac{1}{2}+} : v &= w + \tilde{I} \Pi_{>}^{(1)}(\partial_x \bar{w}, v, v) \\ &+ \tilde{I} \Pi_{>}^{(2)}(\partial_x \bar{w}, w, v), \quad w \in X_{p,\infty-}^{\frac{1}{2}, 1-}\}. \end{aligned} \quad (1.15)$$

We will show that the submanifold \mathcal{W} is well-defined, parametrized by $w \in X_{p,\infty-}^{\frac{1}{2}, 1-}$, and that the trilinear estimates (1.6), which fail for arbitrary input functions in $X_{p,2}^{\sigma, \frac{1}{2}+}$, $p \geq 4$ are actually true for input functions in \mathcal{W} . These together will then allow one to construct the solution $v \in \mathcal{W}$ by a contraction mapping argument. Finally, by inverting the gauge transform, one can construct the solution u to (1.1) in \mathcal{Z} , which is the preimage of \mathcal{W} under the gauge transform. See Sect. 4 for details.

⁹ More precisely, here the solution v is supposed to contain three parts: w , a term paracontrolled by $\partial_x \bar{w}$ and a term paracontrolled by $\partial_x \bar{w} w$ as in (1.13).

¹⁰ Note that the decomposition of v is nonlinear both in w and in the para-controlled terms.

¹¹ More precisely for $\delta > 0$ and $q = \frac{1}{4\delta}$, the right space is $X_{p,q}^{\frac{1}{2}, 1-2\delta}$. See Sect. 3 for precise definitions.

Remark 1.6. Note that in the submanifold \mathcal{W} defined in (1.15), the function $w \in X_{p,\infty}^{\frac{1}{2},1-}$ is not smoother than $v \in X_{p,2}^{\sigma,\frac{1}{2}+}$ in space but rather gains regularity relative to the parabolic weight $|\xi + k^2|$ (see the definition (3.3)) over v . Hence w allows more room to close the estimates for $\tilde{I} \Pi_{>}^{(1)}(\partial_x \bar{w}, v, v)$ and $\tilde{I} \Pi_{>}^{(2)}(\partial_x \bar{w}, w, v)$ when \tilde{I} in principle restricts these two terms to the region where the parabolic weight is large.

Remark 1.7. Nonlinear transformations based on para-products have been used in some other equations, see for example [1, 43, 65]. Our nonlinear transform $v \leftrightarrow w$ is neither a para-linearized normal form nor a gauge transform, and is thus different from the ones appearing in the literature.

Remark 1.8. We conclude this introductory discussion by noting that there is a large body of work that has contributed to our current understanding of the Cauchy problem for (1.1) for data in the Sobolev spaces H^s , $s \geq \frac{1}{2}$ both in the periodic and non-periodic settings; we refer the reader to [3, 16, 17, 37–40, 45, 46, 48, 50, 54, 55, 61–64, 67, 69] and references therein for a more comprehensive treatment.

1.4. Plan of the paper. The paper is organized as follows. In Sect. 2 we recall the periodic gauge transformation used in [27] and perform such transformation to (1.1). Then, we lay out the set up and frequency interactions splitting of the nonlinearities in the gauged derivative Schrödinger equation which will guide our analysis. In Sect. 3 we define and set up our function spaces, prove the main linear estimates and prove an improved divisor bound which is used in some of our estimates. In Sect. 4 we discuss the structure of the solution, identifying the para-controlled terms and the precise solution submanifold \mathcal{W} where v will belong. In Sect. 5 we prove *a priori* bounds for the para-controlled terms. Section 6 constitutes the heart of the paper. Here we find w and prove all the underlying multilinear estimates involved in its construction. In the course of the proof we show in particular that all relevant nonlinearities are well defined as space-time distributions whence the integral equation (1.4) for u will be equivalent to the integral equation formulation of (1.5) (see Sect. 2.4 for details). Finally in Sect. 7 we prove a preservation of regularity result.

1.5. Notations and parameters. We will use the notation

$$\mathbb{P}_0 h = \frac{1}{2\pi} \int_{\mathbb{T}} h \, dx, \quad \mathbb{P}_{\neq 0} h = h - \mathbb{P}_0 h.$$

The space, time and spacetime Fourier transforms are respectively defined as

$$\begin{aligned} \widehat{u}(k) &= \mathcal{F}_x u(k) = \frac{1}{2\pi} \int_{\mathbb{T}} e^{-ikx} u(x) \, dx, \quad \widehat{u}(\xi) = \mathcal{F}_t u(\xi) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\xi t} u(t) \, dt, \\ \widehat{u}(k, \xi) &= \mathcal{F} u(k, \xi) = \frac{1}{(2\pi)^2} \int_{\mathbb{R} \times \mathbb{T}} e^{-i(kx + \xi t)} u(t, x) \, dt \, dx \end{aligned}$$

so \mathcal{F} is reserved for the spacetime Fourier transform. As for \widehat{u} , whether it means space, time or spacetime Fourier transform will be clear from the context. The integral over the set

$$\{(\lambda_1, \dots, \lambda_r) : \lambda_1 \pm \dots \pm \lambda_r = \mu\}$$

for fixed μ will be with respect to the Lebesgue measure $d\lambda_1 \cdots d\lambda_{r-1}$. We denote by $\mathbf{1}_P$ the characteristic function of a property P .

Recall that p_0 is fixed; we will fix a small parameter $0 < \delta \ll 1$ depending on p_0 , and define the other parameters $(b_0, b_1, q_0, q_1, r_0, r_1, r_2)$ as follows:

$$\begin{aligned} b_0 &= 1 - 2\delta, \quad b_1 = 1 - \delta, \quad q_0 = \frac{1}{4\delta}, \quad q_1 = \frac{1}{(4.5)\delta}, \quad \frac{1}{r_0} = \frac{1}{2} + \delta, \\ \frac{1}{r_1} &= \frac{1}{2} + 2\delta, \quad \frac{1}{r_2} = \frac{1}{2} + 3\delta. \end{aligned} \quad (1.16)$$

We also use θ to denote a generic positive quantity that is sufficiently small depending on δ (so θ may have different values at different instances.)

We will fix A as in the statement of Theorem 1.1, and let A_1 be large depending on A , A_2 be large depending on A_1 , etc. All implicit constants below will depend on these A_j 's and the above parameters. The time length T will also be fixed, and small enough depending on these implicit constants.

2. The Gauge Transform and Other Reductions

2.1. The gauge transform. Notice that $\mathbb{P}_0|u|^2$ is conserved under the flow of (1.1). Consider the gauge transform, see [27],

$$v(t, x) = (\mathcal{G}u)(t, x) := (\mathcal{G}_0u)(t, x - 2\mathbb{P}_0|u|^2 \cdot t), \quad (\mathcal{G}_0u)(t, x) := e^{-iG(t, x)} \cdot u(t, x), \quad (2.1)$$

where

$$G = \partial_x^{-1} \mathbb{P}_{\neq 0}(|u|^2) \quad (2.2)$$

is the unique mean-zero antiderivative of $\mathbb{P}_{\neq 0}|u|^2$. This gauge transform is easily inverted, with inverse given by

$$u(t, x) = (\mathcal{G}^{-1}v)(t, x) = e^{iG(t, x)} \cdot v_0(t, x), \quad G = \partial_x^{-1} \mathbb{P}_{\neq 0}|v_0|^2, \quad (2.3)$$

where

$$v_0(t, x) = v(t, x + 2\mathbb{P}_0|v|^2 \cdot t). \quad (2.4)$$

Proposition 2.1. *The maps \mathcal{G} and \mathcal{G}^{-1} are continuous from $C_t^0 H_{p_0}^{\frac{1}{2}}(J)$ to itself for any interval J , and map bounded sets to bounded sets.*

Proof. Notice that $\mathcal{G} = \mathcal{G}_1 \mathcal{G}_0$, where

$$\mathcal{G}_0u = \exp(-i\partial_x^{-1} \mathbb{P}_{\neq 0}|u|^2) \cdot u, \quad \mathcal{G}_0^{-1}u = \exp(i\partial_x^{-1} \mathbb{P}_{\neq 0}|u|^2) \cdot u,$$

and

$$\mathcal{G}_1u(t, x) = u(t, x - 2\mathbb{P}_0|u|^2 \cdot t), \quad \mathcal{G}_1^{-1}u(t, x) = u(t, x + 2\mathbb{P}_0|u|^2 \cdot t).$$

In [27], Lemma 6.2 and Lemma 6.3, it is proved that $\mathcal{G}_0 : H_{p_0}^{\frac{1}{2}} \rightarrow H_{p_0}^{\frac{1}{2}}$ is locally bi-Lipschitz, and that $\mathcal{G}_1 : C_t^0 H_{p_0}^{\frac{1}{2}}(J) \rightarrow C_t^0 H_{p_0}^{\frac{1}{2}}(J)$ is a homeomorphism. Moreover, it is easily checked that

$$\|\mathcal{G}_1 u\|_{C_t^0 H_{p_0}^{\frac{1}{2}}(J)} = \|\mathcal{G}_1^{-1} u\|_{C_t^0 H_{p_0}^{\frac{1}{2}}(J)} = \|u\|_{C_t^0 H_{p_0}^{\frac{1}{2}}(J)},$$

so \mathcal{G} and \mathcal{G}^{-1} map bounded sets to bounded sets. \square

2.2. The transformed equation. We calculate that $v = \mathcal{G}u$ satisfies the equation

$$(\partial_t - i\partial_x^2)v = \mathcal{C}(v, v, v) + \mathcal{Q}(v, \dots, v), \quad (2.5)$$

where the cubic and quintic nonlinearities are defined as

$$\mathcal{F}_x \mathcal{C}(v_1, v_2, v_3)(k) = \sum_{\mathbb{V}_3} k_1 M_3(k, k_1, k_2, k_3) \cdot \overline{\widehat{v}_1(k_1)} \widehat{v}_2(k_2) \widehat{v}_3(k_3), \quad (2.6)$$

and

$$\mathcal{F}_x \mathcal{Q}(v_1, \dots, v_5)(k) = \sum_{\mathbb{V}_5} M_5(k, k_1, \dots, k_5) \cdot \widehat{v}_1(k_1) \overline{\widehat{v}_2(k_2)} \overline{\widehat{v}_3(k_3)} \overline{\widehat{v}_4(k_4)} \widehat{v}_5(k_5). \quad (2.7)$$

The sets \mathbb{V}_3 and \mathbb{V}_5 are defined by

$$\begin{aligned} \mathbb{V}_3 &= \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k_2 + k_3 - k_1 = k, |k_2| \geq |k_3|, k \notin \{k_2, k_3\}\} \cup \{(k, k, k)\}, \\ \mathbb{V}_5 &= \{(k_1, \dots, k_5) \in \mathbb{Z}^5 : k_1 - k_2 + k_3 - k_4 + k_5 = k\}, \end{aligned} \quad (2.8)$$

‘and the coefficients M_j are explicitly defined functions, with $|M_j| \lesssim 1$ for $j \in \{3, 5\}$. They also have the right symmetry so that (2.5) conserves $\mathbb{P}_0|v|^2$. See [27] for the precise formulas.

Remark 2.2. For integers k, k_1, k_2 and k_3 such that $k_2 + k_3 - k_1 = k$, we will rely throughout the proofs on the quantity $\Delta := k^2 + k_1^2 - k_2^2 - k_3^2$.

2.3. Splitting the cubic nonlinearity. We will further split the cubic nonlinearity \mathcal{C} into four parts: a ‘high–high’ part,¹² a ‘low–low’ part, a ‘semilinear’ part and a ‘non-resonant’ part. Decompose \mathbb{V}_3 into four subsets:

$$\begin{aligned} \mathbb{X}_H &= \{(k_1, k_2, k_3) \in \mathbb{V}_3 : |k_3| \geq 2^{-20}|k|\}, \\ \mathbb{X}_L &= \{(k_1, k_2, k_3) \in \mathbb{V}_3 : |k_2| < 2^{-20}|k|\}, \\ \mathbb{X}_S &= \{(k_1, k_2, k_3) \in \mathbb{V}_3 : 2^{-10}|k_1| \leq |k_3| < 2^{-20}|k|\}, \\ \mathbb{X}_N &= \mathbb{V}_3 - (\mathbb{X}_H \cup \mathbb{X}_L \cup \mathbb{X}_S). \end{aligned} \quad (2.9)$$

The following properties of this splitting are elementary and so we omit the proof.

¹² Here ‘high’ and ‘low’ are with respect to the frequencies k_2 and k_3 .

Proposition 2.3. *We have the following properties for the sets \mathbb{X}_* where $* \in \{H, L, S, N\}$:*

- (1) *For $(k_1, k_2, k_3) \in \mathbb{X}_H$ we have $|k_2| \geq |k_3| \geq 2^{-20}|k|$.*
- (2) *For $(k_1, k_2, k_3) \in \mathbb{X}_L$ we have $|k|/2 \leq |k_1| \leq 2|k|$ and $\min(|k|, |k_1|) \geq 2^{18} \max(|k_2|, |k_3|)$.*
- (3) *For $(k_1, k_2, k_3) \in \mathbb{X}_S$ we have $|k|/2 \leq |k_2| \leq 2|k|$, $|k| \geq 2^{20}|k_3|$ and $|k_3| \geq 2^{-10}|k_1|$.*
- (4) *For $(k_1, k_2, k_3) \in \mathbb{X}_N$ we have $|k_2| \geq 2^{-22} \max(|k|, |k_1|)$ and $\min(|k|, |k_1|) \geq 2^{10}|k_3|$.*
- (5) *For $(k_1, k_2, k_3) \in \mathbb{X}_H \cup \mathbb{X}_S$ we have*

$$|k_1| \cdot (\langle k_1 \rangle \langle k_2 \rangle \langle k_3 \rangle)^{-\frac{1}{2}} \lesssim \langle k \rangle^{-\frac{1}{2}}. \quad (2.10)$$

- (6) *For $(k_1, k_2, k_3) \in \mathbb{X}_L \cup \mathbb{X}_N$ we have*

$$|\Delta| \sim \langle k \rangle \langle k_1 \rangle, \quad \text{where } \Delta = k^2 + k_1^2 - k_2^2 - k_3^2 = 2(k - k_2)(k - k_3). \quad (2.11)$$

Remark 2.4. The splitting (2.9), namely $\mathbb{X}_H, \mathbb{X}_L, \mathbb{X}_S, \mathbb{X}_N$ in Proposition 2.3, can be interpreted as a way to compare (k_1, k_2, k_3) with k . In this way, the set \mathbb{X}_L can be understood as a region for the ‘high–low–low’ ($|k| \sim |k_1| \gg |k_2| \geq |k_3|$) interaction. Similarly, $\mathbb{X}_H, \mathbb{X}_S$ and \mathbb{X}_N in principle can be viewed as the regions for ‘high/low–high–high’, ‘low–high–low’ and ‘high–high–low’ interactions. The corresponding trilinear operators restricted to the region \mathbb{X}_* such as \mathcal{C}_* and E_* ($* \in \{H, L, S, N\}$) will be widely used in the following sections.

We also need the following result, which will be used in analyzing the quintic terms in Sect. 6. Once again these properties are elementary. We omit the proof.

Proposition 2.5. *Suppose*

$$k_2 + k' - k_1 = k, \quad (k_1, k_2, k') \in \mathbb{X}_*; \quad k_4 + k_5 - k_3 = k', \quad (k_3, k_4, k_5) \in \mathbb{X}_\#,$$

where $*, \# \in \{H, L, S, N\}$ and where here

$$\Delta = k^2 + k_1^2 - k'^2 - k_2^2 \quad \text{and} \quad \Delta' = k'^2 + k_3^2 - k_5^2 - k_4^2.$$

Moreover let us define,

$$\alpha := \frac{|k_1||k_3|}{\langle \Delta \rangle}, \quad \beta := \frac{|k_1||k_3|}{\langle \Delta \rangle \langle \Delta' \rangle}, \quad \gamma := \frac{|k_1||k_3|}{\langle \Delta' \rangle}.$$

Then we have the followings:

- (1) *Assume $* \in \{H, S\}$ and $\# \in \{L, N\}$. Then either (1a) $|\gamma| \lesssim 1$ or (1b) $* = H$ and $|k_1| \geq 2^{40}|k'|$. In case (1b), (i) if $\# = L$, or if $\# = N$ and $|k_3| \leq 2^{30}|k'|$, then we have that,*

$$|k_1|/2 \leq |k_2| \leq 2|k_1|, \quad |k_1| \geq 2^5 \max_{3 \leq j \leq 5} |k_j|, \quad |\gamma| \lesssim \frac{\langle k_1 \rangle}{\max(\langle k_3 \rangle, \langle k_4 \rangle, \langle k_5 \rangle, \langle k \rangle)},$$

that $k_1 \neq k_2$, and that $\max(|k_3|, |k_4|, |k_5|) = |k_j|$ for $j = 3$ if $\# = L$ and $j \in \{3, 4\}$ if $\# = N$;

(ii) if $\# = N$ and $|k_3| \geq 2^{30}|k'|$, then we have that

$$\begin{aligned} |k_1|/2 &\leq |k_2| \leq 2|k_1|, & |k_3|/2 &\leq |k_4| \leq 2|k_3|, \\ |k_1| &\geq 2^5 \max(|k|, |k_5|), & |\alpha| &\lesssim \frac{\langle k_1 \rangle}{\max(\langle k \rangle, \langle k_5 \rangle, \langle k_1 \pm k_2 \rangle)}. \end{aligned}$$

(2) Assume $*, \# \in \{L, N\}$, then either (2a) $|\alpha| \lesssim 1$ or (2b) $\# = N$ and $|k_3| \geq 2^{40}|k|$.

In case (2b),

(i) if $* = L$ or if $* = N$ and $|k_1| \leq 2^{30}|k|$, then we have

$$|k_3|/2 \leq |k_4| \leq 2|k_3|, \quad |k_3| \geq 2^5 \max_{j \in \{1, 2, 5\}} |k_j|, \quad |\gamma| \lesssim \frac{\langle k_3 \rangle}{\max(\langle k_1 \rangle, \langle k_2 \rangle, \langle k_5 \rangle, \langle k \rangle)},$$

that $k_3 \neq k_4$, and that $\max(|k_1|, |k_2|, |k_5|) = |k_j|$ for $j = 1$ if $* = L$ and, $j \in \{1, 2\}$ if $* = N$;

(ii) if $* = N$ and $|k_1| \geq 2^{30}|k|$, then we have that

$$\begin{aligned} |k_1|/2 &\leq |k_2| \leq 2|k_1|, & |k_3|/2 &\leq |k_4| \leq 2|k_3|, \\ |k_3| &\geq 2^5 \max(|k|, |k_5|), & |\alpha| &\lesssim \frac{\langle k_3 \rangle}{\max(\langle k \rangle, \langle k_5 \rangle)}, \quad |k| \neq |k_5|. \end{aligned}$$

(3) Assume $*, \# \in \{L, N\}$, then we have that

$$|\beta| \lesssim \frac{1}{\langle k \rangle \langle k_3 \pm k_4 \rangle}, \quad \langle k \rangle \gtrsim \langle k_5 \rangle.$$

Now, for $* \in \{H, L, S, N\}$ define \mathcal{C}_* by

$$\mathcal{F}_x \mathcal{C}_*(v_1, v_2, v_3)(k) = \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \cdot \widehat{v_1}(k_1) \widehat{v_2}(k_2) \widehat{v_3}(k_3), \quad (2.12)$$

then we have

$$\mathcal{C} = \mathcal{C}_H + \mathcal{C}_L + \mathcal{C}_S + \mathcal{C}_N. \quad (2.13)$$

2.4. The full setup. When all the relevant nonlinearities are well-defined as spacetime distributions, which we will see in the course of the proof, the integral equation (1.4) for u will be equivalent to the integral equation

$$v(t) = e^{it\partial_x^2} v_0 + I(\mathcal{C}(v, v, v) + \mathcal{Q}(v, \dots, v)), \quad IF(t) = \int_0^t e^{i(t-s)\partial_x^2} F(s) \, ds, \quad (2.14)$$

for v , where the nonlinearities \mathcal{C} and \mathcal{Q} are as in (2.6)–(2.8), and the initial data

$$v_0 = \exp(-i\partial_x^{-1} \mathbb{P}_{\neq 0} |u_0|^2) \cdot u_0, \quad (2.15)$$

which satisfies $\|v_0\|_{H_{p_0}^{\frac{1}{2}}} \leq A_1$ given that $\|u_0\|_{H_{p_0}^{\frac{1}{2}}} \leq A$.

In the proof we will be extending the function v , which is defined on $J = [-T, T]$, to the whole line \mathbb{R}_t ; to this end we fix a smooth function $\varphi(t)$ that is 1 for $|t| \leq 1$ and 0 for $|t| \geq 2$, and define the truncated versions of the linear solution and Duhamel operator

$$\psi(t) = \varphi(t) \cdot e^{it\partial_x^2} v_0, \quad \mathcal{I}F(t) = \varphi(t) \cdot I(\varphi(s) \cdot F(s)). \quad (2.16)$$

For later uses we will also define $\varphi_T(t) = \varphi(T^{-1}t)$.

3. Preparations

In this section we define and set up our function spaces, prove the main linear estimates and prove an improved divisor bound which is used in some of our estimates.

3.1. Function spaces. We begin by properly defining the functions spaces that play a role in our proof. Denote¹³ the Fourier-Lebesgue norms $H_p^s(\mathbb{T})$, where $p \in [2, \infty)$ and $s \in \mathbb{R}$, by

$$\|u\|_{H_p^s} = \|\langle k \rangle^s \widehat{u}(k)\|_{\ell_k^p}. \quad (3.1)$$

In one dimension, H_p^s has the same scaling as the Sobolev space H^γ for

$$\gamma = s + \frac{1}{p} - \frac{1}{2}. \quad (3.2)$$

When $p > 2$ we have $\gamma < s$, which allows the regularity index γ to decrease while keeping $s \geq \frac{1}{2}$.

The associated Fourier restriction norm spaces $X_{p,q}^{s,b}$, where $p \in [2, \infty)$, $q \in [1, \infty)$ and $s, b \in \mathbb{R}$, are then defined by

$$\|u\|_{X_{p,q}^{s,b}} = \|\langle k \rangle^s \langle \xi + k^2 \rangle^b \widehat{u}(k, \xi)\|_{\ell_k^p L_\xi^q}. \quad (3.3)$$

For $2 \leq p_0 < \infty$ fixed and $0 < \delta \ll 1$ small depending on p_0 also fixed, let the parameters $(b_0, b_1, q_0, q_1, r_0, r_1, r_2)$ be defined as in (1.16). We define the four spaces in which the estimates are proved as follows:

$$\begin{aligned} Y_0 &= X_{p_0, r_0}^{\frac{1}{2}, \frac{1}{2}}, & Y_1 &= X_{p_0, r_1}^{\frac{1}{2}, \frac{1}{2}}, \\ Z_0 &= X_{p_0, q_0}^{\frac{1}{2}, b_0}, & Z_1 &= X_{p_0, q_0}^{\frac{1}{2}, b_1}. \end{aligned} \quad (3.4)$$

Note that by Hölder we have

$$X_{p', q'}^{s', b'} \subset X_{p, q}^{s, b}, \quad \text{provided } p \leq p', q \leq q', \quad s + \frac{1}{p} < s' + \frac{1}{p'}, \quad b + \frac{1}{q} < b' + \frac{1}{q'}, \quad (3.5)$$

in particular $Z_0 \subset Y_0 \subset C_t^0 H_{p_0}^{\frac{1}{2}}$. Finally, for any finite interval I and any spacetime norm \mathcal{Y} , define

$$\|u\|_{\mathcal{Y}(I)} = \inf \{\|v\|_{\mathcal{Y}} : v = u \text{ on } I\}. \quad (3.6)$$

¹³ Note that in [27] this same space is denoted by $\widehat{H}_{p'}^s(\mathbb{T})$ where $\frac{1}{p} + \frac{1}{p'} = 1$.

3.2. Linear estimates. We will be using the following notation for a spacetime function F :

$$\tilde{F}(k, \lambda) = \mathcal{X}F(k, \lambda) := \widehat{F}(k, \lambda - k^2), \quad (3.7)$$

where \widehat{F} is the spacetime Fourier transform.

Lemma 3.1. *Define the function*

$$K(\lambda, \sigma) = i \left[\int_{\mathbb{R}} \frac{\widehat{\varphi}(\lambda - \mu)\widehat{\varphi}(\mu - \sigma)}{\mu} d\mu - \widehat{\varphi}(\lambda) \int_{\mathbb{R}} \frac{\widehat{\varphi}(\mu - \sigma)}{\mu} d\mu \right], \quad (3.8)$$

where integrations are defined as principal value limits, then it satisfies

$$|K(\lambda, \sigma)| \lesssim_B \left(\frac{1}{\langle \lambda \rangle^B} + \frac{1}{\langle \lambda - \sigma \rangle^B} \right) \frac{1}{\langle \sigma \rangle} \quad (3.9)$$

for any $B > 0$, and we have

$$\mathcal{I}\tilde{F}(k, \lambda) = \int_{\mathbb{R}} K(\lambda, \sigma) \tilde{F}(k, \sigma) d\sigma. \quad (3.10)$$

Proof. In [21], Lemma 3.3, it is derived that

$$\begin{aligned} \mathcal{I}\tilde{F}(k, \lambda) &= c_0 \int_{\mathbb{R}} \frac{\widehat{\varphi}(\lambda - \mu)}{\mu} d\mu \int_{\mathbb{R}} \widehat{\varphi}(\mu - \sigma) \tilde{F}(k, \sigma) d\sigma \\ &\quad + c_1 \widehat{\varphi}(\lambda) \cdot \int_{\mathbb{R}} \frac{d\mu}{\mu} \int_{\mathbb{R}} \widehat{\varphi}(\mu - \sigma) \tilde{F}(k, \sigma) d\sigma, \end{aligned}$$

where c_0 and c_1 are numerical constants, and integrations are defined as principal value limits. By our convention with Fourier transform, we can calculate that $c_0 = i$ and $c_1 = -i$, which gives the formula (3.8). The bound (3.9) follows easily, using that $\widehat{\varphi}$ is a Schwartz function. \square

Proposition 3.2. *Suppose u is a smooth function such that $u(0) = 0$. Then we have the estimates*

$$\|\varphi_T \cdot u\|_{Y_0} \lesssim T^\theta \|u\|_{Y_1}, \quad \|\varphi_T \cdot u\|_{Z_0} \lesssim T^\theta \|u\|_{Z_1}. \quad (3.11)$$

Proof. First notice that, by (3.5),

$$\|u\|_{X_{\tilde{p}_0, q_1}^{\frac{1}{2}, b_0}} \lesssim \|u\|_{X_{\tilde{p}_0, q_0}^{\frac{1}{2}, b_1}} = \|u\|_{Z_1}.$$

Then, by separating different Fourier modes and conjugating by $e^{\pm itk^2}$ at Fourier mode e^{itkx} , it suffices to prove that for any function $g = g(t)$ satisfying $g(0) = 0$,

$$\|\langle \lambda \rangle^b (\widehat{g} * \widehat{\varphi_T})(\lambda)\|_{L^q} \lesssim T^{\frac{1}{q} - \frac{1}{\tilde{q}}} \|\langle \sigma \rangle^b \widehat{g}(\sigma)\|_{L^{\tilde{q}}}, \quad (3.12)$$

provided $\infty > q > \tilde{q} > 1$ and $b + \frac{1}{\tilde{q}} > 1 > b$. Let $g = g_1 + g_2$ where

$$\widehat{g}_1(\sigma) = \mathbf{1}_{|\sigma| \geq T^{-1}}(\sigma) \widehat{g}(\sigma), \quad \widehat{g}_2(\sigma) = \mathbf{1}_{|\sigma| < T^{-1}}(\sigma) \widehat{g}(\sigma),$$

we will actually prove that

$$\|\langle \lambda \rangle^b (\widehat{g_1} * \widehat{\varphi_T})(\lambda)\|_{L^q} \lesssim T^{\frac{1}{q} - \frac{1}{q}} \|\langle \sigma \rangle^b \widehat{g}(\sigma)\|_{L^{\tilde{q}}}, \quad (3.13)$$

and

$$\|\langle \lambda \rangle^b (\widehat{g_2} * \widehat{\varphi_T})(\lambda)\|_{L^q} \lesssim T^{\frac{1}{q} - \frac{1}{q}} \|\langle \sigma \rangle^b \widehat{g}(\sigma)\|_{L^{\tilde{q}}}. \quad (3.14)$$

To prove (3.13), we can reduce it to the $L^{\tilde{q}} \rightarrow L^q$ bound for the operator

$$\widehat{g}(\sigma) \mapsto \int_{\mathbb{R}} R(\lambda, \sigma) \widehat{g}(\sigma) d\sigma, \quad R(\lambda, \sigma) = \mathbf{1}_{|\sigma| \geq T^{-1}} \cdot T \widehat{\varphi}(T(\lambda - \sigma)) \frac{\langle \lambda \rangle^b}{\langle \sigma \rangle^b}.$$

Since

$$\mathbf{1}_{|\sigma| \geq T^{-1}} \cdot \frac{\langle \lambda \rangle^b}{\langle \sigma \rangle^b} \lesssim \frac{\langle T\lambda \rangle^b}{\langle T\sigma \rangle^b} \lesssim \langle T(\lambda - \sigma) \rangle^b,$$

it follows from Schur's estimate that this $L^{\tilde{q}} \rightarrow L^q$ bound is at most

$$\|T \widehat{\varphi}(T\mu) \langle T\mu \rangle^b\|_{L_{\mu}^{\beta}} \lesssim T^{1 - \frac{1}{\beta}} = T^{\frac{1}{q} - \frac{1}{q}}, \quad \frac{1}{\beta} = 1 + \frac{1}{q} - \frac{1}{\tilde{q}},$$

which proves (3.13).

To prove (3.14), notice that

$$\begin{aligned} (\widehat{g_2} * \widehat{\varphi_T})(\lambda) &= -T \widehat{\varphi}(T\lambda) \int_{|\sigma| \geq T^{-1}} \widehat{g}(\sigma) d\sigma \\ &\quad - \int_{|\sigma| < T^{-1}} T \widehat{g}(\sigma) [\widehat{\varphi}(T\lambda) - \widehat{\varphi}(T(\lambda - \sigma))] d\sigma \end{aligned}$$

since $g(0) = 0$; thus

$$|(\widehat{g_2} * \widehat{\varphi_T})(\lambda)| \lesssim_B T \langle T\lambda \rangle^{-B} \int_{\mathbb{R}} \min(1, |T\sigma|) |\widehat{g}(\sigma)| d\sigma$$

for any $B > 0$. Since by an elementary calculation we can prove

$$\begin{aligned} \int_{\mathbb{R}} \min(1, |T\sigma|) |\widehat{g}(\sigma)| d\sigma &\lesssim \|\langle \sigma \rangle^b \widehat{g}(\sigma)\|_{L^{\tilde{q}}} \cdot \|\min(1, |T\sigma|) \langle \sigma \rangle^{-b}\|_{L^{\gamma}} \\ &\lesssim T^{b - \frac{1}{\gamma}} = T^{b + \frac{1}{q} - 1}, \quad \frac{1}{\gamma} = 1 - \frac{1}{\tilde{q}}, \end{aligned}$$

and that

$$\|T \langle T\lambda \rangle^{-B} \langle \lambda \rangle^b\|_{L^q} \lesssim T^{1 - b - \frac{1}{q}},$$

we deduce (3.14). \square

Remark 3.3. The requirement $u(0) = 0$ is necessary. Below (3.11) will be applied only for those u of form

$$u(t) = \int_0^t (\text{expression}),$$

namely $u = \mathcal{I}(\dots)$ or $u = \mathcal{E}_*^*(\dots)$, see Sect. 4.1 for the definition of \mathcal{E} , so $u(0) = 0$ will always be true.

3.3. *A divisor bound.* Finally, in this subsection we prove and record an improved divisor bound that will be handy later on in some parts of the proof.¹⁴

Lemma 3.4. (1) *Let $\mathcal{R} = \mathbb{Z}$ or $\mathbb{Z}[\omega]$, where $\omega = \exp(2\pi i/3)$, and fix $\varepsilon > 0$. Let $k, q \in \mathcal{R}$ and $\rho > 0$ be such that $|q| \geq |k|^\varepsilon > 0$. Then the number of divisors $r \in \mathcal{R}$ of k that satisfies $|r - q| \leq \rho$ is at most $O_\varepsilon(\rho^\varepsilon)$.*

(2) *Consider the system*

$$\begin{cases} \pm a \pm b \pm c = \text{const.} \\ \mp a^2 \mp b^2 \mp c^2 = \text{const.} \end{cases} \quad (3.15)$$

where the signs are arbitrary, but the signs of $\pm a$ and $\mp a^2$ etc. are always opposite. Assume also that there is no pairing, where a pairing means that (say) $a = b$ and the signs of a and b in (3.15) are opposite. Then the number of solutions that satisfy $|a| \sim N_1$, $|b| \sim N_2$ and $|c| \sim N_3$ is $\lesssim_\varepsilon N^\varepsilon$, where N is the second largest of the N_j 's.

Proof.

(1) It is well-known that \mathcal{R} has unique factorization and satisfies the standard divisor bound: the number of divisors of $k \neq 0$ is at most $O_\varepsilon(|k|^\varepsilon)$. Thus the result is trivial if $\rho \geq |k|^\delta$, where $\delta = \varepsilon^4$. Now suppose $|\rho| \leq |k|^\delta$ (and $|k|$ is large enough), we claim that the number of divisors r is at most $m - 1$, where $m \sim \varepsilon^{-2}$ is an integer.

In fact, suppose d_j , where $1 \leq j \leq m$ are distinct divisors, then by unique factorization we know that k is divisible by $\text{lcm}(d_1, \dots, d_m)$, and hence divisible by

$$\frac{\prod_{j=1}^m d_j}{\prod_{1 \leq i < j \leq m} \gcd(d_i, d_j)} =: k'.$$

On the other hand, since $\gcd(d_i, d_j)$ divides $d_i - d_j$, we know $|\gcd(d_i, d_j)| \leq |d_i - d_j| \leq 2\rho$. As also $|d_j| \gtrsim |k|^\varepsilon$, we conclude that

$$|k| \geq |k'| \gtrsim |k|^{\varepsilon m} |\rho|^{-m^2} \geq |k|^{\varepsilon m - \delta m^2},$$

which is impossible for sufficiently large $|k|$, due to our choices of m and δ .

(2) By symmetry we may assume $N' := \max(N_2, N_3) \leq N_1^{1/10}$ (the other case will follow from the same argument below, using standard divisor bounds). There are three cases to consider: when $b + c - a = \text{const.}$ (and $b^2 + c^2 - a^2 = \text{const.}$), when $a + b - c = \text{const.}$, and when $a + b + c = \text{const.}$

- (a) Suppose $b + c - a = \ell$ is constant and $b^2 + c^2 - a^2$ is also constant. Then $2(\ell - b)(\ell - c) = \ell^2 - (b^2 + c^2 - a^2) =: D$ is also constant, and $D \neq 0$ as there is no pairing. Now choosing $\mathcal{R} = \mathbb{Z}$, $k = D/2$, $q = \ell$ and $\rho \sim N'$ in part (1) yields the result.
- (b) Suppose $a + b - c = \ell$ and $a^2 + b^2 - c^2$ is constant. Then similarly $2(b - c)(\ell - b) =: D$ is a nonzero constant. Considering the divisor $\ell - b$ and choosing $\mathcal{R} = \mathbb{Z}$, $k = D/2$, $q = \ell$ and $\rho \sim N'$ in part (1) yields the result.

¹⁴ We note that a weaker version of Lemma 3.4 here already appeared in [27], Lemma 3.1.

(c) Suppose $a + b + c = \ell$ and $a^2 + b^2 + c^2$ is constant. Then letting $u = 3a - \ell$ and $v = 3b - \ell$, we have that

$$(u - \omega v)(u - \omega^2 v) = u^2 + uv + v^2 = 9(a^2 + b^2 + c^2) - 3\ell^2 =: D$$

is constant. Considering the divisor $u - \omega v$ and choosing $\mathcal{R} = \mathbb{Z}[\omega]$, $k = D$, $q = (\omega + 2)\ell$ and $\rho \sim N'$ in part (1) yields the result. \square

4. Structure of the Solution

In this section we discuss the structure of the solution, identifying the para-controlled terms and the precise solution submanifold \mathcal{W} where v will belong. From now on we will focus on the Eq. (2.14). The submanifold \mathcal{Z} in Theorem 1.1 will be defined as $\mathcal{Z} = \mathcal{G}^{-1}\mathcal{W}$, where \mathcal{G} is the gauge transform (2.1), and \mathcal{W} is a submanifold of $Y_0(J) \subset C_t^0 H_{p_0}^{\frac{1}{2}}(J)$, in which the solution solution v of (2.14) will be constructed. To define \mathcal{W} we need some further preparations.

4.1. Splitting the Duhamel operator. Let $\eta(t)$ be a Schwartz function that satisfies the cancellation condition

$$\widehat{\eta}(1) = 0, \quad \mathcal{H}\widehat{\eta}(1) = 1, \quad (4.1)$$

where \mathcal{H} is the Hilbert transform (principal value convolution by $1/\xi$). For $* \in \{N, L\}$, consider the trilinear operator $E_* := IC_*$. Recall that E_* satisfies that

$$\begin{aligned} \mathcal{F}_x E_*(v_1, v_2, v_3)(k, t) &= \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \int_0^t e^{-i(t-s)k^2} \overline{\widehat{v}_1(k_1, s)} \widehat{v}_2(k_2, s) \widehat{v}_3(k_3, s) ds. \end{aligned} \quad (4.2)$$

As before let $\Delta = k^2 + k_1^2 - k_2^2 - k_3^2$ (we always have $|\Delta| \geq 1$), we will define the modified trilinear operators E_*^Y and E_*^X by

$$\begin{aligned} \mathcal{F}_x E_*^Y(v_1, v_2, v_3)(k, t) &= \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \int_0^t e^{-i(t-s)k^2} \eta(\Delta(t-s)) \overline{\widehat{v}_1(k_1, s)} \widehat{v}_2(k_2, s) \widehat{v}_3(k_3, s) ds \\ & \quad (4.3) \end{aligned}$$

(Y indicates that this term is to be estimated in the Y space) and

$$\begin{aligned} \mathcal{F}_x E_*^X(v_1, v_2, v_3)(k, t) &= \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \\ & \quad \int_0^t e^{-i(t-s)k^2} [1 - \eta(\Delta(t-s))] \overline{\widehat{v}_1(k_1, s)} \widehat{v}_2(k_2, s) \widehat{v}_3(k_3, s) ds \end{aligned} \quad (4.4)$$

(X is for “extra”). Clearly $I\mathcal{C}_* = E_* = E_*^X + E_*^Y$. As with I , we will also define the time truncated versions

$$\begin{aligned}\mathcal{E}_*^Y(v_1, v_2, v_3) &= \varphi(t) \cdot E_*^Y(\varphi(s)v_1, v_2, v_3), \\ \mathcal{E}_*^X(v_1, v_2, v_3) &= \varphi(t) \cdot E_*^X(\varphi(s)v_1, v_2, v_3).\end{aligned}\quad (4.5)$$

Proposition 4.1. *For $* \in \{N, L\}$ we have the expressions:*

$$\begin{aligned}\mathcal{X}\mathcal{E}_*^Y(v_1, v_2, v_3)(k, \lambda) &= \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \\ &\quad \int_{\mathbb{R}} K_{\Delta}^Y(\lambda, \sigma) d\sigma \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} \overline{\tilde{v}_1(k_1, \lambda_1)} \tilde{v}_2(k_2, \lambda_2) \tilde{v}_3(k_3, \lambda_3), \\ \mathcal{X}\mathcal{E}_*^X(v_1, v_2, v_3)(k, \lambda) &= \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \\ &\quad \int_{\mathbb{R}} K_{\Delta}^X(\lambda, \sigma) d\sigma \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} \overline{\tilde{v}_1(k_1, \lambda_1)} \tilde{v}_2(k_2, \lambda_2) \tilde{v}_3(k_3, \lambda_3),\end{aligned}\quad (4.6)$$

where $\Delta = k^2 + k_1^2 - k_2^2 - k_3^2$, and the functions K_{Δ}^Y and K_{Δ}^X satisfy the bounds

$$|K_{\Delta}^Y(\lambda, \sigma)| \lesssim_B \frac{1}{\langle \lambda - \sigma \rangle^B} \min\left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \sigma \rangle}\right) + \frac{1}{\langle \lambda - \sigma \rangle} \min\left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle}\right), \quad (4.7)$$

$$\begin{aligned}|K_{\Delta}^X(\lambda, \sigma)| &\lesssim_B \frac{1}{\langle \lambda \rangle^B \langle \sigma \rangle} + \frac{\langle \sigma - \Delta \rangle}{\langle \lambda - \sigma \rangle^B \langle \sigma \rangle} \min\left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \sigma \rangle}\right) \\ &\quad + \frac{\langle \lambda - \Delta \rangle}{\langle \lambda - \sigma \rangle} \min\left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle}\right)^2.\end{aligned}\quad (4.8)$$

Proof. Fix $* \in \{N, L\}$. Let K_{Δ}^Y be the integral kernel of the linear operator

$$F(s) \mapsto \int_0^t \eta(\Delta(t-s)) F(s) ds$$

on the Fourier side, i.e.

$$\mathcal{F}_t \left(\int_0^t \eta(\Delta(t-s)) F(s) ds \right) (\lambda) = \int_{\mathbb{R}} K_{\Delta}^Y(\lambda, \sigma) \widehat{F}(\sigma) d\sigma,$$

and $K_{\Delta}^X = K - K_{\Delta}^Y$ where K is defined in (3.8). By making Fourier expansion in x twisting by $e^{\pm itk^2}$ on the time-Fourier side at mode k , one can see that (4.6) holds with exactly the same kernels K_{Δ}^Y and K_{Δ}^X .

It then suffices to calculate these kernels; by an argument similar to [21], Lemma 3.3, we have

$$\begin{aligned}K_{\Delta}^Y(\lambda, \sigma) &= i \int_{\mathbb{R}} \widehat{\varphi}(\lambda - \mu) \widehat{\varphi}(\mu - \sigma) \frac{1}{\Delta} (\mathcal{H} \widehat{\eta}) \left(\frac{\mu}{\Delta} \right) d\mu \\ &\quad - i \int_{\mathbb{R}} \widehat{\varphi}(\lambda - \mu) \frac{1}{\Delta} \widehat{\eta} \left(\frac{\mu}{\Delta} \right) (\mathcal{H} \widehat{\varphi})(\mu - \sigma) d\mu.\end{aligned}\quad (4.9)$$

Since $\widehat{\eta}$ and $\widehat{\varphi}$ are Schwartz functions, their Hilbert transforms will decay like $\langle \lambda \rangle^{-1}$, thus

$$\left| \frac{1}{\Delta} (\mathcal{H}\widehat{\eta}) \left(\frac{\mu}{\Delta} \right) \right| \lesssim \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \mu \rangle} \right), \quad |(\mathcal{H}\widehat{\varphi})(\mu - \sigma)| \lesssim \frac{1}{\langle \mu - \sigma \rangle}.$$

Then, by elementary estimates of the integral, the first term on the right hand side of (4.9) is bounded by the first term on the right hand side of (4.7), and the second term on the right hand side of (4.9) is bounded by the second term on the right hand side of (4.7).

As with K_Δ^X , using (4.9) and (3.8) we can calculate

$$\begin{aligned} K_\Delta^X(\lambda, \sigma) &= i \int_{\mathbb{R}} \widehat{\varphi}(\lambda - \mu) \widehat{\varphi}(\mu - \sigma) \left[\frac{1}{\mu} - \frac{1}{\Delta} (\mathcal{H}\widehat{\eta}) \left(\frac{\mu}{\Delta} \right) \right] d\mu \\ &+ i \int_{\mathbb{R}} \widehat{\varphi}(\lambda - \mu) \frac{1}{\Delta} \widehat{\eta} \left(\frac{\mu}{\Delta} \right) (\mathcal{H}\widehat{\varphi})(\mu - \sigma) d\mu - i \widehat{\varphi}(\lambda) \int_{\mathbb{R}} \frac{\widehat{\varphi}(\mu - \sigma)}{\mu} d\mu. \end{aligned} \quad (4.10)$$

The third term on the right hand side of (4.10) is bounded by the first term on the right hand side of (4.8). The first term on the right hand side of (4.10) can be bounded by the second term on the right hand side of (4.8), once we can prove

$$\left| \frac{1}{\mu} - \frac{1}{\Delta} (\mathcal{H}\widehat{\eta}) \left(\frac{\mu}{\Delta} \right) \right| \lesssim \frac{\langle \mu - \Delta \rangle}{\langle \mu \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \mu \rangle} \right)$$

for $|\mu| \geq 1$, but this follows from rescaling and the assumption $\mathcal{H}\widehat{\eta}(1) = 1$. Similarly, the second term on the right hand side of (4.10) can be bounded by the third term on the right hand side of (4.8), due to the estimate

$$\left| \frac{1}{\Delta} \widehat{\eta} \left(\frac{\mu}{\Delta} \right) \right| \lesssim \langle \mu - \Delta \rangle \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \mu \rangle} \right)^2$$

and the fact that $\widehat{\eta}(1) = 0$. \square

Remark 4.2. Note that the first term on the right hand side of (4.7) is bounded by the second term, so we have

$$|K_\Delta^Y| \lesssim \frac{1}{\langle \lambda - \sigma \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right). \quad (4.11)$$

Moreover, by (4.8) we can write $K_\Delta^X = K_\Delta^{X,0} + K_\Delta^{X,+}$, where

$$|K_\Delta^{X,0}| \lesssim \mathbf{1}_{\langle \sigma \rangle \gtrsim \langle \Delta \rangle} \frac{1}{\langle \lambda \rangle^B \langle \Delta \rangle} + \mathbf{1}_{\langle \lambda - \sigma \rangle \gtrsim \langle \sigma - \Delta \rangle} \cdot \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right)^2, \quad (4.12)$$

$$\begin{aligned} |K_\Delta^{X,+}| &\lesssim_B \mathbf{1}_{\langle \sigma \rangle \ll \langle \Delta \rangle} \frac{1}{\langle \lambda \rangle^B \langle \sigma \rangle} + \frac{\langle \sigma - \Delta \rangle}{\langle \lambda - \sigma \rangle^B \langle \sigma \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \sigma \rangle} \right) \\ &+ \mathbf{1}_{\langle \lambda - \sigma \rangle \ll \langle \sigma - \Delta \rangle} \frac{\langle \sigma - \Delta \rangle}{\langle \lambda - \sigma \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right)^2. \end{aligned} \quad (4.13)$$

We will define the terms $\mathcal{E}_*^{X,0}$ and $\mathcal{E}_*^{X,+}$ accordingly, for $* \in \{N, L\}$.

4.2. *The submanifold \mathcal{W} .* We can now define \mathcal{W} as follows.

Definition 4.3. Suppose that A_2, A_3 and T are fixed. Let $J = [-T, T]$. We define

$$\begin{aligned} \mathcal{W} = \{v \in Y_0(J) : & \|v\|_{Y_0(J)} \leq A_3, \text{ and there exists } w \text{ with } \|w\|_{Z_0(J)} \leq A_2, \\ & \text{such that } v = w + E_N^Y(w, w, v) + E_L^Y(w, v, v)\}. \end{aligned} \quad (4.14)$$

This is a submanifold of $Y_0(J) \subset C_t^0 H_{p_0}^{\frac{1}{2}}(J)$. Moreover, we will define the submanifold \mathcal{Z} of $C_t^0 H_{p_0}^{\frac{1}{2}}(J)$ in the statement of Theorem 1.1 by $\mathcal{Z} = \mathcal{G}^{-1}\mathcal{W}$.

We will need the following proposition, whose proof is postponed to Sect. 5.

Proposition 4.4. *For any w which satisfies $\|w\|_{Z_0(J)} \leq A_2$ there is a unique v satisfying $\|v\|_{Y_0(J)} \leq A_3$, such that*

$$v = w + E_N^Y(w, w, v) + E_L^Y(w, v, v). \quad (4.15)$$

This mapping $w \mapsto v = v[w]$ is Lipschitz from the A_2 -ball of $Z_0(J)$ to the A_3 -ball of $Y_0(J)$. The submanifold \mathcal{W} of $Y_0(J)$ is the image of this mapping.

Remark 4.5. Based on Remark 2.4, E_L can be understood as the ‘high–low–low’ ($|k_1| \gg |k_2| \geq |k_3|$) interaction and E_L^Y in principle is E_L restricted¹⁵ to the region where the parabolic weight in frequency is large ($|\xi + k^2| \gtrsim k^2$), as in (1.14) and Sect. 4.1. Similarly E_N^Y can be understood as the ‘high–high–low’ ($|k_2| \gtrsim |k_1| \gg |k_3|$) interaction with the restriction that the parabolic weight is large. Now it is clear that $E_N^Y(w, w, v)$ and $E_L^Y(w, v, v)$ in the ansatz (4.15) can be viewed as two para-controlled terms where w is located at the comparable ‘high’ inputs of the trilinear operator and v is located at the comparable ‘low’ inputs.

4.3. *Reducing to an equation for w .* The next step is to reduce (2.14) to an equation for w . We will construct a function w satisfying $\|w\|_{Z_0(J)} \leq A_2$, such that the function $v = v[w]$ defined by Proposition 4.4 satisfies (2.14). By direct calculation, we see that (2.14) reduces to

$$\begin{aligned} w = e^{it\partial_x^2} v_0 + I\mathcal{Q}(v, \dots, v) + I\mathcal{C}_H(v, v, v) + I\mathcal{C}_S(v, v, v) \\ + I(\mathcal{C}_N(v, v, v) - \mathcal{C}_N(w, w, v)) + I(\mathcal{C}_L(v, v, v) - \mathcal{C}_L(w, v, v)) \\ + E_N^X(w, w, v) + E_L^X(w, v, v) \end{aligned} \quad (4.16)$$

where $v = v[w]$ (we will always assume this below) and satisfies

$$v = w + E_N^Y(w, w, v) + E_L^Y(w, v, v). \quad (4.17)$$

It is now clear that Theorem 1.1 will be a consequence of the following result.

Proposition 4.6. *The mapping that maps w to the right hand side of (4.16) is a contraction mapping from the A_2 -ball of $Z_0(J)$ to itself.*

This proposition will be proved in Sect. 6.

¹⁵ The difference between E_L and E_L^Y —same for E_N —which we call E_L^X as defined in (4.4) will appear in the Eq. (4.16) for w .

5. Proof of Proposition 4.4

In this section we prove *a priori* bounds for the para-controlled terms which will crucially enter in the next section. We start by noting that $Z_0(J) \subset Y_0(J)$. In order to prove Proposition 4.4, it suffices to prove the trilinear estimates

$$\|E_N^Y(v_1, v_2, v_3)\|_{Y_0(J)} \lesssim T^\theta \|v_1\|_{Z_0(J)} \|v_2\|_{Z_0(J)} \|v_3\|_{Y_0(J)}, \quad (5.1)$$

$$\|E_L^Y(v_1, v_2, v_3)\|_{Y_0(J)} \lesssim T^\theta \|v_1\|_{Z_0(J)} \|v_2\|_{Y_0(J)} \|v_3\|_{Y_0(J)}. \quad (5.2)$$

In fact, these would imply that given w which satisfies $\|w\|_{Z_0(J)} \leq A_2$, the mapping

$$v \mapsto w + E_N^Y(w, w, v) + E_L^Y(w, v, v)$$

is a contraction mapping from the A_3 -ball of $Y_0(J)$ to itself. It then has a unique fixed point $v = v[w]$, and the Lipschitz property of the mapping $w \mapsto v$ is also easily checked.

In order to prove (5.1) and (5.2), we will assume that w^+ and v^+ are extensions of w and v respectively, such that $\|w^+\|_{Z_0} \leq 2A_2$ and $\|v\|_{Y_0} \leq 2A_3$. Recall that $\varphi_T(t) = \varphi(T^{-1}t)$, clearly $\varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+)$ and $\varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+)$ are extensions of $E_N^Y(w, w, v)$ and $E_L^Y(w, v, v)$ respectively. Using also Proposition 3.2, we can reduce Proposition 4.4 to the following

Proposition 5.1. *We have the following bounds*

$$\|\mathcal{E}_N^Y(v_1, v_2, v_3)\|_{Y_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Z_0} \|v_3\|_{Y_0}, \quad (5.3)$$

$$\|\mathcal{E}_L^Y(v_1, v_2, v_3)\|_{Y_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Y_0} \|v_3\|_{Y_0}. \quad (5.4)$$

Proof. Let $* \in \{N, L\}$, using the embedding $Z_0 \subset Y_0$, we only need to prove the stronger result

$$\|\mathcal{E}_*^Y(v_1, v_2, v_3)\|_{Y_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Y_0} \|v_3\|_{Y_0}. \quad (5.5)$$

Let $\mathcal{E} = \mathcal{E}_*^Y(v_1, v_2, v_3)$, we may assume the norms on the right hand side are all equal to 1. Recall from (4.6) and (4.11) that

$$|\tilde{\mathcal{E}}(k, \lambda)| \lesssim \sum_{\mathbb{X}_*} |k_1| \min\left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle}\right) \int_{\lambda_2 + \lambda_3 + \lambda_4 - \lambda_1 = \lambda - \Delta} \frac{1}{\langle \lambda_4 \rangle} \prod_{j=1}^3 |\tilde{v}_j(k_j, \lambda_j)|, \quad (5.6)$$

where $\Delta = 2(k - k_2)(k - k_3)$ by (2.11). We may restrict to the dyadic region $\langle k_2 \rangle \sim N_2$ and $\langle k_3 \rangle \sim N_3$ (so $N_2 \gtrsim N_3$), where N_2 and N_3 are powers of two.

Recall that $\frac{1}{r_2} = \frac{1}{2} + 3\delta$ (so $r_2 < r_0$). Notice that

$$\|\langle k_1 \rangle^{\frac{1}{2}} \tilde{v}_1\|_{L_\lambda^1 \ell_k^{r_0}} \lesssim \|\langle k_1 \rangle^{\frac{1}{2}} \langle \lambda_1 \rangle^{b_0} \tilde{v}_1\|_{L_\lambda^{q_0} \ell_k^{r_0}} \lesssim \|\langle k_1 \rangle^{\frac{1}{2}} \langle \lambda_1 \rangle^{b_0} \tilde{v}_1\|_{\ell_k^{r_0} L_\lambda^{q_0}} \lesssim 1$$

by Hölder and Minkowski, and similarly

$$\begin{aligned} \|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \tilde{v}_j\|_{L_\lambda^1 \ell_k^{r_2}} &\lesssim \|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \langle \lambda_j \rangle^{\frac{1}{2}} \tilde{v}_j\|_{L_\lambda^{r_0} \ell_k^{r_2}} \\ &\lesssim \|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \langle \lambda_j \rangle^{\frac{1}{2}} \tilde{v}_j\|_{\ell_k^{r_2} L_\lambda^{r_0}} \lesssim \|\langle k_j \rangle^{\frac{1}{2}} \langle \lambda_j \rangle^{\frac{1}{2}} \tilde{v}_j\|_{\ell_k^{r_0} L_\lambda^{r_0}} \lesssim 1 \end{aligned}$$

for $j \in \{2, 3\}$. We may then fix $(\lambda_1, \lambda_2, \lambda_3)$ which we eventually integrate over, and denote

$$|\tilde{v}_1(k_1, \lambda_1)| = \langle k_1 \rangle^{-\frac{1}{2}} f_1(k_1), \quad |\tilde{v}_j(k_j, \lambda_j)| = N_j^{-(1-\sqrt{\delta})/p_0} f_j(k_j), \quad 2 \leq j \leq 3,$$

where (after a further normalization)

$$\|f_1\|_{\ell_k^{p_0}} \lesssim 1, \quad \|f_j\|_{\ell_k^{r_2}} \lesssim 1 \quad (k = 2, 3), \quad (5.7)$$

and it will suffice to prove that for any fixed $\mu (= \lambda_2 + \lambda_3 - \lambda_1) \in \mathbb{R}$,

$$\begin{aligned} & \left\| \langle k \rangle^{\frac{1}{2}} \langle \lambda \rangle^{\frac{1}{2}} \sum_{\mathbb{X}_*} \langle k_1 \rangle^{\frac{1}{2}} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right) \frac{1}{\langle \lambda - \Delta - \mu \rangle} \prod_{j=1}^3 f_j(k_j) \right\|_{\ell_k^{p_0} L_\lambda^{r_1}} \\ & \lesssim (N_2 N_3)^{(1-\sqrt{\delta})/p_0 - \theta}. \end{aligned} \quad (5.8)$$

In the above summation over $(k_1, k_2, k_3) \in \mathbb{X}_*$, we may first fix Δ and sum over $(k_1, k_2, k_3) \in \mathbb{X}_*$ that correspond to this fixed Δ .

We first assume $* = L$, which is the slightly harder case. Note that $\langle k \rangle \sim \langle k_1 \rangle$, by Lemma 3.4, we can bound the left hand side of (5.8) by

$$\left\| \langle k \rangle \langle \lambda \rangle^{\frac{1}{2}} \sum_{\Delta} F(k, \Delta) \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right) \frac{1}{\langle \lambda - \Delta - \mu \rangle} \right\|_{\ell_k^{p_0} L_\lambda^{r_1}}, \quad (5.9)$$

where

$$F(k, \Delta) = \sum_{\substack{(k_1, k_2, k_3) \in \mathbb{X}_L \\ k^2 + k_1^2 - k_2^2 - k_3^2 = \Delta}} \prod_{j=1}^3 f_j(k_j) \lesssim N_2^\theta \left(\sum_{\substack{(k_1, k_2, k_3) \in \mathbb{X}_L \\ k^2 + k_1^2 - k_2^2 - k_3^2 = \Delta}} \prod_{j=1}^3 f_j(k_j)^{r_2} \right)^{\frac{1}{r_2}}.$$

Using the facts that

$$\min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right) \lesssim \langle \Delta \rangle^{-\frac{1}{2}} \langle \lambda \rangle^{-\frac{1}{2}}, \quad \langle \Delta \rangle \sim \langle k \rangle^2$$

and by Schur's estimate, we can bound

$$\left\| \langle k \rangle \langle \lambda \rangle^{\frac{1}{2}} \sum_{\Delta} F(k, \Delta) \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right) \frac{1}{\langle \lambda - \Delta - \mu \rangle} \right\|_{L_\lambda^{r_1}} \lesssim \|F(k, \Delta)\|_{\ell_\Delta^{r_2}}$$

for each fixed k . By the definition of $F(k, \Delta)$, it then suffices to prove that

$$N_2^\theta \left\| \left(\sum_{\mathbb{X}_L} \prod_{j=1}^3 f_j(k_j)^{r_2} \right)^{\frac{1}{r_2}} \right\|_{\ell_k^{p_0}} \lesssim (N_2 N_3)^{(1-\sqrt{\delta})/p_0} N_2^{-\theta}. \quad (5.10)$$

Let $f_j(k_j)^{r_2} = g_j(k_j)$ and $\beta = (p_0/r_2)$, it suffices to prove (for a possibly different θ) that

$$\left\| \sum_{\substack{|k_2| \sim N_2 \\ |k_3| \sim N_3}} g_1(k_2 + k_3 - k) g_2(k_2) g_3(k_3) \right\|_{\ell_k^\beta} \lesssim (N_2 N_3)^{r_2(1-\sqrt{\delta})/p_0} N_2^{-\theta}.$$

As $\|g_1\|_{\ell_k^\beta} = \|f_1\|_{\ell_k^{p_0}}^{r_2} \lesssim 1$, by Minkowski we can bound the above by

$$\|g_2\|_{\ell_k^1} \|g_3\|_{\ell_k^1} = \|f_2\|_{\ell_k^{r_2}}^{r_2} \|f_3\|_{\ell_k^{r_2}}^{r_2} \lesssim 1 \quad (5.11)$$

using (5.7). This finishes the case $* = L$.

When $* = N$, we will further assume $\langle k \rangle \sim N_0$ and $\langle k_1 \rangle \sim N_1$, then the proof will be the same as above, using the fact that

$$\langle k \rangle^{\frac{1}{2}} \langle k_1 \rangle^{\frac{1}{2}} \sim N_0 N_1 \sim \langle \Delta \rangle^{\frac{1}{2}}.$$

The sum over N_0 and N_1 is then taken care of using the positive power of N_2 on the right hand side of (5.10), and the fact that $N_2 \gtrsim \max(N_0, N_1)$ when $(k_1, k_2, k_3) \in \mathbb{X}_N$. \square

6. Proof of Proposition 4.6

This section constitutes the heart of the paper. Here we find w and prove all the underlying multilinear estimates involved in its construction. In the course of the proof we show in particular that all relevant nonlinearities are well defined as space-time distributions whence the integral equation (1.4) for u will be equivalent to the integral equation formulation of (1.5) from Sect. 2.4.

Given w satisfying $\|w\|_{Z_0(J)} \leq A_2$, let w^+ be an extension of w such that $\|w^+\|_{Z_0} \leq 2A_2$. By the proof of Proposition 4.4 in Sect. 5, we know that there is a unique $v^+ = v^+[w^+]$ such that $\|v^+\|_{Y_0} \leq A_3$, and

$$v^+ = w^+ + \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+) + \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+). \quad (6.1)$$

Moreover this v^+ is an extension of $v = v[w]$. Therefore, recall that $\psi := \varphi(t) e^{it\partial_x^2} v_0$, the function

$$\begin{aligned} z := & \psi + \varphi_T \cdot \mathcal{I}\mathcal{Q}(v^+, \dots, v^+) + \varphi_T \cdot \mathcal{I}\mathcal{C}_H(v^+, v^+, v^+) + \varphi_T \cdot \mathcal{I}\mathcal{C}_S(v^+, v^+, v^+) \\ & + \varphi_T \cdot \mathcal{I}(\mathcal{C}_N(v^+, v^+, v^+) - \mathcal{C}_N(w^+, w^+, v^+)) \\ & + \varphi_T \cdot \mathcal{I}(\mathcal{C}_L(v^+, v^+, v^+) - \mathcal{C}_L(w^+, v^+, v^+)) \\ & + \varphi_T \cdot \mathcal{E}_N^X(w^+, w^+, v^+) + \varphi_T \cdot \mathcal{E}_L^X(w^+, v^+, v^+) \end{aligned} \quad (6.2)$$

will be an extension of the right hand side of (4.16).

6.1. Splitting the formula of z . Now that w^+ , v^+ and z are defined for all time, we can further manipulate the expression of z , as this manipulation sometimes requires inserting time-frequency cutoffs. We will analyze each term in (6.2) separately. The initial data term ψ is trivial. For the other terms, we will remove the φ_T factor in front, and bound the corresponding terms in the stronger space Z_1 ; Proposition 3.2 then allows us to gain a factor T^θ which provides the required smallness.

(1) The term $\mathcal{I}\mathcal{Q}(v^+, \dots, v^+)$. This is a single term, we will name it

$$z_{11} = \mathcal{I}\mathcal{Q}(v^+, \dots, v^+). \quad (6.3)$$

(2) The term $\varphi_T \cdot \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(v^+, v^+, v^+)$. Here decomposing v^+ by (6.1), we can obtain the following terms

$$\begin{aligned} z_{21} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(w^+, w^+, w^+), \\ z_{22} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(\varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+), v^+, v^+), \\ z_{23} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(\varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+), v^+, v^+), \\ z_{24} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(w^+, \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+), v^+), \\ z_{25} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(w^+, \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+), v^+), \\ z_{26} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(w^+, w^+, \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+)), \\ z_{27} &= \mathcal{I}(\mathcal{C}_H + \mathcal{C}_S)(w^+, w^+, \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+)). \end{aligned} \quad (6.4)$$

Here z_{21} is a cubic expression, and the others are quintic expressions.

(3) The term $\mathcal{I}(\mathcal{C}_N(v^+, v^+, v^+) - \mathcal{C}_N(w^+, w^+, v^+))$. Similar to (2), we can obtain the terms

$$\begin{aligned} z_{31} &= \mathcal{I}(\mathcal{C}_N(\varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+), v^+, v^+), \\ z_{32} &= \mathcal{I}(\mathcal{C}_N(\varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+), v^+, v^+), \\ z_{33} &= \mathcal{I}(\mathcal{C}_N(w^+, \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+), v^+), \\ z_{34} &= \mathcal{I}(\mathcal{C}_N(w^+, \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+), v^+). \end{aligned} \quad (6.5)$$

They are all quintic expressions.

(4) The term $\mathcal{I}(\mathcal{C}_L(v^+, v^+, v^+) - \mathcal{C}_L(w^+, v^+, v^+))$. In the same way we get two terms

$$\begin{aligned} z_{41} &= \varphi_T \cdot \mathcal{I}(\mathcal{C}_L(\mathcal{E}_N^Y(w^+, w^+, v^+), v^+, v^+), \\ z_{42} &= \mathcal{I}(\mathcal{C}_L(\varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+), v^+, v^+)). \end{aligned} \quad (6.6)$$

They are both quintic expressions.

(5) The term $\mathcal{E}_N^X(w^+, w^+, v^+) + \mathcal{E}_L^X(w^+, v^+, v^+)$. This term requires a little more care. Let $* \in \{N, L\}$, recall that from Proposition 4.1 and Remark 4.2, we have

$$\begin{aligned} \mathcal{X}\mathcal{E}_*^{X,+}(v_1, v_2, v_3)(k, \lambda) &= \sum_{\mathbb{X}_*} k_1 M_3(k, k_1, k_2, k_3) \int_{\mathbb{R}} K_{\Delta}^{X,+}(\lambda, \sigma) d\sigma \\ &\quad \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} \overline{\tilde{v}_1(k_1, \lambda_1)} \tilde{v}_2(k_2, \lambda_2) \tilde{v}_3(k_3, \lambda_3). \end{aligned} \quad (6.7)$$

We may further decompose this expression into

$$\mathcal{E}_*^{X,+} = \mathcal{E}_*^{X,1} + \mathcal{E}_*^{X,2} + \mathcal{E}_*^{X,3}, \quad (6.8)$$

where in $\mathcal{E}_*^{X,j}$ we make the restriction

$$|\lambda_j| = \max_{1 \leq \ell \leq 3} |\lambda_{\ell}|, \quad |\lambda_j| \gtrsim |\sigma - \Delta|. \quad (6.9)$$

Now if $* = N$ and $j = 3$, or $* = L$ and $j \in \{2, 3\}$, we will make decompose the v^+ corresponding to frequency λ_j using (6.1). This gives the following terms

$$\begin{aligned}
z_{51} &= (\mathcal{E}_N^{X,0} + \mathcal{E}_N^{X,1} + \mathcal{E}_N^{X,2})(w^+, w^+, v^+), \\
z_{52} &= \mathcal{E}_N^{X,3}(w^+, w^+, w^+), \\
z_{53} &= \mathcal{E}_N^{X,3}(w^+, w^+, \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+)), \\
z_{54} &= \mathcal{E}_N^{X,3}(w^+, w^+, \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+)), \\
z_{55} &= (\mathcal{E}_L^{X,0} + \mathcal{E}_L^{X,1})(w^+, v^+, v^+), \\
z_{56} &= \mathcal{E}_L^{X,2}(w^+, w^+, v^+), \\
z_{57} &= \mathcal{E}_L^{X,2}(w^+, \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+), v^+), \\
z_{58} &= \mathcal{E}_L^{X,2}(w^+, \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+), v^+), \\
z_{59} &= \mathcal{E}_L^{X,3}(w^+, v^+, w^+), \\
z_{5A} &= \mathcal{E}_L^{X,3}(w^+, v^+, \varphi_T \cdot \mathcal{E}_N^Y(w^+, w^+, v^+)), \\
z_{5B} &= \mathcal{E}_L^{X,3}(w^+, v^+, \varphi_T \cdot \mathcal{E}_L^Y(w^+, v^+, v^+)).
\end{aligned} \tag{6.10}$$

Some of these are cubic expressions, and some of them are quintic.

(6) An operation on quintic terms. Each of the above $z_{j\ell}$'s is a multilinear expression, either cubic or quintic; we will always list its input functions from left to right. Consider now a general quintic term. Let k and k_j , where $1 \leq j \leq 5$, be the (space) frequencies of the output and input functions, then it will involve a summation

$$\sum_{\pm k_1 \dots \pm k_5 = k} (\text{expression}).$$

As with Lemma 3.4, we say a *pairing* (i, j) happens, if $k_i = k_j$ and the signs of k_i and k_j in the expression $\pm k_1 \dots \pm k_5$ are opposite.

For each tuple (k_j) , we will choose an index $i \in \{1, \dots, 5\}$ as follows: if there is no pairing, then let $i \in \{1, \dots, 5\}$ be such that $|k_i|$ is the maximum; if there is a pairing, say $(1, 2)$, and there is no pairing in $\{3, 4, 5\}$, then let $i \in \{3, 4, 5\}$ be such that $|k_i|$ is the maximum; if there is a pairing in $\{3, 4, 5\}$, say $(3, 4)$, then let $i = 5$. It is clear that we always have $|k_i| \gtrsim |k|$.

This procedure then decomposes this quintic term into five parts; once an i is fixed, and if the input function corresponding to this i in this quintic term happens to be v^+ (instead of w^+), we will decompose this v^+ using (6.1), so that this quintic term is decomposed into a quintic and two septic terms.

(7) Summary. Now we have decomposed z into a superposition of multilinear expressions $\underline{z_{j\ell}}$ (including those coming from step (6) above), either cubic or quintic or septic, with input functions being either w^+ or v^+ . Moreover, if we consider two different w and w' , then we may choose extensions w^+ and $(w')^+$ such that

$$\|(w')^+ - w^+\|_{X_{p,q}^{\frac{1}{2},1}} \leq 2\|w' - w\|_{X_{p,q}^{\frac{1}{2},1}(J)}.$$

Let v^+ and $(v')^+$ be defined from w^+ and $(w')^+$ by (6.1), then we also have

$$\|(v')^+ - v^+\|_{Y_p^{\frac{1}{2}}} \lesssim \|(w')^+ - w^+\|_{X_{p,q}^{\frac{1}{2},1}} \lesssim \|w' - w\|_{X_{p,q}^{\frac{1}{2},1}(J)}.$$

Then z and z' , which are defined by (6.2) using w and w' , satisfy that $z - z'$ is an extension of the difference of the right hand sides of (4.16) corresponding to w and w' . Therefore, in order to prove Proposition 4.6, it will suffice to prove the following result.

Proposition 6.1. *All these terms $z_{j\ell}$, including those coming from step (6) above, satisfy the multilinear estimates*

$$\|z_{j\ell}(v_1, \dots, v_r)\|_{Z_1} \lesssim \|v_1\| \cdots \|v_r\|,$$

where $r \in \{3, 5, 7\}$, and for each i , v_i is measured in the Z_0 norm if the corresponding input function in $z_{j\ell}$ is w^+ , and in the Y_0 norm if the input is v^+ . For example the estimate for z_{42} will be

$$\|z_{42}(v_1, \dots, v_5)\|_{Z_1} \lesssim \|v_1\|_{Z_0} \prod_{j=2}^5 \|v_j\|_{Y_0}.$$

Remark 6.2. We make a further remark about the operation in step (6) above. For some quintic terms $z_{j\ell}$ this operation is necessary; for others it is not. However, even in the latter case, performing this operation will not affect the proof: if $z_{j\ell}$ itself satisfies a multilinear estimate where this input function v^+ is measured in the Y_0 norm, then by Propositions 3.2 and 5.1, after decomposing this v^+ using (6.1), the resulting quintic and septic terms will also satisfy the right multilinear estimate. For example, we will see below that

$$\|z_{26}(v_1, \dots, v_5)\|_{Z_1} \lesssim \prod_{j=1}^4 \|v_j\|_{Z_0} \cdot \|v_5\|_{Y_0}.$$

Then, even after performing this operation (with the chosen index $i = 5$) we still have

$$\begin{aligned} \|z_{26}(v_1, \dots, v_5)\|_{Z_1} &\lesssim \prod_{j=1}^4 \|v_j\|_{Z_0} \cdot \|v_5\|_{Y_0}, \\ \|z_{26}(v_1, \dots, v_4, \varphi_T \cdot \mathcal{E}_N^Y(v_5, v_6, v_7))\|_{Z_1} &\lesssim \prod_{j=1}^6 \|v_j\|_{Z_0} \cdot \|v_7\|_{Y_0}, \\ \|z_{26}(v_1, \dots, v_4, \varphi_T \cdot \mathcal{E}_N^Y(v_5, v_6, v_7))\|_{Z_1} &\lesssim \prod_{j=1}^5 \|v_j\|_{Z_0} \cdot \|v_6\|_{Y_0} \|v_7\|_{Y_0}. \end{aligned}$$

The following subsections are devoted to the proof of Proposition 6.1.

6.2. Cubic terms. In this subsection we treat the cubic terms, which are z_{21} and the cubic z_{5*} terms. First we deal with the z_{21} term in the following proposition.

Proposition 6.3. *Let z_{21} be defined as in (6.4). We have the following bound*

$$\|z_{21}(v_1, v_2, v_3)\|_{Z_1} \lesssim \prod_{j=1}^3 \|v_j\|_{Z_0}. \quad (6.11)$$

Proof. Let $*$ $\in \{H, S\}$, we need to show the following bound

$$\|\mathcal{IC}_*(v_1, v_2, v_3)\|_{Z_1} \lesssim \prod_{j=1}^3 \|v_j\|_{Z_0}. \quad (6.12)$$

We may assume the norms on the right hand side are all equal to 1. Recall from (3.9) and (3.10) that for any $B > 0$,

$$|\widetilde{\mathcal{IC}}_*(v_1, v_2, v_3)(k, \lambda)| \lesssim \sum_{\mathbb{X}_*} |k_1| \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} \left(\frac{1}{\langle \lambda \rangle^B} + \frac{1}{\langle \lambda - \sigma \rangle^B} \right) \frac{1}{\langle \sigma \rangle} \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)|, \quad (6.13)$$

where $\Delta = 2(k - k_2)(k - k_3)$ as before.

It will suffice to prove that

$$\left\| \langle k \rangle^{\frac{1}{2}} \langle \lambda \rangle^{b_1} \sum_{\mathbb{X}_*} |k_1| \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} \frac{1}{\langle \lambda \rangle \langle \lambda - \sigma \rangle} \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)| \right\|_{\ell_k^{p_0} L_\lambda^{q_0}} \lesssim 1 \quad (6.14)$$

by the definition of the Y_1 norm (3.4) and the following inequality

$$\left(\frac{1}{\langle \lambda \rangle^B} + \frac{1}{\langle \lambda - \sigma \rangle^B} \right) \frac{1}{\langle \sigma \rangle} \lesssim \frac{1}{\langle \lambda \rangle \langle \lambda - \sigma \rangle} \quad (6.15)$$

for B large enough.

Recall that $b_0 = 1 - 2\delta$ and $1/q_0 = 4\delta$ and hence similarly to the proof of Proposition 5.1 we have for $j \in \{1, 2, 3\}$,

$$\| \langle k_j \rangle^{\frac{1}{2}} \langle \lambda_j \rangle^\delta \widetilde{v}_j \|_{L_\lambda^1 \ell_k^{p_0}} \lesssim \| \langle k_j \rangle^{\frac{1}{2}} \langle \lambda_j \rangle^{b_0} \widetilde{v}_j \|_{L_\lambda^{q_0} \ell_k^{p_0}} \lesssim \| \langle k_j \rangle^{\frac{1}{2}} \langle \lambda_j \rangle^{b_0} \widetilde{v}_j \|_{\ell_k^{p_0} L_\lambda^{q_0}} \lesssim 1$$

by Hölder and Minkowski. We may then fix $(\lambda_1, \lambda_2, \lambda_3)$ which we eventually integrate over, and denoting

$$|\widetilde{v}_j(k_j, \lambda_j)| = \langle k_j \rangle^{-\frac{1}{2}} \langle \lambda_j \rangle^{-\delta} f_j(k_j) \quad (1 \leq j \leq 3),$$

it will suffice to prove that for any fixed $\mu (= \lambda_2 + \lambda_3 - \lambda_1) \in \mathbb{R}$,

$$\left\| \frac{\langle k \rangle^{\frac{1}{2}}}{\langle \lambda \rangle^\delta} \sum_{\mathbb{X}_*} \frac{|k_1|}{\langle k_1 \rangle^{\frac{1}{2}} \langle k_2 \rangle^{\frac{1}{2}} \langle k_3 \rangle^{\frac{1}{2}}} \frac{1}{\langle \lambda - \Delta - \mu \rangle \langle \lambda_1 \rangle^\delta \langle \lambda_2 \rangle^\delta \langle \lambda_3 \rangle^\delta} \prod_{j=1}^3 |f_j| \right\|_{\ell_k^{p_0} L_\lambda^{q_0}} \lesssim 1. \quad (6.16)$$

Then applying the inequality (2.10), we see that it will suffice to prove

$$\left\| \frac{1}{\langle \lambda \rangle^\delta} \sum_{\mathbb{X}_*} \frac{1}{\langle \lambda - \Delta - \mu \rangle \langle \lambda_1 \rangle^\delta \langle \lambda_2 \rangle^\delta \langle \lambda_3 \rangle^\delta} \prod_{j=1}^3 |f_j| \right\|_{\ell_k^{p_0} L_\lambda^{q_0}} \lesssim 1. \quad (6.17)$$

In the above summation over $(k_1, k_2, k_3) \in \mathbb{X}_*$, we again first fix Δ and sum over $(k_1, k_2, k_3) \in \mathbb{X}_*$ corresponding to this fixed Δ . Using the fact that

$$(\langle \lambda_1 \rangle \langle \lambda_3 \rangle \langle \lambda_3 \rangle \langle \lambda \rangle \langle \lambda - \Delta - \mu \rangle)^\delta \gtrsim \langle \Delta \rangle^\delta \quad (6.18)$$

and the divisor bound¹⁶ in Lemma 3.4, we can bound the left hand side of (6.17) by

$$\left\| \sum_{\Delta} \frac{1}{\langle \lambda - \Delta - \mu \rangle^{1-\delta} \langle \Delta \rangle^\delta} F(k, \Delta) \right\|_{\ell_k^{p_0} L_{\lambda}^{q_0}} \quad (6.19)$$

where

$$F(k, \Delta) = \sum_{\substack{(k_1, k_2, k_3) \in \mathbb{X}_* \\ 2(k-k_2)(k-k_3) = \Delta}} \prod_{j=1}^3 |f_j| \lesssim \langle \Delta \rangle^\theta \left(\sum_{\substack{(k_1, k_2, k_3) \in \mathbb{X}_* \\ 2(k-k_2)(k-k_3) = \Delta}} \prod_{j=1}^3 |f_j|^{p_0} \right)^{\frac{1}{p_0}}. \quad (6.20)$$

By choosing p_0 such that $\delta < \frac{1}{5p_0}$ and $\theta < \delta$, and using Schur's estimate we can bound

$$\left\| \sum_{\Delta} \frac{1}{\langle \lambda - \Delta - \mu \rangle^{1-\delta} \langle \Delta \rangle^\delta} F(k, \Delta) \right\|_{L_{\lambda}^{q_0}} \lesssim \left\| \frac{F(k, \Delta)}{\langle \Delta \rangle^\delta} \right\|_{\ell_{\Delta}^{p_0}}. \quad (6.21)$$

Then we sum over k to obtain that

$$\left\| \sum_{\Delta} \frac{1}{\langle \lambda - \Delta - \mu \rangle^{1-\delta} \langle \Delta \rangle^\delta} F(k, \Delta) \right\|_{\ell_k^{p_0} L_{\lambda}^{q_0}} \lesssim \prod_{j=1}^3 \|f_j\|_{\ell_k^{p_0}} \quad (6.22)$$

by (6.20) and (6.21). Finally we integrate (6.22) over $(\lambda_1, \lambda_2, \lambda_3)$ to conclude this proof. \square

Next let us consider the cubic z_{5*} terms (i.e. $z_{51}, z_{52}, z_{55}, z_{56}$ and z_{59}). The following proposition gives the desired bounds for $z_{51}, z_{52}, z_{55}, z_{56}$ and z_{59} in Proposition 6.1.

Proposition 6.4. *For $* \in \{N, L\}$, $j \in \{0, 1, 2, 3\}$ and $\mathcal{E}_*^{X,j}$ defined in (4.6) and (6.8)–(6.9), we have the following bounds.*

(1) *If $j = 0$, we obtain that*

$$\left\| \mathcal{E}_*^{X,0}(v_1, v_2, v_3) \right\|_{Z_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Y_0} \|v_3\|_{Y_0}. \quad (6.23)$$

(2) *If $j = 1$, we obtain that*

$$\left\| \mathcal{E}_*^{X,1}(v_1, v_2, v_3) \right\|_{Z_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Y_0} \|v_3\|_{Y_0}. \quad (6.24)$$

(3) *If $j \in \{2, 3\}$ and $i \in \{2, 3\} - \{j\}$, we obtain that*

$$\left\| \mathcal{E}_*^{X,j}(v_1, v_2, v_3) \right\|_{Z_1} \lesssim \|v_1\|_{Z_0} \|v_j\|_{Z_0} \|v_i\|_{Y_0}. \quad (6.25)$$

¹⁶ The divisor bound applies when $\Delta \neq 0$; however when $\Delta = 0$ we must have $k = k_1 = k_2 = k_3$ by the definition of \mathbb{V}_3 , so the bound is still true.

Proof. Recall from (4.6) that

$$|\widetilde{\mathcal{E}_*^{X,0}}(v_1, v_2, v_3)(k, \lambda)| \lesssim \sum_{\mathbb{X}_*} |k_1| \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} |K_{\Delta}^{X,0}(\lambda, \sigma)| \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)|, \quad (6.26)$$

and for $j \in \{1, 2, 3\}$

$$|\widetilde{\mathcal{E}_*^{X,j}}(v_1, v_2, v_3)(k, \lambda)| \lesssim \sum_{\mathbb{X}_*} |k_1| \int_{\substack{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta \\ |\lambda_j| = \max_{1 \leq \ell \leq 3} |\lambda_\ell|}} |K_{\Delta}^{X,+}(\lambda, \sigma)| \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)|, \quad (6.27)$$

where $\Delta = 2(k - k_2)(k - k_3)$ as before.

(1) Let us consider the case when $j = 0$ and $* \in \{N, L\}$, and then left hand side of the bound (6.23) can be bounded by

$$\left\| \langle k \rangle^{\frac{1}{2}} \langle \lambda \rangle^{b_1} \sum_{\mathbb{X}_*} |k_1| \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma - \Delta} |K_{\Delta}^{X,0}(\lambda, \sigma)| \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)| \right\|_{\ell_k^{p_0} L_{\lambda}^{q_0}}. \quad (6.28)$$

Recall (4.12) whence it will suffice to prove that

$$\begin{aligned} & \left\| \langle k \rangle^{\frac{1}{2}} \langle \lambda \rangle^{b_1} \sum_{\mathbb{X}_*} |k_1| \int_{\lambda_1, \lambda_2, \lambda_3} \frac{1}{\langle \lambda \rangle^{1+4\delta} \langle \Delta \rangle^{1-4\delta}} \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)| \right\|_{\ell_k^{p_0} L_{\lambda}^{q_0}} \\ & \lesssim \|v_1\|_{Z_0} \|v_2\|_{Y_0} \|v_3\|_{Y_0}. \end{aligned} \quad (6.29)$$

To prove (6.29), we first use Minkowski's inequality and integrate over λ to bound the left hand side of (6.29) by

$$\left\| \langle k \rangle^{\frac{1}{2}} \sum_{\mathbb{X}_*} |k_1| \int_{\lambda_1, \lambda_2, \lambda_3} \frac{1}{\langle \Delta \rangle^{1-4\delta}} \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)| \right\|_{\ell_k^{p_0}}. \quad (6.30)$$

Next, we may fix $(\lambda_1, \lambda_2, \lambda_3)$ which we eventually integrate over. In the summation in (6.30) over $(k_1, k_2, k_3) \in \mathbb{X}_*$, we may first fix Δ and sum over $(k_1, k_2, k_3) \in \mathbb{X}_*$ that correspond to this fixed Δ . Moreover, as before we may restrict to the dyadic region $\langle k_3 \rangle \sim N_2$ and $\langle k_2 \rangle \sim N_3$ (so $N_2 \gtrsim N_3$), where N_2 and N_3 are dyadic numbers. Hence we are reduced to bound

$$\left\| \langle k \rangle^{\frac{1}{2}} \sum_{\Delta} |k_1| \frac{1}{\langle \Delta \rangle^{1-4\delta}} \sum_{\substack{(k_1, k_2, k_3) \in \mathbb{X}_* \\ (k_1 - k_2)(k_1 - k_3) = \Delta \\ |k_2| \sim N_2, |k_3| \sim N_3}} \prod_{j=1}^3 |\widetilde{v}_j(k_j, \lambda_j)| \right\|_{\ell_k^{p_0}} \quad (6.31)$$

where $\langle \Delta \rangle \sim \langle k \rangle \langle k_1 \rangle$ and $N_2 \sim |k_2| \gtrsim \max(|k|, |k_1|)$ (by Proposition 2.3). By the divisor bound in Lemma 3.4 and Hölder's inequality as in the proof of Proposition 5.1 we obtain that

$$(6.31) \lesssim \frac{\|\langle k_1 \rangle^{\frac{1}{2}} \tilde{v}_1\|_{\ell_k^{p_0}} \|\langle k_2 \rangle^{\frac{1}{2p_0}} \tilde{v}_2\|_{\ell_k^{r_2}} \|\langle k_3 \rangle^{\frac{1}{2p_0}} \tilde{v}_3\|_{\ell_k^{r_2}}}{N_2^{\frac{1}{2p_0} - 2\delta - \theta} N_3^{\frac{1}{2p_0}}}, \quad (6.32)$$

where r_2 is defined in (1.16). Then we may integrate over $\lambda_1 \lambda_2$ and λ_3 and sum over (N_2, N_3) . By using the negative power of N_2 (suppose $\delta < 1/(4p_0)$) and the following facts (similarly as before):

$$\begin{aligned} \|\langle k_1 \rangle^{\frac{1}{2}} \tilde{v}_1\|_{L_\lambda^1 \ell_k^{p_0}} &\lesssim \|v_1\|_{Z_0}, \quad \|\langle k_2 \rangle^{\frac{1}{2p_0}} \tilde{v}_2\|_{L_\lambda^1 \ell_k^{r_2}} \lesssim \|v_2\|_{Y_0}, \\ \|\langle k_3 \rangle^{\frac{1}{2p_0}} \tilde{v}_3\|_{L_\lambda^1 \ell_k^{r_2}} &\lesssim \|v_3\|_{Y_0}, \end{aligned} \quad (6.33)$$

we conclude the proof of (6.23).

Before proving parts (2) and (3), let us prove an easier bound for $|K_\Delta^{X,+}|$. Suppose $|\lambda_j| = \max_{1 \leq \ell \leq 3} |\lambda_\ell|$ and $|\lambda_j| \gtrsim |\sigma - \Delta|$. Recall that $b_0 = 1 - 2\delta$ and $b_1 = 1 - \delta$, then we obtain that

$$\frac{\langle \lambda \rangle^{b_1}}{\langle \lambda_j \rangle^{b_0 - \delta}} |K_\Delta^{X,+}| \lesssim \frac{1}{\langle \Delta \rangle^{1-6\delta} \langle \sigma - \lambda \rangle}. \quad (6.34)$$

We may then fix the other two λ_ℓ ($\ell \neq j$) and Δ , and integrate over λ_j and λ to obtain the following bound:

$$\left\| \int_{\lambda_j} \frac{1}{\langle \lambda_1 + \lambda - \lambda_2 - \lambda_3 - \Delta \rangle} \left(\langle \lambda_j \rangle^{b_1} \tilde{v}_j(k_j, \lambda_j) \right) \right\|_{L_\lambda^{q_0}} \lesssim \left\| \langle \lambda_j \rangle^{b_0 - \delta} \tilde{v}_j(k_j, \lambda_j) \right\|_{L_{\lambda_j}^{q_1}} \quad (6.35)$$

by Schur's estimate. For $|\lambda_j| = \max_{\ell \in \{1, 2, 3\}} |\lambda_\ell|$, to prove the parts (2) and (3), it will suffice to consider the norm

$$\left\| \langle k \rangle^{\frac{1}{2}} \sum_{\mathbb{X}_*} |k_1| \frac{1}{\langle \Delta \rangle^{1-6\delta}} \left\| \langle \lambda_j \rangle^{b_0 - \delta} \tilde{v}_j(k_j, \lambda_j) \right\|_{L_{\lambda_j}^{q_1}} \prod_{\ell \in \{1, 2, 3\} - \{j\}} \|\tilde{v}_\ell(k_\ell, \lambda_\ell)\|_{L_{\lambda_\ell}^1} \right\|_{\ell_k^{p_0}} \quad (6.36)$$

by (6.34) and (6.35).

(2) Let us consider the case when $j = 1$. By (6.36), it will suffice to bound

$$\left\| \langle k \rangle^{\frac{1}{2}} \sum_{\mathbb{X}_*} \frac{\langle k_1 \rangle^{\frac{1}{2}}}{\langle k_2 \rangle^{\frac{1}{2p_0}} \langle k_3 \rangle^{\frac{1}{2p_0}}} \frac{1}{\langle \Delta \rangle^{1-6\delta}} \prod_{\ell=1}^3 |f_\ell(k_\ell)| \right\|_{\ell_k^{p_0}}, \quad (6.37)$$

where $f_1(k_1) = \langle k_1 \rangle^{\frac{1}{2}} \|\langle \lambda_1 \rangle^{b_0 - \delta} \tilde{v}_1\|_{L_{\lambda_1}^{q_1}}$ and $f_\ell(k_\ell) = \langle k_\ell \rangle^{\frac{1}{2p_0}} \|\tilde{v}_\ell\|_{L_\lambda^1}$ for $\ell = 2, 3$. Similarly to the proof of Proposition 5.1, we also have the bounds:

$$\|f_1\|_{\ell_k^{p_0}} \lesssim \|v_1\|_{Z_0}, \quad \|f_\ell\|_{\ell_k^{r_2}} \lesssim \|v_\ell\|_{Y_0}$$

for $\ell = 2, 3$. We may use dyadic decomposition on (k_2, k_3) and sum over $(k_1, k_2, k_3) \in \mathbb{X}_*$ for fixed Δ and then sum over Δ and k . Following the same proof as in the part (1), the negative power of N_2 help us bound (6.37) by $\|v_1\|_{Z_0} \|v_2\|_{Y_0} \|v_3\|_{Y_0}$, when $\delta < 1/(6p_0)$. This finishes the proof of (6.24).

(3) Let us consider the case when $j \in \{2, 3\}$ and denote by i the other number in $\{2, 3\}$. Just as we did above, it will suffice to bound

$$\left\| \langle k \rangle^{\frac{1}{2}} \sum_{\mathbb{X}_*} \frac{\langle k_1 \rangle^{\frac{1}{2}}}{\langle k_2 \rangle^{\frac{1}{2p_0}} \langle k_3 \rangle^{\frac{1}{2p_0}}} \frac{1}{\langle \Delta \rangle^{1-6\delta}} \prod_{\ell=1}^3 |f_\ell(k_\ell)| \right\|_{\ell_k^{p_0}}, \quad (6.38)$$

where $f_1(k_1) = \langle k_1 \rangle^{\frac{1}{2}} \|\tilde{v}_1\|_{L_\lambda^1}$, $f_j(k_j) = \langle k_j \rangle^{\frac{1}{2p_0}} \|\langle \lambda_j \rangle^{b_0-\delta} \tilde{v}_j\|_{L_\lambda^{q_1}}$ and $f_i(k_i) = \langle k_i \rangle^{\frac{1}{2p_0}} \|\tilde{v}_i\|_{L_\lambda^1}$. Similarly to the proof of Proposition 5.1, we also have the bounds:

$$\|f_\ell\|_{\ell_k^{p_0}} \lesssim \|v_\ell\|_{Z_0}, \quad \|f_i\|_{\ell_k^{r_2}} \lesssim \|v_i\|_{Y_0}$$

for $\ell = 1, j$. Following the same proof of the part (2), (6.38) can be bounded by $\|v_1\|_{Z_0} \|v_j\|_{Z_0} \|v_i\|_{Y_0}$ when $\delta < 1/(6p_0)$. This finishes the proof of (6.25). \square

6.3. The canonical quintic term. The majority of $z_{j\ell}$ are quintic terms; in fact the majority of them can be treated in the same way, using the following estimate.

Proposition 6.5. *Consider a quintic expression \mathcal{R} that satisfies*

$$\begin{aligned} & |\mathcal{X}\mathcal{R}(v_1, \dots, v_5)(k, \lambda)| \\ & \lesssim \sum_{\pm k_1 \pm \dots \pm k_5 = k} \int_{\mathbb{R}} \frac{d\sigma}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \sigma \rangle^{1-\theta}} \int_{\pm \lambda_1 \pm \lambda_2 \pm \dots \pm \lambda_5 = \sigma - \Xi} \prod_{j=1}^5 |\tilde{v}_j(k_j, \lambda_j)|, \end{aligned} \quad (6.39)$$

where $\Xi := k^2 \mp k_1^2 \mp \dots \mp k_5^2$ (the signs are arbitrary, but the signs of $\pm k_j$ and $\mp k_j^2$ are always opposite). Then after the operation in Sect. 6.1, step (6), the resulting terms satisfy the corresponding multilinear estimates. In particular, suppose the chosen index during this operation is $i = 1$, then we have

$$\|\mathcal{R}(v_1, \dots, v_5)\|_{Z_1} \lesssim \|v_1\|_{Z_0} \prod_{j=2}^5 \|v_j\|_{Y_0}, \quad (6.40)$$

$$\|\mathcal{R}(\varphi_T \cdot \mathcal{E}_N^Y(v_1, v_2, v_3), v_4, \dots, v_7)\|_{Z_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Z_0} \prod_{j=3}^7 \|v_j\|_{Y_0}, \quad (6.41)$$

$$\|\mathcal{R}(\varphi_T \cdot \mathcal{E}_L^Y(v_1, v_2, v_3), v_4, \dots, v_7)\|_{Z_1} \lesssim \|v_1\|_{Z_0} \prod_{j=2}^7 \|v_j\|_{Y_0}. \quad (6.42)$$

Proof. (1) We first prove (6.40). Assume all the norms on the right hand side are 1. By a dyadic decomposition, we may restrict to the region where $\langle k_j \rangle \sim N_j$ for $2 \leq j \leq 5$; by symmetry we may assume $N_2 \geq \dots \geq N_5$. As in the proof of Proposition 5.1 we have

$$\|\langle k_1 \rangle^{\frac{1}{2}} \tilde{v}_1\|_{L_\lambda^1 \ell_k^{p_0}} \lesssim 1$$

and

$$\|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \tilde{v}_j\|_{L_\lambda^1 \ell_k^2} \lesssim \|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \tilde{v}_j\|_{L_\lambda^1 \ell_k^{r_2}} \lesssim 1$$

for $2 \leq j \leq 5$. We may then again fix λ_j for $1 \leq j \leq 5$, which we eventually integrate over, and assume

$$f_1(k_1) = \langle k_1 \rangle^{\frac{1}{2}} |\tilde{v}_1(k_1, \lambda_1)|, \quad f_j(k_j) = N_j^{(1-\sqrt{\delta})/p_0} |\tilde{v}_j(k_j, \lambda_j)| \quad (2 \leq j \leq 5),$$

such that (after a further normalization)

$$\|f_1\|_{\ell_k^{p_0}} \leq 1, \quad \|f_j\|_{\ell_k^2} \lesssim 1 \quad (2 \leq j \leq 5). \quad (6.43)$$

Using also that $\langle k \rangle \lesssim \langle k_1 \rangle$, it then suffices to prove that

$$\begin{aligned} & \left\| \langle \lambda \rangle^{b_1} \sum_{\pm k_1 \pm \dots \pm k_5 = k} \int_{\mathbb{R}} \frac{1}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \Xi - \mu \rangle^{1-\theta}} \prod_{j=1}^5 f_j(k_j) \right\|_{\ell_k^{p_0} L_\lambda^{q_0}} \\ & \lesssim N_2^{-\theta} (N_2 N_3 N_4 N_5)^{(1-\sqrt{\delta})/p_0} \end{aligned} \quad (6.44)$$

for any fixed μ (which is a linear combination of λ_j for $1 \leq j \leq 5$). Using the fact that $b_1 < 1 - \theta$ and Schur's estimate, we can bound for fixed k ,

$$\left\| \langle \lambda \rangle^{b_1} \sum_{\pm k_1 \pm \dots \pm k_5 = k} \int_{\mathbb{R}} \frac{1}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \Xi - \mu \rangle^{1-\theta}} \prod_{j=1}^5 f_j(k_j) \right\|_{L_\lambda^{q_0}} \lesssim \|F(k, \Xi)\|_{\ell_\Xi^{q_1}},$$

where

$$F(k, \Xi) = \sum_{\substack{\pm k_1 \dots \pm k_5 = k \\ \pm k_1^2 \dots \pm k_5^2 = \Xi - k^2}} \prod_{j=1}^5 f_j(k_j).$$

As Ξ is determined by (k_2, k_3, k_4, k_5) and hence the number of different Ξ 's does not exceed $O(N_2^4)$, we can bound the $\ell_\Xi^{q_1}$ norm of $F(k, \Xi)$ by $N_2^{\frac{4}{q_1}}$ times its ℓ_Ξ^∞ norm.

(a) Assume $N_3 \geq N_2^{4\sqrt{\delta}}$. For fixed k and Ξ , by assumption we know that either there is no pairing,¹⁷ or there is a pairing, say $(2, 3)$, and there is no pairing in $\{1, 4, 5\}$, or there are two pairings, say $(2, 3)$ and $(4, 5)$. In the first case, for fixed (k_2, k_3) (or (k_4, k_5)), the number of choices for (k_1, k_4, k_5) (or (k_1, k_2, k_3)) is at most $O(N_2^\theta)$ by Lemma 3.4, so using (6.43) we have that

¹⁷ Here we mean pairings as defined in Lemma 3.4.

$$\begin{aligned} F(k, \Xi) &\lesssim \sum_{k_2, k_3} \sum_{k_1, k_4, k_5} f_1(k_1) f_2(k_2)^2 f_3(k_3)^2 \\ &+ \sum_{k_4, k_5} \sum_{k_1, k_2, k_3} f_1(k_1) f_4(k_4)^2 f_5(k_5)^2 \lesssim N_2^\theta \sup_{|k_1 - k| \lesssim N_2} f_1(k_1). \end{aligned}$$

In the other two cases this estimate can be similarly established. This gives

$$\|F(k, \Xi)\|_{\ell_\Xi^\infty} \lesssim N_2^\theta \left(\sum_{|k_1 - k| \lesssim N_2} |f_1(k_1)|^{p_0} \right)^{\frac{1}{p_0}},$$

and hence the left hand side of (6.44) is bounded by

$$N_2^{\theta + \frac{4}{q_1}} \left\| \left(\sum_{|k_1 - k| \lesssim N_2} |f_1(k_1)|^{p_0} \right)^{\frac{1}{p_0}} \right\|_{\ell_k^{p_0}} \lesssim N_2^{\theta + \frac{4}{q_1} + \frac{1}{p_0}}$$

using (6.43). As $N_3 \geq N_2^{4\sqrt{\delta}}$, $4/q_1 = O(\delta)$, δ is small enough depending on p_0 and θ is small enough depending on δ , this implies (6.44).

(b) Assume $N_3 \leq N_2^{4\sqrt{\delta}}$. Then in estimating $F(k, \Xi)$, we may fix the choices of (k_3, k_4, k_5) and eventually sum over them. In this process we lost at most $N_2^{O(\sqrt{\delta})}$. Then, with (k_3, k_4, k_5) fixed, we know that (k_1, k_2) is uniquely determined by k and Ξ , since by assumption (1, 2) cannot be a pairing. Thus, with (k_3, k_4, k_5) fixed, we have

$$|F(k, \Xi)| \lesssim \sup_{k_2} f_1(k - \ell - k_2) f_2(k_2),$$

where ℓ is a linear combination of (k_3, k_4, k_5) . As

$$\begin{aligned} &\left\| \sup_{k_2} f_1(k - \ell - k_2) f_2(k_2) \right\|_{\ell_k^{p_0}} \\ &\lesssim \left\| \left(\sum_{k_2} f_1(k - \ell - k_2)^{p_0} f_2(k_2)^{p_0} \right)^{\frac{1}{p_0}} \right\|_{\ell_k^{p_0}} = \|f_1\|_{\ell_k^{p_0}} \|f_2\|_{\ell_k^{p_0}} \end{aligned}$$

and $\|f_2\|_{\ell_k^{p_0}} \lesssim \|f_2\|_{\ell_k^2} \lesssim 1$, this bounds the left hand side of (6.44) by $N_2^{\frac{4}{q_1} + O(\sqrt{\delta})}$, which also suffices as δ is small enough depending on p_0 .

(2) Next we will prove (6.41) and (6.42). By (6.39) and (4.6)-(4.7), we can write (where $* \in \{N, L\}$)

$$\begin{aligned} |\tilde{\mathcal{R}}(k, \lambda)| &\lesssim \sum_{\pm k' \pm k_4 \pm \dots \pm k_7 = k} \int_{\mathbb{R}} \frac{d\sigma}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \sigma \rangle^{1-\theta}} \int_{\pm \lambda' \pm \lambda_4 \pm \dots \pm \lambda_7 = \sigma - \Xi} \prod_{j=4}^7 |\tilde{v}_j(k_j, \lambda_j)| \\ &\sum_{\substack{k_2 + k_3 - k_1 = k' \\ (k_1, k_2, k_3) \in \mathbb{X}_*}} |k_1| \int_{\mathbb{R}} T\widehat{\varphi}(T(\lambda' - \mu')) d\mu' \\ &\int_{\mathbb{R}} \min \left(\frac{1}{\langle \Delta' \rangle}, \frac{1}{\langle \mu' \rangle} \right) \frac{d\sigma'}{\langle \mu' - \sigma' \rangle} \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma' - \Delta'} \prod_{j=1}^3 |\tilde{v}_j(k_j, \lambda_j)|. \end{aligned} \tag{6.45}$$

Here $\Xi = k^2 \mp (k')^2 \mp k_4^2 \mp \dots \mp k_7^2$, and $\Delta' = (k')^2 + k_1^2 - k_2^2 - k_3^2$. This can be reduced to

$$\sum_{\pm k_1 \pm \dots \pm k_7 = k} \frac{|k_1|}{|\Delta'|} \int_{\mathbb{R}} R(\lambda, \tau) d\tau \int_{\pm \lambda_1 \pm \dots \pm \lambda_7 = \tau - \Xi'} \prod_{j=1}^7 |\tilde{v}_j(k_j, \lambda_j)|, \quad (6.46)$$

where $\Xi' = k^2 \mp k_1^2 \mp \dots \mp k_7^2$, and the kernel

$$|R(\lambda, \tau)| \lesssim \int_{\mathbb{R}} \frac{d\sigma}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \sigma \rangle^{1-\theta}} \int_{\mathbb{R}} T \widehat{\phi}(T(\xi - \sigma)) \frac{d\xi}{\langle \tau - \xi \rangle}.$$

Here we can verify that $\lambda' - \sigma' = \tau - \sigma$, and ξ is the variable such that $\xi - \sigma = \lambda' - \mu'$ and $\tau - \xi = \mu' - \sigma'$. Using the fact that $|T \widehat{\phi}(T\xi)| \lesssim \langle \xi \rangle^{-1}$, we can easily bound the above by

$$|R(\lambda, \tau)| \lesssim \frac{1}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \tau \rangle^{1-2\theta}}.$$

We may then restrict to the dyadic region $\langle k \rangle \sim N_0$, $\langle k' \rangle \sim N'$ and $\langle k_j \rangle \sim N_j$ for $1 \leq j \leq 7$. Let N^+ be the maximum of all the N_j 's. Then we have $N' \gtrsim N_0$, and $|\Delta'| \sim N' N_1$.

(a) Assume $N_1 \lesssim N'$, we will then measure v_2 in the Y_0 norm. By repeating the above proof and fixing λ_j for $1 \leq j \leq 7$, we may reduce to proving

$$\|F(k, \Xi')\|_{\ell_k^{p_0} \ell_{\Xi'}^{\infty}} \lesssim N_0^{-\frac{1}{2}} N' N_1^{\frac{1}{2}} \cdot (N_2 \dots N_7)^{(1-\sqrt{\delta})/p_0 - (1/q_1)} (N^+)^{-\theta}, \quad (6.47)$$

where

$$F(k, \Xi') = \sum_{\substack{\pm k_1 \pm \dots \pm k_7 = k \\ \pm k_1^2 \pm \dots \pm k_7^2 = k^2 - \Xi'}} \prod_{j=1}^7 f_j(k_j), \quad (6.48)$$

and

$$\|f_1\|_{\ell_k^{p_0}} \leq 1, \quad \|f_j\|_{\ell_k^2} \leq 1 \ (2 \leq j \leq 7). \quad (6.49)$$

First assume $\max(N_2, \dots, N_7) \geq (N^+)^{\sqrt{\delta}}$, then with fixed k, k_1 and Ξ , by using Lemma 3.4 and similar arguments as in the above proof, we can easily show that (whether or not there is any pairing in $\{2, 3, \dots, 7\}$)

$$\sum_{\substack{\pm k_2 \pm \dots \pm k_7 = \text{const.} \\ \pm k_2^2 \pm \dots \pm k_7^2 = \text{const.}}} \prod_{j=2}^7 f_j(k_j) \lesssim \max(N_2, \dots, N_7)^{\theta} \prod_{j=2}^7 \|f_j\|_{\ell_k^2}, \quad (6.50)$$

therefore

$$|F(k, \Xi')| \lesssim \max(N_2, \dots, N_7)^{\theta} \sum_{k_1} f_1(k_1) \lesssim \max(N_2, \dots, N_7)^{\theta} N_1^{1 - (\frac{1}{p_0})}$$

pointwise in (k, Ξ') , thus

$$\|F(k, \Xi')\|_{\ell_k^{p_0} \ell_{\Xi'}^\infty} \lesssim N_0^{\frac{1}{p_0}} N_1^{1 - (\frac{1}{p_0})}.$$

Using the fact that $N' \gtrsim \max(N_0, N_1)$ and $\max(N_2, \dots, N_7) \geq (N^+)^{\sqrt{\delta}}$, this easily implies (6.47). Next assume $\max(N_2, \dots, N_7) \leq (N^+)^{\sqrt{\delta}}$, then $N^+ \sim N'$, and by fixing (k_2, \dots, k_7) we easily deduce that $|F(k, \Xi')| \lesssim (N')^{O(\sqrt{\delta})}$, from which (6.47) follows trivially.

(b) Assume $N_1 \gg N'$, then we must have $* = N$ and $N_1 \sim N_2 \gg N' \gtrsim N_0$. In particular v_2 will be measured in the Z_0 norm, so we may reduce to proving

$$\|F(k, \Xi')\|_{\ell_k^{p_0} \ell_{\Xi'}^\infty} \lesssim N_0^{-\frac{1}{2}} N' N_1 \cdot (N_3 \cdots N_7)^{(1 - \sqrt{\delta})/p_0} (N^+)^{-\theta - \frac{6}{q_1}}, \quad (6.51)$$

where $F(k, \Xi')$ is as (6.48), and

$$\|f_1\|_{\ell_k^{p_0}} \leq 1, \quad \|f_2\|_{\ell_k^{p_0}} \leq 1, \quad \|f_j\|_{\ell_k^2} \leq 1 \quad (3 \leq j \leq 7). \quad (6.52)$$

Here we argue in the same way as case (1), using (6.50), but make the additional observation that for fixed k_1 we must have $|k_2 - k_1| \sim N'$. Therefore

$$\sum_{\substack{\pm k_2 \pm \dots \pm k_7 = \text{const.} \\ \pm k_2^2 \pm \dots \pm k_7^2 = \text{const.}}} \prod_{j=2}^7 f_j(k_j) \lesssim (N^+)^{\theta} \prod_{j=3}^7 \|f_j\|_{\ell_k^2} \cdot \|f_2 \cdot \mathbf{1}_{|k_2 - k_1| \sim N'}\|_{\ell_k^2} \lesssim (N^+)^{\theta} (N')^{\frac{1}{2} - \frac{1}{p_0}},$$

and hence

$$\|F(k, \Xi')\|_{\ell_k^{p_0} \ell_{\Xi'}^\infty} \lesssim N_0^{\frac{1}{p_0}} N_1^{1 - (\frac{1}{p_0})} (N')^{\frac{1}{2} - \frac{1}{p_0}} (N^+)^{\theta}.$$

Using the fact that $N_0 \lesssim N'$, this implies (6.51). \square

Remark 6.6. From the proof above we actually deduce something slightly stronger: the bounds (6.40)–(6.42) remain true if the right hand side of (6.39) gets multiplied by $\langle k^+ \rangle^\theta$ where $|k^+|$ is the maximum of all relevant frequencies, unless $\langle k \rangle \sim \langle k_1 \rangle \gtrsim N_2^{100}$. This fact will be used in the analysis of the z_{5*} terms of Sect. 6.4.

To apply Proposition 6.5, we will verify that $z_{11}, z_{2\ell}$ ($2 \leq \ell \leq 5$), $z_{3\ell}$ ($1 \leq \ell \leq 4$) and $z_{4\ell}$ ($1 \leq \ell \leq 2$) all have the form (6.39). The claim for z_{11} follows from (3.9). For the other terms, let us look at z_{22} as an example. By (3.9) and (4.6) we have, for $z_{22} = z_{22}(v_1, \dots, v_5)$, that

$$\begin{aligned} |\widetilde{z}_{22}(k, \lambda)| &\lesssim_B \sum_{\substack{k_4 + k_5 - k' = k \\ (k', k_4, k_5) \in \mathbb{X}_H \cup \mathbb{X}_S}} |k'| \\ &\quad \int_{\mathbb{R}} \left(\frac{1}{\langle \lambda \rangle^B} + \frac{1}{\langle \lambda - \tau \rangle^B} \right) \frac{d\tau}{\langle \tau \rangle} \int_{\lambda_4 + \lambda_5 - \lambda' = \tau - \Delta} |\widetilde{v}_4(k_4, \lambda_4)| |\widetilde{v}_5(k_5, \lambda_5)| \\ &\quad \sum_{\substack{k_2 + k_3 - k_1 = k' \\ (k_1, k_2, k_3) \in \mathbb{X}_N}} |k_1| \int_{\mathbb{R}} T\widehat{\varphi}(T(\lambda' - \mu')) d\mu' \end{aligned}$$

$$\int_{\mathbb{R}} \min\left(\frac{1}{\langle \Delta' \rangle}, \frac{1}{\langle \mu \rangle}\right) \frac{1}{\langle \mu' - \sigma' \rangle} d\sigma' \int_{\lambda_2 + \lambda_3 - \lambda_1 = \sigma' - \Delta'} \prod_{j=1}^3 |\tilde{v}_j(k_j, \lambda_j)|, \quad (6.53)$$

where $\Delta = k^2 + (k')^2 - k_4^2 - k_5^2$ and $\Delta' = (k')^2 + k_1^2 - k_2^2 - k_3^2$. Note that $|\Delta'| \sim \langle k \rangle \langle k_1 \rangle$, the above can be written as

$$\sum_{k_1 - k_2 - k_3 + k_4 + k_5 = k} \int_{\mathbb{R}} R(\lambda, \sigma) d\sigma \int_{\lambda_1 - \lambda_2 - \lambda_3 + \lambda_4 + \lambda_5 = \sigma - \Xi} \prod_{j=1}^5 |\tilde{v}_j(k_j, \lambda_j)|,$$

where $\Xi = \Delta - \Delta' = k^2 - k_1^2 + k_2^3 + k_3^2 - k_4^2 - k_5^2$, and

$$|R(\lambda, \sigma)| \lesssim \int_{\mathbb{R}} \left(\frac{1}{\langle \lambda \rangle^B} + \frac{1}{\langle \lambda - \tau \rangle^B} \right) \frac{d\tau}{\langle \tau \rangle} \int_{\mathbb{R}} T\widehat{\varphi}(T(\sigma - \xi)) \frac{d\xi}{\langle \xi - \tau \rangle}.$$

Here we can verify that $\lambda' - \sigma' = \sigma - \tau$, and ξ is the variable such that $\sigma - \xi = \lambda' - \mu'$ and $\xi - \tau = \mu' - \sigma'$. The above integral can easily be bounded by $\langle \lambda \rangle^{-1+\theta} \langle \lambda - \sigma \rangle^{-1+\theta}$, so Proposition 6.5 can be applied.

The other $z_{2\ell}$, $z_{3\ell}$ and $z_{4\ell}$ terms can be treated in the same way; in fact the kernel $R(\lambda, \sigma)$ will have exactly the same form, the only difference is that the weight

$$\frac{|k| \cdot |k_1|}{|\Delta'|}$$

will be replaced by different weights depending on which input function gets substituted by \mathcal{E}^Y , and which \mathbb{X}_* subset we are in. For the terms $z_{22}, z_{23}, z_{31}, z_{32}, z_{41}$ and z_{42} one can directly check that this weight is $\lesssim 1$; for the terms z_{24}, z_{25}, z_{33} and z_{34} , this weight is $\lesssim 1$ as it follows from Proposition 2.3 that $|k_2| \gtrsim |k_1|$ when $(k_1, k_2, k_3) \in \mathbb{X}_H \cup \mathbb{X}_S \cup \mathbb{X}_N$. Thus Proposition 6.1 has been proved for these terms.

6.4. Remaining quintic terms. The remaining quintic terms, namely z_{26}, z_{27} and quintic z_{5*} terms, may not have the canonical form (6.39). In fact these terms will be estimated directly without performing the operation in step (6) of Sect. 6.1, see Remark 6.2. For them we need two extra estimates, which we state in the following two propositions.

Proposition 6.7. *Suppose a quintic term \mathcal{R} satisfies*

$$|\tilde{\mathcal{R}}(k, \lambda)| \lesssim \sum_{\pm k_1 \pm \dots \pm k_5 = k} \alpha(k, k_1, \dots, k_5) \int_{\mathbb{R}} \frac{d\sigma}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \sigma \rangle^{1-\theta}} \int_{\pm \lambda_1 \pm \dots \pm \lambda_5 = \sigma - \Xi} \prod_{j=1}^5 |\tilde{v}_j(k_j, \lambda_j)|, \quad (6.54)$$

where as usual $\Xi = k^2 \mp k_1^2 \mp \dots \mp k_5^2$. Then we have the following (below $|k^+|$ will denote the maximum of all relevant frequencies):

(1) Assume $|k_1|/2 \leq |k_2| \leq 2|k_1|$, $|k_1| \geq 2^5 |k_3|$, $|k_3| \sim \max(|k_3|, |k_4|, |k_5|)$ and

$$|\alpha| \lesssim \langle k^+ \rangle^\theta \frac{\langle k_1 \rangle}{\max(\langle k_3 \rangle, \langle k_4 \rangle, \langle k_5 \rangle, \langle k \rangle)}, \quad (6.55)$$

moreover assume (1, 2) is not a pairing. Then we have that

$$\|\mathcal{R}\|_{Z_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Z_0} \|v_3\|_{Z_0} \cdot \|v_4\|_{Y_0} \|v_5\|_{Y_0}; \quad (6.56)$$

(2) Assume $|k_1|/2 \leq |k_2| \leq 2|k_1|$, $|k_3|/2 \leq |k_4| \leq 2|k_3|$, $|k_1| \geq 2^5 \max(|k_5|, |k|)$ and

$$|\alpha| \lesssim \langle k^+ \rangle^\theta \frac{\langle k_1 \rangle}{\max(\langle k_5 \rangle, \langle k \rangle)}. \quad (6.57)$$

Moreover assume (1, 2) is not a pairing, and that, either $|k| \neq |k_5|$, or the stronger bound

$$|\alpha| \lesssim \langle k^+ \rangle^\theta \frac{\langle k_1 \rangle}{\max(\langle k_5 \rangle, \langle k \rangle, \langle \pm k_1 \pm k_2 \rangle)} \quad (6.58)$$

holds. Then we have

$$\|\mathcal{R}\|_{Z_1} \lesssim \|v_1\|_{Z_0} \|v_2\|_{Z_0} \|v_3\|_{Z_0} \|v_4\|_{Z_0} \cdot \|v_5\|_{Y_0}. \quad (6.59)$$

Proof. We may restrict to the region where $\langle k_j \rangle \sim N_j$, where $1 \leq j \leq 5$, and $\langle k \rangle \sim N_0$ and $\langle k^+ \rangle \sim N^+$.

(1) By the same arguments as in the proof of Propositions 5.1 and 6.5, we may fix λ_j ($1 \leq j \leq 5$) and reduce to estimating

$$\begin{aligned} & \left\| \langle \lambda \rangle^{b_1} \sum_{\pm k_1 \pm \dots \pm k_5 = k} \int_{\mathbb{R}} \frac{d\mu}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \Xi - \mu \rangle^{1-\theta}} \prod_{j=1}^5 f_j(k_j) \right\|_{\ell_k^{p_0} L_{\lambda}^{q_0}} \\ & \lesssim (N^+)^{-2\theta} \frac{N' N_3^{\frac{1}{2}} (N_4 N_5)^{(1-\sqrt{\delta})/p_0}}{N_0^{\frac{1}{2}}}, \end{aligned} \quad (6.60)$$

where $N' = \max(N_0, N_3)$, and f_j satisfies that

$$\|f_j\|_{\ell_k^{p_0}} \leq 1, \quad 1 \leq j \leq 3; \quad \|f_j\|_{\ell_k^2} \leq 1, \quad 4 \leq j \leq 5. \quad (6.61)$$

By the same argument as in the proof of Proposition 6.5, we may apply Schur's estimate and reduce to proving

$$\begin{aligned} \|F(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{q_1}} & \lesssim (N^+)^{-2\theta} \frac{N' N_3^{\frac{1}{2}} (N_4 N_5)^{(1-\sqrt{\delta})/p_0}}{N_0^{\frac{1}{2}}}, \quad F(k, \Xi) \\ & := \sum_{\substack{\pm k_1 \pm \dots \pm k_5 = k \\ \pm k_1^2 \pm \dots \pm k_5^2 = k^2 - \Xi}} \prod_{j=1}^5 f_j(k_j). \end{aligned} \quad (6.62)$$

By fixing (k_4, k_5) we get that

$$\|F(k, \Xi)\|_{\ell_{\Xi}^{q_1}} \lesssim \|f_4\|_{\ell_k^1} \|f_5\|_{\ell_k^1} \sup_{\ell, \rho} \|F_{\ell, \rho}(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{q_1}}, \quad F_{\ell, \rho}(k, \Xi)$$

$$:= \sum_{\substack{\pm k_1 \pm k_2 \pm k_3 = k + \ell \\ \pm k_1^2 \pm k_2^2 \pm k_3^2 = k^2 - \Xi + \rho}} \prod_{j=1}^3 f_j(k_j),$$

while since there is no pairing in $\{1, 2, 3\}$, by the divisor estimate in Lemma 3.4, we have

$$|F_{\ell, \rho}(k, \Xi)| \lesssim (N^+)^{\theta} \left(\sum_{\substack{\pm k_1 \pm k_2 \pm k_3 = k + \ell \\ \pm k_1^2 \pm k_2^2 \pm k_3^2 = k^2 - \Xi + \rho}} \prod_{j=1}^3 f_j(k_j)^{p_0} \right)^{\frac{1}{p_0}},$$

and hence $\|F_{\ell, \rho}(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{q_1}} \lesssim \|F_{\ell, \rho}(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{p_0}} \lesssim (N^+)^{\theta}$. Using also Hölder we obtain

$$\|F(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{q_1}} \lesssim (N_4 N_5)^{\frac{1}{2}} (N^+)^{\theta}.$$

Comparing with (6.62) and using that $N' \sim \max(N_0, N_3)$ and $\max(N_4, N_5) \lesssim N_3$, we see that (6.62) is proved, except for the loss $(N^+)^{\theta}$. Clearly this loss can be covered if $N' \gtrsim N_1^{1/10}$; now suppose $\max(N_0, N_3, N_4, N_5) \ll N_1^{1/10}$, then since (1, 2) is not a pairing, we must have $|\Xi| \gtrsim N^+$, which gives

$$\max(|\lambda_1|, \dots, |\lambda_5|, |\lambda|, |\lambda - \Xi - \mu|) \gtrsim N^+,$$

where μ is a linear combination of $\lambda_1, \dots, \lambda_5$. Now, in estimating (6.60) we can gain a power $\langle \lambda \rangle^{(1-\theta)-b_1} \geq \langle \lambda \rangle^{\delta/2}$; in the process of fixing λ_j we can also gain a power $\langle \lambda_j \rangle^{\delta/2}$, as

$$\begin{aligned} & \|\langle \lambda_j \rangle^{\delta/2} \langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \tilde{v}_j\|_{L_{\lambda}^1 \ell_k^2} \\ & \lesssim \|\langle \lambda_j \rangle^{\delta/2} \langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \tilde{v}_j\|_{L_{\lambda}^1 \ell_k^{r_2}} \lesssim \|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \langle \lambda_j \rangle^{\frac{1}{2}} \tilde{v}_j\|_{L_{\lambda}^{r_0} \ell_k^{r_2}} \\ & \lesssim \|\langle k_j \rangle^{(1-\sqrt{\delta})/p_0} \langle \lambda_j \rangle^{\frac{1}{2}} \tilde{v}_j\|_{\ell_k^{r_2} L_{\lambda}^{r_0}} \lesssim \|\langle k_j \rangle^{\frac{1}{2}} \langle \lambda_j \rangle^{\frac{1}{2}} \tilde{v}_j\|_{\ell_k^{p_0} L_{\lambda}^{r_0}}. \end{aligned}$$

Finally, in the process of using Schur's estimate to reduce (6.60) to (6.62), we can also replace the power $\langle \lambda - \Xi - \mu \rangle^{-1+\theta}$ by a slightly higher power to gain a power $\langle \lambda - \Xi - \mu \rangle^{\delta/4}$. In this way we can gain a power of at least $(N^+)^{\delta/4}$ which suffices to cover the $(N^+)^{\theta}$ loss.

(2) If there is no pairing in $\{1, 2, 3\}$, then similar to (1), we may fix λ_j and reduce to proving

$$\begin{aligned} \|F(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{q_1}} & \lesssim (N^+)^{-2\theta} \frac{N' N_3 \cdot N_5^{(1-\sqrt{\delta})/p_0}}{N_0^{\frac{1}{2}}}, \quad F(k, \Xi) \\ & := \sum_{\substack{\pm k_1 \pm \dots \pm k_5 = k \\ \pm k_1^2 \pm \dots \pm k_5^2 = k^2 - \Xi}} \prod_{j=1}^5 f_j(k_j), \end{aligned} \tag{6.63}$$

where $N^+ = \max(N_1, N_3)$, $N' = \max(N_0, N_5)$ and

$$\|f_j\|_{\ell_k^{p_0}} \leq 1, \quad 1 \leq j \leq 4; \quad \|f_5\|_{\ell_k^2} \leq 1. \tag{6.64}$$

Then we may fix k_4 and k_5 and argue as in part (1) to get

$$\|F(k, \Xi)\|_{\ell_k^{p_0} \ell_{\Xi}^{q_1}} \lesssim (N^+)^{\theta} \|f_4\|_{\ell_k^1} \|f_5\|_{\ell_k^1} \lesssim (N^+)^{\theta} N_3^{1 - \frac{1}{p_0}} N_5^{\frac{1}{2}},$$

which implies (6.63) except for the loss $(N^+)^{\theta}$, which can be covered in the same way as part (1) by considering Ξ .

If there is a pairing in $\{1, 2, 3\}$, say $(1, 3)$, then $\frac{1}{2} \leq N_1/N_3 \leq 2$. If $(2, 4)$ is not a pairing, then we can fix (k_3, k_5) and repeat the above argument to get the same (in fact better) estimate; so we may assume $(2, 4)$ is also a pairing. This forces $\Xi = 0$ and $k = k_5$, in particular the stronger bound (6.58) holds. Let $\langle k_1 \pm k_2 \rangle \sim N_6$ and $N'' = \max(N', N_6)$. In this case we will still fix $\lambda_j (1 \leq j \leq 4)$ but will not fix λ_5 . Instead, let μ be a linear combination of $\lambda_j (1 \leq j \leq 4)$ and is thus fixed, and notice that $\Xi = 0$, we have

$$|\tilde{\mathcal{R}}(k, \lambda)| \lesssim N_1^{-1} (N'')^{-1} \sum_{\substack{|k_1| \sim |k_2| \sim N_1 \\ |k_1 \pm k_2| \sim N_6}} \int_{\mathbb{R}} \frac{d\sigma}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \sigma \rangle^{1-\theta}} f_1(k_1) f_2(k_2) f_3(k_1) f_4(k_2) |\tilde{v}_5(k, \pm\sigma \pm \mu)|,$$

where $\|f_j\|_{\ell_k^{p_0}} \leq 1$ for $1 \leq j \leq 4$. This implies that

$$\begin{aligned} & \|\langle \lambda \rangle^{b_1} \tilde{\mathcal{R}}(k, \lambda)\|_{L_{\lambda}^{q_0}} \\ & \lesssim N_1^{-1} (N'')^{-1} \sum_{\substack{|k_1| \sim |k_2| \sim N_1 \\ |k_1 \pm k_2| \sim N_6}} f_1(k_1) f_2(k_2) f_3(k_1) f_4(k_2) \|\langle \lambda_5 \rangle^{\frac{1}{2}} \tilde{v}_5(k, \lambda_5)\|_{L_{\lambda}^{r_0}}, \end{aligned}$$

and hence

$$\begin{aligned} & \|\langle k \rangle^{\frac{1}{2}} \langle \lambda \rangle^{b_1} \tilde{\mathcal{R}}(k, \lambda)\|_{\ell_k^{p_0} L_{\lambda}^{q_0}} \\ & \lesssim N_1^{-1} (N'')^{-1} \|\langle k \rangle^{\frac{1}{2}} \langle \lambda_5 \rangle^{\frac{1}{2}} \tilde{v}_5(k, \lambda_5)\|_{\ell_k^{p_0} L_{\lambda}^{r_0}} \\ & \quad \cdot \sum_{\substack{|k_1| \sim |k_2| \sim N_1 \\ |k_1 \pm k_2| \sim N_6}} f_1(k_1) f_2(k_2) f_3(k_1) f_4(k_2), \end{aligned}$$

while the latter sum is bounded by

$$\sum_{k_1} f_1(k_1) f_3(k_1) \cdot N_6^{1 - \frac{2}{p_0}} \|f_2\|_{\ell_k^p} \|f_4\|_{\ell_k^p} \lesssim (N_1 N_6)^{1 - \frac{2}{p_0}},$$

which gives the desired estimate as $N'' \gtrsim N_6$. \square

Proposition 6.8. *Suppose a quintic term \mathcal{R} satisfies*

$$\begin{aligned} |\tilde{\mathcal{R}}(k, \lambda)| & \lesssim \frac{1}{\langle \lambda \rangle^{1+10\delta}} \sum_{\pm k_1 \pm \dots \pm k_5 = k} \beta(k, k_1, \dots, k_5) \\ & \int_{\mathbb{R}^5} \prod_{j=1}^5 |\tilde{v}_j(k_j, \lambda_j)| d\lambda_1 \dots d\lambda_5, \end{aligned} \tag{6.65}$$

where (as usual $|k^+|$ is the maximum of all relevant frequencies)

$$|\beta| \lesssim \langle k^+ \rangle^{\sqrt{\delta}} \frac{1}{\langle k \rangle \langle \pm k_3 \pm k_4 \rangle}, \quad \langle k \rangle \gtrsim \langle k_5 \rangle \quad (6.66)$$

then we have

$$\|\mathcal{R}\|_{Z_1} \lesssim \prod_{j=1}^5 \|v_j\|_{Y_0}. \quad (6.67)$$

Note that all the norms on the right hand side are Y_0 (in particular the bound is symmetric in v_1 and v_2 , v_3 and v_4).

Proof. As before we will restrict to the region where $\langle k_j \rangle \sim N_j$ for $1 \leq j \leq 5$, $\langle k \rangle \sim N_0$, and $\langle k_3 \pm k_4 \rangle \sim N_6$. Let $N^+ \sim \langle k^+ \rangle$. This time we will not fix λ_j ; instead we first integrate in them. We may assume all the norms on the right hand side are 1. Let

$$N_j^{\frac{1}{2}} \|\tilde{v}_j(k_j, \lambda_j)\|_{L_\lambda^1} = f_j(k_j),$$

then $\|f_j\|_{\ell_k^{p_0}} \lesssim 1$ as $v_j \in Y_0$. By (6.65), it suffices to prove that

$$\begin{aligned} & \left\| \langle \lambda \rangle^{b_1} \langle \lambda \rangle^{-1-10\delta} \sum_{\pm k_1 \pm \dots \pm k_5 = k} \prod_{j=1}^5 f_j(k_j) \right\|_{\ell_k^{p_0} L_\lambda^{q_0}} \\ & \lesssim (N^+)^{-2\sqrt{\delta}} \cdot N_6 (N_0 N_1 N_2 N_3 N_4 N_5)^{\frac{1}{2}}, \end{aligned} \quad (6.68)$$

where

$$\|f_j\|_{\ell_k^{p_0}} \leq 1, \quad 1 \leq j \leq 5. \quad (6.69)$$

By symmetry we may assume $N_1 \leq N_2$ and $N_3 \leq N_4$. By the choice of power of λ , the $L_\lambda^{q_0}$ part is easily estimated, so we only need to bound the $\ell_k^{p_0}$ norm

$$\left\| \sum_{\pm k_1 \pm \dots \pm k_5 = k} \prod_{j=1}^5 f_j(k_j) \right\|_{\ell_k^{p_0}}.$$

By Young's inequality, this is bounded by the $\ell_k^{p_0}$ norm of f_2 (which is ~ 1), multiplied by

$$\sum_{k_1, k_3, k_4, k_5} f_1(k_1) f_3(k_3) f_4(k_4) f_5(k_5).$$

The sum over k_1 and k_5 gives (by Hölder) $(N_1 N_5)^{1-\frac{1}{p_0}}$; when k_3 is fixed the sum over k_4 gives $N_6^{1-\frac{1}{p_0}}$ as $|k_3 \pm k_4| \lesssim N_6$, and finally the sum over k_3 gives $N_3^{1-\frac{1}{p_0}}$. This gives the bound

$$(N_1 N_3 N_5 N_6)^{1-\frac{1}{p_0}} \lesssim (N^+)^{-\frac{1}{2p_0}} N_6 (N_0 N_1 N_2 N_3 N_4 N_5)^{\frac{1}{2}},$$

as $N_1 \leq N_2$, $N_3 \leq N_4$ and $N_0 \gtrsim N_5$. □

Using Propositions 6.7, we can easily deal with the terms z_{26} and z_{27} . For these two terms, by repeating the arguments for z_{22} detailed above, we are led to considering the tuple (k_1, k_2, k') and (k_3, k_4, k_5) , such that

$$\begin{aligned} k_2 + k' - k_1 &= k, \quad (k_1, k_2, k') \in \mathbb{X}_H \cup \mathbb{X}_S; \\ k_3 + k_4 - k_5 &= k', \quad (k_3, k_4, k_5) \in \mathbb{X}_N \cup \mathbb{X}_L, \end{aligned}$$

and a weight

$$\alpha(k, k_1, \dots, k_5) \sim \frac{|k_1||k_3|}{\langle \Delta' \rangle},$$

noticing that $|\Delta'| \geq 1$. By Proposition 2.5, this term can be bounded using either Proposition 6.5, or Proposition 6.7, (1) or (2).

6.4.1. The z_{5} terms* Finally let us consider quintic z_{5*} terms. By (4.6)-(4.7) and (4.13) we write, where $z_{5*} = z_{5*}(v_1, \dots, v_5)$ and $*, \# \in \{N, L\}$, that (strictly speaking z_{57} and z_{58} have a different formula, but taking into account that the set \mathbb{X}_L is symmetric with respect to k_2 and k_3 - apart from the artificial restriction $|k_2| \geq |k_3|$ - they can be treated in exactly the same way):

$$\begin{aligned} |\widetilde{z}_{5*}(k, \lambda)| &\lesssim_B \sum_{\substack{k_2+k'-k_1=k \\ (k_1, k_2, k') \in \mathbb{X}_*}} |k_1| \int_{\mathbb{R}} \left[\frac{1}{\langle \lambda \rangle^B \langle \tau \rangle} \right. \\ &\quad \left. + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle^B \langle \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \tau \rangle} \right) + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right)^2 \right] d\tau \\ &\quad \int_{\lambda_2 + \lambda' - \lambda_1 = \tau - \Delta} \mathbf{1}_{|\lambda'| \gtrsim |\tau - \Delta|} |\widetilde{v}_1(k_1, \lambda_1)| |\widetilde{v}_2(k_2, \lambda_2)| \\ &\quad \sum_{\substack{k_4+k_5-k_3=k' \\ (k_3, k_4, k_5) \in \mathbb{X}_{\#}}} |k_3| \int_{\mathbb{R}} T \widehat{\varphi}(T(\lambda' - \mu')) d\mu' \\ &\quad \int_{\mathbb{R}} \min \left(\frac{1}{\langle \Delta' \rangle}, \frac{1}{\langle \mu' \rangle} \right) \frac{1}{\langle \mu' - \sigma' \rangle} d\sigma' \\ &\quad \int_{\lambda_4 + \lambda_5 - \lambda_3 = \sigma' - \Delta'} |\widetilde{v}_3(k_3, \lambda_3)| |\widetilde{v}_4(k_4, \lambda_4)| |\widetilde{v}_5(k_5, \lambda_5)|, \end{aligned} \tag{6.70}$$

where $\Delta = k^2 + k_1^2 - k_2^2 - (k')^2$ and $\Delta' = (k')^2 + k_3^2 - k_4^2 - k_5^2$. The above can be reduced to

$$\sum_{-k_1+k_2-k_3+k_4+k_5=k} \int_{\mathbb{R}} R(\lambda, \sigma) d\sigma \int_{-\lambda_1+\lambda_2-\lambda_3+\lambda_4+\lambda_5=\sigma-\Xi} \prod_{j=1}^5 |\widetilde{v}_j(k_j, \lambda_j)|, \tag{6.71}$$

where $\Xi = \Delta + \Delta' = k^2 + k_1^2 - k_2^2 + k_3^2 - k_4^2 - k_5^2$, and the kernel¹⁸

$$R(\lambda, \sigma) = \int_{\mathbb{R}} \left[\frac{1}{\langle \lambda \rangle^B \langle \tau \rangle} + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle^B \langle \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \tau \rangle} \right) \right]$$

¹⁸ This kernel depends on k_j and λ_j , but we will write it as $R(\lambda, \sigma)$ for simplicity.

$$\begin{aligned}
& + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right)^2 \left[\right. \\
& \left. + k_1 k_3 |\mathbf{1}_{|\lambda'| \gtrsim |\tau - \Delta|} \int_{\mathbb{R}} T \widehat{\varphi}(T(\lambda' - \mu')) \min \left(\frac{1}{\langle \Delta' \rangle}, \frac{1}{\langle \mu' \rangle} \right) \frac{1}{\langle \mu' - \sigma' \rangle} d\mu' \right] \quad (6.72)
\end{aligned}$$

Here $\lambda' = \tau - \Delta + \lambda_1 - \lambda_2$ and $\sigma' = \lambda_4 + \lambda_5 - \lambda_3 + \Delta'$ are defined in terms of τ and (k_j, λ_j) . First fix τ and integrate in μ' ; this integral is bounded by

$$\int_{\mathbb{R}} \frac{1}{\langle \lambda' - \mu' \rangle} \min \left(\frac{1}{\langle \Delta' \rangle}, \frac{1}{\langle \mu' \rangle} \right) \frac{1}{\langle \mu' - \sigma' \rangle} d\mu',$$

and we separate two cases.

(1) Assume $|\sigma'| \ll |\lambda'|$, then we can calculate that

$$\frac{1}{\langle \Delta' \rangle} \int_{\mathbb{R}} \frac{1}{\langle \lambda' - \mu' \rangle \langle \mu' - \sigma' \rangle} d\mu' \lesssim \frac{1}{\langle \Delta' \rangle} \frac{1}{\langle \lambda' - \sigma' \rangle^{1-\theta}} \sim \frac{1}{\langle \Delta' \rangle} \frac{1}{\langle \lambda' \rangle^{1-\theta}}.$$

Note that $|\lambda'| \gtrsim |\tau - \Delta|$, we can then bound the resulting integral in τ by

$$\begin{aligned}
& \int_{|\tau - \Delta| \lesssim |\lambda_1 - \lambda_2|} \left[\frac{1}{\langle \lambda \rangle^B \langle \tau \rangle} + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle^B \langle \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \tau \rangle} \right) \right. \\
& \left. + \mathbf{1}_{\langle \lambda - \tau \rangle \ll \langle \tau - \Delta \rangle} \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right)^2 \right] \frac{1}{\langle \Delta' \rangle} \frac{1}{\langle \tau - \Delta \rangle^{1-\theta}} d\tau,
\end{aligned}$$

which is then bounded by

$$\frac{1}{\langle \Delta \rangle \langle \Delta' \rangle} (\max_j \langle k_j \rangle)^{\sqrt{\delta}} \langle \lambda \rangle^{-1-10\delta}$$

by actually performing the integration in τ . By bounding the weight

$$\beta = \frac{|k_1| |k_3|}{\langle \Delta \rangle \langle \Delta' \rangle}$$

using Proposition 2.5, we can apply Proposition 6.8 and conclude the estimate for this term.

(2) Assume $|\sigma'| \gtrsim |\lambda'|$, then we can calculate that

$$\int_{\mathbb{R}} \frac{1}{\langle \lambda' - \mu' \rangle \langle \mu' \rangle \langle \mu' - \sigma' \rangle} d\mu' \lesssim \frac{1}{\langle \lambda' \rangle^{1-\theta}} \frac{1}{\langle \lambda' - \sigma' \rangle^{1-\theta}}.$$

Note that $\lambda' - \sigma' = \tau - \sigma$, and using the fact that $|\lambda'| \gtrsim |\tau - \Delta|$, we can bound the resulting integral in τ by

$$\begin{aligned}
& \int_{\mathbb{R}} \left[\frac{\mathbf{1}_{\langle \sigma \rangle \ll \langle \Delta \rangle}}{\langle \lambda \rangle^B \langle \tau \rangle} + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle^B \langle \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \tau \rangle} \right) \right. \\
& \left. + \frac{\langle \tau - \Delta \rangle}{\langle \lambda - \tau \rangle} \min \left(\frac{1}{\langle \Delta \rangle}, \frac{1}{\langle \lambda \rangle} \right)^2 \right] \frac{1}{\langle \tau - \Delta \rangle^{1-\theta}} \frac{1}{\langle \tau - \sigma \rangle^{1-\theta}} d\tau,
\end{aligned}$$

which can be bounded by

$$(\max_j \langle k_j \rangle)^{10\theta} \frac{1}{\langle \lambda \rangle^{1-\theta} \langle \lambda - \sigma \rangle^{1-\theta}} \frac{1}{\langle \Delta \rangle}.$$

By bounding the weight

$$\alpha = \frac{|k_1||k_3|}{\langle \Delta \rangle}$$

using Proposition 2.5, we can apply either Proposition 6.5, or Proposition 6.7, (1) or (2).

In the case we apply Proposition 6.5, we will also use Remark 6.6 to cover the loss $(\max_j \langle k_j \rangle)^{10\theta}$, which can be done unless for some j we have $|k_j| \sim |k| \gtrsim \max_{\ell \neq j} |k_\ell|^{100}$; in this final case we can check that the stronger bound $|\alpha| \lesssim |k|^{-\frac{1}{2}}$ holds, so the loss can still be covered. This completes the proof of Proposition 6.1.

7. Preservation of Regularity

Finally in this section we prove a preservation of regularity result. More precisely, we prove the properties of our solution stated in Remark 1.2. The following proposition is standard:

Proposition 7.1. *Given $s > \frac{1}{2}$ and $2 \leq p_0 < \infty$, all the arguments in the previous sections carry over to $H_{p_0}^s$ (and correspondingly X_{p_0, q_0}^{s, b_j} and $X_{p_0, r_j}^{s, \frac{1}{2}}$ for $j \in \{0, 1\}$). Moreover, in these arguments T still depends only on the $H_{p_0}^{\frac{1}{2}}$ (instead of $H_{p_0}^s$) size of the initial data.*

Proof. This follows from the elementary inequality that

$$\langle k \rangle^{s-(\frac{1}{2})} \lesssim \max_{1 \leq j \leq r} \langle k_j \rangle^{s-(\frac{1}{2})}, \quad \text{if } k = \pm k_1 \pm \cdots \pm k_r, \quad r \in \{3, 5, 7\}.$$

Thus, any previously proved multilinear estimate will continue to be true if the exponent $\frac{1}{2}$ in the output function space is replaced by s , provided that the exponent $\frac{1}{2}$ in one appropriate input function space is replaced by s .

Suppose the initial data has $H_{p_0}^{\frac{1}{2}}$ norm A and $H_{p_0}^s$ norm L , then for $T = T(A)$, all the $X_{p_0, q_0}^{s, b_j}(J)$ —and similarly for $X_{p_0, r_j}^{s, \frac{1}{2}}(J)$ —contraction mappings proved before will still be contraction mapping under the norm

$$\| \cdot \|_{X_{p_0, q_0}^{s, b_j}(J)} + L^{-1} \| \cdot \|_{X_{p_0, q_0}^{s, b_j}(J)}, \quad \text{similarly} \quad \| \cdot \|_{X_{p_0, r_j}^{s, \frac{1}{2}}(J)} + L^{-1} \| \cdot \|_{X_{p_0, r_j}^{s, \frac{1}{2}}(J)}.$$

□

Now consider a smooth initial data u_0 . Proposition 7.1 implies that, if $\|u_0\|_{H_{p_0}^{\frac{1}{2}}} \leq A$, then for $T = T(A)$, we can construct a solution to (1.1) on $J = [-T, T]$ that belongs to $C_t^0 H_{p_0}^s(J)$ for s sufficiently large. This is clearly the classical solution to (1.1). If a sequence of smooth initial data $u_0^{(n)}$ satisfies $\|u_0^{(n)}\|_{H_{p_0}^{\frac{1}{2}}} \leq A$ and $u_0^{(n)} \rightarrow u_0$ in $H_{p_0}^{\frac{1}{2}}$,

then by continuity of the data-to-solution map in $H_{p_0}^{\frac{1}{2}}$ (which follows from the previous proofs), the corresponding solutions $u^{(n)}$ will converge to u in $C_t^0 H_{p_0}^{\frac{1}{2}}$, where u is the solution we construct in Theorem 1.1 with initial data u_0 . This shows that our solution is the unique limit of smooth solutions.

Finally, suppose $p_0 < 4$, then the gauged solution v we construct belongs to the space $X_{p_0, r_0}^{\frac{1}{2}, \frac{1}{2}}(J)$ where $r_0 < 2$. It can be shown that

$$X_{p_0, r_0}^{\frac{1}{2}, \frac{1}{2}}(J) \subset X_{p_0, 2}^{\frac{1}{2}, \frac{1}{2}}(J) \cap X_{p_0, 1}^{\frac{1}{2}, 0}(J)$$

for any interval J of length not exceeding 1, so our solution belongs to the function space defined in [27] in which the authors have proved uniqueness. Therefore when $p_0 < 4$, our solution must coincide with the one constructed in [27], as long as the latter exists.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Alazard, T.: Paralinearization of the Dirichlet–Neumann operator and applications to progressive gravity waves. *Sci. China Math.* **55**(2), 207–220 (2012)
2. Bailleul, I., Bernicot, F.: Heat semigroup and singular PDEs. *J. Funct. Anal.* **270**(9), 3344–3452 (2016)
3. Biagioni, H.A., Linares, F.: Ill-posedness for the derivative Schrödinger and generalized Benjamin–Ono equations. *Trans. Am. Math. Soc.* **353**(9), 3649–3659 (2001)
4. Bényi, A., Oh, T., Pocovnicu, O.: On the probabilistic Cauchy theory of the cubic nonlinear Schrödinger equation on \mathbb{R}^d , $d \geq 3$. *Trans. Am. Math. Soc. Ser. B* **2**, 1–50 (2015)
5. Bényi, A., Pocovnicu, O., Oh, T.: Higher order expansions for the probabilistic local Cauchy theory of the cubic nonlinear Schrödinger equation on \mathbb{R}^3 . *Trans. Am. Math. Soc. Ser. B* **6**, 114–160 (2019)
6. Bourgain, J.: Invariant measures for the 2D defocusing nonlinear Schrödinger equation. *Commun. Math. Phys.* **176**, 421–445 (1996)
7. Bourgain, J.: Global solutions of nonlinear Schrödinger equations. *Am. Math. Soc. Colloq. Pub.* **46**, 99–107 (1999)
8. Bourgain, J.: Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations. *Geom. Funct. Anal.* **3**, 107–156 (1993)
9. Burq, N., Tzvetkov, N.: Random data Cauchy theory for supercritical wave equations. I. Local theory. *Invent. Math.* **173**(3), 449–475 (2008)
10. Burq, N., Thomann, L., Tzvetkov, N.: Long time dynamics for the one dimensional non linear Schrödinger equation. *Ann. Inst. Fourier (Grenoble)* **63**(6), 2137–2198 (2013)
11. Catellier, R., Chouk, K.: Paracontrolled distributions and the 3-dimensional stochastic quantization equation. *Ann. Probab.* **46**(5), 2621–2679 (2018)
12. Champeaux, L.D., Passot, T., Sulem, P.-L.: Remarks on the parallel propagation of small-amplitude dispersive Alfvén waves. *Nonlinear Proc. Geophys.* **6**, 169–178 (1999)
13. Chandra, A., Weber, H.: Stochastic PDEs, regularity structures, and interacting particle systems. *Ann. Fac. Sci. Toulouse Math.* (6) **26**(4), 847–909 (2017)
14. Chanillo, S., Czubak, M., Mendelson, D., Nahmod, A., Staffilani, G.: Almost sure boundedness of iterates for derivative nonlinear wave equations. To appear in *Comm. Anal. Geom.* (2020). [arXiv:1710.09346](https://arxiv.org/abs/1710.09346) [math.AP]
15. Christ, M.: Power series solution of a nonlinear Schrödinger equation, mathematical aspects of nonlinear dispersive equations. *Ann. Math. Stud.* **163**, 131–155 (2007)
16. Colliander, J., Keel, M., Staffilani, G., Takaoka, H., Tao, T.: Global well-posedness for Schrödinger equations with derivative. *SIAM J. Math. Anal.* **33**(3), 649–669 (2001)
17. Colliander, J., Keel, M., Staffilani, G., Takaoka, H., Tao, T.: A refined global well-posedness for Schrödinger equations with derivative. *SIAM J. Math. Anal.* **34**(1), 64–86 (2002)

18. Colliander, J., Oh, T.: Almost sure well-posedness of the cubic nonlinear Schrödinger equation below $L^2(\mathbb{T})$. *Duke Math. J.* **161**(3), 367–414 (2012)
19. Da Prato, G., Debussche, A.: Two-dimensional Navier–Stokes equations driven by a space-time white noise. *J. Funct. Anal.* **196**(1), 180–210 (2002)
20. Da Prato, G., Debussche, A.: Strong solutions to the stochastic quantization equations. *Ann. Probab.* **31**(4), 1900–1916 (2003)
21. Deng, Y.: Invariance of the Gibbs measure for the Benjamin–Ono equation. *J. Eur. Math. Soc. (JEMS)* **17**(5), 1107–1198 (2015)
22. Deng, Y.: Two-dimensional nonlinear Schrödinger equation with random radial data. *Anal. PDE* **5**(5), 913–960 (2012)
23. Dodson, B., Lührmann, J., Mendelson, D.: Almost sure local well-posedness and scattering for the 4D cubic nonlinear Schrödinger equation. *Adv. Math.* **347**, 619–676 (2019)
24. Grigoryan, V., Nahmod, A.: Almost critical well-posedness for nonlinear wave equations with $Q_{\mu\nu}$ null forms in 2D. *Math. Res. Lett.* **21**(2), 313–332 (2014)
25. Grünrock, A.: Bi and Trilinear Schrödinger estimates in one space dimension with applications to cubic NLS and DNLS. *Int. Math. Res. Not.* **41**, 2525–2558 (2005)
26. Grünrock, A.: On the wave equation with quadratic nonlinearities in three space dimensions. *J. Hyperbolic Differ. Equ.* **8**(1), 1–8 (2011)
27. Grünrock, A., Herr, S.: Low regularity local well-posedness of the derivative nonlinear Schrödinger equation with periodic initial data. *SIAM J. Math. Anal.* **39**(6), 1890–1920 (2008)
28. Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, 75 (2015)
29. Gubinelli, M., Imkeller, P., Perkowski, N.: A Fourier analytic approach to pathwise stochastic integration. *Electron. J. Probab.* **21**(2), 37 (2016)
30. Gubinelli, M., Perkowski, N.: Lectures on singular stochastic PDEs. In: *Ensaios Matemáticos [Mathematical Surveys]*, vol. 29, p. 89. Sociedade Brasileira de Matemática, Rio de Janeiro (2015)
31. Gubinelli, M., Koch, H., Oh, T.: Renormalization of the two-dimensional stochastic nonlinear wave equations. *Trans. Am. Math. Soc.* **370**(10), 7335–7359 (2018)
32. Gubinelli, M., Koch, H., Oh, T.: Paracontrolled approach to the three-dimensional stochastic nonlinear wave equation with quadratic nonlinearity. [arXiv:1811.07808](https://arxiv.org/abs/1811.07808) [math.AP]
33. Hairer, M.: Solving the KPZ equation. *Ann. Math.* (2) **178**(2), 559–664 (2013)
34. Hairer, M.: A theory of regularity structures. *Invent. Math.* **198**(2), 269–504 (2014)
35. Hairer, M.: Singular stochastic, P.D.E. Proc. ICM-Seoul **I**, 685–709 (2014)
36. Hairer, M.: Regularity structures and the dynamical Φ_3^4 model. In: Jerison, D., Mazur, B., Mrowka, T., Schmid, W., Stanley, R., Yau, S.T. (eds.) *Current Developments in Mathematics 2014*, pp. 1–49. International Press, Somerville (2016)
37. Hayashi, N.: The initial value problem for the derivative nonlinear Schrödinger equation in the energy space. *Nonlinear Anal.* **20**(7), 823–833 (1993)
38. Hayashi, N., Ozawa, T.: On the derivative nonlinear Schrödinger equation. *Phys. D* **55**(1–2), 14–36 (1992)
39. Hayashi, N., Ozawa, T.: Finite energy solutions of nonlinear Schrödinger equation of derivative type. *SIAM J. Math. Anal.* **25**(6), 1488–1503 (1994)
40. Herr, S.: On the Cauchy problem for the derivative nonlinear Schrödinger equation with periodic boundary condition. *Int. Math. Res. Not.* Article ID 96763, 1–33 (2006)
41. Hirayama, H., Okamoto, M.: Random data Cauchy problem for the nonlinear Schrödinger equation with derivative nonlinearity. *Discrete Contin. Dyn. Syst.* **36**(12), 6943–6974 (2016)
42. Hörmander, L.: *The Analysis of Linear Partial Differential Operators. II*. Grundlehren der Mathematischen Wissenschaften, vol. 257. Springer, Berlin (1983)
43. Hörmander, L.: *The Nash–Moser Theorem and Paradifferential Operators, Analysis et Cetera*, pp. 429–449. Academic Press, Boston (1990)
44. Jenkins, R., Liu, J., Perry, P., Sulem, C.: Global well-posedness for the derivative nonlinear Schrödinger equation. *Commun. Partial Differ. Equ.* **43**(8), 1151–1195 (2018)
45. Jenkins, R., Liu, J., Perry, P., Sulem, C.: Soliton resolution for the derivative nonlinear Schrödinger equation. *Commun. Math. Phys.* **363**(3), 1003–1049 (2018)
46. Jenkins, R., Liu, J., Perry, P., Sulem, C.: Global existence for the derivative nonlinear Schrödinger equation with arbitrary spectral singularities. [arXiv:1804.01506v3](https://arxiv.org/abs/1804.01506v3) (math.AP)
47. Jenkins, R., Liu, J., Perry, P., Sulem, C.: The derivative nonlinear Schrödinger equation, global well-posedness and soliton resolution. *Quart. Appl. Math.* **78**(1), 33–73 (2020)
48. Kaup, D.J., Newell, A.C.: An exact solution for the derivative nonlinear Schrödinger equation. *J. Math. Phys.* **19**(4), 798–801 (1978)
49. Lee, J.-H.: Global solvability of the derivative nonlinear Schrödinger equation. *Trans. Am. Math. Soc.* **314**(1), 107–118 (1989)

50. Miao, C., Wu, Y., Xu, G.: Global well-posedness for Schrödinger equation with derivative in $H^{\frac{1}{2}}(\mathbb{R})$. *J. Differ. Equ.* **251**(8), 2164–2195 (2011)
51. Mio, K., Ogino, T., Minami, K., Takeda, S.: Modified nonlinear Schrödinger equation for Alfvén waves propagating along the magnetic field in cold plasmas. *J. Phys. Soc. Jpn.* **41**, 265–271 (1976)
52. Mjølhus, E.: On the modulational instability of hydromagnetic waves parallel to the magnetic field. *J. Plasma Phys.* **16**, 321–334 (1976)
53. Mosincat, R.: Global well-posedness of the derivative nonlinear Schrödinger equation with periodic boundary condition in $H^{\frac{1}{2}}$. *J. Differ. Equ.* **263**(8), 4658–4722 (2017)
54. Mosincat, R., Yoon, H.: Unconditional uniqueness for the derivative Schrödinger equation on the real line. *Discrete Contin. Dyn. Syst.* **40**(1), 47–80 (2020)
55. Mosincat, R., Oh, T.: A remark on global well-posedness of the derivative nonlinear Schrödinger equation on the circle. *C. R. Math. Acad. Sci. Paris* **353**(9), 837–841 (2015)
56. Mourrat, J., Weber, H.: The dynamic Φ_3^4 model comes down from infinity. *Commun. Math. Phys.* **356**(3), 673–753 (2017)
57. Mourrat, J., Weber, H., Xu, W.: Construction of Φ_3^4 diagrams for pedestrians. In: Gonçalves, P., Soares A. (eds.) *From Particle Systems to Partial Differential Equations. PSPDE 2015*. Springer Proceedings in Mathematics & Statistics, vol 209, pp. 1–46. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-66839-0_1
58. Nahmod, A.R., Oh, T., Rey-Bellet, L., Staffilani, G.: Invariant weighted Wiener measures and almost sure global well-posedness for the periodic derivative NLS. *J. Eur. Math. Soc. (JEMS)* **14**(4), 1275–1330 (2012)
59. Nahmod, A.R., Rey-Bellet, L., Sheffield, S., Staffilani, G.: Absolute continuity of Brownian bridges under certain gauge transformations. *Math. Res. Lett.* **18**(5), 875–887 (2011)
60. Nahmod, A., Staffilani, G.: Almost sure well-posedness for the periodic 3D quintic nonlinear Schrödinger equation below the energy space. *J. Eur. Math. Soc. (JEMS)* **17**(7), 1687–1759 (2015)
61. Ozawa, T.: On the nonlinear Schrödinger equations of derivative type. *Indiana Univ. Math. J.* **45**(1), 137–163 (1996)
62. Pelinovsky, D.E., Shimabukuro, Y.: Existence of global solutions to the derivative NLS equation with the inverse scattering transform method. *Int. Math. Res. Not.* **18**, 5663–5728 (2018)
63. Takaoka, H.: Well-posedness for the one dimensional nonlinear Schrödinger equation with the derivative nonlinearity. *Adv. Differ. Equ.* **4**(4), 561–580 (1999)
64. Takaoka, H.: Global well-posedness for Schrödinger equations with derivative in a nonlinear term and data in low order Sobolev spaces. *Electron. J. Differ. Equ.* **42**, 23 (2001). (electronic)
65. Tao, T.: Global regularity of wave maps. II. Small energy in two dimensions. *Commun. Math. Phys.* **224**(2), 443–544 (2001)
66. Thomann, L.: Random data Cauchy problem for supercritical Schrödinger equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26**(6), 2385–2402 (2009)
67. Tsutsumi, M., Fukuda, I.: On solutions of the derivative nonlinear Schrödinger equation II. *Funkcial. Ekvac.* **24**(1), 85–94 (1981)
68. Vargas, A., Vega, L.: Global well-posedness for 1D Schrödinger equations for data with an infinite L^2 -norm. *J. Math. Pures Appl. (9)* **80**(10), 1029–1044 (2001)
69. Win, Y.: Global well-posedness of the derivative nonlinear Schrödinger equations on T . *Funkcial. Ekvac.* **53**, 51–88 (2010)
70. Yue, H.: Almost sure well-posedness for the cubic nonlinear Schrödinger equation in the super-critical regime on \mathbb{T}^d , $d \geq 3$. [arXiv:1808.00657](https://arxiv.org/abs/1808.00657) [math.AP]

Communicated by W. Schlag