Communications in
Mathematical Physics

Energy Dispersed Large Data Wave Maps in 2 + 1 Dimensions

Jacob Sterbenz[1],[*], Daniel Tataru[2],[**]

[1] Department of Mathematics, University of California, San Diego, CA 92093-0112, USA.
    E-mail: jsterben@math.ucsd.edu
[2] Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.
    E-mail: tataru@math.berkeley.edu

Abstract: In this article we consider large data Wave-Maps from $\mathbb{R}^{2+1}$ into a compact Riemannian manifold $(\mathcal{M}, g)$, and we prove that regularity and dispersive bounds persist as long as a certain type of bulk (non-dispersive) concentration is absent. This is a companion to our concurrent article [21], which together with the present work establishes a full regularity theory for large data Wave-Maps.

Contents

1. Introduction

In this article we consider finite energy large data Wave-Maps from the Minkowski space $\mathbb{R}^{2+1}$ into a compact Riemannian manifold $(\mathcal{M}, g)$. Our main result asserts that regularity and dispersive bounds persist as long as a certain type of bulk concentration is absent. The results proved here are used in the companion article [21] to establish a full regularity theory for large data Wave-Maps.

The set-up we consider is the same as the one in [33], using the so-called extrinsic formulation of the Wave-Maps equation. Precisely, we consider the target manifold $(\mathcal{M}, g)$ as an isometrically embedded submanifold of $\mathbb{R}^N$. Then we can view the $\mathcal{M}$ valued functions as $\mathbb{R}^N$ valued functions whose range is contained in $\mathcal{M}$. Such an embedding always exists by Nash's theorem [18] (see also Gromov [3] and Günther [4]). In this context the Wave-Maps equation can be expressed in a form which involves the second fundamental form $\mathcal{S}$ of $\mathcal{M}$, viewed as a symmetric bilinear form:

$$\mathcal{S} : T\mathcal{M} \times T\mathcal{M} \ \rightarrow \ N\mathcal{M}, \quad \langle S(X, Y), N \rangle = \langle \partial_X N, Y \rangle.$$

For the standard d'Allembertian in $\mathbb{R}^{2+1}$ we use the notation

$$\Box = \partial_t^2 - \Delta_x = -\partial^\alpha \partial_\alpha.$$

The Cauchy problem for the wave maps equation has the form:

$$\Box \phi^a = -\mathcal{S}_{bc}^a(\phi)\partial^\alpha \phi^b \partial_\alpha \phi^c, \quad \phi \in \mathbb{R}^N, \tag{1a}$$

$$\phi(0, x) = \phi_0(x), \quad \partial_t \phi(0, x) = \dot{\phi}_0(x), \tag{1b}$$

where the initial data $(\phi_0, \dot{\phi}_0)$ is chosen to obey the constraint:

$$\phi_0(x) \ \in \ \mathcal{M}, \quad \dot{\phi}_0(x) \ \in \ T_{\phi_0(x)}\mathcal{M}, \quad x \in \mathbb{R}^2.$$

In the sequel, it will be convenient for us to use the notation $\phi[t] = (\phi(t), \partial_t \phi(t))$. The system of Eqs. (1) admits a conserved quantity, namely the Dirichlet energy:

$$E[\phi(t)] := \int_{\mathbb{R}^2} |\partial_t \phi(t)|^2 + |\nabla_x \phi(t)|^2 dx := \| \phi[t] \|_{\dot{H}^1 \times L^2}^2 = E. \tag{2}$$

Finite energy solutions for (1) correspond to initial data in the energy space, namely $\phi[t] \in \dot{H}^1 \times L^2$. We call a Wave-Map "classical" on a bounded time interval $(t_0, t_1) \times \mathbb{R}^2$ if $\nabla_{x,t}\phi(t)$ belongs to the Schwartz class for all $t \in (t_0, t_1)$.

The Wave-Maps equation is also invariant with respect to the change of scale $\phi(t, x) \ \rightarrow \ \phi(\lambda t, \lambda x)$ for any positive $\lambda \in \mathbb{R}$. In $(2 + 1)$ dimensions, it is easy to

see that the energy $E[\phi]$ is dimensionless with respect to this scale transformation. For this reason, the problem we consider is called *energy critical*.

For the evolution (1), a local well-posedness theory in Sobolev spaces $H^s \times H^{s+1}$ for s above scaling, $s > 1$, was established some time ago. See [7] and [9], and references therein. The small data Cauchy-problem in the scale invariant Sobolev space is, by now, also well understood. Following work of the second author [32] for initial data in a scale invariant Besov space, Tao was the first to consider the wave map equation with small energy data. In the case when the target manifold is a sphere, Tao [29] proved global regularity and scattering for small energy solutions. This result was extended to the case of arbitrary compact target manifolds by the second author in [33]. Finite energy solutions were also introduced in [33] as unique strong limits of classical solutions, and the continuous dependence of the solutions with respect to the initial data was established. The case when the target is the hyperbolic plane was handled by Krieger [15]. There is also an extensive literature devoted to the more tractable higher dimensional case; we refer the reader to [8,14,17,28,31], and [20] for more information.

To measure the dispersive properties of solutions ϕ to the Wave-Maps equation, we shall use a variant of the standard dispersive norm S from [33]. This was originally defined in [29] by modifying a construction in [32]. S is used together with its companion space N which has the linear property (precise definitions will be given shortly):

$$\| \phi \|_{S[I]} \lesssim \| \phi \|_{L_t^\infty(L_x^\infty)[I]} + \| \phi[0] \|_{\dot{H}^1 \times L^2} + \| \Box \phi \|_{N[I]}.$$

The main result in [33] asserts that global regularity and scattering hold for the small energy critical problem:

Theorem 1.1. *The wave maps Eq. (1) is globally well-posed for small initial data $\phi[0] \in \dot{H}^1 \times L^2$ in the following sense:*

(i) *Classical Solutions. If the initial data $\phi[0]$ is constant outside of a compact set and C^∞, then there is a global classical solution ϕ with this data.*

(ii) *Finite Energy Solutions. For each small initial data set in $\phi[0] \in \dot{H}^1 \times L^2$ there is a global solution $\phi \in S$, obtained as the unique S limit of classical solutions, so that:*

$$\| \phi \|_S \lesssim \| \phi[0] \|_{\dot{H}^1 \times L^2}. \tag{3}$$

(iii) *Continuous dependence. The solution map $\phi[0] \to \phi$ from a small ball in $\dot{H}^1 \times L^2$ to S is continuous.*

We remark that due to the finite speed of propagation one can also state a local version of the above result, where the small energy initial data is taken in a ball, and the solution is defined in the corresponding uniqueness cone. This allows one to define large data finite energy solutions:

Definition 1.2. *Let I be a time interval. We say that ϕ is a finite energy wave map in I if $\phi[\cdot] \in C(I; \dot{H}^1 \times L^2)$ and, for each $(t_0, x_0) \in I$ and $r > 0$ so that $E[\phi(t_0)](B(x_0, r))$ is small enough, the solution ϕ coincides with the one given by Theorem 1.1 in the uniqueness cone $I \cap \{|x - x_0| + |t - t_0| \leqslant r\}$.*

In this work we consider a far more subtle case, which is a conditional version of the large data problem. It is first important to observe that for general targets the

above theorem cannot be extended to arbitrarily large C^∞ initial data, and that this failure can be attributed to several different mechanisms. For instance any harmonic map $\phi_0 : \mathbb{R}^2 \rightarrow \mathcal{M}$ yields a time independent wave-map which does not decay in time, therefore it does not belong to S. More interesting is that for certain non-convex targets, for example when we take $\mathcal{M} = \mathbb{S}^2$, finite time blow-up of smooth solutions is possible (see [13,19]). In this latter case, the blow-up occurs along a family of rescaled harmonic maps. To avoid such Harmonic-Map based solutions, as well as other possible concentration scenarios, in this article we prove a conditional regularity theorem:

Theorem 1.3 (Energy Dispersed Regularity Theorem). *There exist two functions* $1 \ll F(E)$ *and* $0 < \epsilon(E) \ll 1$ *of the energy* (2) *such that the following statement is true. If* ϕ *is a finite energy solution to* (1) *on the open interval* (t_1, t_2) *with energy* E *and:*

$$\sup_k \| P_k \phi \|_{L^\infty_{t,x}[(t_1,t_2) \times \mathbb{R}^2]} \leqslant \epsilon(E) \tag{4}$$

then one also has:

$$\| \phi \|_{S(t_1,t_2)} \leqslant F(E). \tag{5}$$

Finally, such a solution $\phi(t)$ *extends in a regular way to a neighborhood of the interval* $I = [t_1, t_2]$.

Remark 1.4. In Sect. 4, Theorem 4.1, we shall state a slightly stronger version of this result which uses the language of frequency envelopes from [29]. In particular, we will show the energy dispersion bound (4) implies that a certain range of subcritical Sobolev norms may only grow by a universal energy dependent factor. Put another way, one may interpret this restatement of Theorem 1.3 as saying that *in the energy dispersed scenario, the Wave-Maps equation becomes subcritical in the sense that there is a quasi-conserved norm of **higher** regularity than the physical energy*. This information, coupled with the standard regularity theory for Wave-Maps (e.g. see [33]) provides us with the continuation property.

Remark 1.5. The result in this article is stated and proved in space dimension $d = 2$. However, given its perturbative nature, one would expect to have a similar result in higher dimension $d \geqslant 3$ as well. That is indeed the case. There are two reasons why we have decided to stay with $d = 2$ here. One is to fix the notations. The second, and the more important reason, is to avoid lengthening the paper with an additional argument in Sect. 4, which is the only place in the article where the conservation of energy is used. In higher dimensions, this aspect would have to be replaced by an almost conservation of energy, with errors controlled by the energy dispersion parameter ϵ.

Remark 1.6. The proof of Theorem 1.3 allows us to obtain explicit formulas for $F(E)$ and $\epsilon(E)$. Precisely, in the conclusion of the proof of Corollary 4.4 below, we show that these parameters may be chosen of the form:

$$F(E) = e^{Ce^{E^M}}, \quad \epsilon(E) = e^{-Ce^{E^M}},$$

with C and M sufficiently large.

As a consequence of the frequency envelope version of this result in Theorem 4.1 we can also state a weaker non-conditional version of the above result:

Corollary 1.7. *There exists two functions* $1 \ll F(E)$ *and* $0 < \epsilon(E) \ll 1$ *of the energy* (2) *such that for each initial data* $\phi[0]$ *satisfying:*

$$\sup_k \| P_k \phi[0] \|_{\dot{H}^1 \times L^2} \leqslant \epsilon(E), \tag{6}$$

there exists a unique global finite energy solution $\phi \in S$, *satisfying:*

$$\| \phi \|_S \leqslant F(E), \tag{7}$$

which depends continuously on the initial data. If in addition the initial data is smooth, then the solution is also smooth.

Our main interest in Theorem 1.3 is to combine it with the results of our concurrent work [21], which together implies a full regularity theory for Wave-Maps. In this context, one may view Theorem 1.3 as providing a "compactness continuation" principle, which roughly states that there is the following dichotomy for classical Wave-Maps defined on the open time interval $(t_0, t_1) \times \mathbb{R}^2$:

(1) The solution ϕ continues to a neighborhood of the closed time interval $[t_0, t_1]$ as a classical Wave-Map.
(2) The solution ϕ exhibits a compactness property on a sequence of rescaled times.

In particular, the second case may be used with the energy estimates from [21] to conclude that a portion of any singular Wave-Map must become stationary, and via compactness must therefore rescale to a Harmonic-Map of non-trivial energy. This was known as the *bubbling conjecture* (see the introduction of [21] for more background).

Finally, we would like to remark that results similar in spirit to the ones of this paper and [21] have been recently announced. In the case where $\mathcal{M} = \mathbb{H}^n$, the hyperbolic spaces, global regularity and scattering follows from the program of Tao [22–24,26,30] and [25]. In the case where the target $\mathcal{M}$ is a negatively curved Riemann surface, Krieger and Schlag [16] provide global regularity and scattering via a modification of the Kenig-Merle method [6], which uses as a key component suitably defined Bahouri-Gerard [1] type decompositions.

1.1. A guide to reading the paper. The paper has a "two tier" structure, whose aim is to enable the reader to get quickly to the proof of the main result in Sect. 4. The first tier consists of Sects. 2, 3 and 4, which play the following roles:

Section 2 is where the notations are set-up. In addition, in Proposition 2.3 we review the linear, bilinear, trilinear and Moser estimates concerning the S and N spaces, as proved in [29,33]. The N space we use is the same as in [29,32]. For the S space we begin with the definition in [29] and add to it the Strichartz norm $\underline{S}$ defined later in (148). This modification costs almost nothing, but saves a considerable amount of work in several places.

Section 3 contains new contributions, reaching in several directions:

- *Renormalization.* A main difficulty in the study of wave maps is that the nonlinearity is non-perturbative at the critical energy level. A key breakthrough in the work of Tao [29] was a renormalization procedure whose aim is to remove the nonperturbative part of the nonlinearity. However, despite subsequent improvements in [33], this procedure only applies to the small data problem. We remedy this in Proposition 3.1, introducing a large data version of the renormalization procedure. This

applies without any reference to the energy dispersion bounds. We note that other large data renormalization procedures are available in certain cases, for instance by using the Coulomb or the caloric gauge.

- *S bounds for the paradifferential evolution with a large connection.* After peeling off the perturbative part of the nonlinearity in the wave map equation, one is left with a family of frequency localized linear paradifferential evolutions as in (38). In the case of the small data problem, by renormalization this turns into a small perturbation of the linear wave equation. Here this is no longer possible, as the connection coefficients A^α are large, and this cannot be improved using the energy dispersion. However, what the energy dispersion allows us to do is to produce a large frequency gap m in (38). As it turns out, this is all that is needed in order to have good estimates for Eq. (38).

- *New bilinear and trilinear estimates* which take advantage of the energy dispersion. The main bilinear bound is the L^2 estimate in Proposition 3.4. Ideally one would like to have such estimates for functions in S, but that is too much to ask. Instead we introduce a narrower class $\mathcal{W}$ of "renormalizable" functions ϕ of the form $\phi = U^\dagger w$, where $U \in S$ is a gauge transformation, while for w we control both $\|w\|_S$ and $\|\Box w\|_N$. As a consequence of Proposition 3.4 and the more standard bounds in Proposition 2.3, we later derive the trilinear estimates in Proposition 3.6, which are easy to apply subsequently in the proof of our main theorem.

Section 4 contains the proof of Theorem 4.1, which is a stronger frequency envelope version of Theorem 1.3. This is done via an induction on energy argument. The non-inductive part of the proof is separated into Propositions 4.2 and 4.3, whose aim is to bound in two steps the difference between a wave-map ϕ and a lower energy wave map $\widetilde{\phi}$ whose initial data is essentially obtained by truncating in frequency the initial data for ϕ. The arguments in this section use exclusively the results in Sects. 2, 3.

The second tier of the article contains the proofs of all the results stated in Sects. 2, 3, with the exception of those already proved in [29] and [33]. These are organized as follows:

Section 5's content is as follows:

- *A full description of the S and N spaces.* Some further properties of these spaces are detailed in Proposition 5.4; most of these are from [29] and [33], with the notable exception of the fungibility estimate (159). The bound (159) is proved using only the definition of N.

- *Extension properties for the S space.* In most of our analysis we do not work with the spaces S and N globally, instead we use their restrictions to time intervals, $S[I]$ and $N[I]$. This is not important for N, since the multiplication by a characteristic function of an interval is bounded on N. However, that is not the case for S. One can define the $S[I]$ norm using minimal extensions. But in our case, we also need good control of the energy dispersion and of the high modulation bounds for the extensions. To address this, in Proposition 5.5 we introduce a canonical way to define the extensions which obey the appropriate bounds, and which also produce an equivalent $S[I]$ norm.

- *Strichartz and L^2 bilinear estimates.* Using the U^p and V^p spaces[1] associated to the half-wave evolutions, we first show that solutions to the wave equation $\Box \phi = F$ with a right hand side $F \in N$ satisfy the full Strichartz estimates. The fungibility estimate (159) plays a significant role here, as it allows us to place the solution ϕ in a V^2 type space, see (195). A second goal is to prove L^2 bilinear bounds for products of

---

[1] For further information on the U^p and V^p spaces we refer the reader to [5,11,12].

two such inhomogeneous waves with frequency localization and angular frequency separation, see Lemma 5.10. This is accomplished using the Wolff [34]-Tao [27] type L^p bilinear estimates with $p < 2$.

Section 6 is devoted to the proof of the bilinear null form estimates in Proposition 3.4. A preliminary step, achieved in Lemma 6.1, is to establish the counterpart of the bounds (44) and (46) in the absence of the renormalization factor. The proofs here use only Lemma 5.10 and the estimates in Propositions 2.3, 5.4.

Section 7 contains the proof of the trilinear estimates in Proposition 3.6. There are a number of dyadic decompositions and multiple cases to consider, but this is largely routine, using either Proposition 3.4 or the estimates in Propositions 2.3 and 5.4.

Section 8 is concerned with the construction of the gauge transformation in Proposition 3.1. The discrete inductive construction in [29,33] is replaced with a continuous version which serves to insure that the renormalization matrices $U_{,<k}$ are exactly orthogonal. To allow for wave-maps which are large in S, we need to forego the simpler inductive way of proving S estimates for $U_{,k}$ and instead build them up in a less direct fashion using iterated paradifferential type expansions. On the positive side, this proof uses only the estimates in Propositions 2.3, 5.4.

Section 9 is devoted to the proof of the linear bounds for the paradifferential equation in Proposition 3.2. A key element in this proof is the gauge transformation in Proposition 3.1, combined with the trilinear estimate (25). This would suffice for connections A_α arising from wave maps ϕ which are small in S. However, in our case we need to handle large wave maps, and a different source of smallness is required. This is provided by the large size of the frequency gap m, which leads to energy conservation with small $O(2^{-cm})$ errors. Feeding these almost apriori energy and characteristic energy bounds back in the bilinear and trilinear null form estimates turns out to suffice to estimate the large trilinear contributions, again modulo terms which are small, i.e. $O(2^{-cm})$.

Section 10's goal is to provide the description of finite S norm wave-maps in Proposition 3.9. The renormalization bound is a direct consequence of Proposition 3.2 and Proposition 3.1. The partial fungibility of the S norm is tied to the fungibility of the N norm in the renormalized setting, although the proof is somewhat more technical.

Under the assumption of small energy dispersion, the smallness of high modulations is given by the trilinear estimate (51). After that, the bound (55) combined with Proposition 3.1 lead quickly to the frequency envelope control in (67) via a bootstrap argument.

Section 11 contains the proof of the data truncation result in Proposition 3.5. The argument is self-contained and uses only L^2 type methods.

2. Standard Constructions, Function Spaces, and Estimates

In this section we record the standard portion of the framework we shall use in our primary demonstration of Theorem 1.3. While we aim to keep our account of things self contained, we also assume that the reader is thoroughly familiar with the content of the two papers [29] and [33]. For the sake of completeness, in Sect. 5 below we also include proofs of several results not contained in these two works, but which are needed for the more detailed analysis of this paper. Further notation and estimates that are not needed for the "first tier" of our demonstration of Theorem 1.3, but are needed in later technical sections are also given in Sect. 5 below.

The symbols $\lesssim, \gtrsim, \sim, \ll$, and $\gg$ are defined with their usual meanings. The constants in these notations are allowed to vary from line to line.

2.1. Constants. There will be a number of large and small constants in the present work. For the most part these are flexible, although the specific construction of $F(E)$ and $\epsilon(E)$ from Theorem 1.3 will be sensitive to each other as well as to choices of other constants. Lower case Greek letters such as δ, ϵ, and η will always denote small quantities. We shall employ a *globally* defined string of small constants:

$$\delta_0 \ll \delta_1 \ll \delta_2 \ll 1, \quad \delta_i \leqslant \delta_{i+1}^{100}. \tag{8}$$

As it occurs often in the sequel, we will set $\delta = \delta_2$ throughout. For the convenience of the reader we list here the purposes of these constants, which all measure various fractional frequency gains in our dyadic estimates:

- The base constant δ enters our proof through the various multilinear estimates for the S and N spaces listed below (e.g. in the current section); it measures for instance various dyadic gains in estimates from [29] and [33]. It also influences any portion of our argument which is a direct consequence of these estimates, but has nothing to do with directly bootstrapping large data Wave-Maps. For example, δ also represents various dyadic gains in our gauge construction (see Proposition 3.1).
- The constant δ_1 measures a small fractional gain coming from energy dispersion in L^2 and N-norm null form estimates. It enters our proof through estimates (51) and (52), and variations thereof.
- The constant δ_0 is reserved for slowly varying frequency envelopes, and for the smallest fractional quantities built from the energy dispersion constant ϵ. It enters in the core part of our proof of Theorem 1.3, and is the assumption on the frequency envelopes of Proposition (3.6).

Large quantities, for example C, F, K, and M will be used in various contexts as constants in estimates and the size of norms which are *not* globally defined. We will also often use m to denote a (possibly) large integer which represents various gaps in frequency truncations.

To denote growth and dependence of various estimates on that growth we employ the following notation in the sequel:

Definition 2.1 (Complexity Notation). *We say that a positive function $f(y)$ is of "polynomial type" if $f(y) \leqslant y^M$ for some constant M as $y \to \infty$. We use the notation:*

$$A \lesssim_F B,$$

*if $A \leqslant K(F)B$ for some function K of polynomial type. This notation does **not** fix K from line to line, although K is fixed on any single line where it occurs.*

2.2. Basic harmonic analysis. As usual we denote by ξ and τ the spatial and temporal Fourier variables (resp). We set up both discrete and continuous spatial Littlewood-Paley (LP) multipliers:

$$I - P_{-\infty} = \sum_k P_k, \quad I - P_{-\infty} = \int_{-\infty}^{\infty} P_k dk.$$

For the purposes of trichotomy, these two sets of multipliers are interchangeable, and we will only distinguish them by the use of $\sum$ or $\int$ in identities. However, for the purposes

of proving Moser type estimates or constructing gauge transformations, the integral definition of LP projections is essential. We refer the reader to [33] for an earlier use of continuous LP multipliers, and further information. We often denote by $\phi_k = P_k\phi$. If ϕ is any affinely Schwartz function, the above notation means that we have the identities:

$$\phi - \lim_{|x| \to \infty} \phi(x) = \sum_k \phi_k = \int_{-\infty}^{\infty} \phi_k \, dk.$$

Therefore, care must be taken to add constants back into certain estimates involving very low frequencies.

Many times in the sequel we shall have use for the inequality:

$$\| P_{\mathcal{B}}\phi \|_{L_x^q} \lesssim |\mathcal{B}|^{(\frac{1}{r} - \frac{1}{q})} \| \phi \|_{L_x^r}, \tag{9}$$

where $\mathcal{B} \subseteq \mathbb{R}_\xi^2$ is a frequency box. Furthermore, the P_k multipliers enjoy a commutator structure as follows:

$$P_k(\phi\psi) = \phi\psi_k + L(\nabla_x\phi, 2^{-k}\psi), \tag{10}$$

where the bilinear form L is translation invariant and bounded on all Lebesgue type spaces. Such multilinear expressions occur often in the sequel. We call a multilinear form L of the form:

$$L(\phi^{(1)}, \ldots, \phi^{(k)})(x) = \int \phi^{(1)}(x + y_1) \cdots \phi^{(k)}(x + y_k) d\mu(y_1, \ldots, y_k),$$

where

$$\int |d\mu| \lesssim 1,$$

"disposable". Any disposable operator generates a family of estimates from any single product estimate involving translation invariant norms in the usual way (see [29]).

We will also use the variable notation for frequency envelopes from [33] (see [29] for another definition):

Definition 2.2 (Frequency Envelopes). *A frequency envelope $\{c_k\}$ is called "(σ, Δ)-admissible" if it obeys the bounds:*

$$2^{\sigma(j-k)}c_k \leqslant c_j \leqslant 2^{\Delta(k-j)}c_k, \tag{11}$$

for any $j < k$, where $0 < \sigma \leqslant \Delta$. If $\| \phi \|_Y$ is any non-negative real valued functional, and $\{c_k\}$ is a frequency envelope, we define:

$$\| \phi \|_{Y_c} := \sup_k c_k^{-1} \| P_k\phi \|_Y. \tag{12}$$

There is an exception to this notation for the norm $S[I]$ introduced below, in which case we set:

$$\| \phi \|_{S_c[I]} := \| \phi \|_{L_t^\infty(L_x^\infty)[I]} + \sup_k c_k^{-1} \| \phi \|_{S_k[I]}.$$

Frequency envelopes may be defined in either the discrete or continuous settings. It is easy to see that for any such frequency envelope we have the pair of sum rules (uniformly):

$$\sum_{k'\leqslant k} 2^{Ak'} c_{k'} \lesssim (A-\Delta)^{-1} 2^{Ak} c_k, \quad A > \Delta, \tag{13}$$

$$\sum_{k\leqslant k'} 2^{-ak'} c_{k'} \lesssim (a-\sigma)^{-1} 2^{-ak} c_k, \quad a > \sigma, \tag{14}$$

with similar bounds for integrals. These two inequalities capture the essence of every use we have for the $\{c_k\}$ notation, which is simply to bookkeep (resp.) $Low \times Low \Rightarrow High$ and $High \times High \Rightarrow Low$ frequency cascades.

2.3. Function spaces and standard estimates. We use the function spaces S and N from [32,33] and [29] with only a few minor modifications. The spaces of restrictions of S and N functions to a time interval I are denoted by $S[I]$, respectively $N[I]$, with the induced norms. The first part of our proof does not use the precise structure of these spaces, only the following statement:

Proposition 2.3 (Standard Estimates and Relations: Part I). *Let F, ϕ, and $\phi^{(i)}$ be a collection of test functions, $I \subseteq \mathbb{R}$ any subinterval (including $\mathbb{R}$ itself). Then there exists function spaces $S[I]$ and $N[I]$ with the following properties:*

- *Triangle Inequality for S. Let $I = \cup_i^K I_i$ be a decomposition of I into consecutive intervals, then the following bounds hold (uniform in K):*

$$\|\phi\|_{S[I]} \lesssim \sum_i \|\phi\|_{S[I_i]}. \tag{15}$$

- *Frequency Orthogonality. The spaces $S[I]$ and $N[I]$ are made up of dyadic pieces in the sense that:*

$$\|\phi\|_{S[I]}^2 = \|\phi\|_{L_t^\infty(L_x^\infty)[I]}^2 + \sum_k \|P_k\phi\|_{S[I]}^2, \tag{16}$$

$$\|\phi\|_{N[I]}^2 = \sum_k \|P_k\phi\|_{N[I]}^2. \tag{17}$$

- *Energy Estimates. We have that $L_t^1(L_x^2)[I] \subseteq N[I]$, and also the estimate:*

$$\|\phi_k\|_{S[I]} \lesssim \|\Box\phi_k\|_{N[I]} + \|\phi_k[0]\|_{\dot{H}^1 \times L^2}. \tag{18}$$

- *Core Product Estimates. We have that:*

$$\|\phi_{<k+O(1)}^{(1)} \cdot \phi_k^{(2)}\|_{S[I]} \lesssim \|\phi_{<k+O(1)}^{(1)}\|_{S[I]} \cdot \|\phi_k^{(2)}\|_{S[I]}, \tag{19}$$

$$\|P_k(\phi_{k_1}^{(1)} \cdot \phi_{k_2}^{(2)})\|_{S[I]} \lesssim 2^{-(\max\{k_i\}-k)} \|\phi_{k_1}^{(1)}\|_{S[I]} \cdot \|\phi_{k_2}^{(2)}\|_{S[I]}, \tag{20}$$

$$\|P_k(\phi_{<k+O(1)} \cdot F_k)\|_{N[I]} \lesssim \|\phi\|_{S[I]} \cdot \|F_k\|_{N[I]}, \tag{21}$$

$$\|P_k(\phi_{k_1} \cdot F_{k_2})\|_{N[I]} \lesssim 2^{-\delta(k-k_2)_+} \|\phi_{k_1}\|_{S[I]} \cdot \|F_{k_2}\|_{N[I]}. \tag{22}$$

- *Bilinear Null Form Estimates. We have that:*

$$\| P_k \left(\partial^\alpha \phi_{k_1}^{(1)} \cdot \partial_\alpha \phi_{k_2}^{(2)} \right) \|_{L_t^2(L_x^2)[I]} \lesssim 2^{\frac{1}{2}\min\{k_i\}} 2^{-(\frac{1}{2}+\delta)(\max\{k_i\}-k)} \prod_i \|\phi_{k_i}^{(i)}\|_{S[I]}, \quad (23)$$

$$\| P_k(\partial^\alpha \phi_{k_1}^{(1)} \cdot \partial_\alpha \phi_{k_2}^{(2)}) \|_{N[I]} \lesssim 2^{-\delta(\max\{k_i\}-k)} \prod_i \| \phi_{k_i}^{(i)} \|_{S[I]}. \quad (24)$$

- *Trilinear Null Form Estimate. We have that:*

$$\| P_k(\phi_{k_1}^{(1)} \cdot \partial^\alpha \phi_{k_2}^{(2)} \cdot \partial_\alpha \phi_{k_3}^{(3)}) \|_{N[I]} \lesssim 2^{-\delta(\max\{k_i\}-k)} 2^{-\delta(k_1-\min\{k_2,k_3\})_+} \prod_i \| \phi_{k_i}^{(i)} \|_{S[I]}. \quad (25)$$

- *Moser Estimates. Let G be any bounded function with uniformly bounded derivatives, and $\{c_k\}$ a (δ, Δ)-admissible frequency envelope. Then there exists a universal $K > 0$ such that:*

$$\| G(\phi) \|_{S[I]} \lesssim \| \phi \|_{S[I]}(1 + \| \phi \|_{S[I]}^K), \quad (26)$$

$$\| G(\phi) \|_{S_c[I]} \lesssim \| \phi \|_{S_c[I]}(1 + \| \phi \|_{S[I]}^K). \quad (27)$$

The space N is the same one as used in [29,33]. To obtain the space S we start with the one used in [29,33] and add the control of the Strichartz norms $\underline{S}$ defined in (148). The bound (15) is relatively straightforward; we prove it in Sect. 5. The relations (16) and (17) can be thought of as a part of the definition of the spaces S, N starting from their dyadic versions. The linear estimate (18) was proved in [29]; here we show that we can add the Strichartz component $\underline{S}$ in Corollary 5.9. The bounds (19)–(22) as well as (24),(25) were proved in [29]. The one unit gain in the exponent in (20) is not explicitly stated in [29], but it is implicit in the proof. In our context the proofs of (19), (20) need to be augmented to add the control over the Strichartz norm $\underline{S}$; this is a straightforward matter which is left for the reader. The bound (23) is implicit in [29], but for the reader's convenience we prove it in Sect. 6.

The Moser estimates (26) and (27) were proved in [33]. Adding in the $\underline{S}$ norm is again straightforward. An interesting side remark is that in effect the addition of the $\underline{S}$ norm to S can be taken advantage of to simplify considerably the proof of the Moser estimates in [33]. In particular, one can show that it is possible to take $K = 2$. Since it does not lead to significant improvements in the present article, we leave this as an exercise for the reader.

At several places in our argument, it will be necessary for us to introduce some auxiliary norms. We choose to keep these separate from S defined above for notational purposes:

Definition 2.4 (Auxiliary Energy and $X^{s,b}$ Type Norms). *We define:*

$$\| \phi \|_{\underline{E}[I]} := \| \nabla_{t,x}\phi \|_{L_t^\infty(L_x^2)[I]} + \sup_\omega \| \overline{\nabla}_{t,x}^\omega \phi \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)[I]}, \quad (28)$$

$$\| \phi \|_{\underline{X}_k[I]} := 2^{-\frac{1}{2}k} \| \Box P_k\phi \|_{L_t^2(L_x^2)[I]}. \quad (29)$$

Here the second term in the RHS of (28) represents the energy of ϕ on characteristic hyperplanes, see [32,29]. We also define $\underline{X}[I]$ as the square sum of $\underline{X}_k[I]$, and $\underline{X}_c[I]$ according to (12). Notice that there are no square sums or frequency localizations in the norm $\underline{E}$. The size of this norm depends only on the initial energy of any (global) classical solution to (1).

In the sequel, it will be also be notationally convenient for us to work with the following definition which one should think of as a variant of the $S[I]$ space introduced above. The reader should keep in mind that this not even a quasinorm due to the lack of any good additivity property:

Definition 2.5 (Renormalizable Functions). *Let $C > 0$ be a large parameter. We define a non-linear functional $\mathcal{W}_k$ on S as follows:*

$$\| \phi \|_{\mathcal{W}_k} := \inf_{U \in SO(d)} \left[\left(\| U \|_{S \cap \underline{X}} + \sup_{j \geqslant k} 2^{C(j-k)} \| P_j U \|_{S \cap \underline{X}} \right) \right.$$
$$\left. \cdot \sup_{k'} 2^{|k'-k|} \left(\| P_{k'}(U \phi_k)[0] \|_{\dot{H}^1 \times L^2} + \| P_{k'} \Box (U \phi_k) \|_N \right) \right]. \quad (30)$$

The functionals $\mathcal{W}[I]$, $\mathcal{W}_c[I]$ are also defined as above.

Notice that while the definition of $\mathcal{W}$ is nonlinear, one still has the scaling relation $\| \lambda \phi \|_{\mathcal{W}[I]} = \lambda \| \phi \|_{\mathcal{W}[I]}$. The reader should note that while these bounds are cumbersome to state, they are all natural in light of Propositions 3.1–3.2 below.

3. New Estimates and Intermediate Constructions

In this section we introduce the main technical components of the paper. We begin with the core underlying tools that allow us to handle more complicated constructions. In a later sub-section we derive some further useful results that encapsulate many of the repetitive computations in the sequel.

3.1. Core technical estimates and constructions. The right-hand side in Eq. (1) is non-perturbative even when the energy is small. In the case of larger energies, it becomes quite a bit more difficult to handle things in a perturbative manner. Therefore, we introduce a set of tools which are general enough to handle large data situations. The first two of these work without any additional properties (e.g. energy dispersion), and form the technical heart of the paper. The first is a novel gauge construction that should be of more general use. It should be noted that this construction is stable regardless of the size of the energy or the convexity properties of the target, as its key properties depend only on the compactness of the underlying gauge group.

Proposition 3.1 (The "Diffusion Gauge"). *Let ϕ be a wave-map in a time interval I with energy E, $S[I]$ norm F, and $S[I]$ norm (δ, Δ)-admissible envelope $\{c_k\}$. Let the antisymmetric B be defined by:*

$$(B_b^a)_{<k} = \int_{-\infty}^{k} \left(\mathcal{S}_{bc}^a(\phi) - \mathcal{S}_{ac}^b(\phi) \right)_{<k'-10} \phi_{k'}^c \, dk', \quad (31)$$

where $\mathcal{S}_{bc}^a$ is a smoothly bounded (a, b) symmetric matrix valued vector. We denote the integrand by B_k. Then for each real number k there exists an orthogonal matrix $U_{,<k}$ defined on all of $\mathbb{R}^{2+1}$ with the following properties:

- $U_{,<k}$ *is a Sum of Frequency Localized Pieces in* S_k. *For each real number k there exists a matrix* $U_{,k}$ *such that:*

$$U_{,<k} = \int_{-\infty}^{k} U_{,k'} dk',$$

where each $U_{,k} = U_{,<k} B_k$, *and each* $U_{,k}$ *obeys the bounds:*

$$\| P_{k'} U_{,k} \|_{S \cap \underline{X}} \lesssim_F 2^{-\delta|k-k'|} 2^{-C(k'-k)_+} c_k, \tag{32}$$

$$\| P_{k'} \nabla_{t,x}^J U_{,k} \|_{L_t^1(L_x^1)} \lesssim_F 2^{(|J|-3)k} 2^{-C(k'-k)} c_k, \quad k' > k + 10, \quad |J| \leqslant 2, \tag{33}$$

$$\| P_{k'} \left(U_{,<k-20} \cdot G_k \right) \|_N \lesssim_F 2^{-|k'-k|} \| G_k \|_N, \tag{34}$$

$$\| P_k \left(\Box U_{,k_1} \cdot \psi_{k_2} \right) \|_N \lesssim_F 2^{-|k-k_2|} 2^{-\delta(k_2-k_1)} c_{k_1} \| \psi_{k_2} \|_S, \quad k_1 < k_2 - 10. \tag{35}$$

In addition, if $\tilde{c}_k$ *is a* (δ_0, Δ)-*admissible frequency envelope for the energy* $\| \nabla_{t,x} \phi_k \|_{L_t^\infty(L_x^2)[I]}$ *then we have a similar bound for* $U_{,k}$:

$$\| P_{k'} \nabla_{t,x} U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim_E 2^{-|k-k'|-C(k'-k)_+} \tilde{c}_k. \tag{36}$$

Here $C \geqslant 0$ *is any constant.*
- *The Matrix U Approximately Renormalizes* $A_\alpha = \nabla_\alpha B$. *We have the formula:*

$$U_{,<k}^\dagger \nabla_\alpha U_{,<k} = \nabla_\alpha B_{<k} - \int_{-\infty}^{k} \left[B_{k'}, U_{,<k'}^\dagger \nabla_\alpha U_{,<k'} \right] dk'. \tag{37}$$

This result is proved in Sect. 8. Next, we state a technical proposition that will help us to deal with the non-fungibility of the S norm. The wave map nonlinearity is nonperturbative. However, due to the small energy dispersion, at fixed frequency we are able to perturbatively replace the nonlinearity in the wave map equation with a paradifferential term, i.e. a linear term involving the lower frequencies of the wave map. This term is large, and due to the non-fungibility of the S norm, it cannot be made small on small time intervals. Fortunately, it has another redeeming feature, namely a large frequency gap (see m below). We take advantage of this in Sect. 9 to prove that:

Proposition 3.2 (Gauge Covariant $S[I]$ Estimate). *Let* $\psi_k = P_k \psi$ *be a solution to the linear problem:*

$$\Box \psi_k = -2A_{<k-m}^\alpha \partial_\alpha \psi_k + G, \tag{38}$$

where $A_{<k-m}^\alpha$ *is the* $\mathfrak{so}(N)$ *matrix:*

$$(A_{<k-m}^\alpha)_b^a = \left(\mathcal{S}_{bc}^a(\phi) - \mathcal{S}_{ac}^b(\phi) \right)_{<k-m} \partial^\alpha \phi_{<k-m}^c. \tag{39}$$

Assume that ϕ *is a classical Wave-Map on I with the bounds:*

$$\| \phi \|_{E[I]} + \| \phi \|_{X[I]} + \| \phi \|_{S[I]} \leqslant F. \tag{40}$$

Furthermore, assume that $m \geqslant m(F) > 20$, *for a certain function* $m(F) \sim \ln(F)$ *(to be defined in the proof). Then we have the estimate:*

$$\| \psi_k \|_{\mathcal{W}[I]} \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + \| G \|_{N[I]}. \tag{41}$$

Remark 3.3. As will become apparent in the proof of estimate (41), the *only* use of the large frequency gap parameter m is to be able to bootstrap the RHS involving ψ_k. In the sequel, there will be situations where one already has good $S[I]$ norm bounds on ψ_k, and the task is to provide a renormalization $w_{,k}$ such that $\Box w_{,k}$ has good N norm bounds. Therefore, we state the following:

- Let ψ_k, $A^\alpha_{<k-m}$, and G be defined as in Proposition 3.2. Then by simply assuming that $m > 20$ we have the following estimate for ψ_k:

$$\| \psi_k \|_{\mathcal{W}[I]} \lesssim_F \| \psi_k \|_{S[I]} + \| P_k G \|_{N[I]}. \tag{42}$$

- Furthermore, in the above situation, the renormalization on the LHS of estimate (42) is given by a matrix as in Proposition 3.1 where the pieces B_j are defined from $A^\alpha_{<k-m}$ in the obvious way (this is of course true for estimate (41) as well).

See Remark 9.4 in Sect. 9 for more details.

Next, we state a gauged version of certain improved multilinear estimates for the wave equation. Roughly speaking, these estimates imply that matched frequency interactions in the RHS of (1) behave in a perturbative fashion in the presence of energy dispersion. The heart of these estimates lies in the Wolff-Tao bilinear estimates (see [34] and [27]) for non-parallel waves, and the parallel wave cancelation property of the "Q_0 null structure" which was originally investigated in [7]:

Proposition 3.4 (Improved Matched Frequency Estimates). *Let $\phi^{(i)}_{k_i}$ be functions localized at frequency k_i for $i = 1, 2$. Assume that these functions are normalized as follows:*

$$\| \phi^{(i)}_{k_i} \|_{\mathcal{W}[I]} \leqslant 1, \quad \| \phi^{(1)}_{k_1} \|_{L^\infty_t(L^\infty_x)[I]} \leqslant \eta. \tag{43}$$

Then the following estimates hold:

- *Bilinear L^2 Estimate. We have that:*

$$\| \partial^\alpha \phi^{(1)}_{k_1} \partial_\alpha \phi^{(2)}_{k_2} \|_{L^2_t(L^2_x)[I]} \lesssim 2^{\frac{1}{2}\max\{k_1,k_2\}} \eta^\delta. \tag{44}$$

- *Bilinear N Estimate. Assume that in addition to (43) we also have the high modulation bounds:*

$$\| \Box\phi^{(1)}_{k_1} \|_{L^2_t(L^2_x)[I]} \leqslant 2^{\frac{k_1}{2}} \eta, \quad \| \Box\phi^{(2)}_{k_2} \|_{L^2_t(L^2_x)[I]} \leqslant 2^{\frac{k_2}{2}} \eta. \tag{45}$$

Then the following estimate holds:

$$\| \partial^\alpha \phi^{(1)}_{k_1} \partial_\alpha \phi^{(2)}_{k_2} \|_{N[I]} \lesssim 2^{C|k_1-k_2|} \eta^\delta. \tag{46}$$

This is proved in Sect. 6. Finally, we list a technical result concerning initial data frequency truncation. This does not preserve the space of functions with values in $T\mathcal{M}$, so it has to be followed by a non-linear physical space projection Π back onto $T\mathcal{M}$. We will show that in the energy dispersed case, this operation is very well behaved in the energy norm. Theorems of this type may be useful for other problems involving the need for a "non-linear Littlewood-Paley theory" of functions with values in a manifold:

Proposition 3.5. *For each $E > 0$ there exists $\epsilon_0 > 0$ so that for each initial data set $\phi[0]$ for (1) with energy E and energy dispersion $\epsilon \leqslant \epsilon_0$ and $k, k_* \in \mathbb{Z}$ we have*

$$\| P_k \left(P_{<k_*}\phi[0] - \Pi(P_{<k_*}\phi[0]) \right) \|_{\dot{H}^1 \times L^2} \lesssim_E \epsilon^{\frac{1}{4}} 2^{-\frac{1}{2}|k-k_*|}. \tag{47}$$

This is proved in Sect. 11 using Moser estimates and some integral identities involving the continuous Littlewood-Paley theory developed in [33].

3.2. Derived estimates and intermediate constructions. A corollary of the above Propositions is the following, which will be needed for the proof of our main theorem. The reader should keep in mind that this proposition is merely a bookkeeping device that will allow us to avoid many repetitive calculations in the sequel:

Proposition 3.6 (Improved Multilinear Estimates). *Let $\phi^{(i)}$ be three test functions defined on a time interval I normalized so that:*

$$\| \phi^{(1)} \|_{S[I]} \leqslant 1, \quad \sup_{i=2,3} \| \phi^{(i)} \|_{\mathcal{W}[I]} \leqslant 1. \tag{48}$$

Suppose in addition that $\phi^{(2)}$ has the improved energy dispersion bound on I:

$$\sup_k \| P_k \phi^{(2)} \|_{L^\infty[I]} \leqslant \eta. \tag{49}$$

Finally, let $\{c_k\}$ be any (δ_0, δ_0)-admissible frequency envelope, and $0 \leqslant m$ an additional integer subject to the condition:

$$m \leqslant \sqrt{\delta_1} |\ln(\eta)|. \tag{50}$$

Then one has the following multilinear bounds:

i) *Core Trilinear L^2 Estimate. Suppose along with the above assumptions that $\phi^{(3)}$ has unit $\mathcal{W}_c[I]$ norm for the frequency envelope $\{c_k\}$. Then for any disposable trilinear form L we have the bound:*

$$\| L(\phi^{(1)}, \partial^\alpha \phi^{(2)}, \partial_\alpha \phi^{(3)}) \|_{L^2_t(\dot{H}^{-\frac{1}{2}})_c[I]} \lesssim \eta^{\delta_1}. \tag{51}$$

ii) *Additional Trilinear L^2 Estimate. Suppose again that we have the conditions (48)–(49), and that this time $\phi^{(1)}$ has unit $S_c[I]$ norm for the frequency envelope $\{c_k\}$. Then for any disposable trilinear form L we have the bound:*

$$\| P_k L(\phi^{(1)}, \partial^\alpha \phi^{(2)}, \partial_\alpha \phi^{(3)}) \|_{L^2_t(L^2_x)[I]} \lesssim 2^{\frac{k}{2}} \eta^{\delta_1} \left(c_k + \| P_{<k} \phi^{(1)} \|_{S[I]} \right). \tag{52}$$

iii) *Core Trilinear N Estimate. For a positive integer m and integer k and disposable trilinear form L, define the following trilinear form:*

$$T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) := P_k L(\phi^{(1)}, \partial^\alpha \phi^{(2)}, \partial_\alpha \phi^{(3)})$$
$$- L(\phi^{(1)}_{<k-m}, \partial^\alpha \phi^{(2)}_{<k-m}, \partial_\alpha \phi^{(3)}_k) - L(\phi^{(1)}_{<k-m}, \partial^\alpha \phi^{(2)}_k, \partial_\alpha \phi^{(3)}_{<k-m}). \tag{53}$$

Suppose in addition to the (48)–(49) we also have unit $\mathcal{W}_c[I]$ norm of $\phi^{(3)}$, and furthermore the high modulation bounds:

$$\| \phi^{(2)} \|_{\underline{X}[I]} \leqslant \eta, \quad \| \phi^{(3)} \|_{\underline{X}_c[I]} \leqslant \eta. \tag{54}$$

Then the following trilinear estimate holds:

$$\| T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim \eta^{\delta_1} c_k. \tag{55}$$

iv) *Additional Trilinear N Estimate.* Suppose in addition to (48)–(49) we have unit $S_c[I]$ norm of $\phi^{(1)}$, and in addition the high modulation bounds:

$$\sup_{i=2,3} \| \phi^{(i)} \|_{\underline{X}[I]} \leqslant \eta. \tag{56}$$

Then if T_k^m is defined as on line (53) we have the bound:

$$\| T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim \eta^{\delta_1} \left(c_k + \| P_{<k}\phi^{(1)} \|_{S[I]} \right). \tag{57}$$

Remark 3.7. If the functions $\phi^{(i)}$ admit a common frequency envelope $\{c_k\}$ then we can relax the admissibility condition on $\{c_k\}$ and work with (δ_0, Δ) frequency envelopes. Precisely, for any (δ_0, Δ)-admissible frequency envelope $\{c_k\}$ we have the following:

- If (48) is replaced by

$$\| \phi^{(1)} \|_{S_c[I]} \leqslant 1, \quad \sup_{i=2,3} \| \phi^{(i)} \|_{\mathcal{W}_c[I]} \leqslant 1, \tag{58}$$

  then (51) follows.
- If in addition (54) is replaced by

$$\| \phi^{(2)} \|_{\underline{X}_c[I]} \leqslant \eta, \quad \| \phi^{(3)} \|_{\underline{X}_c[I]} \leqslant \eta, \tag{59}$$

  then the following version of (55) holds:

$$\| T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim \eta^{\delta_1} c_k. \tag{60}$$

Remark 3.8. As will become apparent in the sequel, the *only* use of the renormalized norms $\mathcal{W}[I]$ and the high modulation bounds $\underline{X}[I]$ in the estimates of Proposition 3.6 is to ensure the smallness coming from the parameter η. Thus, under the simpler assumption that the $\phi^{(i)}$ are only normalized so that $\| \phi^{(i)} \|_{S[I]} \leqslant 1$ we have the following:

- If $\phi^{(3)}$ has (δ_0, δ_0)-admissible $S[I]$ norm frequency envelope $\{c_k\}$, then estimate (51) holds with $\eta = 1$.
- If $\phi^{(3)}$ has a (δ_0, δ_0)-admissible $S[I]$ norm frequency envelope $\{c_k\}$, and if we let $m \geqslant 10$ be any integer, then we may replace estimate (55) with the bound:

$$\| T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim 2^{4\delta_0 m} c_k. \tag{61}$$

- If $\phi^{(1)}$ has (δ_0, δ_0)-admissible $S[I]$ norm frequency envelope $\{c_k\}$, and if we let $m \geqslant 0$ be any integer, then we may replace estimate (57) with the bound:

$$\| T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim 2^{4\delta_0 m} \left(c_k + \| P_{<k}\phi^{(1)} \|_{S[I]} \right). \tag{62}$$

For further details, see Remarks 7.1, 7.2 and 7.3 in Sect. 7 below.

Next, we state a result that ties together many of the previous propositions. This is a structure theorem for large data wave-maps which says that in the presence of good $S[I]$ norm bounds one has some additional regularity properties, as well as a crucial "fungibility" property that is central to energy norm inductions.

Proposition 3.9 (Structure of Finite S Norm Wave-Maps). *Let ϕ be a wave-map defined on the interval I with energy E and S norm F. Then the following is true:*

- *Additional Norm Control. We have the bounds:*

$$\| \phi \|_{\underline{X}[I]} + \| \phi \|_{\underline{E}[I]} \lesssim_F 1. \tag{63}$$

- *Renormalization. If $\{c_k\}$ is a (δ_0, Δ)-admissible frequency envelope for $\| \phi \|_{S[I]}$, then we may renormalize our wave-map as follows:*

$$\| \phi \|_{\mathcal{W}_c[I]} \lesssim_F 1. \tag{64}$$

- *Partial Fungibility. If $\| \phi \|_{S[I]} = F$, then there exists a collection of subintervals $I = \cup_{i=1}^K I_i$, such that $K = K(F)$ depends only on F, and such that the following bound holds on each I_i:*

$$\| \phi \|_{S[I_i]} \lesssim_E 1. \tag{65}$$

- *Smallness of High Modulations. Suppose in addition that we have energy dispersion $\sup_k \| P_k \phi \|_{L_t^\infty (L_x^\infty)[I]} \leqslant \epsilon$. Then we also have the estimate:*

$$\| \phi \|_{\underline{X}[I]} \lesssim_F \epsilon^{\delta_1}. \tag{66}$$

- *Frequency Envelope Control. Suppose that ϕ has sufficiently small energy dispersion $\epsilon < \epsilon(F)$. Then if $\{c_k\}$ is a (δ_0, Δ)-admissible $\dot{H}^1 \times L^2$ frequency envelope for $\phi[0]$ we have:*

$$\| \phi_k \|_{S[I]} \lesssim_F c_k. \tag{67}$$

Proposition 3.9 is proved in Sect. 10.

Finally, for the reader's convenience we group together the results which enable us to carry out our bootstrapping arguments:

Proposition 3.10 (Bootstrapping Tool). *Let $I = [a, b]$ be an interval and c a (δ_0, Δ) frequency envelope. Then for each affinely Schwartz function ϕ in I the following properties hold:*

- *Seed S bound. Let $I_n \subset I$ be a decreasing sequence of intervals which converges to the point $t = 0$. Then:*

$$\lim_{n \to \infty} \| \phi \|_{S[I_n]} \lesssim \| \phi[0] \|_{\dot{H}^1 \times L^2}, \quad \lim_{n \to \infty} \| \phi \|_{S_c[I_n]} \lesssim \| \phi[0] \|_{(\dot{H}^1 \times L^2)_c}. \tag{68}$$

- *Continuity Properties. For each subinterval $J \subset I$ we have $\phi \in S[J] \cap S_c[J]$, and its S norm $\| \phi \|_{S[J]}$, its S_c norm $\| \phi \|_{S_c[J]}$, and its energy dispersion norm $\sup_k \| P_k \phi \|_{L_t^\infty (L_x^\infty)[J]}$ all depend continuously on the endpoints of J.*
- *Closure and Extension Property. Let I_n be an increasing sequence of intervals and $\cup I_n = I = (a, b)$. Let ϕ be a classical Wave-Map in I which satisfies the uniform bounds:*

$$\| \phi \|_{S[I_n]} \leqslant F, \quad \sup_k \| P_k \phi \|_{L_t^\infty (L_x^\infty)[I_n]} \leqslant \epsilon,$$

with $\epsilon \leqslant \epsilon(F)$. Then $\phi \in S[I]$, and furthermore it can be extended to a classical Wave-Map in a larger interval $I_1 = [a_1, b_1]$ with $a_1 < a < b < b_1$.

Proof. The first part is a direct consequence of the solvability bound (18) since $\Box\phi \in L^1_t L^2_x[I]$ as well as $\Box\phi \in (L^1_t L^2_x)_c[I]$.

For the second part we first consider the S norm. Let $J_n \subset I$ be a sequence of intervals converging to J. We consider a sequence of rescalings mapping J to J_n,

$$(t, x) \rightarrow (\lambda_n t + t^0_n, \lambda_n x), \quad \lambda_n \rightarrow 1, \quad t_n \rightarrow 0.$$

This allows us to map functions in J_n to functions in J,

$$\phi \rightarrow \phi_n(t, x) = \phi(\lambda_n t + t^0_n, \lambda_n x).$$

Hence using the scale invariance of the S norm, we have

$$\|\phi\|_{S[J_n]} = \|\phi_n\|_{S[J]} \rightarrow \|\phi\|_{S[J]},$$

where in the last step we simply use the fact that convergence in the Schwartz space implies the convergence in $S[J]$.

For S_c norms the proof is similar. The dyadic convergence $\|\phi_k\|_{S[J_n]} \rightarrow \|\phi_k\|_{S[J]}$ follows by the same rescaling argument. This implies the S_c convergence since the tails are small,

$$\lim_{k \to \pm\infty} c_k^{-1} \|\phi_k\|_{S[I]} = 0,$$

which is due to the Schwartz regularity of ϕ. A similar decay of the tails yields the continuity of the energy dispersion norm.

For the last part we observe that by (67), for each (δ_0, Δ) frequency envelope c we obtain a uniform bound for $\|P_k\phi\|_{S_c[I_n]} + \|P_k\phi\|_{\underline{X}_c[I_n]}$. Letting $n \rightarrow \infty$ we directly obtain $P_k\phi \in \underline{X}_c[I]$, which shows that for each k we have $P_k\phi \in S[I]$ and $\|P_k\phi\|_{S[I_n]} \rightarrow \|P_k\phi\|_{S[I]} \lesssim c_k$. Hence ϕ is a Schwartz wave map in $[a, b]$, therefore by the local well-posedness result it admits a Schwartz extension to a larger interval. $\Box$

4. Proof of the Main Result

The purpose of this section is to use the setup of the previous two sections to prove the following result, which easily implies our main Theorem 1.3 as well as Corollary 1.7.

Theorem 4.1 (Frequency Envelope Version of the Main Theorem). *There exist two functions $1 \ll F(E)$ and $0 < \epsilon(E) \ll 1$ of the energy (2) such that if ϕ is a finite energy solution to (1) in a closed interval $I \times \mathbb{R}^2$, where $I = [a, b]$, with energy E and dispersion (4), then estimate (5) holds in $S[I]$. In addition, there exists a universal polynomial $K(F)$ such that if $\{c_k\}$ is any (δ_0, Δ)-admissible frequency envelope for $\phi[0]$, we have the bound:*

$$\|\phi\|_{S_c[I]} \leqslant K(F(E)). \tag{69}$$

*In particular, one may extend ϕ to a finite energy Wave-Map on the open neighborhood $I \subseteq (a - i_0, b + i_0)$ whose additional length i_0 depends **only** on E, $\{c_k\}$, and ϵ.*

We immediately observe that it suffices to prove the result for classical wave-maps. This is due to the small data result in Theorem 1.1, which implies that any finite energy wave map in a closed interval can be approximated in S by classical wave maps. In addition, the S convergence easily implies the convergence of the energy dispersion norm (4).

In the sequel we simply focus on proving (5). Once (5) is known, Proposition 3.9 can be applied and the estimate (69) is an immediate consequence of (67). In fact, it would be tempting to use the more direct analysis employed in the proof of (67) to establish (5) as well in a single go. Such a strategy seems to fail basically due to linearized $Low \times High \Rightarrow High$ frequency interactions. These interactions need to be handled via Proposition 3.2, which in turn requires one to already control S type norms (e.g. in assumption (40)). To avoid this dilemma, we employ a simple induction scheme to reduce things to estimates for Wave-Maps of (slightly) smaller energy. The reader should keep in mind however that modulo this single $Low \times High$ obstruction, our analysis would work to prove (5) and (67) simultaneously. More specifically, the remaining estimates basically boil down to using (44)–(46) to eliminate matched frequency "semilinear" type interactions (this is the *only* place where the energy dispersion (4) really comes in), and (24)–(25) to kill off $High \times High \Rightarrow Low$ frequency cascades.

We now construct the functions $F(E)$ and $\epsilon(E)$ such that (4) and (5) hold. Precisely, we will show that there exists a strictly positive nonincreasing function defined for *all* values of E, $c_0 = c_0(E) \ll 1$, so that if the conclusion of the theorem holds up to energy E then it also holds up to energy $E + c_0$. It is important here that c_0 depends *only* on E and not on the size of $F(E)$ or $\epsilon(E)$, as otherwise we would only be able to conclude the usual first step in an induction on energy proof which is establishing that the set of regular energies is open.[2] Also, we note here the monotonicity of c_0 is only used to conclude that c_0 admits a positive lower bound on any compact set.

According to Theorem 1.1 we know that $\epsilon(E)$ and $F(E)$ can be constructed up to some $E_0 \ll 1$. We now assume that E_0 is fixed by induction, and to increase its range we consider a solution ϕ defined on an interval I with energy $E[\phi] = E_0 + c$, $c \leqslant c_0(E_0)$ and with energy dispersion $\leqslant \epsilon$ (at first this is a free parameter which we may take as small as we like). We will compare ϕ with a solution $\widetilde{\phi}$ with energy E_0. To construct $\widetilde{\phi}$ we reduce the initial data energy of $\phi[0]$ by truncation in frequency. We define the "cut frequency" $k_* \in \mathbb{R}$ according to (this can be done by adjusting the definition of the $P_{<k}$ continuously if necessary):

$$E[\Pi P_{\leqslant k_*} \phi[0]] = E_0.$$

We consider the Wave-Map $\widetilde{\phi}$ with this initial data $\widetilde{\phi}[0] = \Pi P_{\leqslant k_*} \phi[0]$. This Wave-Map exists classically for at least a short amount of time according to Cauchy stability, and where it exists classically we have:

$$E[\widetilde{\phi}(t)] = E_0. \tag{70}$$

Since ϕ has energy dispersion $\leqslant \epsilon$, by (47) it follows that $\widetilde{\phi}$ has energy dispersion $\lesssim_{E_0} \epsilon^{\frac{1}{4}}$ at time $t = 0$. Again by the usual Cauchy stability theory, if ϵ is chosen small

---

[2] In this latter setup, one is then left with the arduous task of eliminating minimal energy blowup solutions. Our strategy is a bit more direct because we accomplish this as well in our construction of c_0, so we are able to avoid a good deal of repetitive analysis.

enough in comparison to the inductively defined parameter $\epsilon(E_0)$ it follows that there exists a non-empty interval J_0 where $\widetilde{\phi}$ satisfies:

$$\sup_k \| P_k \widetilde{\phi} \|_{L_t^\infty(L_x^\infty)[J_0]} \leqslant \epsilon(E_0). \tag{71}$$

Then our induction hypothesis guarantees that we have the dispersive bounds:

$$\| \widetilde{\phi} \|_{S[J_0]} \leqslant F(E_0). \tag{72}$$

The plan is now very simple. On one hand, we try to pass the space-time control of $\widetilde{\phi}$ up to ϕ via linearization around $\widetilde{\phi}$ to control the low frequencies, and conservation of energy and perturbation theory to control the high frequencies. On the other hand, we need to pass the good energy dispersion bounds from ϕ back down to $\widetilde{\phi}$ in order to increase the size of $J \subseteq I$ on which (71) holds until it eventually fills up all of I. To achieve all of this, we proceed via two core estimates:

Proposition 4.2 (Evolution of Low Frequency Errors). *Let ϕ be a Wave-Map defined on an interval J with energy $E + c$ with $0 < c \lesssim 1$ and bounds:*

$$\sup_k \| P_k \phi \|_{L_t^\infty(L_x^\infty)[J]} \leqslant \epsilon, \quad \| \phi \|_{S[J]} \leqslant F. \tag{73}$$

Suppose in addition that $\widetilde{\phi}$ is the Wave-Map with energy E defined by $\widetilde{\phi}[0] = \Pi P_{\leqslant k_*} \phi[0]$, and that $\widetilde{\phi}$ is classical on J with bounds:

$$\sup_k \| P_k \widetilde{\phi} \|_{L_t^\infty(L_x^\infty)[J]} \leqslant \widetilde{\epsilon}, \quad \| \widetilde{\phi} \|_{S[J]} \leqslant \widetilde{F}. \tag{74}$$

Assume also that the two energy dispersion constants are chosen so that:

$$\epsilon \leqslant \widetilde{\epsilon}, \quad \epsilon \leqslant (CF)^{-\delta_0^{-10}}, \quad \widetilde{\epsilon} \leqslant (C\widetilde{F})^{-\delta_0^{-10}}, \tag{75}$$

where we may assume that $F \geqslant \widetilde{F} \geqslant E \geqslant C^{-\frac{1}{2}}$ and C is a sufficiently large constant. Then in addition we have the bound:

$$\| \widetilde{\phi} - P_{\leqslant k_*} \phi \|_{S[J]} \lesssim_F \epsilon^{\delta_0}. \tag{76}$$

Proposition 4.3 (High Frequency Evolution Estimates). *Let ϕ and $\widetilde{\phi}$ be defined as in the last proposition, in particular with the bounds (73) and (74) (resp), and that the dispersion constants obey (75). Then there exists a universal function $c_0(E)$ with $c_0^{-1} \lesssim_E 1$ such that if we assume $c_0 = c_0(E)$ in the definition of $\widetilde{\phi}$ we in addition have the bound:*

$$\| \widetilde{\phi} - \phi \|_{S[J]} \lesssim_{\widetilde{F}} 1. \tag{77}$$

We postpone the proof of the above propositions to show how to use them to conclude our induction. By the seed bound (68) we may assume that in addition to (71) and (72) above we also have:

$$\| \phi \|_{S[J]} \leqslant 2F(E_0 + c),$$

on some interval J. With this setup, and by an application of the continuity property in Proposition 3.10 it suffices to combine Propositions 4.2–4.3 to show the following:

Corollary 4.4. *Assume there exists functions $\epsilon(E)$ and $F(E)$ defined up to E_0 such that* (4) *implies* (5). *Choose $c_0(E_0)$ according to Proposition* 4.3. *Then there exist extensions of $\epsilon(E)$ and $F(E)$ such that for each $0 < c \leqslant c_0(E_0)$ and each classical Wave-Map ϕ in a time interval J with energy $E_0 + c$ and the bounds:*

$$\sup_k \| P_k \phi \|_{L_t^\infty(L_x^\infty)[J]} \leqslant \epsilon(E_0 + c), \quad \| \phi \|_{S[J]} \leqslant 2F(E_0 + c), \tag{78}$$

we have:

$$\| \phi \|_{S[J]} \leqslant F(E_0 + c). \tag{79}$$

Proof. In addition to (78) $\Rightarrow$ (79), we will make the additional assumption that $\widetilde{\phi}$ is defined as a Schwartz wave map in J and satisfies:

$$\sup_k \| P_k \widetilde{\phi} \|_{L_t^\infty(L_x^\infty)[J]} \leqslant \epsilon(E_0), \tag{80}$$

and show that if the extensions to $\epsilon(E)$ and $F(E)$ are chosen correctly then we in addition have the following improvement to (80):

$$\sup_k \| P_k \widetilde{\phi} \|_{L_t^\infty(L_x^\infty)[J]} \leqslant \frac{1}{2}\epsilon(E_0). \tag{81}$$

To see that this is sufficient, we first note that by (71) the bound (80) holds in a smaller interval $J_0 \subset J$. Extending J_0 to a maximal interval in J, denoted still J_0, so that (80) holds, by the closure property in Proposition 3.10 it follows that J_0 must be closed. The same part of Proposition 3.10 shows that $\widetilde{\phi}$ has a Schwartz extension to a neighborhood of J_0. Then by (81) applied in J_0 and the continuity property in Proposition 3.10 it follows that (80) holds in a larger interval. Hence J_0 must be both closed and open in J, and therefore $J_0 = J$.

It remains to find extensions $\epsilon(E)$ and $F(E)$ so that (78) together with (80) imply (79) and (81). Our extensions of $\epsilon(E)$ and $F(E)$ in $(E_0, E_0 + c_0]$ are constant:

$$\epsilon(E) = \epsilon, \quad F(E) = F, \quad E \in (E_0, E_0 + c_0].$$

Let $K_1(F)$ and $K_2(\widetilde{F})$ be the implicit polynomials from lines (76) and (77) (resp.). In order to get the improvement (81) we need that:

$$\epsilon^{\delta_0} \cdot K_1(F) \ll \epsilon(E_0).$$

In order to conclude (79) we need that:

$$K_2(F(E_0)) \ll F.$$

Finally, we also need to choose ϵ and F so that (75) holds, and so that (which is of course redundant):

$$E_0^{\frac{1}{2}} \epsilon^{\frac{1}{2}} \ll \epsilon(E_0),$$

which was used right before line (71) to get things started. All of these goals can easily be satisfied as long as we choose $\sigma = \delta_0^{20}$, with $\delta_0 \ll 1$ sufficiently small, and then first choose F, followed by ϵ, such that:

$$F(E_0) \ll F^\sigma, \quad \epsilon^\sigma \ll \min\{\epsilon(E_0), F^{-1}\}.$$

Notice that this process can be carried on indefinitely, regardless of the size of E, because we have taken care to decouple the step size c_0 from the growth and decay properties of F and ϵ.

We remark that the above proof allows us to estimate the size of $E(F)$ and $\epsilon(F)$. Indeed, what we have obtained are piecewise constant functions $c_0(E)$, $\epsilon(E)$, and $F(E)$ which at the jump points E_n are given by the recurrence relation:

$$E_{n+1} = E_n + c_0(E_n),$$

and which satisfy:

$$c_0(E_n) = cE_n^{-\sigma^{-1}}, \quad F(E_{n+1}) = CF(E_n)^{\sigma^{-1}}, \quad \epsilon(E_{n+1}) = cF(E_n)^{\sigma^{-2}},$$

with sufficiently small σ, c and sufficiently large C. The first relation shows that:

$$E_n \approx n^{\frac{1}{\sigma^{-1}+1}},$$

while the next two give relations of the form:

$$F(E_n) \leqslant C_1^{\sigma^{-n}}, \quad \epsilon(E_n) \geqslant c_1^{\sigma^{-n}}.$$

Together the last two bounds yield estimates for F and ϵ of the form:

$$F(E) \leqslant e^{Ce^{EM}}, \quad \epsilon(E) \geqslant e^{-Ce^{EM}},$$

again with C and M sufficiently large. □

The remainder of this section is devoted to the proof of Propositions 4.2–4.3. This will be done in order because we will use some of the estimates of Proposition 4.2 in our demonstration of Proposition 4.3.

Proof of Proposition 4.2. Denoting:

$$\psi = P_{\leqslant k_*}\phi - \widetilde{\phi}, \tag{82}$$

we will prove the stronger bound:

$$\| \psi \|_{S_c[J]} \leqslant 1, \tag{83}$$

where $\{c_k\}$ is the (δ_0, δ_0)-admissible frequency envelope $c_k = 2^{-\delta_0|k-k_*|}\epsilon^{\delta_0}$. We first consider the initial data for ψ. By an immediate application of Proposition 11.1 and the energy dispersion bound (4.2) we have:

$$\| P_k\psi[0] \|_{\dot{H}^1 \times L^2} \lesssim_E \epsilon^{\frac{1}{4}}2^{-\frac{1}{2}|k-k_*|}. \tag{84}$$

Since ψ is a Schwartz function, this implies that for a small interval $I \subset J$ containing $t = 0$ we have:

$$\| \psi \|_{S_c[I]} \leqslant 1. \tag{85}$$

Using this as a seed bound, by the continuity property in Proposition 3.10 it suffices to prove that (83) holds under the bootstrap assumption:

$$\| \psi \|_{S_c[I]} \leqslant 2. \tag{86}$$

As a preliminary step we use the general renormalization bound (64) as well as the high modulation bound (66), which in light of the estimates on each of lines (73) and (74) imply the set of inequalities:

$$\| \phi \|_{\mathcal{W}[J]} \lesssim_F 1, \quad \| \widetilde{\phi} \|_{\mathcal{W}[J]} \lesssim_{\widetilde{F}} 1, \tag{87}$$

$$\| \phi \|_{\underline{X}[J]} \lesssim_F \epsilon^{\delta_1}, \quad \| \widetilde{\phi} \|_{\underline{X}[J]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_1}. \tag{88}$$

The proof is deduced in a series of steps:

Step 1 (Outline of the proof). The equation for ψ has the form:

$$\Box \psi = -P_{\leqslant k_*} \left(\mathcal{S}(\phi) \partial^\alpha \phi \partial_\alpha \phi \right) + \mathcal{S}(\widetilde{\phi}) \partial^\alpha \widetilde{\phi} \partial_\alpha \widetilde{\phi}. \tag{89}$$

This may be rewritten as follows:

$$\Box \psi = -\mathcal{D}(\widetilde{\phi}, \psi) + \mathcal{C}(\phi), \tag{90}$$

where the difference $\mathcal{D}$ and the generalized commutator $\mathcal{C}$ are defined as follows:

$$\mathcal{D}(\widetilde{\phi}, \psi) = \mathcal{S}(\widetilde{\phi} + \psi) \partial^\alpha (\widetilde{\phi} + \psi) \partial_\alpha (\widetilde{\phi} + \psi) - \mathcal{S}(\widetilde{\phi}) \partial^\alpha \widetilde{\phi} \partial_\alpha \widetilde{\phi}, \tag{91}$$

$$\mathcal{C}(\phi) = \mathcal{S}(\phi_{\leqslant k_*}) \partial^\alpha \phi_{\leqslant k_*} \partial_\alpha \phi_{\leqslant k_*} - P_{\leqslant k_*} \left(\mathcal{S}(\phi) \partial^\alpha \phi \partial_\alpha \phi \right). \tag{92}$$

This form of the equation will be used for proving pure L^2 estimates.

Alternatively, freezing the spatial frequency k and introducing a frequency gap parameter $m \geqslant 20$, we will write (89) in the following paradifferential form:

$$\Box \psi_k + 2 \widetilde{A}^\alpha_{<k-m} \partial_\alpha \psi_k = \mathcal{D}^m_k(\widetilde{\phi}, \psi) + \mathcal{L}^m_k(\widetilde{\phi}, \psi) + \mathcal{C}^m_k(\phi), \tag{93}$$

which will be useful for establishing N estimates. Here we are writing:

$$\widetilde{A}^\alpha_{<k-m} = A^\alpha_{<k-m}(\widetilde{\phi}) := \left(\mathcal{S}(\widetilde{\phi})_{<k-m} - \mathcal{S}^\dagger(\widetilde{\phi})_{<k-m} \right) \partial^\alpha \widetilde{\phi}_{<k-m}. \tag{94}$$

These terms are chosen roughly as follows. The term $\mathcal{D}^m_k$ denotes differences of the form (91) between $\widetilde{\phi}$ and ψ which are frequency localized according to the general T^m structure defined on line (53). In particular, these never contain $Low \times Low \times High$ or $Low \times High \times Low$ interactions. The term $\mathcal{L}^m_k$ contains certain $Low \times Low \times High$ and $Low \times High \times Low$ interactions in $\widetilde{\phi}$ and ψ differences, with the additional structure that ψ is always at Low frequency with a (possibly large) m dependent gap. Finally, the expression $\mathcal{C}^m_k(\phi)$ contains ϕ dependent commutators of the form (92).

With this setup, we prove the following estimates. First, we show that the commutators are always favorable, regardless of m:

$$\| (\mathcal{C}, \mathcal{C}^m) \|_{(L^2_t (\dot{H}^{-\frac{1}{2}}) \cap N)_c[I]} \lesssim_F \epsilon^{\frac{1}{2} \delta_1^2}. \tag{95}$$

Second, under the bootstrapping assumption (86), we will show the first two terms on the RHS of (93) may be estimated as follows:

$$\| \mathcal{D}^m_k(\widetilde{\phi}, \psi) \|_{N_c[I]} \lesssim_{\widetilde{F}} 2^{4\delta_0 m}, \tag{96}$$

$$\| \mathcal{L}^m_k(\widetilde{\phi}, \psi) \|_{N_c[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_0} + 2^{-\delta_0 m}. \tag{97}$$

While the second of these last two estimates is favorable for closing a bootstrap via Proposition 3.2, the first is not. However, via Remark 3.3 the above estimates with $m = 20$ allow us to gain renormalization control of ψ, namely:

$$\| \psi \|_{\mathcal{W}_c[I]} \lesssim_{\widetilde{F}} 1. \tag{98}$$

To close the bootstrap, we now use two additional estimates. The first shows that with (87) and (98), we have improved L^2 control:

$$\| \mathcal{D}(\widetilde{\phi}, \psi) \|_{L^2_t(\dot{H}^{-\frac{1}{2}})_c[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_1}. \tag{99}$$

In particular by this, (95), and the gap condition (75) we have:

$$\| \psi \|_{\underline{X}_c[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_1}. \tag{100}$$

Finally, we show that this last estimate, (98), and (87)–(88) allow the following drastic improvement to (96):

$$\| \mathcal{D}_k^m(\widetilde{\phi}, \psi) \|_{N_c[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_1^2}. \tag{101}$$

The bootstrap is therefore concluded by choosing $m = \delta_1 |\ln(\widetilde{\epsilon})|$ in estimate (97), and applying the linear bound (41) for the paradifferential flow, with the estimates (101), (95) for the right-hand side and (84) for the initial data. We remark that while (41) gives a stronger $\mathcal{W}_c[I]$ control for ψ, in order to conclude the bootstrap we only use a weaker S_c bound which follows from the $\mathcal{W}_c[I]$ bound.

Step 2 (*The algebraic decomposition*). Here we derive the form of the RHS of (93). To uncover this, we shall employ the following generic notation. We let T be a trilinear expression of the form:

$$T\left(\mathcal{S}(\phi^{(1)}), \phi^{(2)}, \phi^{(3)}\right) = L\left(\mathcal{S}(\phi^{(1)}), \partial^\alpha \phi^{(2)}, \partial_\alpha \phi^{(3)}\right), \tag{102}$$

with L disposable, and $\mathcal{S}$ is a smooth function with uniformly bounded derivatives. From this we may define the T dependent expressions $\mathcal{D}$ and $\mathcal{C}$ as on lines (91)–(92).

The frequency localized equation for ψ is:

$$\Box \psi_k = -P_k P_{\leqslant k_*}\left(\mathcal{S}(\phi)\partial^\alpha \phi \partial_\alpha \phi\right) + P_k\left(\mathcal{S}(\widetilde{\phi})\partial^\alpha \widetilde{\phi} \partial_\alpha \widetilde{\phi}\right), \tag{103}$$

which may be written in the form:

$$\Box \psi_k = 2\mathcal{S}(\widetilde{\phi})_{<k-m}\partial^\alpha \widetilde{\phi}_{<k-m}\partial_\alpha \widetilde{\phi}_k - 2P_{\leqslant k_*}\left[\mathcal{S}(\phi)_{<k-m}\partial^\alpha \phi_{<k-m}\partial_\alpha \phi_k\right] + T_{1;k}^m, \tag{104}$$

where we are writing:

$$T_{1;k}^m = T_k^m\left(\mathcal{S}(\widetilde{\phi}), \widetilde{\phi}, \widetilde{\phi}\right) - P_{\leqslant k_*}T_k^m\left(\mathcal{S}(\phi), \phi, \phi\right) = \mathcal{D}_{1;k}^m + \mathcal{C}_{1;k}^m, \tag{105}$$

with T^m defined as on line (53). We now employ the geometric identity for the second fundamental form:

$$\sum_c \mathcal{S}_{ab}^c(\phi)\nabla_{t,x}\phi^c \equiv 0, \tag{106}$$

which follows simply because the constraint on the image of ϕ to lie in $\mathcal{M}$ implies that $\nabla_{t,x}\phi$ lies in $T_\phi\mathcal{M}$. This is valid for $\widetilde{\phi}$ as well, because it is an exact wave-map. Therefore, we have the zero expression:

$$P_k\left(2\mathcal{S}(\widetilde{\phi})^\dagger\partial^\alpha\widetilde{\phi}\partial_\alpha\widetilde{\phi}_{<k-m+2} - 2P_{\leqslant k_*}\left[\mathcal{S}(\phi)^\dagger\partial^\alpha\phi\partial_\alpha\phi_{<k-m+2}\right]\right) = 0, \qquad (107)$$

which if added to the first two terms on RHS of (104) produces:

$$\text{(First two R.H.S.)}(104) = 2\widetilde{A}^\alpha_{<k-m}\partial_\alpha\widetilde{\phi}_k - 2P_{\leqslant k_*}\left[A^\alpha_{<k-m}\partial_\alpha\phi_k\right] + T^m_{2;k}, \quad (108)$$

where both $\widetilde{A}^\alpha$ and $A^\alpha_{<k-m} = A^\alpha_{<k-m}(\phi)$ are defined as on line (94). Here the trilinear form $T^m_{2;k}$ is a difference:

$$T^m_{2;k} = T^m_{2;k}(\widetilde{\phi}) - P_{\leqslant k_*}T^m_{2;k}(\phi) = \mathcal{D}^m_{2;k} + \mathcal{C}^m_{2;k}, \qquad (109)$$

where each individual form is defined as a T^m from line (53) applied separately to the two trilinear expressions on the LHS of (107).

We now assign the generalized difference labels on the RHS of (93) by setting $\mathcal{D}^m_k = \sum_i \mathcal{D}^m_{i;k}$, where the two summands were defined on lines (105) and (109).

To assign $\mathcal{C}^m_k$, we further denote by $\mathcal{C}^m_{3;k}$ the corresponding expression which results from commuting the $P_{\leqslant k_*}$ in the second term on the RHS of line (108). We then set $\mathcal{C}^m_k = \sum_i \mathcal{C}^m_{i;k}$.

With these choices, Eq. (103) may be written in the form:

$$\Box\psi_k = 2A^\alpha_{<k-m}(\widetilde{\phi})\partial_\alpha\widetilde{\phi}_k - 2A^\alpha_{<k-m}(\widetilde{\phi}+\psi)\partial_\alpha(\widetilde{\phi}_k+\psi_k) + \mathcal{D}^m_k + \mathcal{C}^m_k.$$

As a final step we assign:

$$\mathcal{L}^m_k = 2\left(A^\alpha_{<k-m}(\widetilde{\phi}) - 2A^\alpha_{<k-m}(\widetilde{\phi}+\psi)\right)\partial_\alpha(\widetilde{\phi}_k+\psi_k), \qquad (110)$$

and the form of (93) is achieved.

The remainder of the proof shows estimates (95), (96), (97), (99), and (101).

Step 3 (*Estimates for commutators*). Here we demonstrate (95). Let $\mathcal{C}$ be any expression of the form:

$$\mathcal{C} = T\left(\mathcal{S}(P_{\leqslant k_*}\phi), P_{\leqslant k_*}\phi, P_{\leqslant k_*}\phi\right) - P_{\leqslant k_*}T\left(\mathcal{S}(\phi), \phi, \phi\right). \qquad (111)$$

We will prove the general pair of bounds:

$$\|\mathcal{C}\|_{L^2_t(\dot{H}^{-\frac{1}{2}})_c[I]} \lesssim_F \epsilon^{\frac{1}{2}\delta_1}, \quad \|\mathcal{C}\|_{N_c[I]} \lesssim_F \epsilon^{\frac{1}{2}\delta^2_1}. \qquad (112)$$

As a preliminary step we decompose $\mathcal{C} = \mathcal{C}_1 + \mathcal{C}_2$ where:

$$\mathcal{C}_1 = T\left(\mathcal{S}(\phi_{\leqslant k_*}) - \mathcal{S}(\phi)_{\leqslant k_*}, \phi_{\leqslant k_*}, \phi_{\leqslant k_*}\right),$$
$$\mathcal{C}_2 = T\left(\mathcal{S}(\phi)_{\leqslant k_*}, \phi_{\leqslant k_*}, \phi_{\leqslant k_*}\right) - P_{\leqslant k_*}T\left(\mathcal{S}(\phi), \phi, \phi\right).$$

These terms are handled separately:

Step 3A (*Estimates for $\mathcal{C}_1$*). This is based on the Moser type estimate:

$$\|\mathcal{S}(\phi_{\leqslant k_*}) - \mathcal{S}(\phi)_{\leqslant k_*}\|_{S_k[I]} \lesssim_F 2^{-\delta|k-k_*|}. \qquad (113)$$

To prove this, we further decompose the difference as:

$$\mathcal{S}(\phi_{\leqslant k_*}) - \mathcal{S}(\phi)_{\leqslant k_*} = P_{>k_*}\mathcal{S}(\phi_{\leqslant k_*}) + P_{\leqslant k_*}\left(\mathcal{S}'(\phi_{\leqslant k_*}, \phi_{>k_*})\phi_{>k_*}\right),$$

where here $\mathcal{S}'$ is a bounded and smooth function or its arguments which results from the difference $\mathcal{S}(\phi_{\leqslant k_*} + \phi_{>k_*}) - \mathcal{S}(\phi_{\leqslant k_*})$. The bound (113) now follows by directly applying the Moser estimate (27) to the first term on the last line above, and by applying a combination of the product estimate (20) and the Moser estimate (26) to the second.

To conclude the proof of (112) for the term $\mathcal{C}_1$ we need to split the output frequency into two cases: $k \leqslant k_* + 10$ or $k > k_* + 10$. In the first case, we directly use (52) and (57), which together provide (112) in light of (87)–(88) and the additional L^∞ estimate:

$$\| P_{<k}\left[\mathcal{S}(\phi_{\leqslant k_*}) - \mathcal{S}(\phi)_{\leqslant k_*}\right] \|_{L_t^\infty(L_x^\infty)[I]} \lesssim_F 2^{-\delta|k-k_*|}, \quad \text{for } k \leqslant k_* + 10.$$

This last inequality follows from (113), and the fact that the difference $\mathcal{S}(\phi_{\leqslant k_*}) - \mathcal{S}(\phi)_{\leqslant k_*}$ is rapidly decaying outside of a compact set, so in particular one can control the L^∞ norm by summing dyadically from $k = -\infty$. Note that the use of (52) and (57) costs a power of F because these estimates are in normalized form.

In the second case ($k > k_* + 10$), we establish (112) by directly appealing to estimates (51) and (55), which suffice because of (113) and the observation that due to the fact T is translation invariant we have the identity:

$$P_k\mathcal{C}_1 = P_k T\left(P_{k+O(1)}\left[\mathcal{S}(\phi_{\leqslant k_*}) - \mathcal{S}(\phi)_{\leqslant k_*}\right], \phi_{\leqslant k_*}, \phi_{\leqslant k_*}\right).$$

Step 3B (*Estimating the Term $\mathcal{C}_2$*). We first observe that from the definition we have $P_k\mathcal{C}_2 \equiv 0$ whenever $k > k_* + 10$. Thus, we only need to deal with the frequency range $k \leqslant k_* + 10$. We split this range into two regions: either $k_* - m \leqslant k \leqslant k_* + 10$ or $k < k_* - m$. Here m is defined as follows:

$$2^{-m} = \epsilon^{\delta_1}. \tag{114}$$

Note that this definition has nothing to do with the m in the decomposition (93), and is only local to this step. We now estimate separately:

Step 3B.1 (*The range $k_* - m \leqslant k \leqslant k_* + 10$*). We may write:

$$\begin{aligned}
P_k\mathcal{C}_2 = &\, T_k^m\left(\mathcal{S}(\phi)_{\leqslant k_*}, \phi_{\leqslant k_*}, \phi_{\leqslant k_*}\right) - T_k^m\left(\mathcal{S}(\phi), \phi, \phi\right) \\
&+ 2^{-k_*}\left(\widetilde{L}_1\left(\nabla_x\mathcal{S}(\phi)_{<k-m}, \partial_\alpha\phi_{<k-m}, \partial^\alpha\phi_k\right) + \widetilde{L}_2\left(\mathcal{S}(\phi)_{<k-m}, \nabla_x\partial_\alpha\phi_{<k-m}, \partial^\alpha\phi_k\right)\right. \\
&\left.+ \widetilde{L}_3\left(\nabla_x\mathcal{S}(\phi)_{<k-m}, \partial_\alpha\phi_k, \partial^\alpha\phi_{<k-m}\right) + \widetilde{L}_4\left(\mathcal{S}(\phi)_{<k-m}, \partial_\alpha\phi_k, \nabla_x\partial^\alpha\phi_{<k-m}\right)\right),
\end{aligned} \tag{115}$$

where the T_k^m are defined as on line (53) with the additional structure and frequency localizations from the definition of $\mathcal{C}_2$. The $\widetilde{L}_i$ are an additional collection of translation invariant and disposable trilinear forms resulting from the commutator rule (10) applied to the second and third terms on the RHS of line (53). In particular, this commutator is trivial unless $k_* - 10 \leqslant k \leqslant k_* + 10$, so $\widetilde{L}_i \equiv 0$ without this further restriction.

For the first two terms on the RHS (115), we use (87)–(88) which allows us to apply (51) or (55), and these suffice to give (112) in this case because of the frequency gap (114) and the conditions (8) on the δ_i.

It remains to estimate the commutators. From the version of estimate (51) in Remark 3.8, and the fact that:

$$\| \nabla_x\mathcal{S}(\phi)_{<k-m} \|_{S[I]} + \| \nabla_x\phi_{<k-m} \|_{S[I]} \lesssim_F 2^{k-m},$$

we directly have the L^2 bound from line (112) via (114) and the range restriction $k_* - 10 \leqslant k \leqslant k_* + 10$. To prove the N estimate, we similarly only need to show:

$$\| \, 2^{-k_*} \widetilde{L}_i \, \|_{N[I]} \lesssim_F 2^{-m}.$$

To estimate $2^{-k_*} L_1$ in $N[I]$ we use (25) as follows (again using $k = k_* + O(1)$):

$$\| \, 2^{-k_*} L_1 \, \|_{N[I]} \lesssim_F 2^{-k_*} \sum_{k_1, k_2 \leqslant k-m} 2^{k_1} 2^{-\delta(k_1-k_2)_+} \lesssim_F 2^{-m}.$$

The details of these calculations for other L_i are similar and left to the reader.
Step 3B.2 (The range $k < k_* - m$). Here we simply decompose:

$$P_k \mathcal{C}_2 = - \sum_{k_i : \, \max\{k_i\} > k_*} P_k L \left(\mathcal{S}(\phi)_{k_1}, \partial^\alpha \phi_{k_2}, \partial_\alpha \phi_{k_3} \right),$$

so in particular at least one of the second two factors must be in the range $k_i > k_* - 10$. We remark that this sum has a T^m structure of the form (53), so smallness is guaranteed. The main issue is to also recover the exponential falloff in the definition of $\{c_k\}$. This may be achieved via a direct application of estimates (51) and (55) by first introducing a high frequency (δ_0, δ_0)-admissible $\mathcal{W}[I]$ envelope for $P_{>k_*-10}\phi$ which we denote by $\{d_k\}$. In particular, we have $d_k \lesssim_F 2^{\delta_0(k-k_*)}$, so we directly have (112) for the above sum.
Step 4 (Estimates for matched frequency differences). Here we prove (96), (99), and (101). To do this, it is enough to demonstrate the bound:

$$\| \, \mathcal{D}^m(\widetilde{\phi}, \psi) \, \|_{N_c[I]} \lesssim_{\widetilde{F}} 2^{4\delta_0 m}, \tag{116}$$

under the assumptions (86) and (87), the bound:

$$\| \, \mathcal{D}(\widetilde{\phi}, \psi) \, \|_{L_t^2(\dot{H}^{-\frac{1}{2}})_c[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_1}, \tag{117}$$

under the additional assumption (98), and finally the improved estimate:

$$\| \, \mathcal{D}^m(\widetilde{\phi}, \psi) \, \|_{N_c[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_1^2}, \tag{118}$$

assuming all of the above and also using (88) and (100). Here $\mathcal{D}$ is any expression of the form (91) for a general trilinear form T as on line (102), and $\mathcal{D}^m$ denotes a similar expression in terms of T^m from line (53). To some extent these tasks are redundant, so we will make some effort to collapse cases.[3]

First we introduce the following decomposition of differences of trilinear expressions of the form (102):

$$T \left(\mathcal{S}(\widetilde{\phi}), \widetilde{\phi}, \widetilde{\phi} \right) - T \left(\mathcal{S}(\widetilde{\phi} + \psi), \widetilde{\phi} + \psi, \widetilde{\phi} + \psi \right) = R_0 + R_1 + R_2 + R_3 + R_4.$$

---

[3] We note that one can eliminate this redundancy and also simplify some of the case analysis by simply replacing the $S_c[I]$ bootstrap of the current proof with a direct bootstrap with respect to the $\mathcal{W}_c[I]$ space. We will not pursue this here.

Here we have set:

$$R_0 = T\left(\mathcal{S}'(\widetilde{\phi}, \psi)\psi, \widetilde{\phi}, \widetilde{\phi}\right), \tag{119a}$$

$$R_1 = T\left(\mathcal{S}'(\widetilde{\phi}, \psi)\psi, \psi, \widetilde{\phi}\right) + T\left(\mathcal{S}'(\widetilde{\phi}, \psi)\psi, \widetilde{\phi}, \psi\right) + T\left(\mathcal{S}'(\widetilde{\phi}, \psi)\psi, \psi, \psi\right), \tag{119b}$$

$$R_2 = -T\left(\mathcal{S}(\widetilde{\phi}), \psi, \widetilde{\phi}\right), \tag{119c}$$

$$R_3 = -T\left(\mathcal{S}(\widetilde{\phi}), \widetilde{\phi}, \psi\right), \tag{119d}$$

$$R_4 = -T\left(\mathcal{S}(\widetilde{\phi}), \psi, \psi\right). \tag{119e}$$

Here we are using $\mathcal{S}'$ as shorthand for the formula:

$$\mathcal{S}(\widetilde{\phi}) - \mathcal{S}(\widetilde{\phi} + \psi) := \mathcal{S}'(\widetilde{\phi}, \psi)\psi, \tag{120}$$

so that it symbolically represents some additional set of smooth functions obeying the same bounds as the original second fundamental form $\mathcal{S}$. We proceed to estimate the above terms via two subcases:

Step 4A (*Estimates* (116)–(118) *for the terms* R_1–R_4). The bound (116) for the terms R_2–R_4 follows immediately from the estimate (61) of Remark 3.8 in view of the bounds (86) and (87).

The estimates (117) and (118) for terms R_2–R_4 under the additional assumptions (98) and then (100) result directly from estimates (51) and (55).

It remains to prove these estimates for R_1. This will be accomplished with the aid of the following three bounds, which we state in more detail here for their use in the next step:

$$\| P_{>k_*+10}\widetilde{\phi} \|_{S_c[I]} \lesssim 1, \tag{121}$$

$$\| \mathcal{S}'(\widetilde{\phi}, \psi)\psi \|_{S_k[I]} \lesssim_{\widetilde{F}} c_k, \tag{122}$$

$$\| P_{<k}\left[\mathcal{S}'(\widetilde{\phi}, \psi)\psi\right] \|_{S[I]} \lesssim_{\widetilde{F}} c_k, \quad \text{for } k \leqslant k_* + 20. \tag{123}$$

The proof of the first bound follows immediately from the bootstrapping assumption (86). The second bound follows from the first and estimates (86) and (87) after an application of the product bounds (19)–(20) and the Moser estimate (27). The last estimate above follows by summing the second and using the explicit form of the frequency envelope $\{c_k\}$, and also using the fact that the product vanishes for $k = -\infty$.

The proof of (116)–(118) for R_1 is a direct application of estimates (61), (51), and (55) in conjunction with the bounds (122)–(123) above.

Step 4B (*Estimates* (116)–(118) *for* R_0). We first demonstrate (117)–(118). We split things into two output frequency cases: $k \leqslant k_* + 20$ and $k > k_* + 20$. In the first, we combine the bounds (52) and (57) with both (122) and (123) above. Notice that the condition $k \leqslant k_* + 20$ and the specific form of the frequency envelope $\{c_k\}$ below k_* gives the desired result.

To deal with R_0 in the output range $k > k_* + 20$ requires additional work. Note that any estimate of the form (123) is false for $k > k_* + 20$, so the needed frequency envelope control needs to come from the second and third $\widetilde{\phi}$ factors. For this we employ the following version of (121):

$$\| P_{>k_*+10}\widetilde{\phi} \|_{\mathcal{W}_c[I]} \lesssim_{\widetilde{F}} 1, \tag{124}$$

which follows from the assumption (98). To use it, we split R_0 into a sum of pieces (we may drop L from the picture):

$$R_0 = \mathcal{S}'(\widetilde{\phi}, \psi)\psi \cdot \left[\partial^\alpha P_{<k_*+10}\widetilde{\phi}\partial_\alpha P_{<k_*+10}\widetilde{\phi} + 2\partial^\alpha P_{<k_*+10}\widetilde{\phi}\partial_\alpha P_{\geqslant k_*+10}\widetilde{\phi}\right.$$
$$\left. +\partial^\alpha P_{\geqslant k_*+10}\widetilde{\phi}\partial_\alpha P_{\geqslant k_*+10}\widetilde{\phi}\right] = R_{0,1} + R_{0,2} + R_{0,3}. \tag{125}$$

The estimates (117)–(118) for the pieces $P_{>k_*+20}R_{0,2}$ and $P_{>k_*+20}R_{0,3}$ are an immediate consequence of estimate (51) and (55) because in either case we can use $\{c_k\}$ frequency envelope control of the high frequencies of at least one of the $\widetilde{\phi}$ factors. To handle the term $P_{>k_*+20}R_{0,1}$ we again use (51) and (55) along with the bound (122), which provides the needed $\{c_k\}$ factor on account of the forced P_k frequency localization of the first factor in $P_k P_{>k_*+20}R_{0,1}$.

The proof of estimate (116) is similar to (116) above except that we use (121) instead of (124), and (61)–(62) instead of (55) and (57).

Step 5 (*Estimates for ψ at low frequency*). Here we prove (97). From the definition of $\mathcal{L}_k^m$ on line (110), we see that it suffices to prove the two general bounds:

$$\| \widetilde{\phi}_{<k-m}^{(1)} \partial^\alpha \psi_{<k-m} \partial_\alpha \widetilde{\phi}_k^{(2)} \|_{N[I]} \lesssim_{\widetilde{F}} (\widetilde{\epsilon}^{\delta_0} + 2^{-\delta_0 m})c_k, \tag{126}$$

$$\| \left(\mathcal{S}'(\widetilde{\phi}, \psi)\psi\right)_{<k-m} \partial^\alpha \widetilde{\phi}_{<k-m}^{(1)} \partial_\alpha \widetilde{\phi}_k^{(2)} \|_{N[I]} \lesssim_{\widetilde{F}} (\widetilde{\epsilon}^{\delta_0} + 2^{-\delta_0 m})c_k, \tag{127}$$

where $\| \widetilde{\phi}^{(i)} \|_{S[I]} \lesssim_{\widetilde{F}} 1$, and where the $\widetilde{\phi}^{(i)}$ also has high frequency improvement (121). We split into two cases:

Step 5A (*The range $k < k_*+10$*). For estimate (126) we use (21) and (24) which together give:

$$(\text{L.H.S.})(126) \lesssim_{\widetilde{F}} \sum_{j<k-m} c_j \lesssim_{\widetilde{F}} c_{k-m} \lesssim_{\widetilde{F}} 2^{-\delta_0 m}c_k.$$

For estimate (127) we use (25) and (122):

$$(\text{L.H.S.})(127) \lesssim_{\widetilde{F}} \sum_{j,k_1<k-m} 2^{-\delta(j-k_1)_+}c_j \lesssim_{\widetilde{F}} \sum_{j<k-m}(k-m-j)c_j \lesssim_{\widetilde{F}} 2^{-\delta_0 m}c_k.$$

Step 5B (*The range $k > k_* + 10$*). Here we use the fact that the $\widetilde{\phi}^{(i)}$ also have high frequency improvement (121), which already incorporates $\{c_k\}$. We only need to gain a small factor. By (21) and (24), and the fact that $\sum_k \| \psi \|_{S_k[I]} \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_0}$ we immediately have (126).

The proof of (127) follows from (122) and (25), and the summation:

$$(\text{L.H.S.})(127) \lesssim_{\widetilde{F}} c_k \sum_{j<k-m} |k_* - j|c_j \lesssim_{\widetilde{F}} \widetilde{\epsilon}^{\delta_0}c_k.$$

Notice that the exponential falloff of $\widetilde{\phi}^{(1)}$ was used in the range $k_1 > k_*$ to reduce the factor of $(k-m-j)$ from the previous step to $|k_*-j|$ here, and thus avoid a logarithmic divergence. This concludes our demonstration of Proposition 4.2. $\square$

Proof of Proposition 4.3. We denote the difference to be estimated by:

$$\phi^{high} = \phi - \widetilde{\phi},$$

which represents the evolution of high frequencies in ϕ. This solves the equation:

$$\Box\phi^{high} = -\mathcal{S}(\phi)\partial^\alpha\phi\partial_\alpha\phi + \mathcal{S}(\widetilde{\phi})\partial^\alpha\widetilde{\phi}\partial_\alpha\widetilde{\phi}. \tag{128}$$

A natural attempt is to argue directly as in the preceding proof, namely to replace this nonlinear equation for ϕ^{high} with a linear paradifferential equation plus a nonlinear perturbative term. However, if we do that directly we encounter some difficulties.

Precisely, the initial data $\phi^{high}[0]$ has size on the order of $c_0 = c_0(E)$. Solving the linear paradifferential equation we lose a constant which depends on the $S[I]$ size of the coefficients, namely at least $K(\widetilde{F}(E))$. Thus the solution for the approximate linear flow will have size $c_0(E)K(\widetilde{F}(E))$, and the key point is that we cannot expect this to be small. This would cause the nonlinear effects to be truly non-perturbative, and therefore outside the scope of the current paper.

One fix to this would be to allow c_0 to depend on $\widetilde{F}(E)$ instead of E. However, this would weaken our conclusion to the point where the induction on energy argument only works to show that the set of energies where one has regularity is *open*. While this is the usual first step in an induction on energy strategy, it still leaves one to deal with the heart of the matter which is the task of showing that there is no finite upper bound to the set of regular energies. Our path here will be to establish this latter claim more directly.

As a first step in our argument, we subdivide the time interval I into consecutive subintervals I_k, and we can insure that on each such subinterval we have the partial fungibility:

$$\|\widetilde{\phi}\|_{S[I_k]} \lesssim E. \tag{129}$$

This is possible by estimate (65). This estimate remedies the bootstrapping argument within the first interval because by design ϕ^{high} has small initial energy (see (130) below). However, one might expect that in each subinterval I_i the energy of ϕ^{high} may grow by a $K(E)$ factor, and the number of intervals where (129) is true unfortunately depends on $\widetilde{F}(E)$. Thus a brute force bound would allow the energy of ϕ^{high} to grow by a $K(F(E))$ factor, which would bring us back to the core difficulty. However, such a brute force approach does not make any good use of the fact that both $\widetilde{\phi}$ and ϕ are true Wave-Maps, and therefore exactly conserve their energy.

In order to take advantage of this last observation, we compute that for each fixed $t \in I$:

$$E(\phi) = E(P_{>k_*}\phi) + E(P_{\leqslant k_*}\phi) + 2\langle P_{>k_*}\phi, P_{\leqslant k_*}\phi\rangle \;\geqslant\; E(P_{>k_*}\phi) + E(P_{\leqslant k_*}\phi),$$

where $\langle\cdot,\cdot\rangle$ denotes the $\dot{H}^1 \times L^2$ inner product. The last inequality holds because $P_{\leqslant k_*}P_{>k_*}$ is a nonnegative operator because it has a nonnegative symbol. By Proposition 4.2 just proved, we have the fixed time bound in I:

$$\|(P_{\leqslant k_*}\phi - \widetilde{\phi})[t]\|_{\dot{H}^1\times L^2} \;\lesssim_F\; \epsilon^{\delta_0}.$$

Hence we have:

$$|E(P_{\leqslant k_*}\phi) - E(\widetilde{\phi})| + |E(\phi - \widetilde{\phi}) - E(P_{>k_*}\phi)| \;\lesssim_F\; \epsilon^{\delta_0}.$$

Thus

$$\begin{aligned}
E(\phi^{high}) &\leqslant E(P_{>k_*}\phi) + \epsilon^{\delta_0}K(F(E)),\\
&\leqslant E(\phi) - E(P_{\leqslant k_*}\phi) + \epsilon^{\delta_0}K(F(E)),\\
&\leqslant E(\phi) - E(\widetilde{\phi}) + 2\epsilon^{\delta_0}K(F(E)),\\
&\leqslant c_0(E) + 2\epsilon^{\delta_0}K(F(E)). \tag{130}
\end{aligned}$$

This calculation shows that if ϵ is small enough with respect to the function $K(F)$ which appears implicity on the RHS of estimate (76) then we have a good uniform bound on the energy of ϕ^{high}. The argument now proceeds in a series of steps:

Step 1 (*The bootstrapping construction, and the main estimates*). We now fix the interval $I_i \subseteq I$ and consider an S-norm bootstrap for the ϕ^{high} on subintervals $J \subseteq I_i$, where we may assume J is centered about $t = 0$. We seek to prove the bound:

$$\| \phi^{high} \|_{S[J]} \leqslant 1. \tag{131}$$

Due to (130) we have:

$$\| \phi^{high}[0] \|_{\dot{H}^1 \times L^2} \lesssim c_0. \tag{132}$$

Hence by the second part of Proposition 3.10 we obtain the seed bound:

$$\| \phi^{high} \|_{S[J_0]} \lesssim c_0,$$

for a small enough interval $J_0 \subset J$. Taking this into account and also the continuity of the S norm in Proposition 3.10, it suffices to prove (131) under the additional bootstrap assumption:

$$\| \phi^{high} \|_{S[J]} \leqslant 2. \tag{133}$$

Combining (133) with (129) we obtain:

$$\| \phi \|_{S[J]} + \| \widetilde{\phi} \|_{S[J]} \lesssim_E 1. \tag{134}$$

By Proposition 3.9 this gives:

$$\| \phi \|_{\mathcal{W}[J]} + \| \widetilde{\phi} \|_{\mathcal{W}[J]} \lesssim_E 1, \quad \| \phi \|_{\underline{X}[J]} + \| \widetilde{\phi} \|_{\underline{X}[J]} \lesssim_E \widetilde{\epsilon}^{\delta_1}. \tag{135}$$

We rewrite the bounds (132), (133) and (135) using frequency envelopes. Precisely, we can find a common $(\delta_0/2, \delta_0/2)$-admissible normalized frequency envelope c_k so that $c_{k_*} = 1$ and the following bounds hold:

$$\| \phi^{high}[0] \|_{(\dot{H}^1 \times L^2)_c} \lesssim_E c_0, \tag{136a}$$

$$\| \phi^{high} \|_{S_c[J]} \lesssim_E 1, \tag{136b}$$

$$\| \phi \|_{\mathcal{W}_c[J]} + \| \widetilde{\phi} \|_{\mathcal{W}_c[J]} \lesssim_E 1, \tag{136c}$$

$$\| \phi \|_{\underline{X}_c[J]} + \| \widetilde{\phi} \|_{\underline{X}_c[J]} \lesssim_E \widetilde{\epsilon}^{\delta_1}. \tag{136d}$$

From these four bounds, together with the energy dispersion on lines (73)–(74), we will obtain the following vastly improved frequency envelope S bound for ϕ^{high}:

$$\| \phi^{high} \|_{S_{\lambda c}[J]} \lesssim_E 1, \quad \lambda = c_0 + \widetilde{\epsilon}^{\delta_0 \delta_1^2}. \tag{137}$$

The second term on the right is small ($\ll 1$) due to (75), therefore the desired conclusion (131) follows if c_0 is chosen appropriately small, $c_0 \ll_E 1$.

It remains to show that (73)–(74) together with (136) imply (137). By estimate (83), we may reduce this demonstration to the frequency range $k > k_* - 10$. The mechanics of our argument is to decompose the P_k frequency localized version of (128) as follows:

$$\Box \phi_k^{high} + 2A^\alpha(\phi)_{<k-m} \partial_\alpha \phi_k^{high} = \mathcal{T}_k^m(\widetilde{\phi}) + \mathcal{T}_k^m(\phi) + \mathcal{L}_k^m(\widetilde{\phi}, \phi^{high}), \tag{138}$$

where the large gap parameter m is consistent with Proposition 3.6:

$$2^{-m} = \widetilde{\epsilon}^{\delta_1}. \tag{139}$$

Here the terms $\mathcal{T}_k^m$ are matched frequency trilinear expressions of the form (53), while the term $\mathcal{L}_k^m$ denotes certain trilinear expressions between $\widetilde{\phi}$ and ϕ^{high} which contain at least one (m dependent) low or high frequency factor with improved exponential bounds.

Our first round of estimates shows that:

$$\| \mathcal{T}_k^m(\widetilde{\phi}) \|_{N[J]} \lesssim_E \widetilde{\epsilon}^{\delta_1^2} c_k, \quad \| \mathcal{T}_k^m(\phi) \|_{N[J]} \lesssim_E \epsilon^{\delta_1^2} c_k. \tag{140}$$

Our second round of estimates gives the exponential control:

$$\| \mathcal{L}_k^m(\widetilde{\phi}, \phi^{high}) \|_{N[J]} \lesssim_E 2^{-\frac{1}{4}\delta_0 m} 2^{-\frac{1}{2}\delta_0 |k-k_*|} = \widetilde{\epsilon}^{\frac{1}{4}\delta_0 \delta_1} 2^{-\frac{1}{2}\delta_0 |k-k_*|}. \tag{141}$$

An application of (41) using (136a) and (140)–(141) then implies (137).

Step 2 (Algebraic derivation of (138)). We first write the frequency localized equation for ϕ^{high} as follows:

$$\begin{aligned}
\Box \phi_k^{high} = &- 2\mathcal{S}(\phi)_{<k-m} \partial^\alpha \phi_{<k-m} \partial_\alpha \phi_k + 2\mathcal{S}(\widetilde{\phi})_{<k-m} \partial^\alpha \widetilde{\phi}_{<k-m} \partial_\alpha \widetilde{\phi}_k \\
&+ \mathcal{T}_{1;k}^m(\phi) + \mathcal{T}_{1;k}^m(\widetilde{\phi}),
\end{aligned}$$

where the $\mathcal{T}_{1;k}^m$ are trilinear forms as defined on line (53). Adding to this a zero expression similar to (107) (i.e. without $P_{\leqslant k_*}$ on the second factor), and further decomposing the result into principle terms and $\mathcal{T}^m$ interactions, we have:

$$\Box \phi_k^{high} = -2A^\alpha(\phi)_{<k-m} \partial_\alpha \phi_k + 2A^\alpha(\widetilde{\phi})_{<k-m} \partial_\alpha \widetilde{\phi}_k + \mathcal{T}_k^m(\phi) + \mathcal{T}_k^m(\widetilde{\phi}).$$

Then Eq. (138) is achieved by setting:

$$\mathcal{L}_k^m = - \left(A^\alpha(\widetilde{\phi} + \phi^{high})_{<k-m} - A^\alpha(\widetilde{\phi})_{<k-m} \right) \partial_\alpha \widetilde{\phi}_k. \tag{142}$$

Step 3 (Control of matched frequency interactions). Here we prove (140). This is an immediate consequence of (55) using (136c)–(136d).

Step 4 (Control of separated frequency interactions). Here we prove the estimate (141). We decompose line (142) as a sum of two terms $\mathcal{L}_k^m = R_1 + R_2$ where:

$$\begin{aligned}
R_{1;k} &= -\mathcal{S}(\widetilde{\phi})_{<k-m} \partial^\alpha \phi_{<k-m}^{high} \partial_\alpha \widetilde{\phi}_k, \\
R_{2;k} &= \left[\mathcal{S}'(\widetilde{\phi}, \phi^{high}) \phi^{high} \right]_{<k-m} \partial^\alpha \phi_{<k-m} \partial_\alpha \widetilde{\phi}_k,
\end{aligned}$$

where $\mathcal{S}$ now denotes the antisymmetrization of the original second fundamental form, and $\mathcal{S}'$ is defined as on line (120). Recall that we are restricted to the conditions $k \geqslant k_* - 10$ and $m \geqslant 20$. We proceed to estimate each term separately. In doing so, we repeatedly use the following estimates:

$$\| P_k \widetilde{\phi} \|_{S[J]} \lesssim 2^{-\delta_0(k-k_*)}, \quad k \geqslant k_*, \tag{143}$$

$$\| P_k \phi^{high} \|_{S[J]} \lesssim 2^{-\delta_0(k_*-k)}, \quad k \leqslant k_*, \tag{144}$$

$$\| P_k \left[\mathcal{S}'(\widetilde{\phi}, \phi^{high}) \phi^{high} \right] \|_{S[J]} \lesssim_E 2^{-\delta_0(k_*-k)}, \quad k \leqslant k_*, \tag{145}$$

$$\| \mathcal{S}'(\widetilde{\phi}, \phi^{high}) \phi^{high} \|_{S[J]} \lesssim_E 1. \tag{146}$$

Estimates (143)–(144) are simply a weaker restatement of (83) for the convenience of the reader. Estimates (145)–(146) follow from (144), (133)–(134), and the Moser and product estimates (19)–(20) and (26)–(27) after a standard summation argument.

Step 4A (*Estimating $R_{1;k}$*). After an application of (21)–(22) to peel off the first factor, it suffices to show the bound:

$$\| \partial^\alpha \phi^{high}_{<k-m} \partial_\alpha \widetilde{\phi}_k \|_{N[J]} \lesssim_E 2^{-\frac{1}{4}\delta_0 m} 2^{\frac{1}{2}\delta_0(k_*-k)}.$$

If $k < k_*$ this follows at once from (134) and (144), and summing over (24).

If $k > k_*$, we use (143), and (144) or (133) in (24) to obtain after summation:

$$\| \partial^\alpha \phi^{high}_{<k-m} \partial_\alpha \widetilde{\phi}_k \|_{N[J]} \lesssim_E 2^{-\delta_0(k-k_*)} \sum_{j<k-m} 2^{-\delta_0(k_*-j)_+}.$$

If $k_* < k < k_* + m$ then the expression on the right gives $2^{-\delta_0 m}$ which suffices. If $k > k_* + m$ then the expression on the right gives $|k - k_*| 2^{-\delta_0(k-k_*)}$ which is again sufficient for (141).

Step 4B (*Estimating $R_{2;k}$*). In this final step we show the estimate:

$$\| R_{2;k} \|_{N[J]} \lesssim_E 2^{-\frac{1}{4}\delta_0 m} 2^{\frac{1}{2}\delta_0(k_*-k)}.$$

In the case when $k \leqslant k^*$, using (145), (134), and the trilinear estimate (25) we have the sum:

$$\| R_{2;k} \|_{N[J]} \lesssim_E \sum_{j,k_1<k-m} 2^{-\delta(j-k_1)_+} 2^{-\delta_0(k_*-j)} \lesssim_E 2^{-\delta_0 m} 2^{-\delta_0|k-k_*|}.$$

Finally, in the case where $k > k_*$ we use (134), (143), and (145)–(146) in conjunction with (25) to achieve the sum:

$$\| R_{2;k} \|_{N[J]} \lesssim_E 2^{-\delta_0(k-k_*)} \sum_{j,k_1<k-m} 2^{-\delta(j-k_1)_+} 2^{-\delta_0(k_*-j)_+}.$$

If $k_* < k < k_* + m$ then the expression on the right gives $2^{-\delta_0 m}$ which is enough for (141). If $k > k_* + m$ then the expression on the right gives $|k - k_*| 2^{-\delta_0(k-k_*)}$ which is again sufficient. This concludes our demonstration of Proposition 4.3.   □

5. The Iteration Spaces: Basic Tools and Estimates

This is a continuation of Sect. 2, and our purpose is to fill in any gap between the notation and additional structure of basic function spaces used in this paper and the spaces developed in [32,33 and 29].

5.1. Space-time and angular frequency cutoffs. As usual we denote by Q_j the multiplier with symbol:

$$q_j(\tau, \xi) = \varphi\left(2^{-j} \, ||\tau| - |\xi||\right),$$

where φ truncates smoothly on a unit annulus. We denote by $Q_j^{\pm}$ the restriction of this multiplier to the upper or lower time frequency space. At times we also denote by $Q_{|\tau| \lesssim |\xi|} = Q_{<C}$ for some $C > 0$.

We denote by $\kappa \in K_l$ a collection of caps of diameter $\sim 2^{-l}$ providing a finitely overlapping cover of the unit sphere. According to this decomposition, we cut up the spatial frequency domain according to:

$$P_k = \sum_{\kappa \in K_l} P_{k,\kappa}.$$

These decompositions often occur in conjunction with modulation cutoffs on the order of $j = k - 2l$, and a central principle is that the corresponding multipliers $Q_{<k-2l}^{\pm} P_{k,\pm\kappa}$ are uniformly disposable.

5.2. The S and N function spaces.

Definition 5.1 (Dyadic Iteration Space). *For each integer k we define the following frequency localized norm:*

$$\| \phi \|_{S_k} := \| \nabla_{t,x} \phi_k \|_{L_t^{\infty}(L_x^2)} + \| \nabla_{t,x} \phi_k \|_{X_{\infty}^{0,\frac{1}{2}}} + \| \phi_k \|_{\underline{S}} + \sup_{j < k - 20} \| \phi \|_{S[k;j]}.$$

$$(147)$$

In general, the fixed frequency space $X_p^{s,b}$ is defined as:

$$\| P_k \phi \|_{X_p^{s,b}}^p := 2^{psk} \sum_j 2^{pbj} \| Q_j P_k \phi \|_{L_t^2(L_x^2)}^p,$$

with the obvious definition for $X_{\infty}^{s,b}$. Here we define the "physical space Strichartz" norms:

$$\| \phi_k \|_{\underline{S}} := \sup_{(q,r):\frac{2}{q}+\frac{1}{r} \leqslant \frac{1}{2}} 2^{(\frac{1}{q}+\frac{2}{r}-1)k} \| \nabla_{t,x} \phi_k \|_{L_t^q(L_x^r)}, \qquad (148)$$

the "modulational Strichartz" norms:

$$\| \phi \|_{S[k;j]} := \sup_{\pm} \left(\sum_{\kappa \in K_l} \| Q_{<k-2l}^{\pm} P_{k,\pm\kappa} \phi \|_{S[k,\kappa]}^2 \right)^{\frac{1}{2}}, \quad l = \frac{k-j}{2} > 10, \quad (149)$$

and the "angular Strichartz" space in terms of the three components:

$$\| \phi \|_{S[k,\kappa]} := 2^k \sup_{\omega \notin 2\kappa} dist(\omega, \kappa) \| \phi \|_{L_{t_\omega}^{\infty}(L_{w_\omega}^2)} + 2^k \| \phi \|_{L_t^{\infty}(L_x^2)}$$

$$+ 2^{\frac{1}{2}k} |\kappa|^{-\frac{1}{2}} \inf_{\sum_\omega \phi^\omega = \phi} \sum_\omega \| \phi^\omega \|_{L_{t_\omega}^2(L_{x_\omega}^{\infty})}. \qquad (150)$$

The first component on the RHS above will often be referred to as NFA. We define S as
the space of functions ϕ in $\mathbb{R}^{2+1}$ with $\nabla_{x,t}\phi \in C(\mathbb{R}; L_x^2)$ and finite norm:*

$$\| \phi \|_S^2 = \| \phi \|_{L_t^\infty(L_x^\infty)}^2 + \sum_k \| \phi \|_{S_k}^2,$$

and also use the frequency envelope convention from Sect. 2 to define S_c.

To measure the derivatives of functions in S we introduce a related space DS:

Definition 5.2 (Differentiated S functions). *We define the norm:*

$$\| \phi \|_{DS_k} := \| \phi_k \|_{L^\infty(L_x^2)} + \| \phi_k \|_{X_\infty^{0,\frac{1}{2}}} + \| \phi_k \|_{D\underline{S}_k} + 2^k \sup_{j<k-20} \| \phi \|_{S[k;j]}, \quad (151)$$

where the $D\underline{S}$ norm is as in (148) but without the gradient:

$$\| \phi \|_{D\underline{S}_k} := \sup_{(q,r):\frac{2}{q}+\frac{1}{r}\leqslant\frac{1}{2}} 2^{(\frac{1}{q}+\frac{2}{r}-1)k} \| \phi \|_{L_t^q(L_x^r)}. \quad (152)$$

*The DS space is defined as the space of functions for which the square sum of the DS_k
norms is finite:*

$$\| \phi \|_{DS}^2 = \sum_k \| \phi \|_{DS_k}^2.$$

We remark that by definition we have:

$$\| \phi \|_{S_k} \approx \| \nabla_{x,t}\phi \|_{DS_k}. \quad (153)$$

Definition 5.3 (Dyadic Source Term Space). *For each integer k we define the following
frequency localized norm:*

$$\| F \|_{N_k} := \inf_{F_A+F_B+\sum_{l,\kappa} F_C^{l,\kappa}=F} \left(\| P_k F_A \|_{L_t^1(L_x^2)} + \| P_k F_B \|_{X_1^{0,-\frac{1}{2}}} \right.$$

$$\left. + \sum_\pm \sum_{l>10} \left(\sum_\kappa \inf_{\omega\notin 2\kappa} dist(\omega,\kappa)^{-2} \| Q_{<k-2l}^\pm P_{k,\pm\kappa} F_C^{l,\kappa} \|_{L_{t_\omega}^1(L_{x_\omega}^2)}^2 \right)^{\frac{1}{2}} \right). \quad (154)$$

*We will often refer to the last component on the RHS above as NFA, and the norm applied
to a fixed $Q^\pm F_C^{l,\kappa}$ as NFA$[\pm\kappa]$.*

For any *closed* interval $I = [t_0, t_1]$ we define spaces $\underline{X}[I], \underline{X}_c[I], \underline{E}[I], L_t^\infty(L_x^\infty)[I]$,
etc. as the restriction of these classical L^p based norms to the time slab $I \times \mathbb{R}_x^2$. We also
need a similar procedure for the non-local S and N spaces. As usual we define $S_k[I]$,
$S[I]$, $S_c[I]$, $N[I]$, etc. in terms of minimal extension.[4] For example:

$$\| \phi \|_{S_k[I]} = \inf_\Phi \left\{ \| \Phi \|_{S_k} \mid \Phi \equiv \phi \text{ on } I \times \mathbb{R}^2 \right\}. \quad (155)$$

---

[4] We will modify this procedure somewhat below by an equivalent norm, but for the most part they are
interchangeable.

On an *open* time interval (t_0, t_1) we may also define localized norms by taking $\|\phi\|_{S(t_0,t_1)} = \sup_{I \subseteq (t_0,t_1)} \|\phi\|_{S[I]}$. This definition will only be important for us as a convenience when stating results like Theorem 1.3, so the reader is safe to ignore the distinction and always assume that I denotes a *closed* time interval. We now state a continuation of Proposition 2.3:

Proposition 5.4 (Standard Estimates and Relations: Part 2). *Let F, ϕ, and $\phi^{(i)}$ be a collection of test functions, $I \subseteq \mathbb{R}$ any subinterval (including $\mathbb{R}$ itself). Then the following list of properties for the $S[I]$ and $N[I]$ spaces hold:*

- *Time Truncation of S. Let χ_I be the characteristic function of I. Then*

$$\|\phi\|_{DS_k[I]} \approx \|\chi_I \phi\|_{DS_k} \lesssim \|\phi\|_{DS_k}, \tag{156}$$

$$\|\phi\|_{S_k[I]} \approx \|\chi_I \nabla_{x,t} \phi\|_{DS_k}. \tag{157}$$

- *Time Truncation of N. Let $I = \cup_i^K I_i$ be a decomposition of I into consecutive intervals, and let χ_I, χ_{I_i} be the corresponding sharp time cutoffs. Then the following bounds hold (uniform in K):*

$$\|\chi_I F\|_N \lesssim \|F\|_N, \tag{158}$$

$$\sum_i \|\chi_{I_i} F\|_N^2 \lesssim \|F\|_N^2. \tag{159}$$

Furthermore, for any Schwartz function F the quantity $\|\chi_I F\|_N$ is continuous in the endpoints of I.

- *Basic S and N Relations. We have that:*

$$\|\phi_{k_1}^{(1)} \cdot \phi_{k_2}^{(2)}\|_{DS[I]} \lesssim 2^{(k_1-k_2)} \|\phi_{k_1}^{(1)}\|_{DS[I]} \cdot \|\phi_{k_2}^{(2)}\|_{S[I]}, \quad k_1 < k_2 - 10, \tag{160}$$

$$\langle \phi, F_k \rangle \lesssim \|\phi\|_{DS} \cdot \|F_k\|_N, \tag{161}$$

$$\|\phi_k\|_S \lesssim \|\nabla_{t,x}\phi_k\|_{X_1^{0,\frac{1}{2}}}, \tag{162}$$

$$\|F_k\|_{X_\infty^{0,-\frac{1}{2}}} \lesssim \|F_k\|_N. \tag{163}$$

- $L_t^q(L_x^r)$ *and Disposability Estimates. We have that:*

$$\|Q_j \phi_k\|_{L_t^2(L_x^\infty)} \lesssim 2^{-(j-k)_+} 2^{-\frac{1}{2}j} \|\phi_k\|_S, \tag{164}$$

$$\|\phi_k\|_{L_t^\infty(L_x^2)} + \|Q_{\leqslant j}\phi_k\|_{L_t^\infty(L_x^2)} + \|Q_j \phi_k\|_{L_t^\infty(L_x^2)} \lesssim 2^{-k} \|\phi_k\|_S, \tag{165}$$

$$\|Q_{\leqslant j}\phi_k\|_{L_t^\infty(L_x^\infty)} + \|Q_j \phi_k\|_{L_t^\infty(L_x^\infty)} \lesssim \|\phi_k\|_S. \tag{166}$$

- *Fine Product Estimates. We have that:*

$$\|\phi_{k_1}^{(1)} \cdot \phi_{k_2}^{(2)}\|_{L_t^2(L_x^2)} \lesssim |\kappa|^{\frac{1}{2}} 2^{-\frac{1}{2}k_1} \|\phi_{k_1}^{(1)}\|_{S[k_1,\kappa]} \cdot \sup_{\omega \in \kappa} \|\phi_{k_2}^{(2)}\|_{L_{t_\omega}^\infty(L_{x_\omega}^2)}, \tag{167}$$

$$\|P_{<j-10}Q_{<j-10}\phi^{(1)} \cdot \phi_{k_2}^{(2)}\|_{S[k_2;j]} \lesssim \|\phi^{(1)}\|_{L_t^\infty(L_x^\infty)} \cdot \|\phi_{k_2}^{(2)}\|_S, \tag{168}$$

$$\|\nabla_{t,x} P_k Q_j (\phi_{k_1}^{(1)} \cdot \phi_{k_2}^{(2)})\|_{X_1^{0,\frac{1}{2}}} \lesssim 2^{\delta(k-\max\{k_i\})} 2^{\delta(j-\min\{k_i\})} \|\phi_{k_1}^{(1)}\|_S \|\phi_{k_2}^{(2)}\|_S, \tag{169}$$

$$\|P_k(Q_j F_{k_1} \cdot \phi_{k_2})\|_N \lesssim 2^{\delta(k-\max\{k_i\})} 2^{\delta(j-\min\{k_i\})} \|F_{k_1}\|_{X_\infty^{0,-\frac{1}{2}}} \|\phi_{k_2}\|_S, \tag{170}$$

where in estimates (169)–(170) we are assuming $j \leqslant \min\{k_i\}$.

Estimates (156)–(160) are proved next. The rest of the above bounds are standard, and with the exception of (168) which is Lemma 9.1 in [33], may be found in [29]. For the convenience of the reader we give the detailed citations ([29] page numbers). Estimate (161) is estimate (94) on p. 487. Estimate (162) is Lemma 8 on p. 483. Estimate (163) is Lemma 10 on p. 487. Estimates (164)–(166) are listed in estimates (81)–(84) on p. 483. Estimate (167) is (by duality) the estimate in Step 2 on p. 479, and it also follows more or less immediately by inspecting the third term on line (150) above. Estimate (169) is Lemma 13 on p. 515. Estimate (170) is Lemma 12 on p. 501.

Proof of estimates (156)–(157) *and* (15). Without any loss of generality we replace χ_I by $\chi = \chi_{t<0}$. Our main observation here will be that the multipliers Q_j applied to χ act like time-frequency cutoffs onto dyadic sets $|\tau| \sim 2^j$. For each of these we have the Strichartz type estimate:

$$\| Q_j \chi \|_{L^2_t(L^\infty_x)} \lesssim 2^{-\frac{1}{2}j}. \tag{171}$$

Therefore, one can look upon the estimate (156) as some version of the product bound (19). We rescale to $k = 0$, and set $\phi_0 = P_0\phi$.

We begin with the proof of (156). The $D\underline{S}$ bound in this estimate is immediate. Therefore we focus on proving the $X^{s,b}$ and $S[0, \kappa]$ sum portions of the estimate. This is split into cases:

Step 1 (*Controlling the $X_\infty^{0,\frac{1}{2}}$ norm*). Freezing Q_j, our goal is to show that:

$$\| Q_j (\chi \cdot \phi_0) \|_{L^2_t(L^2_x)} \lesssim 2^{-\frac{j}{2}} \| \phi_0 \|_{DS}. \tag{172}$$

We now split into subcases.

Step 1.A (*χ at low modulation*). In this case we look at the contribution of the product $Q_j(Q_{<j-10}\chi \cdot \phi_0)$. We may freely insert the multiplier $Q_{[j-5,j+5]}$ in front of ϕ_0. Then (172) is immediate from L^∞ control of $Q_{<j-10}\chi$.

Step 1.B (*χ at high modulation*). In this case we'll rely on the even stronger L^2 bound:

$$\| Q_{\geqslant j-10}\chi \cdot \phi_0 \|_{L^2_t(L^2_x)} \lesssim \| Q_{\geqslant j-10}\chi \|_{L^2_t(L^\infty_x)} \| \phi_0 \|_{L^\infty_t(L^2_x)} \lesssim 2^{-\frac{1}{2}j} \| \phi_0 \|_S, \tag{173}$$

which results from summing over (171). In particular, isolating the LHS of this last line at frequency Q_j we have (172) for this term.

Step 2 (*Controlling the $S[0; j]$ norms*). Freezing $j < -20$ we need to demonstrate:

$$\| Q_{<j} (\chi \cdot \phi_0) \|_{S[0;j]} \lesssim \| \phi_0 \|_{DS}.$$

Step 2.A (*χ at low modulation*). The contribution of $Q_{<j-10}\chi \cdot Q_{<j}\phi_0$ is bounded via estimate (168). Notice that χ is automatically at zero spatial frequency.

Step 2.B (*χ at high modulation*). Adding over estimate (173) we have:

$$\| Q_{<j}(Q_{\geqslant j-10}\chi \cdot \phi_0) \|_{X_1^{0,\frac{1}{2}}} \lesssim \| \phi_0 \|_S,$$

which is sufficient via the differentiated version of (162).

To wrap things up here, we need to demonstrate the bounds (157) and (15). Beginning with $\phi \in S[I]$ we consider an extension $\phi \in S$ with comparable norm. Then by (156) and (153) we have the chain of inequalities:

$$\| \nabla_{x,t}\phi \|_{DS_k[I]} \lesssim \| \chi_I \nabla_{x,t}\phi_k \|_{DS_k} \lesssim \| \nabla_{x,t}\phi_k \|_{DS_k} \approx \| \phi \|_{S_k}.$$

It remains to prove the converse. We begin with the energy norm, observing that for $\phi \in S[I]$, for $I = [-i_0, i_0]$, we have:

$$\| \phi_k[\pm i_0] \|_{\dot{H}^1 \times L^2} \lesssim \| \nabla_{x,t} \phi_k \|_{DS_k[I]}.$$

We extend ϕ to $I^{\pm} = [\pm i_0, \pm \infty)$ as a solution to the homogeneous wave equation with data $\phi[\pm i_0]$ and use (153) to compute:

$$\begin{aligned}
\| \phi \|_{S_k[I]} \lesssim \| \phi \|_{S_k} &\approx \| \nabla_{x,t} \phi \|_{DS_k} \lesssim \| \chi_I \nabla_{x,t} \phi \|_{DS_k} + \| (1 - \chi_I) \nabla_{x,t} \phi \|_{DS_k} \\
&\lesssim \| \nabla_{x,t} \phi \|_{DS_k[I]} + \| \phi_k[\pm i_0] \|_{\dot{H}^1 \times L^2} \\
&\lesssim \| \nabla_{x,t} \phi \|_{DS_k[I]}.
\end{aligned}$$

The proof of (157) is concluded.

Finally, we use (156) and (157) to prove (15):

$$\| \phi \|_{S_k[I]} \approx \| \nabla_{x,t} \phi \|_{DS_k[I]} \lesssim \sum_i \| \chi_{I_i} \nabla_{x,t} \phi \|_{DS_k} \lesssim \sum_i \| \phi \|_{S_k[I_i]}.$$

The proof is concluded. $\square$

Proof of estimate (159). Since:

$$\| F \|_N^2 \approx \sum_k \| P_k F \|_{N_k}^2,$$

it suffices to show that the similar relation holds for the N_k spaces:

$$\sum_n \| \chi_{I_n} F \|_{N_k}^2 \lesssim \| F \|_{N_k}^2. \tag{174}$$

The space N_k is an atomic space, therefore is suffices to prove (174) for each atom.

Step 1 ($L_t^1(L_x^2)$ *atoms*). For these we directly have the stronger relation:

$$\sum_n \| \chi_{I_n} F_k \|_{L_t^1(L_x^2)} \lesssim \| F_k \|_{L_t^1(L_x^2)}.$$

Step 2 ($\dot{X}_1^{0,-\frac{1}{2}}$ *atoms*). For F localized at frequency 2^k we will prove the relation:

$$\sum_n \| \chi_{I_n} F_k \|_{L_t^1(L_x^2) + \dot{X}_1^{0,-\frac{1}{2}}}^2 \lesssim \| F_k \|_{\dot{X}_1^{0,-\frac{1}{2}}}^2.$$

Without any restriction in generality we can assume that F_k is also localized in modulation at 2^j. By rescaling we can take $j = 0$. At modulation 1 the $\dot{X}_1^{0,-\frac{1}{2}}$ is equivalent to the $L_t^2(L_x^2)$ norm. Then the last bound would follow from the stronger estimate:

$$\sum_n \| Q_{<-4}(\chi_{I_n} Q_0 F_k) \|_{L_t^1(L_x^2)}^2 + \| Q_{>-4}(\chi_{I_n} Q_0 F_k) \|_{L_t^2(L_x^2)}^2 \lesssim \| Q_0 F_k \|_{L_t^2(L_x^2)}^2. \tag{175}$$

We trivially have:

$$\sum_n \| \chi_{I_n} Q_0 F_k \|_{L_t^2(L_x^2)}^2 \lesssim \| Q_0 F_k \|_{L_t^2(L_x^2)}^2,$$

therefore it remains to prove the $L_t^1(L_x^2)$ bound on line (175). We do this in two cases:

Step 2.A (*Small intervals*). We parse the collection of intervals I_n into two sub-collections, intervals J_n such that $|J_n| \geqslant 1$, and intervals K_n such that $|K_n| < 1$. In the latter case we may drop the outer $Q_{<-4}$ and simply use Hölder's inequality to estimate:

$$\sum_n \| \chi_{K_n} Q_0 F_k \|^2_{L^1_t(L^2_x)} \lesssim \sum_n \| \chi_{K_n} Q_0 F_k \|^2_{L^2_t(L^2_x)},$$

so the estimate follows as above.

Step 2.B (*Large intervals*). In this case we break the first term on LHS (175) up as follows:

$$\sum_m \| Q_{<-4}(\chi_{J_m} Q_0 F_k) \|^2_{L^1_t(L^2_x)} = \sum_m \| Q_{<-4}(Q_{[-10,10]} \chi_{J_m} \cdot Q_0 F_k) \|^2_{L^1_t(L^2_x)}$$
$$\lesssim \sum_m \| Q_{[-10,10]} \chi_{J_m} \cdot Q_0 F_k \|^2_{L^1_t(L^2_x)}. \qquad (176)$$

Denoting $J_m = [a_m, b_m]$, for $Q_{[-10,10]} \chi_{J_m}$ we have the pointwise bounds:

$$|Q_{[-10,10]} \chi_{J_m}(t)| \leqslant (1 + |t - a_m|)^{-N} + (1 + |t - b_m|)^{-N}.$$

Hence by Cauchy-Schwartz we obtain:

$$\text{L.H.S.}(176) \lesssim \sum_m \| (1 + |t - a_m|)^{-\frac{N}{2}} Q_0 F_k \|^2_{L^2_t(L^2_x)} + \| (1 + |t - b_m|)^{-\frac{N}{2}} Q_0 F_k \|^2_{L^2_t(L^2_x)}$$
$$\lesssim \int_{\mathbb{R}} \sum_m \left((1 + |t - a_m|)^{-N} + (1 + |t - b_m|)^{-N} \right) \| Q_0 F_k(t) \|^2_{L^2_x} dt.$$

Since the intervals J_m are disjoint and of size at least 1, the last sum above is bounded by $\lesssim 1$, therefore we obtain:

$$\text{L.H.S.}(176) \lesssim \| Q_0 F_k \|^2_{L^2_t(L^2_x)}.$$

Step 3 (*NFA atoms*). In this case we can express F_k as:

$$F_k = F_k^+ + F_k^- = \sum_{\pm, \kappa} F_{k, \pm \kappa}^{\pm},$$

where $F_{k, \pm \kappa}^{\pm}$ is supported in the wedge carved by the multiplier $Q_{<k-2j}^{\pm} P_{k, \pm \kappa}$, with $j > 10$, and furthermore:

$$\sum_{\kappa} \| F_{k, \kappa}^{\pm} \|^2_{NFA[\pm\kappa]} \leqslant 1.$$

Without loss of generality we may assume we are in the $+$ case, and we rescale to $k = 2j$, and so in particular $k > 20$.

By summing over (163) we have the $L^2_t(L^2_x)$ bound:

$$\| F_k^+ \|_{L^2_t(L^2_x)} \lesssim 1. \qquad (177)$$

The *NFA*[κ] norms are translation invariant, and are defined using characteristic $L^1_{t_\omega}(L^2_{x_\omega})$ norms. Thus they directly satisfy the inequality:

$$\sum_n \| \chi_{I_n} F_{k, \kappa}^+ \|^2_{NFA[\kappa]} \leqslant \| F_{k, \kappa}^+ \|^2_{NFA[\kappa]}. \qquad (178)$$

We write:

$$\chi_{I_n} F_k^+ = Q_{>0}(\chi_{I_n} F_k^+) + Q_{<0}^-(\chi_{I_n} F_k^+) + \sum_\kappa Q_{<0}^+(\chi_{I_n} F_{k,\kappa}^+),$$

and estimate the first component in $\dot{X}_1^{0,-\frac{1}{2}}$, the second in $L_t^1(L_x^2)$, and the third in *NFA*. We have from line (177):

$$\sum_n \| Q_{>0}(\chi_{I_n} F_k^+) \|_{\dot{X}_1^{0,-\frac{1}{2}}}^2 \lesssim \sum_n \| \chi_{I_n} F_k^+ \|_{L_t^2(L_x^2)}^2 \lesssim \| F_k^+ \|_{L_t^2(L_x^2)}^2 \lesssim 1.$$

Next, using the restriction on the Fourier support of F_k^+, we have for any single interval the bound:

$$\| Q_{<0}^-(\chi_{J_n} F_k^+) \|_{L_t^1(L_x^2)} \lesssim \| Q_{[-20,20]} \chi_{J_n} \cdot F_k^+ \|_{L_t^1(L_x^2)}.$$

Using this and (177), one may proceed as in Step 2.A and Step 2.B above. On the other hand, by (178) and the disposability of $Q_{<0}^+$ on the Fourier support of the multiplier $P_{k,\kappa}$ we have:

$$\sum_{n,\kappa} \| Q_{<0}^+(\chi_{I_n} F_{k,\kappa}^+) \|_{NFA[\kappa]}^2 \lesssim \sum_{n,\kappa} \| \chi_{I_n} F_{k,\kappa}^+ \|_{NFA[\kappa]}^2 \lesssim \sum_\kappa \| F_{k,\kappa}^+ \|_{NFA[\kappa]}^2 \lesssim 1.$$

$\square$

Proof of estimate (160). This is a minor variation of (21), and the proof is similar to that of estimate (156) above. We rescale to $k_2 = 0$, discard I, and set $\| \phi_{k_1}^{(1)} \|_{DS} = \| \phi_0^{(2)} \|_S = 1$. Using the fact that:

$$\| Q_{<k_1+10} \phi_{k_1}^{(1)} \|_S \lesssim 2^{k_1} \| \phi_{k_1}^{(1)} \|_{DS}, \quad \| P_0 f \|_{DS} \lesssim \| f \|_S,$$

along with the usual S algebra estimate (19), we control the low modulation contribution of the first factor. Furthermore, it is always possible to gain in the $D\underline{S}$ component by using Bernstein's inequality and energy estimates for the high frequency factor.

We are reduced to bounding the contribution of $Q_{>k_1+10} \phi_{k_1}^{(1)} \cdot \phi_0^{(2)}$. As a general tool we have the L^2 bound:

$$\| Q_{>j-10} \phi_{k_1}^{(1)} \cdot \phi_0^{(2)}) \|_{L_t^2(L_x^2)} \lesssim \| Q_{>j-10} \phi_{k_1}^{(1)} \|_{L_t^2(L_x^\infty)} \| \phi_0^{(2)} \|_{L_t^\infty(L_x^2)} \lesssim 2^{k_1} 2^{-\frac{1}{2}j}.$$

In particular, via the differentiated version of (162), if the output modulation is $j < k_1$ we have both:

$$\| Q_j(Q_{>k_1+10} \phi_{k_1}^{(1)} \cdot \phi_0^{(2)}) \|_{X_\infty^{0,\frac{1}{2}}} + \| Q_{>k_1+10} \phi_{k_1}^{(1)} \cdot \phi_0^{(2)} \|_{S[0;j]} \lesssim 2^{k_1}.$$

On the other hand, if the output modulation is $j > k_1$, then by again using the above general L^2 estimate, it suffices to show:

$$\| Q_j(Q_{[k_1+10, j-10]} \phi_{k_1}^{(1)} \cdot \phi_0^{(2)}) \|_{X_\infty^{0,\frac{1}{2}}} + \| Q_{[k_1+10, j-10]} \phi_{k_1}^{(1)} \cdot \phi_0^{(2)} \|_{S[0;j]} \lesssim \| \phi_{k_1}^{(1)} \|_{L_x^\infty(L_x^\infty)},$$

and then conclude via an application of Bernstein's inequality. For the first term on the LHS, we may freely insert a $Q_{[j-5, j+5]}$ multiplier in front of the second factor, which suffices. For the second term, we directly use (168). $\square$

5.3. Extension and restriction for S and N functions. In the sequel we will build up estimates through an iterative process by which we first prove bounds in weak spaces (such as $P_k L^\infty$, $\underline{X}$, and $\underline{E}$), and then show that these may be used in conjunction with bootstrapping to establish uniform bounds in much stronger spaces (such as S). This process unfortunately leads to some technical difficulties regarding compatibility of extensions in various norms. To tame this difficulty, we will make use of a variable but universal extension process. Because this feature is more of a technicality in our proof, we state here for the convenience of the reader where such extensions are necessary in the sequel:

- The primary use of compatible extensions is in the proof of Proposition 3.4, most importantly in the proof of estimate (46). Here we are forced to use several norms simultaneously in a single estimate that involves space-time frequency cutoffs. As will become apparent soon, in such a situation choosing extensions needs to be done carefully because it is not immediate that this can be done in a way that retains smallness of the various component norms.
- Universal extensions are also used in a key way in the proof of Proposition 3.2 because we need to know that extensions still enjoy good characteristic energy estimates when these estimates are only known on a finite interval. This extended control needs to be used in conjunction with S norm control in estimates requiring space-time frequency cutoffs (see Lemma 9.3 in Sect. 9).
- A secondary use of compatible extensions occurs because we do not include $\underline{X}$ as a component of S defined above. Doing this allows us to quote standard product estimates from [29] modulo physical space Strichartz components. The price one pays is that $\underline{X}$ bounds are established separately, and one then needs to include this a-posteriori into extension estimates. For example, this feature is used at the beginning of the proof of Proposition 3.1 to extend the connection B with good S and $\underline{X}$ bounds.

Proposition 5.5 (Existence of S Extensions/Restrictions). *Let ϕ be any affinely Schwartz function defined on an interval $I = [-i_0, i_0]$.*

- *Canonical extension. For every $0 < \eta \leqslant 1$ there exists a canonical extension $\Phi^{I,\eta}$ which is compactly supported in time and for which the following estimates are true:*

$$\| P_k \Phi^{I,\eta} \|_S \lesssim \| P_k \phi \|_{S[I]}, \tag{179}$$

$$\| P_k \square \Phi^{I,\eta} \|_N \lesssim \| P_k \phi \|_{\underline{E}[I]} + \| P_k \square \phi \|_{N[I]}, \tag{180}$$

$$\| P_k \Phi^{I,\eta} \|_{L_t^\infty(L_x^\infty)} \lesssim \eta^{-\frac{5}{2}} \| P_k \phi \|_{L_t^\infty(L_x^\infty)[I]} + \eta^{\frac{1}{2}} \| P_k \phi \|_{\underline{X}[I]}, \ |I| \geqslant 2^{-k}\eta^2, \tag{181}$$

$$\| P_k \Phi^{I,\eta} \|_{\underline{X}} \lesssim \eta^{\frac{1}{2}} \| P_k \phi \|_{\underline{E}[I]} + \| P_k \phi \|_{\underline{X}[I]}, \tag{182}$$

$$\| P_k \Phi^{I,\eta} \|_{\underline{E}} \lesssim \| P_k \phi \|_{\underline{E}[I]}, \tag{183}$$

$$\| P_k \square \Phi^{I,\eta} \cdot \psi_j \|_N \lesssim 2^{k-j} \| P_k \phi \|_{\underline{E}[I]} \| \psi_j \|_S + \| P_k \square \phi \cdot \psi_j \|_{N[I]}, \tag{184}$$

where the last bound holds under the additional condition that $j > k + 10$.

- *Secondary extension. For every $0 < \eta \leqslant 1$ there exists an extension $\widetilde{\Phi}^{I,\eta}$ which is compactly supported in time and such that* (179),(180) *hold. Furthermore, for this extension the following improvement of* (181) *is valid:*

$$\| P_k \widetilde{\Phi}^{I,\eta} \|_{L_t^\infty(L_x^\infty)} \lesssim \eta^{-\frac{1}{2}} \| P_k \phi \|_{L_t^\infty(L_x^\infty)[I]} + \eta \| P_k \phi \|_{\underline{E}[I]}. \tag{185}$$

The canonical extension above will be used most of the time. Its only disadvantage is that in order to control the $L_t^\infty(L_x^\infty)$ norm of this extension we need to also control the $\underline{X}$ norm. In the rare (single) case where this is missing, we use the secondary extension.

Proof of Proposition 5.5. The canonical extension will be defined dyadically for each ϕ_k. By rescaling we only work with $k = 0$.

Step 1 (*The canonical extension and estimates* (179)–(180), (182)–(184)). The obvious candidate Φ^I for the extension is obtained by solving the homogeneous wave equation to the left of $-i_0$ and to the right of i_0, with Cauchy data $P_0\phi[-i_0]$, respectively $P_0\phi[i_0]$. Denoting the complement of I by $I^- \cup I^+$, we have:

$$\Box\Phi^I = 0, \quad \Phi^I[\pm i_0] = P_0\phi[\pm i_0], \quad \text{in } I^\pm.$$

It is relatively easy to verify that the extension Φ^I satisfies all the properties (179)–(180), (182)–(184). However, there is a core issue with (181), as this bound can easily fail because nonconcentration at time $\pm i_0$, say, does not guarantee nonconcentration at all later times. To avoid this problem, we truncate Φ^I outside a compact set and define:

$$\Phi^{I,\eta} = \chi_I^\eta \Phi^I,$$

where χ_I^η is a smooth cutoff with $|\partial_t^k \chi_I^\eta| \lesssim \eta^k$, such that $\chi_I^\eta \equiv 1$ on I and vanishing outside of the extended interval $\widetilde{I} = [-i_0 - \eta^{-1}, i_0 + \eta^{-1}]$. Furthermore, in $I^\pm$ we have the identity:

$$\Box\Phi^{I,\eta} = 2\partial_t(\chi_I^\eta) \cdot \partial_t\Phi^I + \partial_t^2(\chi_I^\eta) \cdot \Phi^I.$$

This allows us to estimate:

$$\| P_0\Box\Phi^{I,\eta} \|_{L_t^1(L_x^2)[I^\pm]} \lesssim \| P_0\phi[\pm i_0] \|_{\dot{H}^1 \times L^2}, \tag{186}$$

which in turn leads to:

$$\| P_0\Phi^{I,\eta} \|_{S[I^\pm]} + \| P_0\Box\Phi^{I,\eta} \|_{N[I^\pm]} \lesssim \| P_0\phi[\pm i_0] \|_{\dot{H}^1 \times L^2}.$$

Then the bound (179) follows from (15), while (183) follows from energy estimates for $\Phi^{I,\eta}$ in $I^\pm$. The bound (180) is also straightforward, while for (182) we need to compute:

$$\| P_0\Box\Phi^{I,\eta} \|_{L_t^2 L_x^2[I^\pm]} \lesssim \eta^{\frac{1}{2}} \| P_0\phi[\pm i_0] \|_{\dot{H}^1 \times L^2}.$$

To prove (184) we use Bernstein to estimate:

$$\| P_0\Box\Phi^{I,\eta} \cdot \psi_j \|_{L_t^1(L_x^2)[I^\pm]} \lesssim \| P_0\Box\Phi^{I,\eta} \|_{L_t^1(L_x^\infty)[I^\pm]} \cdot \| \psi_j \|_{L_t^\infty(L_x^2)[I^\pm]}$$
$$\lesssim \| P_0\Box\Phi^{I,\eta} \|_{L_t^1(L_x^2)[I^\pm]} \cdot 2^{-j} \| P_0\psi_j \|_S,$$

and conclude with (186).

Step 2 (*The $L_t^\infty(L_x^\infty)$ estimate* (181)). We now turn our attention to the most interesting part, namely (181). The desired bound follows from a reverse dispersive estimate for the 2D wave equation:

$$\| e^{\pm it|D_x|} P_0 f \|_{L_x^\infty} \lesssim \sqrt{1+t} \, \| P_0 f \|_{L_x^\infty},$$

provided that we can first establish the "elliptic" estimate (setting $P_0\phi = \phi_0$):

$$\| \partial_t\phi_0 \|_{L_t^\infty(L_x^\infty)[I]} \lesssim \eta^{-2} \| \phi_0 \|_{L_t^\infty(L_x^\infty)[I]} + \eta \| \phi_0 \|_{\underline{X}[I]}, \tag{187}$$

provided that $|I| \geqslant \eta^2$. Without loss of generality we may assume we are in the worst case scenario $|I| = \eta^2$. We begin with the Poincare type inequality:

$$\| \partial_t \phi_0 - (\partial_t \phi_0)^{av} \|_{L_t^\infty [I]} \lesssim \eta \| \partial_t^2 \phi_0 \|_{L_t^2 [I]},$$

where $(\partial_t \phi_0)^{av} = \eta^{-2} \int_I \partial_t \phi_0 dt$, so in particular:

$$\| (\partial_t \phi_0)^{av} \|_{L_x^\infty [I]} \lesssim \eta^{-2} \| \phi_0 \|_{L_t^\infty (L_x^\infty)[I]}.$$

Therefore, taking the sup over all space in the second to last line above we have:

$$\| \partial_t \phi_0 \|_{L_t^\infty (L_x^\infty)[I]} \lesssim \eta^{-2} \| \phi_0 \|_{L_t^\infty (L_x^\infty)[I]} + \eta \| \Box \phi_0 \|_{L_t^2 (L_x^\infty)[I]} + \eta \| \Delta \phi_0 \|_{L_t^2 (L_x^\infty)[I]}.$$

The proof of (187) is now concluded via Bernstein in space for the second term on the RHS above, and Cauchy-Schwartz in time along with the fact that ΔP_0 is bounded on L_x^∞ to control the third.

Step 3 (*The Secondary Extension*). We next turn our attention to the second extension. Again we set $k = 0$. The additional difficulty we face here is that we no longer have pointwise bounds for $\partial_t P_0 \phi(\pm i_0)$. We split the function Φ^I outside I into two parts:

$$\Phi^I = \Phi_0^I + \Phi_1^I,$$

corresponding to the two different components of its Cauchy data at $\pm i_0$:

$$\Box \Phi_0^I = 0, \quad \Phi_0^I [\pm i_0] = (P_0 \phi(\pm i_0), 0) \quad \text{in } I^\pm,$$

respectively:

$$\Box \Phi_1^I = 0, \quad \Phi_1^I [\pm i_0] = (0, \partial_t P_0 \phi(\pm i_0)) \quad \text{in } I^\pm.$$

Then we define the extension $\widetilde{\Phi}^{I,\eta}$ by truncating the two components on different scales:

$$\widetilde{\Phi}^{I,\eta} = \chi_I^\eta \Phi_0^I + \chi_I^{1/\eta} \Phi_1^I.$$

For the first component we argue as before. For the second, we begin with a fixed time L^2 bound:

$$\| P_0 \Phi_1^I (t) \|_{L_x^2} \lesssim |t \mp i_0| \cdot \| \partial_t P_0 \phi(\pm i_0) \|_{L_x^2} \quad \text{in } I^\pm, \tag{188}$$

which follows at once from integrating the quantity $\partial_t P_0 \Phi_1^I$ and energy estimates. This leads to:

$$\| P_0 \Box (\chi_I^{1/\eta} \Phi_1^I) \|_{L_t^1 (L_x^2)[I^\pm]} \lesssim \| \partial_t P_0 \phi(\pm i_0) \|_{L_x^2},$$

which helps us establish bounds of the type (179)–(180). On the other hand, the improved pointwise bound (185) follows simply by using Bernstein's inequality in (188) to give:

$$\| P_0 \Phi_1^I (t) \|_{L_x^\infty} \lesssim |t \mp i_0| \cdot \| \partial_t P_0 \phi(\pm i_0) \|_{L_x^2} \quad \text{in } I^\pm.$$

The proof of the proposition is thus concluded. $\quad \Box$

5.4. Strichartz and Wolff type bounds. In this section we prove the estimate (18) for the $\underline{S}$ component on line (148), as well as a key L^2 bilinear estimate for transverse waves which takes advantage of the small energy dispersion. The tools we use for these purposes are the $V^p_{\pm|D_x|}$ and $U^p_{\pm|D_x|}$ spaces associated to the two half-wave evolutions.

Precisely, $V^p_{\pm|D_x|}$ is the space of right continuous L^2_x valued functions with bounded p-variation along the half-wave flow:

$$\| u \|_{V^p_{\pm|D_x|}} = \| e^{\mp it|D_x|} u(t) \|_{V^p(L^2_x)},$$

or in expanded form:

$$\| u \|_{V^p_{\pm|D_x|}} := \| u \|^p_{L^\infty_t(L^2_x)} + \sup_{t_k \nearrow} \sum_{k \in \mathbb{Z}} \| u(t_{k+1}) - e^{\pm i(t_{k+1}-t_k)|D_x|} u(t_k) \|^p_{L^2_x},$$

where the supremum is taken over all increasing sequences t_k. We note that if $p < \infty$ then V^p functions can have at most countably many discontinuities as L^2_x valued functions.

On the other hand the slightly smaller space $U^p_{\pm|D_x|}$ is defined as the atomic space generated by a family $\mathcal{A}_p$ of atoms a which have the form:

$$a(t) = e^{\pm it|D_x|} \sum_k 1_{[t_k,t_{k+1})} u^{(k)},$$

where the sequence t_k is increasing and:

$$\sum_k \| u^{(k)} \|^p_{L^2_x} \leqslant 1.$$

Precisely, we have:

$$U^p_{\pm|D_x|} = \{ u = \sum c_k a_k; \ \sum_k |c_k| < \infty, a_k \in \mathcal{A}_p \}.$$

The above sum converges uniformly in L^2_x; it also converges in the stronger $V^p_{\pm|D_x|}$ topology. The $U^p_{\pm|D_x|}$ norm is defined by:

$$\| u \|_{U^p_{\pm|D_x|}} := \inf\{ \sum_k |c_k|; \ u = \sum_k c_k a_k, \ a_k \in \mathcal{A}_p \}.$$

These spaces are related as follows:

$$U^p_{\pm|D_x|} \subset V^p_{\pm|D_x|} \subset U^q_{\pm|D_x|}, \quad 1 \leqslant p < q \leqslant \infty. \tag{189}$$

The first inclusion is straightforward. The second is not, and plays a role similar to the Christ-Kiselev lemma. These spaces were first introduced in unpublished work of the second author, and have proved their usefulness as scale invariant substitutes of $X^{s,\frac{1}{2}}$ type spaces in several problems, see [2,5,11,12].

We use these spaces first in the context of the Strichartz estimates, which for frequency localized homogeneous half-waves can be expressed in the form:

$$\| e^{\pm it|D_x|} u_k \|_{L^q_t(L^r_x)} \lesssim 2^{-(\frac{1}{q}+\frac{2}{r}-1)k} \| u_k \|_{L^2_x}, \quad \frac{2}{q} + \frac{1}{r} \leqslant \frac{1}{2}.$$

Applying this bound for each segment in $U^q_{\pm|D_x|}$ atoms, one directly obtains embeddings of these spaces into Strichartz spaces:

Lemma 5.6. *The following estimates hold:*

$$\| \phi_k \|_{L_t^q(L_x^r)} \lesssim 2^{-(\frac{1}{q}+\frac{2}{r}-1)k} \| \phi_k \|_{U_{\pm|D_x|}^q}, \quad \frac{2}{q}+\frac{1}{r} \leqslant \frac{1}{2}. \tag{190}$$

The second place where these spaces come into play is in the context of bilinear $L_t^2(L_x^2)$ estimates for transversal waves. The classical estimate here (see for instance [10]) has the form:

$$\| e^{\pm it|D_x|}u_{k_1}^{(1)} \cdot e^{\pm it|D_x|}u_{k_2}^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{1}{2}\min\{k_1,k_2\}}\theta^{-\frac{1}{2}} \| u_{k_1}^{(1)} \|_{L_x^2} \| u_{k_2}^{(2)} \|_{L_x^2}, \tag{191}$$

provided the $u_{k_i}^{(i)}$ have angular separation in frequency, namely $|\theta_1 - \theta_2| > \theta$ in the $(++)$ or $(--)$ cases, and $|\pi + (\theta_1 - \theta_2)| > \theta$ in the $(+-)$ or $(-+)$ cases. In subsequent work, Wolff [34] was able to replace the $L_t^2(L_x^2)$ bound on the left with $L_t^p(L_x^p)$ for $p > \frac{5}{3}$. The endpoint $p = \frac{5}{3}$ was later obtained by Tao [27]. Our aim here is to first use the Wolff-Tao estimate to strengthen the classical $L_t^2(L_x^2)$ bound in a way which takes advantage of the small energy dispersion, and then phrase it in the set-up of the $V_{\pm|D_x|}^2$ spaces:

Lemma 5.7. *Let $\phi_{k_i}^{(i)} \in V_{\pm|D_x|}^2$ be two test functions which have angular separation in frequency, namely $|\theta_1 - \theta_2| > \theta$ in the $(++)$ or $(--)$ cases, and also $|\pi + (\theta_1 - \theta_2)| > \theta$ in the $(+-)$ or $(-+)$ cases. Then for $c < \frac{3}{29}$ we have:*

$$\| \phi_{k_1}^{(1)}\phi_{k_2}^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{1}{2}\max\{k_i\}}\theta^{-1} \| \phi_{k_1}^{(1)} \|_{V_{\pm|D_x|}^2} \| \phi_{k_2}^{(2)} \|_{V_{\pm|D_x|}^2}^{1-c} \left(2^{-k_2} \| \phi_{k_2}^{(2)} \|_{L_t^\infty(L_x^\infty)} \right)^c. \tag{192}$$

We remark that we did not make an effort to optimize c, the balance of the frequencies, or the power of θ, as these play no role in the present paper.

Proof. Without loss of generality, let us assume we are in the $(++)$ case. If both $\phi_{k_i}^{(i)}$ were free waves, then Wolff's estimate (with Tao's endpoint) would yield (see [27] Prop. 17.2):

$$\| \phi_{k_1}^{(1)}\phi_{k_2}^{(2)} \|_{L_t^{\frac{5}{3}}(L_x^{\frac{5}{3}})} \lesssim 2^{\frac{1}{5}\max\{k_i\}}\theta^{-1} \| \phi_{k_1}^{(1)}(0) \|_{L_x^2} \| \phi_{k_2}^{(2)}(0) \|_{L_x^2}.$$

Applying this for each intersection of two segments in a product of atoms, we obtain:

$$\| \phi_{k_1}^{(1)}\phi_{k_2}^{(2)} \|_{L_t^{\frac{5}{3}}(L_x^{\frac{5}{3}})} \lesssim 2^{\frac{1}{5}\max\{k_i\}}\theta^{-1} \| \phi_{k_1}^{(1)} \|_{U_{|D_x|}^{\frac{5}{3}}} \| \phi_{k_2}^{(2)} \|_{U_{|D_x|}^{\frac{5}{3}}}. \tag{193}$$

On the other hand by (190) with $(q, r) = (6, 6)$ we have:

$$\| \phi_{k_1}^{(1)}\phi_{k_2}^{(2)} \|_{L_t^3(L_x^3)} \lesssim 2^{\frac{k_1+k_2}{2}} \| \phi_{k_1}^{(1)} \|_{U_{|D_x|}^6} \| \phi_{k_2}^{(2)} \|_{U_{|D_x|}^6}. \tag{194}$$

Interpolating (193) with (194) (it is bilinear interpolation but it suffices to do it for atoms, so it only involves l^p and L^p spaces) we obtain:

$$\| \phi_{k_1}^{(1)}\phi_{k_2}^{(2)} \|_{L_t^p(L_x^p)} \lesssim 2^{(2-\frac{3}{p})\max\{k_i\}-\frac{3}{26}|k_1-k_2|}\theta^{-1} \| \phi_{k_1}^{(1)} \|_{U_{|D_x|}^2} \| \phi_{k_2}^{(2)} \|_{U_{|D_x|}^2}, \quad p = \frac{13}{7}.$$

We want $V_{|D_x|}^2$ instead, so we use the embedding (189) with $U_{|D_x|}^{2+}$ and $L_t^{\frac{13}{7}+}(L_x^{\frac{13}{7}+})$ in this last estimate, which gives the bound:

$$\| \phi_{k_1}^{(1)} \phi_{k_2}^{(2)} \|_{L_t^p(L_x^p)} \lesssim 2^{(2-\frac{3}{p})\max\{k_i\}-\frac{3}{26}|k_1-k_2|\theta-1} \| \phi_{k_1}^{(1)} \|_{V_{|D_x|}^2} \| \phi_{k_2}^{(2)} \|_{V_{|D_x|}^2}, \quad p > \frac{13}{7}.$$

On the other hand by using an $L_t^\infty(L_x^\infty)$ bound we get:

$$\| \phi_{k_1}^{(1)} \phi_{k_2}^{(2)} \|_{L_t^6(L_x^6)} \lesssim 2^{\frac{1}{2}k_1} \| \phi_{k_1}^{(1)} \|_{V_{|D_x|}^2} \| \phi_{k_2}^{(2)} \|_{L_t^\infty(L_x^\infty)}.$$

Then (192) is obtained interpolating the last two lines.   □

In this article we work with the S and N spaces. The next lemma relates them to the $V_{\pm|D_x|}^2$ spaces.

Lemma 5.8. *Let $\phi_k[0] \in \dot{H}^1 \times L^2$ and $F_k \in N$. Then the solution ϕ_k to $\Box\phi_k = F_k$ with initial data $\phi_k[0]$ satisfies:*

$$2^{\frac{1}{2}k}\| Q_{\geqslant k-10}\nabla_{t,x}\phi_k \|_{L_t^2(L_x^2)} + \sum_{\pm} \| Q_{<k-10}^{\pm}\nabla_{t,x}\phi_k \|_{V_{\pm|D_x|}^2} \lesssim \| \phi_k[0] \|_{\dot{H}^1\times L^2} + \| F_k \|_N.$$

$$(195)$$

Proof. By rescaling we may assume that $k = 0$, and we'll relabel ϕ_k and F_k as ϕ, F with the implicit understanding that they are both at unit frequency.

The estimate for $Q_{\geqslant -10}\nabla_{t,x}\phi$ is immediate from the structure of S and the estimate (18) (note that this was shown in [29] for all portions of the norm (147) except $\underline{S}$).

The linear wave evolution in the energy space $\dot{H}^1 \times L^2$ is given by the multiplier:

$$S(t) = \begin{pmatrix} \cos(t|D_x|) & |D_x|^{-1}\sin(t|D_x|) \\ -|D_x|\sin(t|D_x|) & \cos(t|D_x|) \end{pmatrix}.$$

For any increasing sequence t_j we can use the energy component of (18) (again established in [29]) and (159) to estimate:

$$\sum_j \| \phi[t_{j+1}] - S(t_{j+1} - t_j)\phi[t_j] \|_{\dot{H}^1\times L^2}^2 \lesssim \sum_j \| 1_{[t_j, t_{j+1}]}F \|_N^2 \lesssim \| F \|_N^2.$$

Diagonalizing, one may write the $L^2 \times L^2$ evolution as:

$$\begin{pmatrix} |D_x| & 0 \\ 0 & 1 \end{pmatrix} S(t)\phi[t_0] = \mathcal{U}^* \begin{pmatrix} e^{it|D_x|} & 0 \\ 0 & e^{-it|D_x|} \end{pmatrix} \mathcal{U} \begin{pmatrix} |D_x|\phi(t_0) \\ \partial_t\phi(t_0) \end{pmatrix},$$

where $\mathcal{U} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}$. Thus, the LHS of the previous difference formula may be rotated via $\mathcal{U}$ to yield:

$$\| \phi[t] - S(t-s)\phi[s] \|_{\dot{H}^1\times L^2}^2 = \frac{1}{2}\| (\partial_t + i|D_x|)\phi(t) - e^{i(t-s)|D_x|}(\partial_t + i|D_x|)\phi(s) \|_{L^2}^2$$

$$+ \frac{1}{2}\| (\partial_t - i|D_x|)\phi(t) - e^{-i(t-s)|D_x|}(\partial_t - i|D_x|)\phi(s) \|_{L^2}^2.$$

Hence taking the supremum over all increasing sequences t_k we obtain the pair of bounds:

$$\| (\partial_t + i|D_x|)\phi \|^2_{V^2_{|D_x|}} + \| (\partial_t - i|D_x|)\phi \|^2_{V^2_{-|D_x|}} \lesssim \| \phi[0] \|^2_{\dot{H}^1 \times L^2} + \| F \|^2_N.$$

We conclude (195) by noting that one has the following "elliptic" estimate:

$$\| \nabla_{t,x} Q^\pm_{<-10} P_0 \phi \|_Y \lesssim \| (\partial_t \pm i|D_x|)\phi \|_Y,$$

for any translation invariant space-time norm Y, which is valid because the convolution kernel of the frequency localized ratio $\nabla_{t,x}(\partial_t \pm i|D_x|)^{-1} Q^\pm_{<-10} P_0$ is in $L^1_t(L^1_x)$.   □

As a quick application of these ideas, notice that if (q, r) is any pair of indices in the range of (148), we must have $q \geqslant 4$. Hence from (190) and (195), and some Sobolev embeddings interpolated with the $L^\infty_t(L^2_x)$ estimate for $\nabla_{t,x}\phi$ to control the first member on the LHS of (195), we obtain:

Corollary 5.9. *Let $\phi[0] \in \dot{H}^1 \times L^2$ and $F \in N$. Then the solution ϕ to $\Box\phi = F$ with initial data $\phi[0]$ satisfies*

$$\| \phi \|_{\underline{S}} \lesssim \| \phi[0] \|_{\dot{H}^1 \times L^2} + \| F \|_N. \tag{196}$$

This proves (18), therefore completing the linear theory in the S and N spaces, as needed in view of our modification of Tao's [29] definition of the S space, namely by adding the $\underline{S}$ norm to it.

In a similar manner, we can combine the bounds (192) and (195) to obtain:

Lemma 5.10. *Let $\phi^{(i)}_{k_i}$ be two test functions normalized so that:*

$$\| \phi^{(i)}_{k_i} \|_S + \| \Box\phi^{(i)}_{k_i} \|_N \leqslant 1, \quad j = 1, 2, \quad \| \phi^{(2)}_{k_2} \|_{L^\infty_t(L^\infty_x)} \leqslant \eta.$$

Assume in addition that the localizations $Q^\pm_{<k_i-10}\phi^{(i)}_{k_i}$ have the angular separation $|\theta_1 - \theta_2| > \theta$ in the $(++)$ or $(--)$ cases, and $|\pi + (\theta_1 - \theta_2)| > \theta$ in the $(+-)$ or $(-+)$ cases. Then for $c < \frac{3}{29}$ one has:

$$\| \phi^{(1)}_{k_1} \phi^{(2)}_{k_2} \|_{L^2_t(L^2_x)} \lesssim 2^{-\frac{3}{2} \min\{k_i\}} \eta^c \theta^{-1}. \tag{197}$$

Proof. By an application of Lemma 5.7, we need only consider the case where one factor is at high modulation, i.e. a factor of $Q_{\geqslant k_i-10}\phi^{(i)}_{k_i}$. In this case, if the other factor has the improved $L^\infty_t(L^\infty_x)$ bound, estimate (197) is immediate on account of the $L^2_t(L^2_x)$ estimate on the LHS of (195). On the other hand, if the factor at high modulation is also the one with improved $L^\infty_t(L^\infty_x)$ control, then using $L^6_t(L^6_x)$ Strichartz for the first factor we have:

$$\| \phi^{(1)}_{k_1} Q_{\geqslant k_2-10}\phi^{(2)}_{k_2} \|_{L^{\frac{3}{2}}_t(L^{\frac{3}{2}}_x)} \lesssim 2^{-\frac{1}{2}k_1} 2^{-\frac{3}{2}k_2}.$$

Interpolating this with the pointwise bound $|\phi^{(1)}_{k_1} Q_{\geqslant k_2-10}\phi^{(2)}_{k_2}| \lesssim \eta$ we again have (197).   □

6. Bilinear Null Form Estimates

In this section we prove the estimates (23), (44), and (46). The first of these is essentially standard, being implicitly contained in the calculations of [29]. We provide the proof here for the sake of completeness:

Proof of estimate (23). We begin with the estimate:

$$\| P_k Q_{<j} F \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{j}{2}} \| F \|_N. \tag{198}$$

To see this, notice that if $\square_0^{-1}$ inverts the wave equation with zero Cauchy data, we immediately have from (18) the inequality:

$$\| P_k Q_{j'} \square_0^{-1} F \|_{X_\infty^{1,\frac{1}{2}}} \lesssim \| F \|_N,$$

which implies the fixed frequency estimate:

$$\| P_k Q_{j'} F \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{j'}{2}} \| F \|_N.$$

Summing this last line over all $j' < j$ (198) is achieved. We now split the proof of estimate (23) into two cases:

Step 1 (*Low × High interaction*). In this case we assume that $k_1 < k_2 - 10$. The case $k_2 < k_1 - 10$ can be handled via a similar argument. For relatively low modulations we have from estimates (24) and (198):

$$\| Q_{<k_1+10}(\partial^\alpha \phi_{k_1}^{(1)} \partial_\alpha \phi_{k_2}^{(2)}) \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{k_1}{2}} \| \phi_{k_1}^{(1)} \|_S \cdot \| \phi_{k_2}^{(2)} \|_S.$$

Therefore, it suffices to look at output modulations larger than k_1+10. In this case we split the modulations of the low frequency term according to $\phi_{k_1}^{(1)} = Q_{<k_1}\phi_{k_1}^{(1)} + Q_{\geqslant k_1}\phi_{k_1}^{(1)}$. For the first term we have that:

$$\begin{aligned}
\| Q_{\geqslant k_1+10}(Q_{<k_1}\partial^\alpha \phi_{k_1}^{(1)} \cdot \partial_\alpha \phi_{k_2}^{(2)}) \|_{L_t^2(L_x^2)} &\lesssim \| Q_{<k_1}\partial^\alpha \phi_{k_1}^{(1)} \cdot Q_{>k_1}\partial_\alpha \phi_{k_2}^{(2)} \|_{L_t^2(L_x^2)} \\
&\lesssim \| \nabla_{t,x}\phi_{k_1}^{(1)} \|_{L_t^\infty(L_x^\infty)} \| Q_{>k_1}\nabla_{t,x}\phi_{k_2}^{(2)} \|_{L_t^2(L_x^2)} \\
&\lesssim 2^{\frac{k_1}{2}} \| \nabla_{t,x}\phi_{k_1}^{(1)} \|_{L_t^\infty(L_x^2)} \| \nabla_{t,x}\phi_{k_2}^{(2)} \|_{X_\infty^{0,\frac{1}{2}}},
\end{aligned}$$

which suffices. For the high modulations of the first factor in the previous decomposition, we estimate:

$$\| Q_{\geqslant k_1+10}(Q_{>k_1}\partial^\alpha \phi_{k_1}^{(1)} \cdot \partial_\alpha \phi_{k_2}^{(2)}) \|_{L_t^2(L_x^2)} \lesssim \| Q_{>k_1}\nabla_{t,x}\phi_{k_1}^{(1)} \|_{L_t^2(L_x^\infty)} \| \nabla_{t,x}\phi_{k_2}^{(2)} \|_{L_t^\infty(L_x^2)}.$$

We then conclude using (164) for the first factor.

Step 2 (*High × High interaction*). In this case we consider the frequency interaction $|k_1 - k_2| < 5$, and without loss of generality we may also assume that $k_1 \geqslant k_2$. By using estimates (198) and (24), we may reduce to considering the case of output modulation larger than $k + \delta(k_1 - k) + 10$, where δ is from the RHS of (24) (this ultimately forces a harmless redefinition of δ to suit line (23)). For this remaining piece, we will show that:

$$\| P_k Q_{\geqslant k+\delta(k_1-k)+10}(\partial^\alpha \phi_{k_1}^{(1)} \cdot \partial_\alpha \phi_{k_2}^{(2)}) \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{k}{2}} 2^{-\frac{1}{2}\delta(k_1-k)} \| \phi_{k_1}^{(1)} \|_S \cdot \| \phi_{k_2}^{(2)} \|_S.$$

The key observation here is that the output modulation combined with the output spatial frequency localization guarantees that at least one term in the product is at modulation greater than $k + \delta(k_1 - k) - 20$. Without loss of generality we may assume this is the first term in the product, and we estimate via Bernstein:

$$\| P_k Q_{\geqslant k+\delta(k_1-k)+10}(Q_{\geqslant k+\delta(k_1-k)-20}\partial^\alpha \phi^{(1)}_{k_1} \cdot \partial_\alpha \phi^{(2)}_{k_2}) \|_{L^2_t(L^2_x)}$$

$$\lesssim 2^k \sum_{j > k+\delta(k_1-k)-20} \| Q_j \nabla_{t,x} \phi^{(1)}_{k_1} \|_{L^2_t(L^2_x)} \cdot \| \nabla_{t,x}\phi^{(2)}_{k_2} \|_{L^\infty_t(L^2_x)}$$

$$\lesssim 2^{\frac{k}{2}} 2^{-\frac{1}{2}\delta(k_1-k)} \| \nabla_{t,x}\phi^{(1)}_{k_1} \|_{X^{0,\frac{1}{2}}_\infty} \cdot \| \nabla_{t,x}\phi^{(2)}_{k_2} \|_{L^\infty_t(L^2_x)}.$$

This concludes our demonstration of (23). $\square$

Our next step is to prepare for the proof of Proposition 3.4. It will first be useful to have a version of these estimates under simpler assumptions:

Lemma 6.1. *a) Let $\phi^{(i)}_{k_i}$ be functions localized at frequency k_i. Assume that these functions are normalized as follows:*

$$\| \phi^{(i)}_{k_i} \|_{S[I]} + \| \Box\phi^{(i)}_{k_i} \|_{N[I]} \leqslant 1, \quad \| \phi^{(1)}_{k_1} \|_{L^\infty_t(L^\infty_x)[I]} \leqslant \eta. \tag{199}$$

Then the following bilinear L^2 estimate holds:

$$\| \partial^\alpha \phi^{(1)}_{k_1} \partial_\alpha \phi^{(2)}_{k_2} \|_{L^2_t(L^2_x)[I]} \lesssim 2^{\frac{1}{2}\max\{k_1,k_2\}}\eta^\delta. \tag{200}$$

b) Assume that in addition to (199) we also have the high modulation bounds:

$$\| \Box\phi^{(1)}_{k_1} \|_{L^2_t(L^2_x)[I]} \leqslant 2^{\frac{k_1}{2}}\eta, \quad \| \Box\phi^{(2)}_{k_2} \|_{L^2_t(L^2_x)[I]} \leqslant 2^{\frac{k_2}{2}}\eta. \tag{201}$$

Then the following estimate holds:

$$\| \partial^\alpha \phi^{(1)}_{k_1} \partial_\alpha \phi^{(2)}_{k_2} \|_{N[I]} \lesssim 2^{C|k_1-k_2|}\eta^\delta. \tag{202}$$

Proof. We may assume that the interval length is such that $|I| \geqslant 2^{-\min\{k_i\}}\eta^{2\delta}$, as otherwise the desired bounds follow from integrating energy estimates.

We begin by taking extensions of $\phi^{(1)}_{k_1}$ and $\phi^{(2)}_{k_2}$ according to Proposition 5.5 in such a way that the $L^\infty_t(L^\infty_x)$ bound in (199) is preserved; in the case of part (b), we also insure that (201) is preserved. This is achieved using the second extension in Proposition 5.5 in case (a), respectively the canonical extension in Proposition 5.5 in case (b). Doing this requires balancing the parameter η in Proposition 5.5, and has the effect of replacing the η in both (199) and (201) with a small power of η ($\eta^{\frac{1}{8}}$ should suffice). This is harmless given the small constant δ which we seek to obtain in both (200) and (202).

We fix m to be a large spatial frequency separation parameter. In the course of proving (200) and (202), we will decompose into several frequency ranges. In all cases we will show a bound of the form:

$$\text{L.H.S.} \lesssim 2^{Cm}\eta^c + 2^{-\delta m},$$

where c, δ are suitably small constants depending only on the estimates in Propositions 2.3, 5.4, and 5.10 above, and where C is a suitable large constant. In what follows we call any bound of this type a "suitable bound". By choosing m appropriately, and by (globally) redefining the small parameter δ one may produce the RHS of estimates (200) and (202) from such bounds.

Step 1 (*The unbalanced case* $|k_1 - k_2| \geqslant m$). Here we neglect the pointwise bound in (199) as well as the high modulation bound in (201). From estimate (23) we immediately have that:

$$\| \partial^\alpha \phi_{k_1}^{(1)} \partial_\alpha \phi_{k_2}^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{\frac{1}{2} \max\{k_1, k_2\}} 2^{-\frac{1}{2}m}, \quad |k_1 - k_2| \geqslant m,$$

which is a suitable L^2 bound. Similarly, from (24) we obtain a suitable N bound.

Hence, in what follows it suffices to consider the range $|k_1 - k_2| < m$. For the remainder of the proof we let $k_2 = 0$. We split into cases depending on the modulations of the factors and the output.

Step 2 (*The factor* $\phi_{k_2}^{(2)}$ *at high modulation*). Here we first prove a suitable L^2 bound:

$$\| \partial^\alpha \phi_{k_1}^{(1)} Q_{>-10m} \partial_\alpha \phi_0^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{Cm} \eta + 2^{-m}. \tag{203}$$

For moderate modulations of the first factor, i.e. for $Q_{<10m} \phi_{k_1}^{(1)}$, we use (199) to place it in L^∞:

$$\| \partial^\alpha Q_{<10m} \phi_{k_1}^{(1)} \|_{L_t^\infty(L_x^\infty)} \lesssim 2^{Cm} \eta,$$

while the second factor is placed in L^2 via the general embedding:

$$\| Q_{>j} \phi_{k_i}^{(i)} \|_{L_t^2(L_x^2)} \lesssim 2^{-\frac{1}{2}j} \| \phi_{k_i}^{(i)} \|_{X_\infty^{0, \frac{1}{2}}}. \tag{204}$$

For high modulations of the first factor, i.e. for $Q_{>10m} \phi_{k_1}^{(1)}$, we reverse the roles and bound the first factor in L^2:

$$\| \partial^\alpha Q_{>10m} \phi_{k_1}^{(1)} \|_{L_t^2(L_x^2)} \lesssim 2^{-5m} \| \phi_{k_1}^{(1)} \|_S, \tag{205}$$

while the second factor has a $\lesssim 1$ bound in L^∞ thanks to (166), which leads again to a suitable bound.

In this case it is even easier to obtain the suitable N bound because we have access to the high modulation assumption (201). We prove:

$$\| \partial^\alpha \phi_{k_1}^{(1)} Q_{>-10m} \partial_\alpha \phi_0^{(2)} \|_N \lesssim 2^{Cm} \eta. \tag{206}$$

This follows from (24) combined with:

$$\| Q_{>-10m} \phi_0^{(2)} \|_S \lesssim 2^{5m} \| \Box \phi_0^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{5m} \eta,$$

where the first inequality follows from (162).

Step 3 (*The factor* $\phi_{k_1}^{(1)}$ *at high modulation*). Here we can also prove a suitable L^2 bound, namely:

$$\| \partial^\alpha Q_{>-10m} \phi_{k_1}^{(1)} Q_{\leqslant -10m} \partial_\alpha \phi_0^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{Cm} \eta^{\frac{1}{4}} + 2^{-m}. \tag{207}$$

Reusing (205) we can dispense with the very high modulations in $\phi_{k_1}^{(1)}$ and replace the first factor with $\partial^\alpha Q_{[-10m,10m]}\phi_{k_1}^{(1)}$. This time we cannot directly use the $L_t^\infty(L_x^\infty)$ estimate for $\phi_{k_1}^{(1)}$. However, by applying (204) and using the $L_t^6(L_x^6)$ Strichartz estimate contained in (196) we have that:

$$\| \partial^\alpha Q_{[-10m,10m]}\phi_{k_1}^{(1)} Q_{\leqslant -10m}\partial_\alpha\phi_0^{(2)} \|_{L_t^{\frac{3}{2}}(L_x^{\frac{3}{2}})} \lesssim 2^{Cm}.$$

Next, using (199) and (166) we directly have:

$$\| \partial^\alpha Q_{[-10m,10m]}\phi_{k_1}^{(1)} Q_{\leqslant -10m}\partial_\alpha\phi_0^{(2)} \|_{L_t^\infty(L_x^\infty)} \lesssim 2^{Cm}\eta.$$

Interpolating these last two estimates yields (207). It is important to notice that in the above estimates one loses a polynomial in 2^m because the multipliers $P_0 Q_{\leqslant -10m}$ and $P_{k_1}Q_{[-10m,10m]}$ are not uniformly disposable on L^p. However, a short calculation shows that the resulting convolution kernels have $L_t^1(L_x^1)$ bounds on the order of 2^{Cm} which is acceptable.

As in the previous step we also have a suitable N bound:

$$\| \partial^\alpha Q_{>-10m}\phi_{k_1}^{(1)}\partial_\alpha Q_{\leqslant -10m}\phi_0^{(2)} \|_N \lesssim 2^{Cm}\eta. \tag{208}$$

Step 4 (*Low frequency output*). This is the case when $k_1 = k_2 + O(1)$, and we seek to estimate $P_k(\partial^\alpha\phi_{k_1}^{(1)}\partial_\alpha\phi_0^{(2)})$ for $k < -m$. Then we can use (23), respectively (24) to obtain a $\lesssim 2^{-\delta m}$ suitable bound in L^2, respectively N. Here δ is the previously defined constant from Proposition 2.3.

Step 5A (*Both $\phi_{k_i}^{(i)}$ at low modulation, output at low modulation $< -2m$ and high frequency $k > -m$*). In this case, to show (200) we prove the bound:

$$\| P_k Q_{<-2m}\left(\partial^\alpha Q_{\leqslant -10m}\phi_{k_1}^{(1)}\partial_\alpha Q_{\leqslant -10m}\phi_0^{(2)}\right) \|_{L_t^2(L_x^2)} \lesssim 2^{-\delta m}, \tag{209}$$

where the δ is the same as in Propositions 2.3 and 5.4. This estimate again uses only the S bounds on $\phi_{k_1}^{(1)}$ and $\phi_{k_2}^{(2)}$ and the localization conditions $|k_1| \leqslant m$ and $k_2 = 0$. To show (209), by (198) it suffices to prove the following set of bounds which together also imply (202) in the present case:

$$\| P_k Q_{<-2m}\Box\left(Q_{\leqslant -10m}\phi_{k_1}^{(1)} \cdot Q_{\leqslant -10m}\phi_0^{(2)}\right) \|_{X_1^{0,-\frac{1}{2}}} \lesssim m2^{-\delta m}, \tag{210}$$

$$\| P_k Q_{<-2m}\left(\Box Q_{\leqslant -10m}\phi_{k_1}^{(1)} \cdot Q_{\leqslant -10m}\phi_0^{(2)}\right) \|_N \lesssim 2^{-\delta m}, \tag{211}$$

$$\| P_k Q_{<-2m}\left(Q_{\leqslant -10m}\phi_{k_1}^{(1)} \cdot \Box Q_{\leqslant -10m}\phi_0^{(2)}\right) \|_N \lesssim 2^{-\delta m}. \tag{212}$$

The first estimate above follows from (169), while the second and third both follow from (170). Note that while the multiplier $Q_{\leqslant -2m}$ is not disposable on N (e.g. on the *NFA* atoms), one may first replace it by $Q_{<0}$, and separately estimate the contribution of $Q_{[-2m,0]}$ as an $X_1^{0,-\frac{1}{2}}$ atom via (163) at an $O(m)$ loss. A similar method using (162) allows one to handle the interior $Q_{\leqslant -10m}$ multipliers, which are not disposable on S, with another $O(m)$ loss.

Step 5B (*Both* $\phi_{k_i}^{(i)}$ *at low modulation, output at high frequency and high modulation*). In this step, which is the heart of the matter, we establish the single bound:

$$\| P_{\geqslant -m} Q_{\geqslant -2m} \left(\partial^\alpha Q_{\leqslant -10m} \phi_{k_1}^{(1)} \partial_\alpha Q_{\leqslant -10m} \phi_0^{(2)} \right) \|_{L_t^2(L_x^2)} \lesssim 2^{Cm} \eta^c. \qquad (213)$$

Here c is the same small constant from the RHS of line (197). To use that estimate, we only need to establish angular separation of the two factors. This is a standard "geometry of the cone" calculation, and one finds that the angle between the two factors must satisfy $|\theta| \gtrsim 2^{-m}$ in the (++) or (−−) cases, and $|\theta - \pi| \gtrsim 2^{-m}$ in the (+−) or (−+) cases (see for example Lemma 11 in Sect. 13 of [29]). By decomposing the product on the LHS of (213) into $O(2^{Cm})$ angular sectors such that each product has these separation properties, and by repeatedly applying estimate (197) on each interaction we have (213). The proof of the lemma is concluded. $\quad\square$

Proof of Proposition 3.4. For this we use Lemma 6.1. We begin using the extensions (this will be modified somewhat shortly) and the same parameter m as in the proof of Lemma 6.1. We start with several simplifications. The key observation is that in the proof of Lemma 6.1 we have used the bound on $\| \Box \phi_{k_i}^{(i)} \|_N$ just once, namely in Step 5B. All other cases carry over to the proof of Proposition 3.4. Consequently, it suffices to estimate the expression:

$$P_k Q_{>-2m} R = P_k Q_{>-2m} (\partial^\alpha \phi_{k_1}^{(1)} \partial_\alpha \phi_{k_2}^{(2)}),$$

in both L^2 and N under the assumptions $k_2 = 0$, $|k_1| \leqslant m$, and $|k| \leqslant m + 2$.

Furthermore, the contribution $P_k Q_{>-2m} ((1 - \chi_I) R)$ of R in the exterior of I is estimated directly by Lemma 6.1 because the extensions provided by Proposition 5.5 enjoy estimate (186) in the exterior of I. Hence, we only need consider the expression $P_k Q_{>-2m} (\chi_I R)$. For this we will establish the pair of suitable bounds:

$$\| P_k Q_{>-2m} (\chi_I R) \|_{L_t^2(L_x^2)} \lesssim 2^{Cm} \eta^\delta + 2^{-5m}, \quad \| P_k Q_{>-2m} (\chi_I R) \|_N \lesssim 2^{Cm} \eta^\delta + 2^{-4m},$$

with δ as in Lemma 6.1. We remark that due to the frequency and modulation localization of R, the second N bound follows from the first L^2 bound albeit with a readjusted C. Therefore, we drop the modulation and spatial frequency localization and simply prove that:

$$\| P_k R \|_{L_t^2(L_x^2)[I]} \lesssim 2^{Cm} \eta^\delta + 2^{-5m}. \qquad (214)$$

For this we use the renormalization. On the interval I, we may decompose $\phi_{k_i}^{(i)}$ as follows:

$$\phi_{k_i}^{(i)} = (U_{,<k_i}^{(i)})^\dagger w_{,k_i}^{(i)},$$

where by using the definition (30) we may assume that the component pieces separately obey the estimates:

$$\| P_k w_{k_i}^{(i)} \|_{S[I]} + \| P_k \Box w_{k_i}^{(i)} \|_{N[I]} \leqslant 2^{-|k-k_i|} A^{-1},$$
$$\| U_{,<k_i}^{(i)} \|_S + \sup_{k>k_i} 2^{C(k_i-k)} \| P_k U_{,<k_i}^{(i)} \|_S \leqslant A,$$

for a possibly large constant A. By normalization, we may without loss of generality assume that $A = 1$, as any bounds for these two quantities will always appear as a product. Since $\| \phi_{k_1}^{(1)} \|_{L_t^\infty(L_x^\infty)[I]} \leqslant \eta$, we obtain a similar relation for $w_{,k_1}^{(1)}$, namely:

$$\| P_k w_{,k_1}^{(1)} \|_{L_t^\infty(L_x^\infty)[I]} \leqslant \eta.$$

Furthermore, by using Proposition 5.5, we may extend the $w_{k_i}^{(i)}$ so that all of the above listed bounds are global, albeit with a fractional modification of η. Thus, we may drop the interval I, and again work globally.

We decompose the null-form R (first on I, then by extension) into $R = R_1 + R_2 + R_3 + R_4$ where:

$$R_1 = -\partial^\alpha (U_{,<k_1}^{(1)})^\dagger \cdot w_{,k_1}^{(1)} \partial_\alpha (U_{,<0}^{(2)})^\dagger \cdot w_{,0}^{(2)},$$
$$R_2 = \partial^\alpha (U_{,<k_1}^{(1)})^\dagger \cdot w_{,k_1}^{(1)} \cdot \partial_\alpha \phi_0^{(2)},$$
$$R_3 = \partial^\alpha \phi_{k_1}^{(1)} \cdot \partial_\alpha (U_{,<0}^{(2)})^\dagger \cdot w_{,0}^{(2)},$$
$$R_4 = (U_{,<k_1}^{(1)})^\dagger \cdot \partial^\alpha w_{,k_1}^{(1)} \cdot (U_{,<0}^{(2)})^\dagger \cdot \partial_\alpha w_{,0}^{(2)}.$$

We successively consider each of these terms.

Step 1 (*Estimating the term R_1*). Using the S bounds for $U_{,<k_1}^{(1)}$ and $U_{,<0}^{(2)}$ in the bilinear L^2 null form estimate (23), after dyadic summation we obtain:

$$\| \partial^\alpha (U_{,<k_1}^{(1)})^\dagger \partial_\alpha (U_{,<0}^{(2)})^\dagger \|_{L_t^2(L_x^2)} \lesssim m.$$

Note that the RHS loss is the effect of summing over frequencies $k' \leqslant k_1 \leqslant m$ on the first factor. We combine this with the pointwise bound on $w_{,k_1}^{(1)}$ to achieve:

$$\| R_1 \|_{L_t^2(L_x^2)} \lesssim \| w_{,k_1}^{(1)} \|_{L_t^\infty(L_x^\infty)} \lesssim m\eta.$$

Step 2 (*Estimating the term R_2*). This is essentially the same as in the previous step. Here we use the S bounds for $U_{,<k_1}^{(1)}$ and for $\phi_0^{(2)}$ in conjunction with (23), and we again use the pointwise bound for $w_{,k_1}^{(1)}$.

Step 3 (*Estimating the term R_3*). We begin by splitting $\phi_{k_1}^{(1)}$ into a low and a high modulation part. For the high modulation part we have from (204) the L^2 bound:

$$\| Q_{>10m} \partial^\alpha \phi_{k_1}^{(1)} \|_{L_t^2(L_x^2)} \lesssim 2^{-5m}.$$

Furthermore, by summing over the energy estimate for $U_{,<0}^{(2)}$ and using the decay of high frequencies we have the pointwise bound:

$$\| \nabla_{t,x} U_{,<0}^{(2)} \|_{L_t^\infty(L_x^\infty)} \lesssim \sum_k 2^k \| P_k \nabla_{t,x} U_{,<0}^{(2)} \|_{L_t^\infty(L_x^2)} \lesssim 1. \tag{215}$$

Combining these two estimates with the pointwise bound for $w_{,0}^{(2)}$ we can estimate the corresponding part of R_3, call it R_{31}, by:

$$\| R_{31} \|_{L_t^2(L_x^2)} \lesssim 2^{-5m}.$$

It remains to consider the contribution of the low modulation part $Q_{<10m}\phi_{k_1}^{(1)}$ in R_3, which we will label by R_{32}. Using the S bounds for $\phi_{k_1}^{(1)}$ and $U_{,<0}^{(2)}$ along with (23), after dyadic summation we obtain the usual L^2 estimate:

$$\| \partial^\alpha Q_{<10m}\phi_{k_1}^{(1)} \cdot \partial_\alpha (U_{,<0}^{(2)})^\dagger \|_{L_t^2(L_x^2)} \lesssim 1.$$

On the other hand, from the Strichartz control (148) and the boundedness of the gauge we have:

$$\| w_{,0}^{(2)} \|_{L_t^6(L_x^6)} \lesssim \| \phi_0^{(2)} \|_{L_t^6(L_x^6)} \lesssim 1,$$

therefore we obtain a low index space-time L^p bound for R_{32}, namely:

$$\| R_{32} \|_{L_t^{\frac{3}{2}}(L_x^{\frac{3}{2}})} \lesssim 1.$$

On the other hand, from the pointwise bound (43) for $\phi_{k_1}^{(1)}$ we obtain:

$$\| \partial^\alpha Q_{<10m}\phi_{k_1}^{(1)} \|_{L_t^\infty(L_x^\infty)} \lesssim 2^{Cm}\eta.$$

Combining this with (215) and the pointwise bound for $w_{,0}^{(2)}$ we have:

$$\| R_{32} \|_{L_t^\infty(L_x^\infty)} \lesssim 2^{Cm}\eta.$$

Interpolating the last two lines we obtain:

$$\| R_{32} \|_{L_t^2(L_x^2)} \lesssim 2^{Cm}\eta^{\frac{1}{4}}.$$

Step 4 (*Estimating the term R_4*). We start by dividing the main part of the product into all spatial frequencies:

$$\partial^\alpha w_{,k_1}^{(1)} \cdot \partial_\alpha w_{,0}^{(2)} = \sum_{j_i} \partial^\alpha P_{j_1} w_{,k_1}^{(1)} \cdot \partial_\alpha P_{j_2} w_{,0}^{(2)}.$$

Using the bound (200) if $j_1, j_2 < 10m$, and (23) otherwise in conjunction with the $2^{-|j_i-k_i|}$ frequency separation gains for $P_{j_i} w_{,k_i}^{(i)}$ we have:

$$\| \partial^\alpha w_{,k_1}^{(1)} \cdot \partial_\alpha w_{,0}^{(2)} \|_{L_t^2(L_x^2)} \lesssim 2^{Cm}\eta^\delta + 2^{-5m}.$$

This estimate is directly transferred to R_4 due to the pointwise bounds on the gauge factors. $\square$

7. Proof of the Trilinear Estimates

In this section we will prove estimates (51)–(57). In all cases the desired bounds follow easily from a combination of the standard estimates (23)–(25) for widely separated frequencies, and the improved matched frequency estimates (44) and (46).

Proof of estimate (51). The proof will be accomplished in a series of steps whose goal is to reduce things to the matched frequency bilinear estimate (44).
Step 1 (*Disposal of the $\phi^{(1)}$ Factor*). As a first step we will show the general estimate:

$$\| \phi \cdot F \|_{L_t^2(\dot{H}^{-\frac{1}{2}})_c[I]} \lesssim \| \phi \|_{S[I]} \cdot \| F \|_{L_t^2(\dot{H}^{-\frac{1}{2}})_c[I]}, \tag{216}$$

where $\{c_k\}$ is any (δ_0, δ_0)-admissible frequency envelope. To prove this, we split it into the three main frequency interactions.

In the $Low \times High$ case we immediately have:

$$\| P_k(\phi_{<k-10} \cdot F) \|_{L_t^2(\dot{H}_x^{-\frac{1}{2}})[I]} \lesssim \| \phi \|_{S[I]} \cdot 2^{-\frac{k}{2}} \| P_{[k-5,k+5]}F \|_{L_t^2(L_x^2)[I]},$$

which is sufficient.

In the $High \times Low$ case, we freeze the dyadic frequency of F and we have a similar estimate:

$$\| P_k(\phi \cdot F_{k'}) \|_{L_t^2(\dot{H}_x^{-\frac{1}{2}})[I]} \lesssim 2^{\frac{k'-k}{2}} c_{k'} \| \phi \|_{S[I]} \cdot \| F \|_{L_t^2(\dot{H}_x^{-\frac{1}{2}})_c[I]},$$

for any $k' \leqslant k - 10$. Summing this over all such $k' \leqslant k - 10$ and using (13) we have (216) in this case.

In the $High \times High$ case we freeze the frequency of the inputs and output to estimate:

$$\| P_k(\phi_{k_1} \cdot F_{k_2}) \|_{L_t^2(L_x^2)} \lesssim 2^{k-k_1} \| \phi_{k_1} \|_{S[I]} \cdot 2^{\frac{k_2}{2}} c_{k_2} \| F \|_{L_t^2(\dot{H}^{-\frac{1}{2}})_c[I]}, \tag{217}$$

which follows easily from Bernstein's inequality (9) and the bound (165). Multiplying this last line by $2^{-\frac{1}{2}k}$, and then summing over all k_1 and k_2 such that $|k_1 - k_2| \leqslant 20$ and $k_1 \geqslant k - 10$, and then using (14), we arrive at the estimate (216) for this case.
Step 2 (*The Bilinear Estimate*). In light of estimate (216) above, it suffices to show that:

$$\| \partial^\alpha \phi^{(2)} \partial_\alpha \phi^{(3)} \|_{L_t^2(\dot{H}^{-\frac{1}{2}})_c[I]} \lesssim \eta^{\delta_1}, \tag{218}$$

assuming the conditions of estimate (51). This will be done in two steps.
Step 2A (*Reduction to Matched Frequencies*). Our first step is to peel off all frequency interactions that cannot be treated by estimate (44). In all of these interactions, we will exploit the fact that there is a wide separation in the frequency. This is measured by choosing a large integer $m_0 = m_0(\eta)$ such that:

$$2^{-\frac{1}{2}\delta m_0} = \eta^{\delta_1}, \tag{219}$$

where we remind the reader that δ is the small dyadic savings from the standard L^2 bilinear estimate on line (23), and because of the definition of δ_1 we have:

$$m_0 \lesssim \sqrt{\delta_1} |\ln(\eta)|. \tag{220}$$

Our goal in this step is to show the following fixed frequency estimate:

$$\sum_{\substack{k_i \\ \max\{|k_i-k|\}\geqslant m_0}} 2^{-\frac{1}{2}k} \| P_k \left(\partial^\alpha \phi_{k_2}^{(2)} \partial_\alpha \phi_{k_3}^{(3)} \right) \|_{L_t^2(L_x^2)[I]} \lesssim 2^{-\frac{1}{2}\delta m_0} c_k \| \phi^{(2)} \|_{S[I]} \| \phi^{(3)} \|_{S_c[I]},$$

(221)

which in light of (219) suffices to establish (218) for all frequency interactions except for the case $k = k_1 + O(m_0) = k_2 + O(m_0)$. By an application of estimate (23), the two sum rules (13)–(14), and the definition (11) we immediately have:

$$(\text{L.H.S.})(221) \lesssim \sum_{\substack{k_i \\ \max\{|k_i-k|\}\geqslant m_0}} 2^{-\frac{1}{2}k} 2^{\frac{1}{2}\min\{k_i\}} 2^{-(\frac{1}{2}+\delta)(\max\{k_i\}-k)} c_{k_3} \lesssim 2^{-(\delta-\delta_0)m_0} c_k,$$

which by using (219) and the definition of the δ_i suffices to establish (221).

Step 2B (*The Matched Frequency Case*). We have now reduced estimate (218) to showing the matched frequency bound:

$$\sum_{\substack{k_i \\ \max\{|k_i-k|\}<m_0}} 2^{-\frac{1}{2}k} \| P_k \left(\partial^\alpha \phi_{k_2}^{(2)} \partial_\alpha \phi_{k_3}^{(3)} \right) \|_{L_t^2(L_x^2)[I]} \lesssim \eta^{\delta_1} c_k.$$

Due to the fact that there are only $O(m_0) \leqslant |\ln(\eta)|$ terms in this sum, it suffices to show:

$$\| P_k \left(\partial^\alpha \phi_{k_2}^{(2)} \partial_\alpha \phi_{k_3}^{(3)} \right) \|_{L_t^2(L_x^2)[I]} \lesssim \eta^{2\delta_1} 2^{\frac{1}{2}k} c_k, \quad \max\{|k_i - k|\} \lesssim \sqrt{\delta_1} |\ln(\eta)|.$$

But this last estimate follows immediately from (44) and the definition of the δ_i. □

Proof of estimate (52). This estimate was essentially established in the previous proof. We split the estimate into a sum of two pieces:

$$(\text{L.H.S.})(52) \leqslant \| P_k \left[P_{<k+10}\phi^{(1)} \partial^\alpha \phi^{(2)} \partial_\alpha \phi^{(3)} \right] \|_{L_t^2(L_x^2)[I]}$$

$$+ \| P_k \left[P_{\geqslant k+10}\phi^{(1)} \partial^\alpha \phi^{(2)} \partial_\alpha \phi^{(3)} \right] \|_{L_t^2(L_x^2)[I]}.$$

For the first term we simply use (51). For the second term, we use the following version of (217) above:

$$\| P_k(\phi_{k_1} \cdot F_{k_2}) \|_{L_t^2(\dot{H}^{-\frac{1}{2}})} \lesssim 2^{\frac{k-k_1}{2}} c_{k_1} \| \phi \|_{S_c[I]} \cdot \| F \|_{L_t^2(\dot{H}^{-\frac{1}{2}})[I]},$$

for $|k_1 - k_2| \leqslant 5$, along with (218). This suffices via the sum rule (14). □

Remark 7.1. It is possible to prove the frequency envelope estimate (55) with $\eta = 1$ in the case where there is no energy dispersion. As the previous step shows, one may first reduce to a bilinear estimate. Then the desired bound follows from summation over (23) using the sum rules (13)–(14). The details are standard and left to the reader.

Proof of estimate (55). The proof will be accomplished in a series of steps whose goal is to reduce things to the bilinear estimate (46).

Step 0 (*A Preliminary Reduction*). The first order of business is to reduce estimate (55) to the case where we replace the condition on line (50) with a maximal case:

$$m = \max\{\sqrt{\delta_1}|\ln(\eta)|, 10\}. \tag{222}$$

We claim that a proof of (55) with this choice of m implies (55) for any other choice of m where we turn (222) into an $\leqslant$ inequality. The only caveat is that we must replace the multiplier P_k in the definition of (53) by a version $\widetilde{P}_k$ with a slightly fattened support, so that one obtains the quasi-idempotence identity $\widetilde{P}_k P_{[k-5,k+5]} = P_{[k-5,k+5]}$. To see this, simply notice that one has the reshuffling identity:

$$T_k^{m_0}(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) = T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) - \widetilde{T}_{1;k}^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) - \widetilde{T}_{2;k}^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}),$$

for any $10 \leqslant m_0 \leqslant m$, where the $\widetilde{T}_{i;k}^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)})$ are the trilinear forms obtained from applying the definition of T_k^m, with $\widetilde{P}_k$ instead of P_k, to the second and third terms (resp.) on the RHS of (53) in the definition of $T_k^{m_0}$.

Step 1 (*Removal of the Commutator*). We are now trying to prove (55) under the condition (222). Our next step is to use (10) to write (53) in the form:

$$\begin{aligned} T_k^m(\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) &= P_k\left(\phi^{(1)}\partial^\alpha\phi^{(2)}\partial_\alpha\phi^{(3)}\right) - P_k\left(\phi_{<k-m}^{(1)}\partial^\alpha\phi_{<k-m}^{(2)}\partial_\alpha\phi^{(3)}\right) \\ &\quad - P_k\left(\phi_{<k-m}^{(1)}\partial^\alpha\phi^{(2)}\partial_\alpha\phi_{<k-m}^{(3)}\right) - (C_1 + C_2 + C_3 + C_4), \end{aligned} \tag{223}$$

where the commutator terms C_1, C_2, C_3 and C_4 have the form (here the L_i refers to the effect of the commutator, and not the L in the definition of (53) which has been dropped):

$$\begin{aligned} C_1 &= 2^{-k} L_1(\nabla_x\phi_{<k-m}^{(1)}, \partial^\alpha\phi_{<k-m}^{(2)}, \partial_\alpha\widetilde{P}_k\phi^{(3)}), \\ C_2 &= 2^{-k} L_2(\phi_{<k-m}^{(1)}, \nabla_x\partial^\alpha\phi_{<k-m}^{(2)}, \partial_\alpha\widetilde{P}_k\phi^{(3)}), \\ C_3 &= 2^{-k} L_3(\nabla_x\phi_{<k-m}^{(1)}, \partial^\alpha\widetilde{P}_k\phi^{(2)}, \partial_\alpha\phi_{<k-m}^{(3)}), \\ C_4 &= 2^{-k} L_4(\phi_{<k-m}^{(1)}, \partial^\alpha\widetilde{P}_k\phi^{(2)}, \nabla_x\partial_\alpha\phi_{<k-m}^{(3)}). \end{aligned}$$

Here $\widetilde{P}_k$ is the same as in the previous step, and we remind the reader that the L_i are disposable. The goal of this step is to prove the estimates:

$$\| C_i \|_{N[I]} \lesssim \eta^{\delta_1} c_k, \tag{224}$$

which suffice to establish (55) for all but the first three terms on the RHS of the equation for T_k^m above. It suffices to work with the case of $i = 3, 4$; the cases $i = 1, 2$ are similar but simpler because the frequency envelope is on the high term.

For the trilinear form C_3 we decompose into all possible frequencies and use (25), which gives:

$$\| C_3 \|_{N[I]} \lesssim 2^{-k} \sum_{k_1, k_3 < k-m} 2^{k_1} 2^{-\delta(k_1-k_3)_+} c_{k_3} \lesssim 2^{-m} c_{k-m} \lesssim 2^{-(1-\delta_0)m} c_k,$$

which suffices to show (224) in light of the definition (222) for m.

To prove the bound (224) for C_4 we only split $\phi^{(3)}$ into separate frequencies, and we use (21) and (24) to bound:

$$\| C_4 \|_{N[I]} \lesssim 2^{-k} \sum_{k_3 < k-m} 2^{k_3} c_{k_3} \lesssim 2^{-m} c_{k-m} \lesssim 2^{(1-\delta_0)m} c_k.$$

Step 2 (*Reduction to Matched Frequencies*). We are now trying to bound the sum of the first three terms on the RHS of (223) above. Here we write:

$$\text{(First three terms on R.H.S.)}(223) = A_1 + A_2 + A_3 + B_1 + B_2,$$

where A_1, A_2 and A_3 account for the unmatched frequency interactions:

$$A_1 = \sum_{k_1 \geqslant k-m} \sum_{\max\{k_2,k_3\} \geqslant k+m} P_k \left(\phi^{(1)}_{k_1} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \phi^{(3)}_{k_3} \right),$$

$$A_2 = \sum_{k_1 \geqslant k-m} \sum_{\min\{k_2,k_3\} \leqslant k-2m} P_k \left(\phi^{(1)}_{k_1} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \phi^{(3)}_{k_3} \right),$$

$$A_3 = \sum_{\max\{k_2,k_3\} \geqslant k+m} P_k \left(\phi^{(1)}_{<k-m} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \phi^{(3)}_{k_3} \right),$$

while B_1 and B_2 account for the matched frequency interactions:

$$B_1 = \sum_{k-2m < k_2,k_3 < k+m} P_k \left(\phi^{(1)}_{\geqslant k-m} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \phi^{(3)}_{k_3} \right),$$

$$B_2 = \sum_{k-m < k_2,k_3 < k+m} P_k \left(\phi^{(1)}_{<k-m} \partial^\alpha \phi^{(2)}_{k_2} \partial_\alpha \phi^{(3)}_{k_3} \right).$$

The goal of this step is to prove the set of estimates:

$$\| A_i \|_{N[I]} \lesssim 2^{-\frac{1}{2}\delta m} c_k, \tag{225}$$

which is sufficient to establish (55) for these terms because of the definition (222).

To prove (225) for the term A_1 we use (25). The two highest frequencies can only differ by $O(1)$, therefore we get three distinct contributions if the highest pairs are {12}, {13}, or {23} respectively:

$$\| A_1 \|_{N[I]} \lesssim \sum_{k_2 \geqslant k+m} \sum_{k_3 \leqslant k_2} 2^{\delta(k_3-k_2)} 2^{\delta(k-k_2)} c_{k_3} + \sum_{k_3 \geqslant k+m} \sum_{k_2 \leqslant k_3} 2^{\delta(k_2-k_3)} 2^{\delta(k-k_3)} c_{k_3}$$

$$+ \sum_{k_3 \geqslant k+m} \sum_{k_1=k-m}^{k_3} 2^{\delta(k-k_3)} c_{k_3} \lesssim m 2^{-\frac{3}{4}\delta m} c_{k+m} \lesssim 2^{-\frac{1}{2}\delta m} c_k.$$

In the case of the term A_2 we must have either the condition $\max\{k_2, k_3\} > k - 10$, or the conditions $\max\{k_2, k_3\} \leqslant k - 10$ and $k_1 > k - 10$. This gives two distinct contributions using estimate (25), which after summing out the k_1 index may be (resp) written as:

$$\| A_2 \|_{N[I]} \lesssim S_1 + S_2,$$

with:

$$S_1 = \sum_{\min\{k_2,k_3\}<k-2m} \sum_{\max\{k_2,k_3\}>k-10} 2^{\delta(k-\max\{k_2,k_3\})} 2^{\delta(\min\{k_2,k_3\}-k+m)} c_{k_3},$$

$$S_2 = \sum_{\min\{k_2,k_3\}<k-2m} \sum_{\max\{k_2,k_3\}\leqslant k-10} 2^{\delta(\min\{k_2,k_3\}-k)} c_{k_3}.$$

For the sum S_1 we split into cases depending on which index is minimal, and then sum out k_2 which yields:

$$S_1 \lesssim \sum_{k_3<k-2m} 2^{\delta(k_3-k+m)} c_{k_3} + 2^{-\delta m} \sum_{k_3>k-10} 2^{\delta(k-k_3)} c_{k_3} \lesssim 2^{-\delta m}(c_{k-2m}+c_k) \lesssim 2^{-\frac{1}{2}\delta m} c_k.$$

For the sum S_2 we again split into cases depending on which index is minimal:

$$S_2 \lesssim \sum_{k_3<k-2m} (k-k_3) 2^{\delta(k_3-k)} c_{k_3} + \sum_{k_2<k-2m} \sum_{k_3=k_2}^{k-10} 2^{\delta(k_2-k)} c_{k_3}.$$

For the first sum on the RHS above we get $2^{-\delta m} c_{k-2m}$ which is acceptable. For the second sum we further split the range into $k_3 < k - 2m$ and $k - 2m \leqslant k_3 < k - 10$. In the first case we again get $2^{-\delta m} c_{k-2m}$, while in the second we are left with $m2^{-2\delta m} \sup_{k-2m\leqslant k_3<k-10} c_{k_3}$, which again suffices.

Finally, in the term A_3 we must have $|k_2 - k_3| < 10$ and only the frequency 2^k part of the null form $\partial^\alpha \phi_{k_2}^{(2)} \partial_\alpha \phi_{k_3}^{(3)}$ will contribute. Then we use (24) for the null form, combined with (21):

$$\| A_3 \|_{N[I]} \lesssim \sum_{k_3>k+m} 2^{\delta(k-k_3)} c_{k_3} \lesssim 2^{-\delta m} c_{k+m} \lesssim 2^{-\frac{1}{2}\delta m} c_k.$$

Step 3 (*The Matched Frequency Estimate*). After the last step, it remains to bound the remaining two terms B_i. In both cases, by an application of either (21) or (22), we only need to show the more general matched frequency estimate:

$$\| \partial^\alpha \phi_{[k-O(m),k+O(m)]}^{(2)} \partial_\alpha \phi_{[k-O(m),k+O(m)]}^{(3)} \|_{N[I]} \lesssim \eta^{\delta_1} c_k, \tag{226}$$

under the conditions of Proposition 3.6. Using the bound on m (222) and the definition (11), it suffices to establish the fixed frequency estimate:

$$\| \partial^\alpha \phi_{k_2}^{(2)} \partial_\alpha \phi_{k_3}^{(3)} \|_{N[I]} \lesssim 2^{Cm} \eta^\delta c_{k_3},$$

where we are restricting $|k_2 - k_3| \lesssim m$. This follows immediately from (46). □

Remark 7.2. We remark here that one may prove estimate (61) by a quick application of the above work. To see this, notice the above proof up to Step 3 does not use Proposition 3.4. Thus, we are left with showing estimate (226) in this case, and by inspection of Step 2 we may assume the gap between k_2 and k_3 is no larger than $3m$. By directly applying estimate (24) we have (61) in this case.

Proof of estimate (57). The proof of this estimate follows from some simple manipulations of the bounds used to produce (55). A quick review of the previous proof shows that all bounds were achieved with RHS $\lesssim \eta^{\frac{1}{2}\sqrt{\delta_1}\delta}$. Thus, by a direct application of those bounds and using the (δ_0, δ_0) variance condition on $\{c_k\}$ we have:

$$\| T_k^m(P_{<k+m}\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim \eta^{\delta_1} \left(c_k + \| P_{<k}\phi^{(1)} \|_{S[I]} \right),$$

where m is from line (222).

To bound the contribution with a $P_{\geqslant k+m}\phi^{(1)}$ factor we directly apply (25) which yields the sum:

$$\| T_k^m(P_{>k+m}\phi^{(1)}, \phi^{(2)}, \phi^{(3)}) \|_{N[I]} \lesssim \sum_{k_i:k_1>k+m} 2^{-\delta(\max\{k_i\}-k)} 2^{-\delta(k_1-\min\{k_2,k_3\})_+} c_{k_1}$$
$$\lesssim \sum_{k_1:\, k_1>k+m} 2^{-\delta(k_1-k)} c_{k_1} \lesssim 2^{-(\delta-\delta_0)m} c_k,$$

which suffices.   □

Remark 7.3. To prove (62) we follow a similar procedure as in the previous proof, except this time applied to estimate (61) instead of (55). Here it suffices to split cases according to $P_{<k+10}\phi^{(1)}$ or $P_{\geqslant k+10}\phi^{(1)}$ contributions. The details are left to the reader.

8. The Gauge Transformation

In this section we prove Proposition 3.1. The proof is divided into several portions which deal with different aspects of the problem.

8.1. Bounds for B.

Here we transfer the bounds from ϕ to B. Precisely, we have:

Lemma 8.1. *Let ϕ be a wave map as in Proposition 3.1. Then the matrix B defined on line (31) has an antisymmetric extension off the interval I, which satisfies the following global bounds:*

$$\| \nabla_{t,x} B_k \|_{L_t^\infty(L_x^2)} \lesssim_E \widetilde{c}_k, \tag{227}$$

$$\| B_k \|_{S\cap\underline{X}} \lesssim_F c_k, \tag{228}$$

$$\| \Box B_{k'} \cdot \psi_k \|_N \lesssim_F 2^{\delta(k'-k)} c_{k'} \| \psi_k \|_S, \quad k' < k-10. \tag{229}$$

Proof. By definition we have that:

$$B_k = \mathcal{S}(\phi)_{<k-10}\phi_k,$$

where $\mathcal{S}$ is the antisymmetric part of the original second fundamental form. The bound (227) on I follows from the same bound for ϕ combined with Leibniz's rule. Furthermore, an $S[I]$ norm bound as on line (228) follows from the algebra property (19) combined with the Moser estimate (27).

For the estimates involving $\Box B_k$ we remark that the function $\mathcal{S}(\phi)$ solves a similar wave equation to ϕ on the interval I, which we write schematically as:

$$\Box \mathcal{S}(\phi) = \mathcal{F}(\phi)\partial^\alpha \phi \partial_\alpha \phi.$$

By the version of estimate (51) in Remark 3.8, we have the pair of bounds:

$$\| \Box P_k \mathcal{S}(\phi) \|_{L_t^2(L_x^2)[I]} + \| \Box P_k \phi \|_{L_t^2(L_x^2)[I]} \lesssim_F 2^{\frac{k}{2}} c_k. \tag{230}$$

By Leibniz's rule we have:

$$\Box B_k = \Box \mathcal{S}(\phi)_{<k-10} \phi_k + 2 \partial^\alpha \mathcal{S}(\phi)_{<k-10} \partial_\alpha \phi_k + \mathcal{S}(\phi)_{<k-10} \Box \phi_k. \tag{231}$$

Hence using (230) for the first and last term, and again using Remark 3.8 for the null form in the middle term, we obtain an $\underline{X}[I]$ bound as on line (228).

It remains to prove the estimate (229) localized to I. We again use the expression (231) for $\Box B_k$. We multiply the RHS of this line by a function ψ_j of frequency $j > k+10$. The contribution of the middle term can be estimated by (25):

$$\| \partial^\alpha \mathcal{S}(\phi)_{<k-10} \partial_\alpha \phi_k \psi_j \|_{N[I]} \lesssim_F 2^{-\delta(j-k)} c_k \| \psi_j \|_{S[I]}.$$

In both other cases, by the $S \cdot N$ algebra property, it suffices to prove the estimate:

$$\| P_k \left(\phi^{(1)} \partial^\alpha \phi^{(2)} \partial_\alpha \phi^{(3)} \right) \psi_j \|_{N[I]} \lesssim_F 2^{-\delta(j-k)} c_k \| \psi_j \|_{S[I]},$$

for any set of test functions $\phi^{(i)}$ with $S[I]$ norm of size F, and frequency envelopes $\{c_k\}$. To show this, we let T_k^{10} be the trilinear form defined on line (53), built up out of the $\phi^{(i)}$ in the above estimate. Then by a combination of (61) and estimate (22) we have:

$$\| T_k^{10} \cdot \psi_j \|_{N[I]} \lesssim_F 2^{-\delta(j-k)} c_k \| \psi_j \|_{S[I]}.$$

It remains to show a bound of the form:

$$\| \psi_j \cdot \phi^{(1)}_{<k-10} \partial^\alpha \phi^{(2)}_{<k-10} \partial_\alpha \phi^{(3)}_k \|_{N[I]} \lesssim_F 2^{-\delta(j-k)} c_k \| \psi_j \|_{S[I]},$$

which encapsulates the remainder from T_k^{10} (there are two such remainders, but they are essentially symmetric). This bound follows immediately by applying the algebra estimate (19) to the first two terms, and then summing the resulting trilinear via estimate (25).

To conclude our proof of the estimates (227)–(229) we simply need to extend off the interval I in a simultaneous way. To do this we use the canonical extension defined in Proposition 5.5. $\square$

8.2. The gauge construction. Here we construct the gauge transformation U and obtain estimates on U in S and $\underline{X}$. For comparison purposes we note that in the small data results of [29,33] the function U is constructed iteratively by setting:

$$U_{<k} = \sum_{k'<k} U_{k'}, \quad U_k = U_{<k-C} B_k,$$

with $k \in \mathbb{Z}$. This insures that U_k are localized at frequency 2^k, while the smallness of ϕ in S is used to prove that $U^\dagger U - I$ is small.

Such a construction is no longer satisfactory here, as ϕ can be large in S and thus U may fail to be almost orthogonal. Instead we switch to a continuous version of the

above construction where we seek U and its "frequency localized" version $U_{,<k}$ in the integrated form:

$$U = \int_{-\infty}^{\infty} U_{,k} dk, \quad U_{,<j} = \int_{-\infty}^{j} U_{,k} dk,$$

where each $U_{,k}$ is defined by:

$$U_{,k} = U_{,<k} B_k.$$

In other words, $U_{,<k}$ solves the Cauchy problem:

$$\frac{d}{dk} U_{,<k} = U_{,<k} B_k, \quad U_{,<-\infty} = I_N. \tag{232}$$

Owing to the antisymmetry of the B_k, solutions to this ODE enjoy the conservation law $U_{,<k} U_{,<k}^\dagger = I_N$, so they are automatically *exactly* orthogonal. However, the price one pays is that the exact frequency localization of each $U_{,k}$ is lost. In spite of this, we will prove that $U_{,k}$ is approximately localized at frequency 2^k modulo rapidly decreasing tails:

$$\| P_j U_{,k} \|_S \lesssim_F 2^{-\delta|k-j|} c_k. \tag{233}$$

Note however that arbitrarily high frequencies are immediately introduced, and their evolution is not easy to track. In particular a bootstrap argument for the above S norm bound would seem to fail due to the lack of smallness of the B_k's. We proceed with the proof in several steps aimed at building up to the full S norm estimate by using the conservation law of (232) in a crucial way:

Step 1 ($L_t^\infty(L_x^\infty)$ *and* $L_t^\infty(L_x^2)$ *bounds for* $U_{,k}$). We will work exclusively with the energy frequency envelope $\{\widetilde{c}_k\}$ for B in this step. Without loss of generality we may assume that this is bounded by the S norm frequency envelope $\{c_k\}$. We start with the pointwise and energy bounds:

$$\| B_k \|_{L_t^\infty(L_x^\infty)} \lesssim_E \widetilde{c}_k, \ \| \nabla_{t,x} B_k \|_{L_t^\infty(L_x^\infty)} \lesssim_E 2^k \widetilde{c}_k, \ \| \nabla_{t,x} B_k \|_{L_t^\infty(L_x^2)} \lesssim_E \widetilde{c}_k, \tag{234}$$

all derived from (227). We claim that this implies the following energy type for $U_{,k}$ itself:

$$\| P_{k'} \nabla_{t,x} U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim_E 2^{-|k'-k|} 2^{C(k-k')_+} \widetilde{c}_k. \tag{235}$$

To show this, notice that by construction of $U_{,<k}$ we immediately obtain:

$$\| U_{,k} \|_{L_t^\infty(L_x^\infty)} \lesssim_E \widetilde{c}_k, \quad \| U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim_E 2^{-k} \widetilde{c}_k.$$

We estimate $\nabla_{t,x} U_{,k}$ by differentiating (232):

$$\frac{d}{dk} \nabla_{t,x} U_{,<k} = \nabla_{t,x} U_{,<k} \cdot B_k + U_{,<k} \nabla_{t,x} B_k, \quad \nabla_{t,x} U_{,<-\infty} = 0. \tag{236}$$

In view of the second estimate on line (234), we have good bounds for the second term on the RHS of the above expression, and we wish to transfer these to $\nabla_{t,x} U_{,<k}$. In order to do this, we employ the following device that will be used repeatedly in the sequel:

Lemma 8.2 (Unitary Variation of Parameters). *Let $V_{,<k}$ be given by the ODE:*

$$\frac{d}{dk}V_{,<k} = V_{,<k}B_k + W_{,k} \qquad V_{,<-\infty} = W_{,-\infty} = 0, \tag{237}$$

where B_k is antisymmetric and the forcing term $W_{,k}$ is arbitrary. Then in any mixed Lebesgue space $L_t^q(L_x^r)$ we have the following bound:

$$\| V_{,<k} \|_{L_t^q(L_x^r)} \lesssim \int_{-\infty}^k \| W_{,k'} \|_{L_t^q(L_x^r)} dk'. \tag{238}$$

Proof. We write the formula for $V_{,<k}$ via variation of parameters as follows:

$$V_{,<k} = \int_{-\infty}^k \mathcal{P}(k,k')W_{,k'} dk',$$

where $\mathcal{P}$ is the propagator of the unitary problem:

$$\frac{d}{dk}\mathcal{P}(k,k') = B_k\mathcal{P}(k,k'), \quad \mathcal{P}(k',k') = I_N.$$

In particular, $\| \mathcal{P}(k,k') \|_{L_t^\infty(L_x^\infty)} \lesssim 1$. The proof is concluded via an application of Minkowski's integral inequality. $\quad\square$

We now use estimate (238) to integrate (236), which yields:

$$\| \nabla_{t,x}U_{,<k} \|_{L_t^\infty(L_x^\infty)} \lesssim_E 2^k\widetilde{c}_k,$$

through a direct application of the sum rule (13). From the differentiated equation for $U_{,k}$ this shows that:

$$\| \nabla_{t,x}U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim_E \widetilde{c}_k.$$

Repeating the process for all possible spatial derivatives of $\nabla_{t,x}U_{,<k}$ we get the inductive bounds:

$$\| \nabla_x^J\nabla_{t,x}U_{,<k} \|_{L_t^\infty(L_x^\infty)} \lesssim_E 2^{(|J|+1)k}\widetilde{c}_k, \quad \| \nabla_x^J\nabla_{t,x}U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim_E 2^{|J|k}\widetilde{c}_k. \tag{239}$$

The second relation shows in particular that:

$$\| P_{k'}\nabla_{t,x}U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim_E 2^{C(k-k')}\widetilde{c}_k,$$

for any positive constant C, and therefore by integration that:

$$\| P_{k'}\nabla_{t,x}U_{,<k} \|_{L_t^\infty(L_x^2)} \lesssim_E 2^{C(k-k')}\widetilde{c}_k,$$

which suffices for (235) if $k' \geq k$.

It remains to bound the low frequencies in $U_{,k}$, and so we write:

$$P_{k'}U_{,k} = P_{k'}(P_{[k-10,k+10]}U_{<k} \cdot B_k), \quad k' < k - 20.$$

Using Bernstein's inequality (9) we obtain:

$$\| P_{k'}\nabla_{t,x}U_{,k} \|_{L_t^\infty(L_x^2)} \lesssim 2^{k'}\| \nabla_{t,x}(P_{[k-10,k+10]}U_{,<k}B_k) \|_{L_t^\infty(L_x^1)} \lesssim_E \widetilde{c}_k 2^{k'-k}.$$

Hence (235) is proved.

Step 2 (*Strichartz bounds for $U_{,k}$*). This section largely mimics the previous one, so we will be more terse here. By (228) we have the Strichartz bounds:

$$\| B_k \|_{D\underline{S}_k} \lesssim_F 2^{-k} c_k, \quad \| \nabla_{t,x} B_k \|_{D\underline{S}_k} \lesssim_F c_k,$$

where we recall that $D\underline{S}_k$ is the space of Strichartz admissible $L_t^q(L_x^r)$ norms from line (152) with appropriate dyadic weight (note that this norm does *not* include frequency localization, which will be notationally useful here).

Using the bounds for B_k with Eq. (236) or its derivatives, we directly have:

$$\| \nabla_x^J \nabla_{t,x} U_{,k} \|_{D\underline{S}_k} \lesssim_F 2^{|J|k} c_k, \quad |J| \geqslant 0.$$

By using this last set of estimates for high frequencies, and (235) and Bernstein's inequality for low frequencies, we have:

$$\| P_{k'} \nabla_{t,x} U_{,k} \|_{D\underline{S}_k} \lesssim_F 2^{-|k'-k|} 2^{C(k'-k)_+} c_k.$$

In particular, one has the inequality:

$$\| P_{k'} \nabla_{t,x} U_{,k} \|_{L_t^4(L_x^\infty)} \lesssim_F 2^{\frac{3}{4}k} 2^{-|k'-k|} c_k, \tag{240}$$

which will be useful later in this section. Finally, by interpolating this last bound with (235) and recalling the definition from line (148), we have the following $\underline{S}$ norm portion of estimate (233):

$$\| P_{k'} U_{,k} \|_{\underline{S}} \lesssim_F 2^{-\frac{1}{4}|k'-k|} c_k.$$

Step 3 (*High modulation bounds for $U_{,k}$*). Here we will show that:

$$\| P_j \Box U_{,k} \|_{L_t^2(L_x^2)} \lesssim_F 2^{\frac{k}{2}} 2^{-(\frac{1}{2}+\delta)|k-j|} c_k. \tag{241}$$

Differentiating the equation for $U_{,<k}$ we obtain the evolution equation for $\Box U_{,<k}$:

$$\frac{d}{dk} \Box U_{,<k} = \Box U_{,<k} B_k + U_{,<k} \Box B_{,k} + 2\partial^\alpha U_{,<k} \partial_\alpha B_k. \tag{242}$$

Our first goal will be to use Lemma 8.2 to show that:

$$\| \Box U_{,<k} \|_{L_t^2(L_x^2)} \lesssim_F 2^{\frac{k}{2}} c_k. \tag{243}$$

By estimate (238) and the $\underline{X}$ control for B_k from line (228), it suffices to have the null-form bound:

$$\| \partial^\alpha U_{,<k} \partial_\alpha B_k \|_{L_t^2(L_x^2)} \lesssim_F 2^{\frac{k}{2}} c_k. \tag{244}$$

Expanding the term on the LHS of this last line we have:

$$\partial^\alpha U_{,<k} \partial_\alpha B_k = \int_{-\infty}^k \partial^\alpha (U_{,<k'} B_{k'}) \partial_\alpha B_k dk',$$
$$= \int_{-\infty}^k U_{,<k'} \partial^\alpha B_{k'} \partial_\alpha B_k dk' + \int_{-\infty}^k \partial^\alpha U_{,<k'} B_{k'} \partial_\alpha B_k dk'. \tag{245}$$

Estimate (244) for the first term on the RHS of this last line follows by summing over the bound (23). For the second term on the RHS of the last line above we may take a product of two $L_t^4(L_x^\infty)$ estimates for the terms at frequency k' and $, < k'$, and one energy type bound for B_k. This again yields (244). A similar argument allows us to prove the analog of estimates (244) and (243) for higher spatial derivatives:

$$\| \nabla_x^J (\partial^\alpha U_{,<k} \partial_\alpha B_k) \|_{L_t^2(L_x^2)} \lesssim_F 2^{(\frac{1}{2}+|J|)k} c_k,$$
$$\| \nabla_x^J \Box U_{,<k} \|_{L_t^2(L_x^2)} \lesssim_F 2^{(\frac{1}{2}+|J|)k} c_k. \tag{246}$$

Turning our attention to U_k we have the identity:

$$\Box U_{,k} = U_{,<k} \Box B_k + \Box U_{,<k} B_k + 2\partial^\alpha U_{,<k} \partial_\alpha B_k. \tag{247}$$

By estimates (243), (244), and the analogous bound for B_k from line (228) we directly have:

$$\| \Box U_{,k} \|_{L_t^2(L_x^2)} \lesssim_F 2^{\frac{k}{2}} c_k,$$

while the estimates on line (246) combined with the energy and $L_t^\infty(L_x^\infty)$ bounds for derivatives of $U_{,<k}$ proved in the first step allow us to prove:

$$\| \nabla_x^J \Box U_{,k} \|_{L_t^2(L_x^2)} \lesssim_F 2^{(\frac{1}{2}+|J|)k} c_k.$$

This suffices to give (241) for all but the low frequencies.

It remains to obtain improved low frequency bounds, i.e. prove (241) in the case when $j < k - 10$. The first two terms in (247) are easy to estimate, combining the $L_t^2(L_x^2)$ bound for one factor with the $L_t^\infty(L_x^2)$ energy type bound for the other, while using Bernstein's inequality at low frequency.

The third term on the RHS of (247) has already been estimated before using (245), but now we need to be more careful to gain from small j. The first term on the RHS of (245), call it T_1, can at low frequency be split into three contributions, $P_j T_1 = T_{11} + T_{12} + T_{13}$, where

$$T_{11} = P_j \int_{k-4}^k P_{<j+4} U_{,<k'} \cdot P_{<j+8}(\partial^\alpha B_{k'} \partial_\alpha B_k) dk',$$
$$T_{12} = P_j \int_{k-4}^k \int_{j+4}^\infty P_l U_{,<k'} \cdot P_{[l-4,l+4]}(\partial^\alpha B_{k'} \partial_\alpha B_k) dl dk',$$
$$T_{13} = P_j \int_{-\infty}^{k-4} P_{[k-10,k+10]} U_{,<k'} \cdot \partial^\alpha B_{k'} \partial_\alpha B_k dk'.$$

We explain the estimates for each of these terms. In the case of T_{11} we bound $P_{<j+4} U_{,<k'}$ in $L_t^\infty(L_x^\infty)$ and then apply (23) for the remaining null form. In the case of T_{12} we use (235) to bound $P_l U_{,<k'}$ in $L_t^\infty(L_x^2)$, (23) for the remaining null form, and then conclude with Bernstein's inequality. Finally, the bound T_{13} is obtained in the same way as in the case of T_{12}.

Finally, we need to prove the low frequency part of the estimate (241) for the second term on line (245) above, which we denote by T_2. This cannot be done directly,

because there is no extra room in the application of Strichartz estimates to use Bernstein's inequality. Therefore we reexpand as follows:

$$T_2 = T_{21} + T_{22} = \int_{-\infty}^{k} \int_{-\infty}^{k_1} U_{,<k_2} \partial^\alpha B_{k_2} B_{k_1} \partial_\alpha B_k dk_2 dk_1$$
$$+ \int_{-\infty}^{k} \int_{-\infty}^{k_1} \partial^\alpha U_{,<k_2} B_{k_2} B_{k_1} \partial_\alpha B_k dk_2 dk_1.$$

The first term T_{21} on the RHS above has a structure very similar to the whole of T_1 above. The only new development is that extra factor of B_{k_1}, but it is harmless due to the fact that its frequency is always greater than the differentiated term $\partial^\alpha B_{k_2}$. Therefore, one can use the same methods as in the previous paragraph to bound this term (one could as well use the procedure we are about to describe for bounding the second term T_{22}). To handle T_{22} above, we split it further as:

$$P_j T_{22} = T_{221} + T_{222} = P_j \int_{-\infty}^{k-8} \int_{-\infty}^{k_1} P_{>k-20} \partial^\alpha U_{,<k_2} \cdot B_{k_2} B_{k_1} \partial_\alpha B_k dk_2 dk_1$$
$$+ P_j \int_{k-8}^{k} \int_{-\infty}^{k_1} \partial^\alpha U_{,<k_2} B_{k_2} B_{k_1} \partial_\alpha B_k dk_2 dk_1.$$

For the first term above we put the two (i.e. first and fourth) high frequency terms in $L_t^\infty(L_x^2)$, while the middle two terms are both estimated with $L_t^4(L_x^\infty)$; then we use Bernstein's inequality. One is forced to lose in the low frequencies this way, but this is made up for by the arbitrary gain in the difference $(k - k_2)$ coming from estimate (235):

$$\| T_{221} \|_{L_t^2(L_x^2)} \lesssim_F c_k \int_{-\infty}^{k-8} \int_{-\infty}^{k_1} 2^j \cdot 2^{C(k_2-k)} \cdot 2^{-\frac{1}{4}k_2} \cdot 2^{-\frac{1}{4}k_1} dk_2 dk_1 \lesssim_F 2^{\frac{1}{2}j} 2^{\frac{1}{2}(j-k)} c_k.$$

To bound the term T_{222} we put both the k_2 indexed terms in $L_t^4(L_x^\infty)$, and the other two factors in $L_t^\infty(L_x^2)$ while using Bernstein's inequality at low frequency. This gives the inequality:

$$\| T_{222} \|_{L_t^2(L_x^2)} \lesssim_F c_k \int_{k-8}^{k} \int_{-\infty}^{k_1} 2^j \cdot 2^{\frac{1}{2}k_2} \cdot 2^{-k_1} dk_2 dk_1 \lesssim_F 2^{\frac{1}{2}j} 2^{\frac{1}{2}(j-k)} c_k.$$

This completes our demonstration of the estimate (241).

Step 4 (High frequency bounds for $U_{,k}$). Here we show that the high frequencies in $U_{,k}$ can be estimated in a much more favorable way:

$$\| \nabla_{t,x}^J P_{k'} U_{,k} \|_{L_t^1(L_x^1)} \lesssim_F 2^{(-3+|J|)k} 2^{-C(k'-k)} c_k, \quad k' > k + 10, \qquad (248)$$

where $|J| \leqslant 2$. For this we expand with $D = \{k_5 < k_4 < k_3 < k_2 < k_1 < k\}$:

$$P_{k'} U_{,k} = P_{k'} \int_D U_{,<k_5} B_{k_5} B_{k_4} B_{k_3} B_{k_2} B_{k_1} B_k \, dk_5 dk_4 dk_3 dk_2 dk_1.$$

Due to the frequency localizations we can replace $U_{,<k_5}$ by $P_{[k'-10,k'-10]} U_{,<k_5}$, for which we may use the $L_t^\infty(L_x^\infty)$ bound derived from (235). Hence by the Strichartz estimates alone for the B_{k_i}'s we obtain:

$$\| P_{k'} U_{,k} \|_{L_t^1(L_x^1)} \lesssim_F c_k \int_D 2^{C(k_5-k')-\frac{1}{4}(k_5+k_4+k_3+k_2)-k_1-k} c_k dk_5 dk_4 dk_3 dk_2 dk_1.$$

The bound (248) with $J = 0$ follows after integration. The case $|J| = 1$ is treated similarly. A minor variation is needed in the case $|J| = 2$ when two time derivatives occur. There one writes $\partial_t^2 = \Box + \Delta_x$, using either (228) or (246) for the factor containing the d'Alembertian.

Step 5 (Full S norm bounds for $U_{,k}$). Here we prove that:

$$\| P_{k'} U_{,k} \|_S \lesssim_F 2^{-\delta |k-k'|} c_k. \tag{249}$$

In view of the previous step it suffices to consider the case $k' < k + 10$.

Here we encounter the main difficulty compared to [29,33]. The inductive bound used there grows exponentially in k due to lack of smallness, so it is useless. Bootstrapping fails for a similar reason. Instead we consider iterated expansions. There are two bounds we need to prove, namely for $\| P_{k'} Q_j U_{,k} \|_{X_\infty^{1,\frac{1}{2}}}$ and $\| P_{k'} Q_j U_{,k} \|_{S[k';j]}$. Due to the high modulation bound (241) and the high frequency bound (248) it suffices to consider the case $j < k' - 20 < k - 10$. The key technical step asserts that in either case we can bound the contribution of $U_{,<j-20}$ using only pointwise and high modulation bounds:

Lemma 8.3. *Let $j < k - 10$. Then the following estimate holds for test functions u and ϕ:*

$$\| Q_j (u_{<j-10}\phi_k) \|_{X_\infty^{1,\frac{1}{2}}} + \| Q_{<j}(u_{<j-10}\phi_k) \|_{S[k;j]} \lesssim \| u \|_{L_t^\infty(L_x^\infty) \cap \underline{X}} \cdot \| \phi_k \|_S. \tag{250}$$

Proof. We write

$$u_{<j-10}\phi_k = Q_{>j-10}u_{<j-10} \cdot \phi_k + Q_{<j-10}u_{<j-10} \cdot \phi_k.$$

For the first term we obtain an $L_t^2(L_x^2)$ bound, which by (162) suffices for both norms on the left in (250):

$$\begin{aligned}
\| Q_{>j-10}u_{<j-10} \cdot \phi_k \|_{L_t^2(L_x^2)} &\lesssim \| Q_{>j-10}u_{<j-10} \|_{L_t^2(L_x^\infty)} \| \phi_k \|_{L_t^\infty(L_x^2)} \\
&\lesssim 2^j \| Q_{>j-10}u_{<j-10} \|_{L_t^2(L_x^2)} \| \phi_k \|_{L_t^\infty(L_x^2)} \\
&\lesssim 2^{-\frac{j}{2}-k} \| \Box u \|_{L_t^2(\dot{H}^{-\frac{1}{2}})} \| \nabla_{t,x}\phi_k \|_{L_t^\infty(L_x^2)}.
\end{aligned}$$

For the second term we consider separately the two cases. On one hand:

$$Q_j(Q_{<j-10}u_{<j-10} \cdot \phi_k) = Q_j(Q_{<j-10}u_{<j-10} \cdot Q_{>j-10}\phi_k).$$

Therefore we directly have:

$$\begin{aligned}
\| Q_j(Q_{<j-10}u_{<j-10} \cdot \phi_k) \|_{X_\infty^{1,\frac{1}{2}}} &\lesssim 2^{\frac{j}{2}+k} \| Q_{<j-10}u_{<j-10} \|_{L_t^\infty(L_x^\infty)} \| Q_{>j-10}\phi_k \|_{L_t^2(L_x^2)} \\
&\lesssim \| u \|_{L_t^\infty(L_x^\infty)} \| \phi_k \|_{X_\infty^{1,\frac{1}{2}}}.
\end{aligned}$$

On the other hand, by a direct application of estimate (168) we have:

$$\| Q_{<j}(Q_{<j-10}u_{<j-10} \cdot \phi_k) \|_{S[k;j]} \lesssim \| u \|_{L_t^\infty(L_x^\infty)} \| \phi_k \|_S.$$

The proof of the lemma is concluded. $\square$

We now return to the main proof, and consider the two bounds we need in order to bound $U_{,k}$ in S, namely:

$$\| P_{k'} Q_j U_{,k} \|_{X_\infty^{1,\frac{1}{2}}} + \| P_{k'} Q_{<j} U_{,k} \|_{S[k';j]} \lesssim_F 2^{-\frac{1}{4}|k-k'|} c_k \qquad j < k' - 20 < k - 10.$$

For each fixed modulation index j, we expand $U_{,k}$ in the form:

$$U_{,k} = U_{,<j-20} B_k + \int_{j-20}^{k} U_{,<j-20} B_{k_1} B_k dk_1 + \int_{j-20}^{k} \int_{j-20}^{k_1} U_{,<k_2} B_{k_2} B_{k_1} B_k dk_2 dk_1. \tag{251}$$

Step 5A (*Contribution of the first term in* (251)). We write:

$$U_{,<j-20} B_k = P_{<j-10} U_{,<j-20} \cdot B_k + P_{>j-10} U_{,<j-20} \cdot B_k.$$

The first component has output at frequency k, and its contribution is accounted for due to Lemma 8.3. The second can have both high and low frequency output, so we need to split it further.

For the high frequency output we estimate:

$$\| P_{>j-10} U_{,<j-20} \cdot B_k \|_{L_t^2(L_x^2)} \lesssim \| P_{>j-10} U_{,<j-20} \|_{L_t^2(L_x^\infty)} \| B_k \|_{L_t^\infty(L_x^2)} \lesssim_F 2^{-\frac{j}{2}-k} c_k,$$

where the $L_t^2(L_x^\infty)$ norm is estimated by interpolating the (summed version of the) energy bound (235) with the $L_t^1(L_x^1)$ high frequency bound (248) for $U_{,<j-20}$, and by using Bernstein's inequality.

In the case of low frequency output $k' < k - 20$, the first factor is further restricted to high frequencies so we may bound:

$$\| P_{k'}(P_{>k-10} U_{,<j-20} \cdot B_k) \|_{L_t^2(L_x^2)} \lesssim \| P_{>k-10} U_{,<j-20} \|_{L_t^2(L_x^\infty)} \| B_k \|_{L_t^\infty(L_x^2)}$$
$$\lesssim_F 2^{-C(k-j)} 2^{-\frac{3}{2}k} c_k, \tag{252}$$

where we have followed the same procedure as in the previous estimate. The restriction $j < k'$ then suffices to produce (249) for this term.

Step 5B (*Contribution of the second term in* (251)). We need to split this into several subcases:

Step 5B.1 (*Contribution of high frequencies in* $U_{<j-20}$). This term may have both low and high frequency output. In the case of high frequencies we estimate directly in $L_t^2(L_x^2)$ using Strichartz estimates as follows:

$$\| P_{>j-10} U_{,<j-20} \cdot B_{k_1} B_k \|_{L_{t,x}^2} \lesssim \| P_{>j-10} U_{,<j-20} \|_{L_t^4(L_x^\infty)} \| B_{k_1} \|_{L_t^4(L_x^\infty)} \| B_k \|_{L_t^\infty(L_x^2)}$$
$$\lesssim_F 2^{-\frac{j+k_1}{4}-k} c_k,$$

where the $j - 20 < k_1 < k$ integration is now straightforward and yields a RHS expression of the form $\lesssim_F 2^{-k} 2^{-\frac{1}{2}j} c_k$ which suffices.

In the case of low frequency output where $k' < k - 20$, we further split the integrand as follows:

$$P_{k'}(P_{>j-10} U_{,<j-20} \cdot B_{k_1} B_k) = P_{k'}(P_{>j-10} U_{,<j-20} \cdot P_{>k-10} B_{k_1} \cdot B_k)$$
$$+ P_{k'}(P_{>j-10} U_{,<j-20} \cdot P_{<k-10} B_{k_1} \cdot B_k).$$

The first term is estimated as above with a gain of $2^{-\frac{1}{4}(k-j)}$ due to the restriction on k_1 (which in particular restricts the range of integration for this term). This suffices to show (249) for this term. To handle the second term, we use the fact that the first factor is now forced to be at large frequency, which gives an $L_t^2(L_x^2)$ bound as on line (252) above. Notice that the additional integration in $j - 20 < k_1 < k$ may be absorbed via the factor of $2^{-C(k-j)}$.

Step 5B.2 (*Contribution of low frequencies but high modulations in* $U_{<j-20}$). In this case the only possible low frequency contribution comes when $|k_1 - k| < 10$. Therefore we may proceed as above using the high modulation bound (243) for the first factor as follows:

$$\| P_{k'}(Q_{>j-10}P_{<j-10}U_{,<j-20} \cdot B_{k_1}B_k) \|_{L_t^2(L_x^2)}$$
$$\lesssim \| Q_{>j-10}P_{<j-10}U_{,<j-20} \|_{L_t^4(L_x^\infty)} \| B_{k_1} \|_{L_t^4(L_x^\infty)} \| B_k \|_{L_t^\infty(L_x^2)}$$
$$\lesssim_F 2^{-\frac{j+k_1}{4}-k}c_k,$$

and the integral in k_1 is the same as above depending on whether $|k' - k| < 10$ or $k' < k - 10$. In either case one gains a RHS factor of $\lesssim_F 2^{-\frac{1}{4}(k-k')}2^{-k}2^{-\frac{1}{2}j}c_k$.

Step 5B.3 (*Contribution of low frequencies and low modulations in* $U_{,<j-20}$). Here we deal with the expression $Q_{<j-10}P_{<j-10}U_{,<j-20} \cdot B_{k_1}B_k$. We consider two subcases:

Step 5B.3.a (*Contribution of the range* $k_1 > k - 10$). Under this restriction, we may group the product $B_{k_1}B_k$ as a single term, which we further decompose into all frequencies $k' < k + 10$. For each such localized term we have from the algebra bound (20) the estimate:

$$\| P_{k'}(B_{k_1}B_k) \|_S \lesssim 2^{-|k'-k|} \| B_{k_1} \|_S \| B_k \|_S \lesssim_F 2^{-|k'-k|}c_k.$$

Therefore, in the range $j < k'$ the resulting term may be estimated in essentially the same way the first term on the RHS of (251) was estimated in Step 5A above with the additional simplification that the low frequency gains are already implicit in the $P_{k'}(B_{k_1}B_k)$ localization.

Step 5B.3.b (*Contribution of the range* $j - 20 < k_1 < k - 10$). In this case the output is automatically at frequency 2^k. Notice that if we argue as in the previous case then we run into trouble with the k_1 integration. Instead, we observe that one has access to the additional localization:

$$Q_j \left(Q_{<j-10}P_{<j-10}U_{,<j-20} \cdot B_{k_1}B_k \right) = Q_j \left(Q_{<j-10}P_{<j-10}U_{,<j-20} \cdot Q_{<j+4}(B_{k_1}B_k) \right),$$

and according to estimate (169) we may bound the entire contribution of the second factor as:

$$\| Q_{<j+4}(B_{k_1}B_k) \|_S \lesssim 2^{-\delta|j-k_1|} \| B_{k_1} \|_S \| B_k \|_S, \quad j < k_1 \leqslant k.$$

This provides the needed additional gain that enables us to integrate with respect to k_1.

Step 5C (*Contribution of the last term in* (251)). As in the previous step, we need to split into two further subcases depending on the range of integration:

Step 5C.1 (*Contribution of the range* $k_1 > k - 10$). A direct application of Strichartz bounds gives the estimate:

$$\| B_{k_2}B_{k_1}B_k \|_{L_t^2(L_x^2)} \lesssim \| B_{k_2} \|_{L_t^4(L_x^\infty)} \| B_{k_1} \|_{L_t^4(L_x^\infty)} \| B_k \|_{L_t^\infty(L_x^2)} \lesssim_F 2^{-k}2^{-\frac{k_1+k_2}{4}}c_k.$$

The integration with respect to k_1, k_2 over the region $j - 20 < k_2 < k_1 < k$ with the additional restriction that $k_1 = k + O(1)$ is straightforward and yields the RHS term $\lesssim_F 2^{-\frac{1}{4}(k-j)} 2^{-k} 2^{-\frac{1}{2}j} c_k$ which suffices to produce (249) for this term in light of the restriction $j < k'$.

Step 5C.2 (*Contribution of the range* $j - 20 < k_1 < k - 10$). In this case with high frequency output we may proceed as in the previous step. Notice that integration over the full range $j - 20 < k_2 < k_1 < k$ with no additional work still yields a RHS of the form $\lesssim_F 2^{-k} 2^{-\frac{1}{2}j} c_k$.

The contribution of this range with low frequency output forces the first term in the product to have localization in the range $k + O(1)$. One may again proceed as in the last case of Step 5A above to produce an $L_t^2(L_x^2)$ estimate via (252). Notice that the integration in both k_1 and k_2 is safely absorbed by the factor $2^{-C(k-j)}$. This concludes our demonstration of the estimate (249).

Step 6 (*Proof of the bound* (34)). By the algebra estimates (21) and (22) it suffices to do this for $|k' - k| > 20$. We rescale to $k = 0$. There are two cases:

Step 6.A (*Low frequencies;* $k' < -20$). Here we may further localize the transformation matrix to $P_{[-10,10]}U_{,<-20}$. Therefore, we have access to (33). For the lower modulations in G_0 we estimate via Bernstein:

$$\| P_{k'}(P_{[-10,10]}U_{,<-20} \cdot Q_{<20}G_0) \|_{L_t^1(L_x^2)}$$
$$\lesssim 2^{k'} \| P_{[-10,10]}U_{,<-20} \|_{L_t^2(L_x^2)} \| Q_{<20}G_0 \|_{L_t^2(L_x^2)}.$$

This suffices by estimate (198) in Sect. 6 above.

For the high modulation contribution, we split $Q_{>20}G_0 = G^{(1)} + G^{(2)}$, a sum (resp.) of an $L_t^1(L_x^2)$ atom and an $X_1^{0,-\frac{1}{2}}$ atom. For $G^{(1)}$ the bound (34) follows by taking $P_{[-10,10]}U_{,<-20}$ in $L_t^\infty(L_x^2)$ and using Bernstein.

For the $X_1^{0,-\frac{1}{2}}$ atom $G^{(2)}$, we may assume we are working with a single modulation $Q_j G^{(2)}$, where $j > 20$. For modulations $Q_{<j-10}P_{[-10,10]}U_{,<-20}$, estimate (34) follows by again putting the first factor in $L_t^\infty(L_x^2)$ and using Bernstein to estimate the product as a $X_1^{0,-\frac{1}{2}}$ atom with a $2^{k'}$ gain.

For high modulations of the first factor, we estimate:

$$\| P_{k'}(Q_{>j-10}P_{[-10,10]}U_{,<-20} \cdot Q_j G_0^{(2)}) \|_{L_t^1(L_x^2)}$$
$$\lesssim 2^{k'} 2^{-j} \| \partial_t Q_{>j-10}P_{[-10,10]}U_{,<-20} \|_{L_t^2(L_x^2)} \| Q_j G_0^{(2)} \|_{L_t^2(L_x^2)},$$

which is sufficient to place the second factor in $X_1^{0,-\frac{1}{2}}$.

Step 6.B (*High frequencies;* $k' > 20$). Here we may further localize the transformation matrix to $P_{[k'-5,k'+5]}U_{,<-20}$. Therefore, we again have access to (33). In this case we may proceed exactly as above, using at each step the same estimates, which in every case suffice due to the exponential decay in (33) for large frequencies.

Step 7 (*Proof of the bound* (35)). Here we establish the estimate:

$$\| P_k(\Box U_{,k_1} \cdot \psi_{k_2}) \|_N \lesssim_F 2^{-|k-k_2|} 2^{-\delta(k_2-k_1)} c_{k_1} \| \psi \|_S, \quad k_1 < k_2 - 10. \quad (253)$$

We use the expansion (242) for $\Box U_{,<k_1}$ to write the expression we are estimating via a linear combination of the following three terms:

$$G_1 = \Box U_{,<k_1} \cdot B_{k_1}\psi_{k_2}, \quad G_2 = \partial^\alpha U_{,<k_1}\partial_\alpha B_{k_1} \cdot \psi_{k_2}, \quad G_3 = U_{,<k_1}\Box B_{k_1} \cdot \psi_{k_2}.$$

We prove the bound (253) separately for each of these in reverse order. Without loss of generality we will assume that $\| \psi_{k_2} \|_S = 1$.

Step 7A (*Estimating the term G_3*). We directly have from (229) and (34) the product estimate:

$$\| P_k(U_{,<k_1} \cdot \Box B_{k_1} \cdot \psi_{k_2}) \|_N \lesssim_F 2^{-|k-k_2|} 2^{-\delta(k_2-k_1)} c_{k_1}.$$

Step 7B (*Estimating the term G_2*). In this case the bound (253) follows by applying the trilinear null-form estimate (25) along with the bound (249) shown for the first factor in the previous section. One can again split into medium, high and low output frequency cases as in the previous step. The details are left to the reader. Notice that the gains from frequencies higher than k_2 in the first factor are essential for maintaining the separation $2^{-\delta(k_2-k_1)}$.

Step 7C (*Estimating the term G_1*). We break this term into two further contributions:

$$G_1 = G_{11} + G_{12} = P_{<k_2-10}\Box U_{,<k_1} \cdot B_{k_1}\psi_{k_2} + P_{>k_2-10}\Box U_{,<k_1} \cdot B_{k_1}\psi_{k_2}.$$

The first term has output localized to frequency k_2, and we estimate it directly via Strichartz estimates and (243):

$$\| G_{11} \|_{L_t^1(L_x^2)} \lesssim \| \Box U_{,<k_1} \|_{L_t^2(L_x^2)} \| B_{k_1} \|_{L_t^4(L_x^\infty)} \| \psi_{k_2} \|_{L_t^4(L_x^\infty)} \lesssim_F 2^{-\frac{1}{4}(k_2-k_1)} c_{k_1}.$$

The second term G_{12} can have both high and low frequency outputs. When the output is in the range $k < k_2 + 10$ we use (248) and Bernstein's inequality to bound it as follows:

$$\| P_k G_{12} \|_{L_t^1(L_x^2)} \lesssim 2^k \| P_{>k_2-10}\Box U_{,<k_1} \|_{L_t^1(L_x^1)} \| B_{k_1} \|_{L_t^\infty(L_x^\infty)} \| \psi_{k_2} \|_{L_t^\infty(L_x^\infty)}$$
$$\lesssim_F 2^k 2^{-k_1} 2^{-C(k_2-k_1)} c_{k_1},$$

which suffices to show (253) in this case. When G_{12} has output in the high range $k > k_2 + 10$ we have further high frequency localization of the first factor and we may estimate via the same procedure:

$$\| P_k G_{12} \|_{L_t^1(L_x^2)} \lesssim 2^k \| P_{[k-5,k+5]}\Box U_{,<k_1} \|_{L_t^1(L_x^1)} \| B_{k_1} \|_{L_t^\infty(L_x^\infty)} \| \psi_{k_2} \|_{L_t^\infty(L_x^\infty)},$$
$$\lesssim_F 2^k 2^{-k_1} 2^{-C(k-k_1)} c_{k_1},$$

which is again sufficient to show (253) in this case. This concludes our demonstration of Proposition 3.1.

9. The Linear Paradifferential Flow

We now proceed with the proof of Proposition 3.2. The main difficulty here is that we do not necessarily have smallness of the constant from line (40), which would otherwise make estimate (41), consequence of Propositions 2.3 and 3.1. Instead of proceeding directly, we shall follow a more measured approach of building up our estimate piece by piece. Since this is a lengthy argument, we begin with a brief outline.

The first step of the proof is to take advantage of the antisymmetry of A_α, which makes our paradifferential equation almost conservative. Precisely, the only nontrivial contributions to energy estimates arise from terms where one derivative falls on the

coefficients. But such terms are small due to the large frequency gap m. Consequently, we are able to prove a favorable estimate:

$$\| \psi_k \|_{E[I]} \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + 2^{\delta m} \| G_k \|_{N[I]} + 2^{-\delta m} \| \psi_k \|_{S[I]}, \qquad (254)$$

for the energy (28) on both time slices and characteristic surfaces.

We still need an estimate on the S norm of ψ_k, for which we renormalize Eq. (38) using an orthogonal gauge transformation $U_{,<k-m}$ obtained by Proposition 3.1. The function $w_{,k} = U_{,<k-m} \psi_k$ solves a perturbed wave equation of the form:

$$\Box w_{,k} = \mathcal{R}^{pert}_{,<k-m} \psi_k + U_{,<k-m} G_k. \qquad (255)$$

In the analysis of the small data problem in [29,33] one uses a perturbative bound of the form:

$$\| \mathcal{R}^{pert}_{,<k-m} \psi_k \|_{N[I]} \lesssim_F \| \psi_k \|_{S[I]},$$

where the implicit constant is quadratic or better in F, for F small. This is no longer sufficient here. Instead, we observe that we can rebalance the above estimate and use only the energy norm of ψ_k to estimate the bulk of the LHS above. Thus, we prove that for $0 \leqslant m_0 \ll m$ we have:

$$\| P_j(\mathcal{R}^{pert}_{,<k-m} \psi_k) \|_{N[I]} \lesssim_F 2^{-|j-k|} \left(2^{-\delta_1 m_0} \| \psi_k \|_{S[I]} + 2^{2m_0} \| \psi_k \|_{E[I]} \right). \qquad (256)$$

By the linear solvability bound (18) we have:

$$\| w_{,k} \|_{S[I]} \lesssim \| w_{,k}[0] \|_{\dot{H}^1 \times L^2} + \| \Box w_{,k} \|_{N[I]}.$$

Since both $U_{,<m-k}$ and $(U_{,<m-k})^{-1} = U^{\dagger}_{,<m-k}$ are in S with norm $\lesssim_F 1$, by the S algebra property and estimate (34) we have the gauge removal bounds:

$$\| \psi_k \|_{S[I]} \lesssim_F \| w_{,k} \|_{S[I]}, \quad \| P_{k'}(U_{,<k-m} G_k) \|_{N[I]} \lesssim_F 2^{-|k-k'|} \| G_k \|_{N[I]}. \qquad (257)$$

On the other hand using the energy component of (32) we obtain:

$$\| P_j w_{,k}[0] \|_{\dot{H}^1 \times L^2} \lesssim_F 2^{-|j-k|} \| \psi_k[0] \|_{\dot{H}^1 \times L^2}. \qquad (258)$$

Summing up the estimates on the last four lines we obtain the S bound for ψ_k:

$$\| \psi_k \|_{S[I]} \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + \| G_k \|_{N[I]} + 2^{-\delta_1 m_0} \| \psi_k \|_{S[I]} + 2^{2m_0} \| \psi_k \|_{E[I]}. \qquad (259)$$

Now all we have to do is combine this with (254), carefully balancing the constants. Assuming that $m_0 = m_0(F)$ for a large enough $m_0(F) \sim \ln(F)$, the third term on the right can be absorbed on the left to obtain:

$$\| \psi_k \|_{S[I]} \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + \| G_k \|_{N[I]} + \| \psi_k \|_{E[I]}.$$

Substituting (254) for the third term on the RHS of this last line we arrive at:

$$\| \psi_k \|_{S[I]} \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + 2^{\delta m} \| G_k \|_{S[I]} + 2^{-\delta m} \| \psi_k \|_{S[I]},$$

so now assuming $m > m(F)$ for a larger $m(F) \sim \ln(F)$, the last term on the RHS is again absorbed on the left:

$$\| \psi_k \|_{S[I]} \lesssim_F \| \psi_k[0] \|_{\dot{H}^1 \times L^2} + \| G_k \|_{S[I]}.$$

To conclude the proof of (41) we need to improve the S bound above to a $\mathcal{W}$ bound. Returning to $w_{,k}$, we have the estimate:

$$\| P_j \Box w_{,k} \|_{N[I]} + \| P_j w_{,k}[0] \|_{\dot{H}^1 \times L^2} \lesssim_F 2^{-|j-k|} \left(\| \psi_k[0] \|_{\dot{H}^1 \times L^2} + \| G_k \|_{S[I]} \right).$$

This follows from (256), (258), and the second member on line (257).

It remains to prove the two main estimates above, namely (254) and (256). In the proof we shall make use of three auxiliary lemmas whose proofs we postpone until the end of this section. The first one is used to estimate perturbative expressions which are small due to the large frequency gap m.

Lemma 9.1 (Some auxiliary estimates). *Let A_α be the connection one-form defined on line* (39) *above with estimates* (40)*. Then the following bounds hold:*

$$\| A_\alpha \|_{L_t^\infty(L_x^\infty)} \lesssim_F 2^{k-m}, \tag{260}$$

$$\| A_\alpha \psi_k \|_{DS[I]} \lesssim_F 2^{-m} \| \psi_k \|_{S[I]}. \tag{261}$$

Also, for three test functions $\phi^{(i)}$ normalized with $S \cap \underline{E}[I]$ size one, the following list of multilinear estimates holds:

$$\| \phi^{(1)} \partial^\alpha \phi_{<k-m}^{(2)} \partial_\alpha \phi_{<k-m}^{(3)} \cdot \psi_k \|_{N[I]} \lesssim 2^{-\delta m} \| \psi_k \|_{S[I]}, \tag{262}$$

$$\| P_{<k-m} \left(\phi^{(1)} \partial^\alpha \phi^{(2)} \partial_\alpha \phi^{(3)} \right) \psi_k \|_{N[I]} \lesssim 2^{-\delta m} \| \psi_k \|_{S[I]}, \tag{263}$$

$$\| \nabla_{t,x} \phi_{<k-m}^{(1)} \partial^\alpha \phi_{<k-m}^{(2)} \partial_\alpha \psi_k \|_{N[I]} \lesssim 2^{-\delta m} 2^k \| \psi_k \|_{S[I]}, \tag{264}$$

$$\| \nabla_{t,x} \phi_{<k-m}^{(1)} \partial^\alpha \phi_{<k-m}^{(2)} \partial_\alpha \phi_{<k-m}^{(3)} \cdot \psi_k \|_{N[I]} \lesssim 2^{-\delta m} 2^k \| \psi_k \|_{S[I]}. \tag{265}$$

In proving energy estimates we need to restrict integration to half-spaces. This is where the next lemma comes handy:

Lemma 9.2 (Half-space duality estimate). *Let $\psi_k \in S$ and $H_k \in N$ be frequency localized functions. Then for any time-slab I, any unit vector ω, and any spatial point $x_0 \in \mathbb{R}^2$ the following truncated duality estimate holds uniformly:*

$$\left| \iint_{I \cap \{t > \omega \cdot (x - x_0)\}} H_k \cdot \psi_k \, dx dt \right| \lesssim \| H_k \|_{N[I]} \cdot \| \psi_k \|_{DS[I]}. \tag{266}$$

Finally, for the bulk of the estimate (256) we need the following lemma, which improves upon the trilinear bound (25) in the case of balanced low frequencies $|k_1 - k_2| \ll m$, $k_1, k_2 < k_3 - m$:

Lemma 9.3 (An improved trilinear estimate). *There exists a universal constant $C > 0$ such that for any integer $m \geqslant 0$ and $S[I]$ unit normalized test functions $\phi_{k_i}^{(i)}$ with $k_i \leqslant k - m$, and ψ_k any additional test function defined on I, one has the following imbalanced trilinear estimate:*

$$\| \phi_{k_1}^{(1)} \partial_\alpha \phi_{k_2}^{(2)} \partial^\alpha \psi_k \|_{N[I]} \lesssim 2^{C|k_1-k_2|} \left(2^{-\delta^2 m} \| \psi_k \|_{S[I]} + 2^m \| \psi_k \|_{\underline{E}[I]} \right). \tag{267}$$

Assuming these estimates, we give a proof of (254) and (256) in a series of steps. To close the argument properly, we will employ our chain of small constants (8) (although their use here is independent of their use in other sections).

Step 1 (A-priori control of the energy norm of ψ_k: proof of (254)). We begin by writing Eq. (38) for ψ_k on the interval I in a covariant form:

$$\Box_A \psi_k = \mathcal{R}^{en}_{<k-m} \psi_k + G_k, \tag{268}$$

where $\Box_A = (\partial + A)^\alpha (\partial + A)_\alpha$ is the gauge covariant wave equation with the connection A_α is given by the formula on line (39) and the function $\mathcal{R}^{en}_{<k-m}$ has the form:

$$\mathcal{R}^{en}_{<k-m} = \partial^\alpha \mathcal{S}(\phi)_{<k-m} \partial_\alpha \phi_{<k-m} + \mathcal{S}(\phi)_{<k-m} P_{<k-m} \left(\mathcal{S}(\phi) \partial^\alpha \phi \partial_\alpha \phi \right) + A^\alpha A_\alpha. \tag{269}$$

Note that in the RHS of this last line, the matrix $\mathcal{S}(\phi)$ is either the pure second fundamental form $\mathcal{S}(\phi)^c_{ab}$, or its antisymmetric version as it appears in the formula for A_α. The distinction will not be important for us here. Also, notice that we have used the Wave-Map equation for ϕ on the interval I, which we may do by the assumptions of Proposition 3.2.

To obtain the energy estimates we proceed via a simple integration-by-parts argument. First, we form the gauge-covariant energy momentum density:

$$Q_{\alpha\beta}[\psi_k] = ({}^A\nabla_\alpha \psi_k)^{\dagger} {}^A\nabla_\beta \psi_k - \frac{1}{2} g_{\alpha\beta} ({}^A\nabla^\gamma \psi_k)^{\dagger} {}^A\nabla_\gamma \psi_k.$$

Here we are writing ${}^A\nabla_\alpha = \partial_\alpha + A_\alpha$. A quick calculation shows that (notice that this identity crucially uses the antisymmetry of A_α, which is the main source of the cancelation that makes (254) possible):

$$\nabla^\alpha Q_{\alpha\beta}[\psi_k] = (\Box_A \psi_k)^{\dagger} {}^A\nabla_\beta \psi_k + (F_{\gamma\beta} \psi_k)^{\dagger} {}^A\nabla^\gamma \psi_k, \tag{270}$$

where $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta]$ is the curvature of A_α. Next, we form the linear momentum one-form $P_\alpha[\psi_k] = Q_{\alpha 0}[\psi_k]$. Integrating $\nabla^\alpha P_\alpha[\psi_k] = \nabla^\alpha Q_{\alpha 0}[\psi_k]$ over all possible half spaces of the form $[0, t_0] \cap \{t > \omega \cdot (x - x_0)\}$ we have the bound:

$$\| {}^A\nabla_{t,x} \psi_k \|^2_{L^\infty_t(L^2_x)[I]} + \sup_\omega \| {}^A\nabla_{t,x} \psi_k \|^2_{L^\infty_{t_\omega}(L^2_{x_\omega})[I]} \lesssim \| {}^A\nabla_{x,t} \psi_k(0) \|^2_{L^2_x} + I_1 + I_2, \tag{271}$$

where

$$I_1 = \sup_{I, \omega, x_0} \left| \iint_{I \cap \{t > \omega \cdot (x - x_0)\}} (\Box_A \psi_k)^{\dagger} {}^A\nabla_0 \psi_k \, dx dt \right|,$$

$$I_2 = \sup_{I, \omega, x_0} \left| \iint_{I \cap \{t > \omega \cdot (x - x_0)\}} (F_{0\gamma} \psi_k)^{\dagger} {}^A\nabla^\gamma \psi_k \, dx dt \right|.$$

Our task is to estimate I_1 and I_2 and to show that we can replace covariant differentiation by regular differentiation in (271). For the right hand side of (271) we claim that both:

$$\| {}^A\nabla_{x,t} \psi_k(0) \|^2_{L^2_x} \lesssim_F \| \nabla_{x,t} \psi_k(0) \|^2_{L^2_x}, \tag{272}$$

$$I_1 + I_2 \lesssim_F 2^{-2\delta m} \| \psi_k \|^2_{S[I]} + 2^{2\delta m} \| G_k \|^2_{S[I]}. \tag{273}$$

The proof of (272) is an immediate consequence of expanding the covariant derivative ${}^A\nabla$ and using the triangle inequality, followed by the $L_t^\infty(L_x^\infty)$ bound for A_α in (260).

To obtain (273), we use the half-space duality estimate (266) and Young's inequality for the term involving G_k. For the other terms, we again use half-space duality, and then conclude with an application of the estimates (157), (261)–(265). It suffices to establish the bounds:

$$\| \mathcal{R}^{en}_{<k-m}\psi \|_{N[I]} \lesssim_F 2^{-\delta m}\| \psi_k \|_{S[I]}, \quad \| F_{0\gamma} \cdot {}^A\nabla^\gamma \psi_k \|_{N[I]} \lesssim_F 2^{-\delta m}2^k\| \psi_k \|_{S[I]}.$$

The first estimate above follows from applying (262)–(263) to each of the terms in $\mathcal{R}^{en}_{<k-m}$. The second estimate follows from the bounds (264)–(265) applied to the definition of the curvature. Notice that these two multilinear estimates suffice because there are never any terms in I_2 with a single factor containing more than one derivative thanks to the skew symmetry of the curvature.

The bound (254) will now follow once we can rid ourselves of the gauge covariant derivatives ${}^A\nabla_{t,x}$ on the LHS of (271) in favor of the usual derivatives $\nabla_{t,x}$. This can be done with a successive application of the two estimates:

$$\| A\psi_k \|_{L_t^\infty(L_x^2)[I]} \lesssim_F 2^{-m}\| \nabla_{t,x}\psi_k \|_{L_t^\infty(L_x^2)[I]},$$
$$\sup_\omega \|A\psi_k \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)[I]} \lesssim_F \|\psi_k\|_{L_t^\infty(L_x^\infty)[I]} \lesssim \| \nabla_{t,x}\psi_k \|_{L_t^\infty(L_x^2)[I]}.$$

The first of these follows immediately from the bound (260), while the second uses the characteristic energy estimates we are assuming for ϕ. We remark that using the first bound above requires m to be large enough, i.e. $2^m \gg_F 1$.

Step 2 (*The S bound for ψ: Proof of* (256)). The first thing we need to do is to rewrite Eq. (38) in a gauged formulation (we have no further use for (268)). As usual, we write:

$$A^\alpha_{<k-m} = \partial^\alpha B_{<k-m} + D^\alpha_{<k-m},$$

where the RHS is given by the integrated terms:

$$B_{<k-m} = \int_{k'<k-m} \left(S^a_{cb}(\phi) - S^b_{ca}(\phi) \right)_{\leqslant k'-10} \phi^c_{k'}dk',$$
$$D^\alpha_{<k-m} = \int_{k'<k-m} \left(S^a_{cb}(\phi) - S^b_{ca}(\phi) \right)_{k'-10<\cdot<k-m} \partial^\alpha \phi^c_{k'}dk'$$
$$- \int_{k'<k-m} \partial^\alpha \left(S^a_{cb}(\phi) - S^b_{ca}(\phi) \right)_{\leqslant k'-10} \phi^c_{k'}dk'.$$

The connection $\partial^\alpha B_{<k-m}$ is of the form in Proposition 3.1, and we define the $SO(N)$ matrix $U = U_{,<k-m}$ accordingly. We also set $C_\alpha = \nabla_\alpha B - U^\dagger \nabla_\alpha U$, which is given by the second term on the RHS of formula (37). Finally, we denote by $w_{,k} = U_{,<k-m}\psi_k$. Then w_k obeys the gauged equation (255) with:

$$\mathcal{R}^{pert}_{,<k-m}\psi_k = -2U_{,<k-m}(C^\alpha + D^\alpha)\partial_\alpha \psi_k + \Box U_{,<k-m} \cdot \psi_k. \tag{274}$$

The second term on the right is easy to estimate using (35), which yields:

$$\| P_{k'}(\Box U_{,<k-m} \cdot \psi_k) \|_{N[I]} \lesssim_F 2^{-|k-k'|}2^{-\delta_1 m}\| \psi_k \|_{S[I]}. \tag{275}$$

It remains to estimate the first term in $\mathcal{R}_{,<k-m}^{pert}\psi_k$, for which we will show the bound:

$$\| P_{k'}\left(U_{,<k-m}(C^\alpha + D^\alpha)\partial_\alpha \psi_k\right)\|_{N[I]} \lesssim_F 2^{-|k-k'|}\left(2^{-\delta_1 m_0}\|\psi_k\|_{S[I]} + 2^{2m_0}\|\psi_k\|_{E[I]}\right),$$
(276)

for $0 < m_0 \ll m$. We will prove this in a further series of steps.

Step 2A (*Removal of the gauge and high frequency connection*). Here we write C_{low}^α for the second term on the RHS of line (37) with each gauge factor replaced by $P_{<k-10}U_{,<k'}$. Thus $C_{low}^\alpha = P_{<k-10}C_{low}^\alpha$. Notice that the connection D^α also has frequency $< k - m + 10$. Therefore, from estimate (34) we have:

$$\| P_{k'}\left(U_{,<k-m}(C_{low}^\alpha + D^\alpha)\partial_\alpha \psi_k\right)\|_{N[I]} \lesssim_F 2^{-|k-k'|}\|(C_{low}^\alpha + D^\alpha)\partial_\alpha \psi_k\|_{N[I]}.$$

Furthermore, we claim the remainder estimate:

$$\| P_{k'}\left(U_{,<k-m}(C^\alpha - C_{low}^\alpha)\partial_\alpha \psi_k\right)\|_{N[I]} \lesssim_F 2^{-|k-k'|}2^{-Cm}\|\psi_k\|_{S[I]}.$$

Setting $R = U_{,<k-m}(C^\alpha - C_{low}^\alpha)$, this follows at once from Bernstein's inequality and the improved bounds:

$$\| R \|_{L_t^1(L_x^1)[I]} \lesssim_F 2^{-2k}2^{-Cm}, \quad \| P_j R \|_{L_t^1(L_x^1)[I]} \lesssim_F 2^{-2k}2^{-C(j-k)}2^{-Cm}.$$

These estimates are a consequence of the improved estimate (33), and the fact that at least one of the gauge factors in the $C^\alpha - C_{low}^\alpha$ integral is localized to $P_{>k-10}U_{,<k'}$.

Step 2B (*Estimation of the main term*). The purpose of this step is to prove the remaining estimate:

$$\| (C_{low}^\alpha + D^\alpha)\partial_\alpha \psi_k \|_{N[I]} \lesssim_F 2^{-\delta_1 m_0}\| \psi_k \|_{S[I]} + 2^{2m_0}\| \psi_k \|_{E[I]}.$$
(277)

We'll do this separately for each of the two terms on the left.

Step 2B.1 (*Estimation of D^α term*). The plan is to use Lemma 9.3. To do this we need to separate the connection D^α into two pieces, one with essentially matched frequencies and one with wide frequency separation. We write $D^\alpha = D_{(\delta)}^\alpha + \widetilde{D}^\alpha$, where

$$D_{(\delta)}^\alpha = \int_{k'<k-m} \left(S_{cb}^a(\phi) - S_{ca}^b(\phi)\right)_{[k'-10,k'+c\delta^2 m_0]} \partial^\alpha \phi_{k'}^c dk'$$
$$- \int_{k'<k-m} \partial^\alpha \left(S_{cb}^a(\phi) - S_{ca}^b(\phi)\right)_{[k'-c\delta^2 m_0, k'-10]} \phi_{k'}^c dk'.$$

Here $c \ll 1$ is an additional small constant. By a direct application of estimate (267) we have:

$$\| D_{(\delta)}^\alpha \partial_\alpha \psi_k \|_{N[I]} \lesssim_F 2^{Cc\delta^2 m_0}\left(2^{-\delta^2 m_0}\| \psi_k \|_{S[I]} + 2^{m_0}\| \psi_k \|_{E[I]}\right).$$
(278)

For c small enough in relation to C we have (277) for this term. The remainder term is in the range where the standard trilinear estimate (25) gives additional savings. A quick computation shows that for this term we in fact have:

$$\| \widetilde{D}^\alpha \partial_\alpha \psi_k \|_{N[I]} \lesssim_F 2^{-c\delta^3 m_0}\| \psi_k \|_{S[I]}.$$
(279)

The details of the dyadic summation are left to the reader.

Step 2B.2 (*Estimation of C^α_{low} term*). We follow the same strategy as in the previous argument. We split $C^\alpha_{low} = C^\alpha_{(\delta)} + \widetilde{C}^\alpha$, where

$$C^\alpha_{(\delta)} = \int_{-\infty}^{k-m} \left[B_{k'}, P_{<k-10} U^{\dagger}_{,<k'} \nabla^\alpha P_{[k'-c\delta^2 m_0, k'+c\delta^2 m_0]} P_{<k-10} U_{,<k'} \right] dk'.$$

The factors $P_{<k-10} U_{,<k'}$ are bounded on N via estimate (21), and can therefore be neglected. Again, by summing over the bound (267) with the help of (32) we have the analog of (278) (but this time with a factor of $2^{2Cc\delta^2 m_0}$ instead) for the contraction $C^\alpha_{(\delta)} \partial_\alpha \psi_k$. Similarly, we have the analog of (279) for the contraction $\widetilde{C}^\alpha \partial_\alpha \psi_k$, which also uses estimate (33).

Remark 9.4. The above process can also be used to show that if one already has $\| \psi_k \|_{S[I]}$ norm control, then one may conclude normalization bounds $\| \psi_k \|_{\mathcal{W}[I]}$ under the much less restrictive assumption that $m \geqslant 20$. In this case, one simply skips all of Step 1 above, and carry out Step 2 without introducing at all the terms $C^\alpha_{(\delta)}$ and $D^\alpha_{(\delta)}$.

Proof of Lemma 9.1. The estimate (260) follows from the energy bounds for ϕ combined with Bernstein's inequality. On the other hand (261) is a consequence of (160) and (157).

Estimate (262) follows from an application of (21)–(22), and then summation over the trilinear bound (25). The relevant detail is that one has the dyadic sum:

$$\sum_{k_2,k_3: \, k_i < k-m} 2^{-\delta(k-\min\{k_2,k_3\})} \lesssim 2^{-\delta m}.$$

Estimate (263) is a more elaborate use of such summations, but it is standard and left to the reader.

Consider now (264). For modulations at most comparable to the frequencies in the first factor we can replace the time derivative with a frequency factor and prove the estimate (264) by summing over (25). The relevant detail is that one has the dyadic sum:

$$\sum_{k_1,k_2: \, k_i < k-m} 2^{k_2} 2^{-\delta(k_2-k_1)_+} \lesssim 2^{-m} 2^k.$$

It remains to bound the expression when the first factor is at high modulation. In this case we take a product of the two bounds:

$$\| \nabla_{t,x} Q_{|\xi| \ll |\tau|} \phi^{(1)}_{<k-m} \|_{L^2(L^\infty)} \lesssim 2^{\frac{1}{2}(k-m)},$$
$$\| \partial^\alpha \phi^{(2)}_{<k-m} \partial_\alpha \psi_k \|_{L^2_x(L^2_x)} \lesssim 2^{\frac{1}{2}(k-m)} \| \psi_k \|_{S[I]},$$

the first of which follows from summation over (164) and the second of which follows from summation over (23). The estimate (265) follows from similar reasoning and is left to the reader. □

Proof of Lemma 9.2. The bound we seek is scale invariant, so without loss of generality we may assume that $k = 0$, and we may rotate and center the estimate so that $\omega = (1, 0)$ and $x_0 = (0, 0)$. In light of (161) we see that the main point of (266) is to be able to drop half space cutoffs of the form $\chi_{t<0}$ and $\chi_{t<x^1}$. The required boundedness of cutoffs with discontinuities across space-like hypersurfaces was already shown in (156). Therefore,

we seek an analog of (156) in the null case. Due to the frequency localization of both factors on the LHS of (266), it suffices to prove the following product estimate:

$$\| P_0 \left(\chi_{t<x^1} \cdot \psi_0 \right) \|_{DS} \lesssim \| \psi_0 \|_{DS}. \tag{280}$$

To save notation we will write $\chi = \chi_{t<x^1}$. Our point of view will be to observe that χ is a singular solution to the wave equation, so one can hope that (280) is in some sense a version of the standard product estimate (19). While this is true, the demonstration requires a bit of care because the *PW* norm of $P_k(\chi)$ does not gain the usual weight from L^1 summation over angles, even though its Fourier support is well localized in the angular variable. In fact, a quick calculation shows that:

$$\widetilde{\chi}(\tau, \xi) = \begin{cases} \frac{c_+}{\xi_1} \delta(\tau + |\xi|) \delta(\xi_2), & \xi_1 > 0; \\ \frac{c_-}{\xi_1} \delta(\tau - |\xi|) \delta(\xi_2), & \xi_1 < 0. \end{cases}$$

Here $c_\pm$ are appropriate constants depending on ones in the definition of the Fourier transform. The above formulas show that the $(+)$ wave portion of χ is a measure concentrated on the ray $(1, -1, 0)$, and opposite for the $(-)$ wave portion. We have the frequency localized *PW* type bound:

$$\| Q^\pm P_k \chi \|_{L^2_{t_{(1,0)}}(L^\infty_{x_{(1,0)}})} \lesssim 2^{-\frac{1}{2}k}. \tag{281}$$

Finally, note that due to the frequency localization in (280), we may replace the cutoff with $Q_{<10} P_{<10}(\chi)$. Also, if ϕ_0 is at high modulation 10 then $P_0 \left(\chi_{t<x^1} \cdot \psi_0 \right)$ is at comparable modulation, therefore (280) is immediate due to the L^∞ estimate for χ. We now proceed to prove (280) in a series of steps:

Step 1 (*Controlling the Strichartz norms*). Due to the boundedness of χ, we easily have:

$$\| P_0(P_{<10}\chi \cdot \psi_0) \|_{L^q_t(L^r_x)} \lesssim \| \psi_0 \|_{DS}.$$

Step 2 (*Controlling the $X^{s,b}$ norm*). Our first order of business is to bound the $X_\infty^{0,\frac{1}{2}}$ part of the norm (151). Freezing the outer modulation, our goal is to show that:

$$\| Q_j P_0 \left(P_{<10}\chi \cdot \psi_0 \right) \|_{L^2_t(L^2_x)} \lesssim 2^{-\frac{j}{2}} \| \psi_0 \|_{DS}. \tag{282}$$

We now split into subcases.

Step 2.A (*Output far from cone*). In this step we consider the contribution of output modulations $j > 20$. In this case, we may further localize the product to $Q_j P_0(P_{<10}\chi \cdot Q_{j+O(1)}\psi_0)$. Estimate (282) follows immediately from L^∞ control of χ.

Step 2.B (χ *at low frequency* ($\leq j - 10$)). In this case ϕ_0 must be at modulation 2^j, therefore we consider the contribution of the expression $Q_j P_0(P_{<j-10}\chi \cdot Q_{j+O(1)}\psi_0)$. Then (282) is immediate from the L^∞ control of χ.

Step 2.C (χ *at medium frequency, ψ at larger modulation*). In this case we consider the contribution of the term $Q_j P_0(P_{[j-10,10]}\chi \cdot Q_{\geqslant j-20}\psi_0)$. Again, only the boundedness of χ is used.

Step 2.D (χ *at medium frequency, ψ at low modulation*). The contribution of $Q_j P_0(P_{[j-10,10]}\chi \cdot Q_{<j-20}\psi_0)$ is considered here. This is the main term. Without loss

of generality, we may assume that we are in a (++) interaction, which we decompose into all possible angular sectors of cap size $|\kappa| \sim 2^{\frac{1}{2}j-10}$, respectively $|\kappa'| \sim 2^{\frac{1}{2}j}$:

$$Q_j P_0(Q^+ P_{[j-10,10]}\chi \cdot Q^+_{<j-20}\psi_0) = \sum_{j-10<k<10} \sum_{\kappa,\kappa'} Q_j P_{0,\kappa'}\left(Q^+ P_k(\chi) \cdot P_{0,\kappa} Q^+_{<j-20}\psi_0\right).$$

The main difficulty here is that we cannot really sum over k, because χ is only in an ℓ^∞ type Besov space. However, using Lemma 11 of [29] we see that the above sum is both essentially *diagonal* in κ, κ', and essentially *frequency disjoint* in its contribution of angles for each fixed k. Precisely, two sectors κ, κ' and a frequency k can provide nonzero output if and only if:

$$\mathrm{dist}(\kappa,\kappa') \sim 2^{\frac{j}{2}-10}, \quad \mathrm{dist}(\kappa,(-1,0)) \sim 2^{\frac{j-k}{2}}.$$

In particular the sector κ' centered at $(1,0)$ does not yield any output. Taking this into account we may bound:

$$\| Q_j P_0(Q^+ P_{j-10\leqslant\cdot<10}\chi \cdot Q^+_{<j-20}\psi_0) \|^2_{L^2_t(L^2_x)}$$

$$\lesssim \sum_{k=j-10}^{10} \sum_{\substack{\mathrm{dist}(\kappa,(-1,0))\sim 2^{\frac{j-k}{2}}\\ \mathrm{dist}(\kappa,\kappa')\sim 2^{\frac{j}{2}-10}}} \| P_{0,\kappa'}\left(Q^+ P_k(\chi) \cdot P_{0,\kappa}Q^+_{<j-20}\psi_0\right) \|^2_{L^2_t(L^2_x)}$$

$$\lesssim \sum_{k=j-10}^{10} \sum_{\substack{\mathrm{dist}(\kappa,(-1,0))\sim 2^{\frac{j-k}{2}}}} \| Q^+ P_k(\chi) \|^2_{L^2_{t_{(1,0)}}(L^\infty_{x_{(1,0)}})}\| P_{0,\kappa}Q^+_{<j-20}\psi_0 \|^2_{L^\infty_{t_{(1,0)}}(L^2_{x_{(1,0)}})}$$

$$\lesssim \sum_{k=j-10}^{10} \sum_{\substack{\mathrm{dist}(\kappa,(-1,0))\sim 2^{\frac{j-k}{2}}}} 2^{-k} \sup_{\mathrm{dist}(\omega,\kappa)\sim 2^{\frac{j-k}{2}}} \| P_{0,\kappa}Q^+_{<j-20}\psi_0 \|^2_{L^\infty_{t_\omega}(L^2_{x_\omega})}$$

$$\lesssim 2^{-j} \sum_\kappa \| P_{0,\kappa}Q^+_{<j-20}\psi_0 \|^2_{S[0,\kappa]}.$$

From the definition of the S norm (147), this suffices to prove (282).

Step 3 (*Controlling the square sum of $S[0,\kappa]$ norms*). Again freezing $j < -10$ we need to demonstrate that:

$$\sup_\pm \sum_\kappa \| Q^\pm_{<j}P_{0,\pm\kappa}(P_{<10}\chi \cdot \psi_0) \|^2_{S[0,\kappa]} \lesssim \| \psi_0 \|^2_S, \tag{283}$$

where angular sector size is $|\kappa| \sim 2^{\frac{1}{2}j}$. The subcases repeat Case 2 above with little difference, and are mostly left to the reader:

Step 3.A (χ *at low frequency*). This is the contribution of the expression

$$Q_{<j}P_0(P_{<j-10}\chi \cdot Q_{<j+O(1)}\psi_0).$$

In this case (283) is immediate from the L^∞ control of χ.

Step 3.B (χ *at medium frequency, ψ at larger modulation*). As before, this is the term $Q_{<j}P_0(P_{j-10,10]}\chi \cdot Q_{\geqslant j-20}\psi_0)$ for which we have a stronger L^2 bound:

$$\| Q_{<j}P_0(P_{[j-10,10]}\chi \cdot Q_{\geqslant j-20}\psi_0) \|_{L^2_t L^2_x} \lesssim 2^{-\frac{1}{2}j}\| \psi_0 \|_S.$$

Step 3.C (χ *at medium frequency,* ψ *at low modulation*). Here we consider the contribution of $Q_{<j}P_0(P_{[j-10,10]}\chi \cdot Q_{<j-20}\psi_0)$. This is again the main term. Without loss of generality we may assume that we are in a (++) interaction in terms of output and ψ_0 modulation (in particular, from the estimate in Step 2 above we may dispense with the case (+) output and (−) input from ψ_0), and we again use Lemma 11 of [29] to decompose into a diagonal sum over caps of size $|\kappa| \sim 2^{\frac{1}{2}j-C}$, respectively $|\kappa'| \sim 2^{\frac{1}{2}j}$:

$$Q_{<j}^+ P_0(Q^+ P_{[j-C,C]}\chi \cdot Q_{<j-C}^+ \psi_0) = \sum_{\text{dist}(\kappa,\kappa') \sim 2^{\frac{j}{2}}} Q_{<j}^+ P_{0,\kappa'}\left(P_{[j-C,C]}(\chi) \cdot P_{0,\kappa} Q_{<j-C}^+ \psi_0\right).$$

Notice that we do not need to frequency localize the factor $P_{[j-10,10]}\chi$ to obtain this diagonally, which is a good thing because the rougher bounds on the output modulation and that of ψ_0 do not win us disjoint angular contributions in the k-sum of $P_k\chi$. Plugging the above decomposition into the LHS side of estimate, (283) the RHS bound follows at once from L^∞ control of χ. $\quad\square$

Proof of Lemma 9.3. We begin by extending ψ_k via the universal extension in Proposition 5.5 in such a way that we simultaneously maintain the $\underline{E}$ and S norm control. The functions $\phi_{k_i}^{(i)}$ are similarly extended. Thus, it suffices to prove the bound on all of space-time.

The constant C will be fixed in the proof in just a moment. Let $m \geqslant 0$ be any fixed integer. Without loss of generality, we may assume that $k = 0$. Furthermore, we may also assume that $|k_1 - k_2| < \delta m$, for otherwise the estimate follows immediately from an application of the standard trilinear bound (25), and taking $C > 1$ on the RHS of (267). The proof will be accomplished in a series of steps:

Step 1 (*Reduction to a bilinear estimate*). In this step we consider the contribution of $\phi_{k_1}^{(1)} Q_{<k_2-\delta m}\left(\partial_\alpha \phi_{k_2}^{(2)} \partial^\alpha \psi_0\right)$. By an application of the estimates (163), (24), and (170) we easily have that for $j < k_2 - \delta m$ (which also implies $j < k_1$) and $\delta \ll 1$ sufficiently small:

$$\| \phi_{k_1}^{(1)} Q_j \left(\partial_\alpha \phi_{k_2}^{(2)} \partial^\alpha \psi_0\right) \|_N \lesssim 2^{-\delta(k_1-j)} \| \phi_{k_1}^{(1)} \|_S \| \phi_{k_2}^{(2)} \|_S \cdot \| \psi_0 \|_S.$$

Summing over all $j < k_2 - \delta m$ we directly have (267) for this component. It remains to estimate the contribution of $\phi_{k_1}^{(1)} Q_{>k_2-\delta m}\left(\partial_\alpha \phi_{k_2}^{(2)} \partial^\alpha \psi_0\right)$. We peel off the factor $\phi_{k_1}^{(1)}$ from the trilinear estimate via the bound (21). It remains to prove the bilinear bound

$$\|Q_{>k_2-\delta m}\left(\partial_\alpha \phi_{k_2}^{(2)} \partial^\alpha \psi_0\right)\|_N \lesssim_F \left(2^{-\delta^2 m}\| \psi_k \|_{S[I]} + 2^m \| \psi_k \|_{\underline{E}[I]}\right). \qquad (284)$$

Step 2 (ψ_0 *is far from the cone*). In this step we consider the contribution of $Q_{\geqslant k_2-\delta m}\left(\partial_\alpha \phi_{k_2}^{(2)} Q_{\geqslant 0}\partial^\alpha \psi_0\right)$. We will prove that the remaining null-form is an $X_1^{0,-\frac{1}{2}}$ atom. In the present case, we freeze the output modulation j and then estimate:

$$\| Q_j \left(\partial_\alpha \phi_{k_2}^{(2)} Q_{\geqslant 0}\partial^\alpha \psi_0\right) \|_{L_t^2(L_x^2)} \lesssim \| \nabla_{t,x}\phi_{k_2}^{(2)} \|_{L_t^\infty(L_x^\infty)} \cdot \| Q_{\geqslant 0}\nabla_{t,x}\psi_0 \|_{L_t^2(L_x^2)},$$

$$\lesssim 2^{k_2}\| \phi_{k_2}^{(2)} \|_S \cdot \| \psi_0 \|_S.$$

Multiplying both sides of this bound by $2^{-\frac{j}{2}}$ and then summing over all dyadic $j \geqslant k_2 - \delta m$ we arrive at:

$$\| Q_{\geqslant k_2 - \delta m} \left(\partial_\alpha \phi_{k_2}^{(2)} Q_{\geqslant 0} \partial^\alpha \psi_0 \right) \|_N \lesssim 2^{\frac{k_2}{2} + \delta m} \| \phi_{k_2}^{(2)} \|_S \cdot \| \psi_0 \|_S,$$

which suffices due to the condition $k_2 < -m$.

Step 3 ($\phi_{k_2}^{(2)}$ is far from the cone). In this step we consider the contribution of $Q_{\geqslant k_2 - \delta m} \left(Q_{>k_2 - 8\delta m} \partial_\alpha \phi_{k_2}^{(2)} Q_{<0} \partial^\alpha \psi_0 \right)$. In this case, we again freeze the output modulation j and proceed to bound:

$$\| Q_j \left(Q_{>k_2 - 8\delta m} \partial_\alpha \phi_{k_2}^{(2)} Q_{<0} \partial^\alpha \psi_0 \right) \|_{L_t^2(L_x^2)}$$
$$\lesssim \| Q_{>k_2 - 8\delta m} \nabla_{t,x} \phi_{k_2}^{(2)} \|_{L_t^2(L_x^\infty)} \cdot \| \nabla_{t,x} \psi_0 \|_{L_t^\infty(L_x^2)}. \tag{285}$$

By summing over all $j > k_2 - 8\delta m$ in estimate (164) we have that:

$$\| Q_{>k_2 - 8\delta m} \nabla_{t,x} \phi_{k_2}^{(2)} \|_{L_t^2(L_x^\infty)} \lesssim 2^{\frac{1}{2}k_2 + 4\delta m} \| \phi_{k_2}^{(2)} \|_S.$$

Substituting this into the RHS of (285), multiplying the result by $2^{-\frac{1}{2}j}$, and then summing over all $j \geqslant k_2 - \delta m$ we have the estimate:

$$\| Q_{\geqslant k_2 - \delta m} \left(Q_{>k_2 - 8\delta m} \partial_\alpha \phi_{k_2}^{(2)} Q_{<0} \partial^\alpha \psi_0 \right) \|_N \lesssim 2^{5\delta m} \| \phi_{k_2}^{(2)} \|_S \cdot \| \psi_0 \|_E.$$

Step 4 (The core contribution). In this step we consider the contribution of the expression $Q_{\geqslant k_2 - \delta m} \left(Q_{<k_2 - 8\delta m} \partial_\alpha \phi_{k_2}^{(2)} Q_{<0} \partial^\alpha \psi_0 \right)$. This is the main case, and requires a decomposition into angular sectors of cap size $|\kappa| \sim 2^{-4\delta m}$. Without loss of generality we may assume we are in the $(++)$ configuration. The other cases $(--)$, $(+-)$, and $(-+)$ are the same with only minor modifications and are therefore left to the reader. We break the entire contribution into a $Q_{\geqslant k_2 - \delta m}$ localized sum of two principle terms T_1 and T_2, where

$$T_1 = \sum_{\kappa \in K_l} Q_{<k_2 - 8\delta m}^+ P_{k_2, \kappa} \partial_\alpha \phi_{k_2}^{(2)} \cdot (I - P_{0, 2\kappa}) Q_{<0}^+ \partial^\alpha \psi_0,$$

$$T_2 = \sum_{\kappa \in K_l} Q_{<k_2 - 8\delta m}^+ P_{k_2, \kappa} \partial_\alpha \phi_{k_2}^{(2)} \cdot P_{0, 2\kappa} Q_{<0}^+ \partial^\alpha \psi_0.$$

To help state the estimates, we introduce the following weaker version of the *NFA** portion of the $S[k, \kappa]$ norm from line (150):

$$\| \psi \|_{\widetilde{S}_k} := \sup_{\substack{l > 10 \\ \kappa \in K_l}} \| P_k \psi \|_{\widetilde{S}[k, \kappa]},$$

where

$$\| \psi \|_{\widetilde{S}[k, \kappa]} := \sup_{\pm} \sup_{\omega \in \frac{1}{2}\kappa} 2^k |\kappa| \cdot \| Q_{<k}^\pm (I - P_{k, \pm\kappa}) \psi \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)}.$$

Notice that we do not use the more eccentric Q_j multipliers for $j < k - 10$ in this definition, and there is no square-summing over angles. The reason this notation is useful

is that we have the relation: $\| \psi_k \|_{\widetilde{S}} \lesssim \| \psi_k \|_E$. This is shown through an application of the estimate:

$$\| Q_{<k}^\pm (I - P_{k,\pm\kappa})\psi_k \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)} \lesssim 2^{-k}|\kappa|^{-1} \cdot \| \nabla_{t,x}\psi_k \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)}. \tag{286}$$

Such an inequality may be proved by decomposing the multiplier $Q_{<k}^\pm(I - P_{k,\pm\kappa})$ into a dyadic sum of angular sectors of increasing size and spread from $\pm\kappa$. Without loss of generality, we may assume we are in the "+" case, and we decompose $Q_{<k}^+(I - P_{k,\kappa}) = \sum_{1 > 2^j > |\kappa|} Q_{<k}^+ P_{k,\kappa_j}$, where each sector size is $|\kappa_j| \sim 2^j$ with distance $\text{dist}(\kappa, \kappa_j) \sim 2^j$. For each of these sectors we use the uniform multiplier bounds:

$$\| Q_{<k}^+ P_{k,\kappa_j}\psi_k \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)} \lesssim 2^{-k}|\kappa_j|^{-1} \cdot \| \nabla_{t,x}\psi_k \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)},$$

which is an easy consequence of the fact that the kernels associated to the operators:

$$\mathcal{L} = 2^k|\kappa_j|\nabla_{t,x}^{-1} Q_{<k}^+ P_{k,\kappa_j},$$

are uniformly in $L_t^1(L_x^1)$. The estimate (286) now follows from simply summing over this last bound overall all dyadic $1 < |\kappa_j|^{-1} < |\kappa|^{-1}$.

Returning to the main thread, we first bound the term T_1 above. In this case, we are going to lose a large constant because the sum is not well localized in the second factor and therefore we cannot use orthogonality with respect to κ. Furthermore, we will not bother to gain anything from the null-structure, because the frequency localization of this term eliminates parallel interactions. To compensate for the large number of non-orthogonal sectors, we may use the $\widetilde{S}$ norm for the second factor. Using the product estimate (167) we may bound:

$$\| Q_j T_1 \|_{L_t^2(L_x^2)} \lesssim \sum_{\kappa \in K_l} |\kappa|^{\frac{1}{2}} 2^{-\frac{1}{2}k_2} \| Q_{<k_2-8\delta m}^+ P_{k,\kappa} \partial_\alpha \phi_{k_2}^{(2)} \|_{S[k_2,\kappa]}$$
$$\cdot \sup_{\omega \in \kappa} \| (I - P_{0,2\kappa}) Q_{<0}^+ \partial^\alpha \psi_0 \|_{L_{t_\omega}^\infty(L_{x_\omega}^2)}$$
$$\lesssim 2^{\frac{1}{2}k_2} \| \phi_{k_2}^{(2)} \|_S \cdot 2^{4\delta m} \| \psi_0 \|_{\widetilde{S}}.$$

Multiplying both sides of this last estimate by the factor $2^{-\frac{1}{2}j}$ and then summing over all $j > k_2 - \delta m$ we have:

$$\| Q_{>k_2-\delta m} T_1 \|_N \lesssim 2^{5\delta m} \| \phi_{k_2}^{(2)} \|_S \cdot \| \psi_0 \|_{\widetilde{S}},$$

which is sufficient.

Our final task here is to bound the term $Q_{>k_2-\delta m} T_2$ in the space $X_1^{0,\frac{1}{2}}$. Notice that because of the angular and (++) localization, as well as the fact that $j > k_2 - \delta m$, for each $Q_j T_2$ we may freely insert the multiplier $Q_{>j-10}$ in front of the second factor, because the complement vanishes (see Lemma 11 of [29]). In this case the resulting sum is both diagonal and orthogonal in κ, so freezing $Q_j T_2$ we have with the aid of Bernstein's inequality (9) the estimate:

$$\| Q_j T_2 \|_{L_t^2(L_x^2)}^2 \lesssim \sum_{\kappa \in K_l} \| Q_{<k_2-8\delta m}^+ P_{k,\kappa} \partial_\alpha \phi_{k_2}^{(2)} \|_{L_t^\infty(L_x^\infty)}^2 \cdot \| P_{0,2\kappa} Q_{[j-10,0]}^+ \partial^\alpha \psi_0 \|_{L_t^2(L_x^2)}^2$$
$$\lesssim \left(2^{k_2-2\delta m} \| \phi_{k_2}^{(2)} \|_S \cdot 2^{-\frac{1}{2}j} \| \psi_0 \|_S \right)^2.$$

Multiplying the root of this inequality by the factor $2^{-\frac{1}{2}j}$ and then summing over all $j > k_2 - \delta m$ we finally have:

$$\| Q_{>k_2-\delta m} T_2 \|_N \lesssim 2^{-\delta m} \| \phi_{k_2}^2 \|_S \cdot \| \psi_0 \|_S.$$

This concludes our proof of estimate (267). $\square$

10. Structure of Finite S Norm Wave-Maps and Energy Dispersion

In this section we prove Proposition 3.9. There is almost nothing to do for (63). The $\underline{X}$ bound follows from the reduced version of (51) in Remark 3.8, while the $\underline{E}$ bound follows from energy estimates on null surfaces.

10.1. Renormalization. Here we establish the renormalization bound (64). Our starting point is the construction of the renormalization matrix U in Proposition 3.1. The frequency localized wave-map equation for ϕ is given by:

$$\Box \phi_k = -P_k \left(\mathcal{S}(\phi) \partial^\alpha \phi \partial_\alpha \phi \right). \tag{287}$$

For each index m the RHS of this expression can be written in terms of the trilinear from T_k^m from line (53) as follows:

$$P_k \left(\mathcal{S}(\phi) \partial^\alpha \phi \partial_\alpha \phi \right) = 2\mathcal{S}(\phi)_{<k-m} \partial^\alpha_{<k-m} \phi \partial_\alpha \phi_k + T_{1;k}^m \left(\mathcal{S}(\phi), \partial^\alpha \phi, \partial_\alpha \phi \right).$$

Using the identity (106) we have:

$$P_k \left(\mathcal{S}^\dagger(\phi) \partial^\alpha \phi \partial_\alpha \phi_{<k-m+2} \right) = 0.$$

Thus, we may further write:

$$\mathcal{S}(\phi)_{<k-m} \partial^\alpha_{<k-m} \phi \partial_\alpha \phi_k = \left(\mathcal{S}(\phi)_{<k-m} - \mathcal{S}(\phi)^\dagger_{<k-m} \right) \partial^\alpha \phi_{<k-m} \partial_\alpha \phi_k + T_{2;k}^m,$$

where $T_{2;k}^m$ is obtained by applying the decomposition (53) to the previous line. Therefore, we have written the original frequency localized wave-map equation in the form:

$$\Box \phi_k = -2 A^\alpha_{<k-m} \partial_\alpha \phi_k + \sum_i T_{i;k}^m, \tag{288}$$

where the $T_{i;k}^m$ are trilinear forms as on line (53) with $O(m)$ gap indices. By an application of estimates (42) and (61) with $m = 20$ we have the bound:

$$\| \phi_k \|_{\mathcal{W}[I]} \lesssim_F c_k,$$

where $\{c_k\}$ is some $S[I]$ frequency envelope for ϕ_k. This proves (64).

10.2. Partial fungibility of the S norm. Here we prove that there is always a decomposition of intervals $I = \cup_{i,l}^{K(F)} I_{il}$, where $K(F)$ is some polynomial in the $S[I]$ norm of ϕ, and where (65) holds in each subinterval. Our starting point is the series of frequency localized equations (288). For a fixed ϕ_k we use (288) with $m = 20$. As in the previous section, we can find a renormalization $w_{,k} = U_{,<k-20}\phi_k$ on all of I such that:

$$\| P_{k'}\Box w_{,k} \|_{N[I]} \lesssim_F 2^{-|k-k'|}c_k. \tag{289}$$

Let $\eta \ll 1$ to be chosen later. By the fungibility property (159) (and continuity) there exists a polynomial K_1 in $F\eta^{-1}$ such that $I = \cup_i^{K_1} I_i$ such that:

$$\| P_{k'}\Box w_{,k} \|_{N[I_i]} \leqslant 2^{-\frac{1}{2}|k-k'|}\eta c_k^i,$$

where $\{c_k^i\}$ are now some unit normalized frequency envelope which may depend on the interval I_i. We label each time interval as $I_i = [t_i, t_{i+1}]$, and on each of these time slabs we write $w_{,k} = w_{,k}^{free} + w_{,k}^{source}$, where $w_{,k}^{free}$ is a free wave with data $w_{,k}[t_i]$. By the previous line and the energy estimate (18) we have on I_i the bound:

$$\| P_{k'}w_{,k}^{source} \|_{S[I_i]} \lesssim 2^{-\frac{1}{2}|k-k'|}\eta c_k^i. \tag{290}$$

Consequently, for the corresponding part $U_{,<k-20}^{\dagger}w_{,k}^{source}$ of ϕ_k we obtain:

$$\| U_{,<k-20}^{\dagger}w_{,k}^{source} \|_{S[I_i]} \lesssim_F \eta c_k^i.$$

By choosing η as the reciprocal of an appropriate polynomial in F, we have:

$$\| U_{,<k-20}^{\dagger}w_{,k}^{source} \|_{S[I_i]} \leqslant c_k^i. \tag{291}$$

It remains to bound the free wave contribution $U_{,<k-20}^{\dagger}w_{,k}^{free}$ on each of the intervals I_i, or on some further subdivision thereof.

Unfortunately we do not directly know that $U_{,<k-20}^{\dagger}$ is manageable on I_i. However, we do have from estimate (36) and the energy bound for ϕ that:

$$\| P_{k'}w_{,k}^{free}[t_i] \|_{\dot{H}\times L^2} \lesssim_E 2^{-|k-k'|}c_k,$$

uniformly with respect to i where we may choose the unit frequency envelope $\{c_k\}$ to be the same as on line (289) above. In particular, we have the uniform control:

$$\| P_{k'}w_{,k}^{free} \|_{S[I_i]} \lesssim_E 2^{-|k-k'|}c_k. \tag{292}$$

Now we turn our attention to the $U_{,<k-20}$'s. Given a large parameter m to be chosen later, we consider the sections $P_{[j-m,j+m]}U_{,<k-20}$ of $U_{,<k-20}$. Recall that from Remark 3.3 each $U_{,<k-20}$ is built up out of the same connection (31), and therefore the bounds (32) for each $U_{,<k-20}$ may be taken in terms of the *same* frequency envelope. Hence, except for a polynomial in mF number of indices j we already have:

$$\sup_k \| P_{[j-m,j+m]}U_{,<k-20} \|_{S[I]} \leqslant 1. \tag{293}$$

Such indices j are called "good j's"; the remainder (of which we have at most a polynomial in mF) are called "bad j's". We also introduce the corresponding parts of $U^\dagger_{,<k-20}w^{free}_{,k}$:

$$\phi_k(j) = P_{[j-m,j+m]}U^\dagger_{,<k-20} \cdot w^{free}_{,k}.$$

The goal of the argument is now to choose a polynomial in mF collection of subintervals I_{il}, partitioning the I_i, such that on each there is uniform control over all k and j:

$$\| P_{k'}\phi_k(j) \|_{S[I_{il}]} \lesssim_E 2^{-\frac{1}{2}\delta|k-k'|}c^{jl}_k, \tag{294}$$

for some additional set of unit normalized frequency envelopes $\{c^{jl}_k\}$. For good j's this is straightforward in view of (292) and (293). Since there are $\lesssim mF$ bad j's, it suffices to consider such a fixed bad j. The equation for each fixed $\phi_k(j)$ is:

$$\Box\phi_k(j) = P_{[j-m,j+m]}\Box U^\dagger_{,<k-20} \cdot w^{free}_{,k} + 2\partial^\alpha P_{[j-m,j+m]}U^\dagger_{,<k-20}\partial_\alpha w^{free}_{,k}.$$

Therefore, by a direct application of the estimates (292), (32), (24), and (33)–(35) we have on all of I_i the bound:

$$\| \Box P_{k'}\phi_k(j) \|_{N[I_i]} \lesssim_{mF} 2^{-\delta|k-k'|}c_k,$$

and from the energy norm control giving (292) and estimate (36) we also have the uniform energy control:

$$\| P_{k'}\phi_k(j)[t] \|_{\dot{H}^1\times L^2} \lesssim_E 2^{-|k-k'|}c_k.$$

Thus, by again using the property (159) we obtain the desired partition $\{I_{il}\}$ of I, with estimate (294) uniformly, at a cost of at most $\lesssim_{mF} 1$ subdivisions.

To conclude the proof we need to estimate $U^\dagger_{,<k-20}w^{free}_{,k}$ on each subinterval $J = I_{il}$, which is now fixed with the property that (294) holds. We split $U^\dagger_{,<k-20}$ into:

$$U^\dagger_{,<k-20} = P_{<k-m}U^\dagger_{,<k-20} + P_{[k-m,k+m]}U^\dagger_{,<k-20} + P_{>k+m}U^\dagger_{,<k-20}.$$

For the high frequency part we use (33) in conjunction with the product bounds (19)–(20) to obtain:

$$\| P_{k'}(P_{>k+m}U^\dagger_{,<k-20}w^{free}_{,k}) \|_{S[I]} \lesssim_F 2^{-Cm}2^{-|k'-k|}c_k,$$

which suffices provided m is large enough, $m \sim \ln F$. For the medium frequency part we can use directly (294) with $j = k$. Thus, we are reduced to providing good $S[J]$ norm bounds for the quantities $P_{<k-m}U^\dagger_{,<k-20} \cdot w^{free}_{,k}$ which are localized at frequency 2^k. We do this in a series of steps depending on what component of the $S[J]$ norm is being considering:

Step 1 (Energy and Strichartz norm control). For any of the Strichartz norms we immediately have from Leibniz's rule, estimates (292) and (36) the bound:

$$\| \nabla_{t,x} \left(P_{<k-m}U^\dagger_{,<k-10} \cdot w^{free}_{,k} \right) \|_{D\underline{S}[J]} \lesssim_E c_k,$$

which is sufficient.

Step 2 ($X_\infty^{0,\frac{1}{2}}$ *norm control*). Fix a modulation Q_j. Without loss of generality we will assume that $j < k$, as the complimentary region is easier to treat using the high modulation bounds in (32) and (33). We decompose as follows:

$$Q_j \left(P_{<k-m} U_{<k-20}^\dagger \cdot w_{,k}^{free} \right) = \begin{cases} Q_j \phi_k(j) + Q_j R_{,k}, & j < k - 2m; \\ Q_j \phi_k(k-2m) + Q_j R_{,k}, & j > k - 2m, \end{cases} \quad (295)$$

where

$$R_{,k} = \begin{cases} P_{<j-m} U_{<k-20}^\dagger \cdot w_{,k}^{free} + P_{[j+m,k-m]} U_{<k-20}^\dagger \cdot w_{,k}^{free}, & j < k - 2m; \\ P_{<k-3m} U_{<k-20}^\dagger \cdot w_{,k}^{free}, & j > k - 2m. \end{cases}$$

By estimate (294) we already control the first terms on the RHS of (295), so we only need to bound the contribution of $Q_j P_k R_{,k}$. This is given by the following analog of Lemma 8.3:

Lemma 10.1. *Let $j < k - 10$ and $m > 10$ be an integer. Then the following estimates hold for test functions $u = u_{<k-10}$ and ϕ_k:*

$$\|Q_j(u_{<j-m}\phi_k)\|_{X_\infty^{1,\frac{1}{2}}} + \|Q_{<j}(u_{<j-m}\phi_k)\|_{S[k;j]} \lesssim \left(\|u\|_{L_t^\infty(L_x^\infty)} + 2^{-\delta m}\|u\|_{\underline{X}} \right) \|\phi_k\|_S,$$

$$\|Q_j(P_{>j+m}u_{<k-10}\phi_k)\|_{X_\infty^{1,\frac{1}{2}}} + \|Q_{<j}(P_{>j+m}u_{<k-10}\phi_k)\|_{S[k;j]} \lesssim 2^{-\delta m}\|u\|_S\|\phi_k\|_S.$$

Proof. The proof of the first bound is immediate from (168) and the product bounds (19)–(20) in conjunction with the following easy estimate for very high modulations:

$$\| Q_{>j-\frac{1}{2}m} u_{<j-m} \|_S \lesssim 2^{-\frac{1}{4}m} \| u \|_{\underline{X}}.$$

The second estimate is just a summed version of (169) which also incorporates (162). □

Using a combination of the estimates in this last lemma, and (292), we have:

$$\| Q_j P_k R_{,k} \|_{X_\infty^{0,\frac{1}{2}}} \lesssim_E \left(1 + 2^{-\delta m} Q_2(F) \right) c_k,$$

which suffices.

Step 3 ($S[k; j]$ *norm control*). This is immediate from the decomposition (295), the estimate (294), and Lemma 10.1.

10.3. The role of the energy dispersion. By applying estimate (64) and then using (51) on Eq. (287) we have (66).

Suppose now that $\{c_k\}$ is a frequency envelope for the initial data of ϕ in $\dot{H} \times L^2$. Then by the seed bounds (68) we have full control:

$$\| \phi \|_{S_c[J]} \leqslant K_1(F), \quad (296)$$

on some sufficiently small subinterval $J \subseteq I$. Here K_1 is a universal polynomial that will be chosen in a moment. The goal now is to bootstrap this control and show that if:

$$\| \phi \|_{S_c[J]} \leqslant 2K_1(F), \quad (297)$$

then we have (296). By Proposition 3.10 we may continue and finally close this last estimate on all of I.

By applying estimates (64) and (66) to (297), we have:

$$\| \phi \|_{\mathcal{W}_c[J]} \leqslant K_2(F)K_1(F), \quad \| \phi \|_{\underline{X}_c[J]} \leqslant \epsilon^{\delta_1} K_2(F)K_1(F),$$

for a universal polynomial K_2.

Next, choose the gap $m \lesssim \ln(F)$ in Eq. (288) in a way that is consistent with the assumptions of Proposition 3.2, and apply estimate (41) to (288), while using (55) via the last two bounds. This gives:

$$\| \phi_k \|_{S[I]} \leqslant K_3(F) \left(1 + \epsilon^{\delta_1^2} K_2(F)K_1(F)\right) c_k.$$

The proof is concluded by choosing $K_1 = 2K_3$ and assuming ϵ is sufficiently small.

11. Initial Data Truncation

Here we prove that for each initial data set with small energy dispersion we can continuously regularize it. In a sufficiently small tubular neighborhood $V(\mathcal{M})$ of the surface $\mathcal{M} \subset \mathbb{R}^N$ we introduce a projection operator:

$$\Pi : V(\mathcal{M}) \rightarrow \mathcal{M}.$$

This also induces a projection operator on the tangent bundle:

$$\Pi : TV(\mathcal{M}) \rightarrow T\mathcal{M},$$

which is a product of Π in $\mathbb{R}^N$ and Euclidean linear orthogonal projection onto each fiber in the second factor. Given an initial data set:

$$\phi[0] = (\phi_0, \phi_1) : \mathbb{R}^2 \rightarrow T\mathcal{M},$$

which belongs to $\dot{H}^1 \times L^2$, we regularize it as follows:

$$\phi_{,<k}[0] = \Pi(P_{<k}\phi[0]).$$

The following result asserts that if $\phi[0]$ has small energy dispersion then its regularizations are well defined, and stay close to the corresponding Littlewood-Paley projections:

Proposition 11.1. *For each $E > 0$ there exists $\epsilon_0 > 0$ so that for each initial data set $\phi[0]$ for (1) with energy E and energy dispersion $\epsilon \leqslant \epsilon_0$ and $k, k_* \in \mathbb{Z}$ we have:*

$$\|P_k(P_{<k_*}\phi[0] - \phi_{,<k_*}[0])\|_{\dot{H}^1 \times L^2} \lesssim_E \min\{\epsilon|\ln \epsilon|^2, 2^{-|k-k_*|}\}. \tag{298}$$

Proof. By rescaling we assume that $k_* = 0$. We begin with two simple Moser type estimates which we will repeatedly use in the sequel. Precisely, for each smooth and bounded function G with bounded derivatives we have:

$$\| \nabla_x^J G(P_{<k}\phi_0) \|_{L_x^\infty} \lesssim_E 2^{|J|k}, \tag{299}$$

and

$$\| \nabla_x^J G(P_{<k}\phi_0) \|_{L_x^2} \lesssim_E 2^{(|J|-1)k}, \quad |J| \geqslant 1, \tag{300}$$

which are easily proved using the chain rule and Bernstein's inequality.

We first show that if ϵ is small enough then the projection $\Pi P_{<0}\phi[0]$ is well defined:

Lemma 11.2. *Under the assumptions of Proposition 11.1 we have:*

$$dist\,(P_{<0}\phi_0, \mathcal{M}) \lesssim_E \epsilon|\log\epsilon|.$$

Proof. By translation invariance, it suffices to show that:

$$I = \int_{|x|\leqslant 1} |P_{<0}\phi_0(0) - \phi_0(x)|dx \lesssim_E \epsilon|\log\epsilon|. \tag{301}$$

We use a positive parameter m and a Littlewood-Paley decomposition to estimate I as follows:

$$I \lesssim \|\,\nabla_x P_{<-m}\phi_0\,\|_{L_x^\infty} + \|\,P_{[-m,m]}\phi_0\,\|_{L_x^\infty} + \|\,P_{>m}\phi_0\,\|_{L_x^2}.$$

Using Sobolev embeddings for the first term, energy dispersion for the second, and the $\dot{H}^1$ norm for the third we obtain:

$$I \lesssim 2^{-m}E + m\epsilon + 2^{-m}E.$$

Then (301) is obtained by choosing $2^m = \epsilon^{-1}$. $\quad\square$

To continue the proof of the proposition, we remark that Π can be expressed as:

$$\Pi(\phi^{(1)}, \phi^{(2)}) = \left(G(\phi^{(1)}), H(\phi^{(1)})\phi^{(2)}\right), \tag{302}$$

where G is some smooth extension of Π to all of $\mathbb{R}^N$, and H is some extension of the fiber projection composed with G. Note that both G and H may be chosen as bounded functions with bounded derivatives. We separately estimate the high frequencies, middle frequencies and low frequencies of the difference $P_{<k_*}\phi[0] - \phi_{,<k_*}[0]$.

Step 1 (*High frequency bounds, the contribution of $k > 0$*). For the high frequencies we do not use the fact at all that $\phi[0]$ takes values in $T\mathcal{M}$. Instead, we use (299) to directly estimate:

$$\|\,P_k G(P_{<0}\phi_0)\,\|_{L_x^2} \lesssim_E 2^{-(C+1)k},$$

where C is a large integer. Similarly we have:

$$\|\,P_k\,(H(P_{<0}\phi_0)P_{<0}\phi_1)\,\|_{L_x^2} \lesssim_E 2^{-Ck}.$$

Thus we obtain:

$$\|\,P_k(P_{<0}\phi[0] - \phi_{,<0}[0])\,\|_{\dot{H}^1\times L^2} \lesssim_E 2^{-Ck}. \tag{303}$$

Step 2 (*Low frequencies bounds, the contribution of $k < 0$*). Here we take advantage of the identity $\Pi\phi[0] = \phi[0]$. Then we can write:

$$P_k\left(P_{<0}\phi[0] - \phi_{,<0}[0]\right) = P_k\left(\Pi\phi[0] - \Pi(P_{<0}\phi[0])\right) := \psi[0].$$

To estimate the last difference we use an integral expansion as follows:

$$\begin{aligned}
\psi_0 &= P_k \int_0^\infty \frac{d}{dk_1}G(P_{<k_1}\phi_0)dk_1, \\
&= P_k \int_0^\infty G'(P_{<k_1}\phi_0)P_{k_1}\phi_0 dk_1, \\
&= P_k \int_0^\infty P_{>k_1-10}G'(P_{<k_1}\phi_0) \cdot P_{k_1}\phi_0 dk_1.
\end{aligned}$$

Next, we use Bernstein's inequality and (300) to estimate:

$$
\begin{aligned}
\| P_k \left(P_{>k_1-10} G'(P_{<k_1}\phi_0) \cdot P_{k_1}\phi_0 \right) \|_{L_x^2} &\lesssim 2^k \| P_{>k_1-10} G'(P_{<k_1}\phi_0) \cdot P_{k_1}\phi_0 \|_{L_x^1} \\
&\lesssim 2^k \| P_{>k_1-10} G'(P_{<k_1}\phi_0) \|_{L_x^2} \| P_{k_1}\phi_0 \|_{L_x^2} \\
&\lesssim_E 2^{k-2k_1}.
\end{aligned}
\tag{304}
$$

Hence after integration with respect to $k_1 \geqslant 0$ we obtain:

$$
\| \psi_0 \|_{L_x^2} \lesssim_E 2^k.
$$

A similar computation shows that:

$$
\psi_1 = P_k \int_0^\infty H'(P_{<k_1}\phi_0) P_{k_1}\phi_0 \cdot P_{<k_1}\phi_1 dk_1 + P_k \int_0^\infty H(P_{<k_1}\phi_0) P_{k_1}\phi_1 dk_1.
$$

Then proceeding as above, we may estimate both integrands on the RHS in terms of $\lesssim_E 2^{k-k_1}$, which upon integration over all $k_1 \geqslant 0$ yields a similar bound:

$$
\| \psi_1 \|_{L_x^2} \lesssim_E 2^k.
$$

Thus we have proved that:

$$
\| P_k \left(P_{<0}\phi[0] - \phi_{,<0}[0] \right) \|_{\dot{H}^1 \times L^2} \lesssim_E 2^k.
\tag{305}
$$

Step 3 (Intermediate frequency bounds, the contribution of $-m < k < m$). Here m is some fixed large integer. The goal of the argument here is to show the estimate:

$$
\| P_k \left(P_{<0}\phi[0] - \phi_{,<0}[0] \right) \|_{\dot{H}^1 \times L^2} \lesssim_E m^2\epsilon + 2^{-m}, \quad |k| \leqslant m.
\tag{306}
$$

This is used with m chosen so that $2^{-m} \approx \epsilon$. Due to the identity $\Pi\phi[0] = \Pi\phi[0]$ we can rewrite (306) in the form

$$
\| P_k \left(P_{<0}\Pi\phi[0] - \Pi P_{<0}\phi[0] \right) \|_{\dot{H}^1 \times L^2} \lesssim_E m^2\epsilon + 2^{-m}.
\tag{307}
$$

This is a direct consequence of the following paradifferential relation:

Lemma 11.3. *Let Π be as in (302), and $D\Pi$ be its differential. Then for each $\psi[0] \in \dot{H}^1 \times L^2$ with energy E and energy dispersion ϵ and each $k \in \mathbb{R}$ we have*

$$
\| P_k \Pi\psi[0] - D\Pi(P_{<k-m}\psi[0]) P_k\psi[0] \|_{\dot{H}^1 \times L^2} \lesssim_E m^2\epsilon + 2^{-m}, \quad m > 4,
\tag{308}
$$

where P_k can be substituted by any multiplier whose symbol has similar size, localization and regularity.

We remark that in (308) there is no geometry left. That is to say, $\psi[0]$ in (308) need not satisfy the identity $\Pi\psi[0] = \psi[0]$.

It is easy to see that (308) implies (307). Indeed, if $k \geq 2$ then the first term $P_k P_{<0}\Pi\phi[0]$ in (307) does not contribute, while for the second we use (308) with $\psi[0] = P_{<0}\phi[0]$. On the other hand if $k < 2$ then for the first term $P_k P_{<0}\Pi\phi[0]$ in (307) we use (308) with P_k replaced by $P_k P_{<0}$ and $\psi[0] = \phi[0]$, while for the second term we again use (308) with $\psi[0] = P_{<0}\phi[0]$. It remains to prove the lemma.

Proof of Lemma 11.3. We write

$$P_k \Pi \psi[0] - D\Pi(P_{<k-m}\psi[0])P_k\psi[0] = (w_0, w_1),$$

with

$$w_0 = P_k G(\psi_0) - (\nabla G)(P_{<k-m}\psi_0)P_k\psi_0,$$
$$w_1 = P_k(H(\psi_0)\psi_1) - (H(P_{<k-m}\psi_0)P_k\psi_1 + (\nabla H)(P_{<k-m}\psi_0)P_k\psi_0 P_{<k-m}\psi_1).$$

Then we need to prove that

$$\|w_0\|_{\dot{H}^1} + \|w_1\|_{L^2} \lesssim_E m^2\epsilon + 2^{-m}.$$

We observe that the expression for ∇w_0 coincides with the expression for w_1 with $H = \nabla G$ and $\psi_1 = \nabla\psi_0$. Hence it suffices to prove the bound for w_1. Furthermore, the last term in w_1 is directly estimated as

$$\|(\nabla H)(P_{<k-m}\psi_0)P_k\psi_0 P_{<k-m}\psi_1\|_{L^2} \lesssim \|P_k\psi_0\|_{L^\infty}\|\psi_1\|_{L^2} \lesssim_E \epsilon.$$

It remains to show that

$$\|P_k(H(\psi_0)\psi_1) - H(P_{<k-m}\psi_0)P_k\psi_1\|_{L^2} \lesssim_E m^2\epsilon + 2^{-m}. \tag{309}$$

We use the integral representation

$$P_k(H(\psi_0)\psi_1) = P_k \int_{-\infty}^{\infty} \frac{d}{dk_1}(H(P_{<k_1}\psi_0)P_{<k_1}\psi_1)dk_1$$
$$= P_k \int_{-\infty}^{\infty} H(P_{<k_1}\psi_0)P_{k_1}\psi_1 + \nabla H(P_{<k_1}\psi_0)P_{k_1}\psi_0 P_{<k_1}\psi_1 dk_1.$$

The integrals from $-\infty$ to $k-m$, respectively from $k+m$ to ∞ can be bounded $\lesssim_E 2^{-m}$ as in Step 1, respectively Step 2 above. For the integral from $k-m$ to $k+m$ we consider the two terms in the integrand separately. The second term is estimated directly,

$$\|\nabla H(P_{<k_1}\psi_0)P_{k_1}\psi_0 P_{<k_1}\psi_1\|_{L^2} \lesssim \|P_{k_1}\psi_0\|_{L^\infty}\|P_{<k_1}\psi_1\|_{L^2} \lesssim_E \epsilon.$$

Thus so far we have

$$P_k(H(\psi_0)\psi_1) = P_k \int_{k-m}^{k+m} H(P_{<k_1}\psi_0)P_{k_1}\psi_1 dk_1 + O_{L^2}(m\epsilon + 2^{-m}).$$

The remaining integrand is further expanded,

$$H(P_{<k_1}\psi_0)P_{k_1}\psi_1 = H(P_{<k-m}\psi_0)P_{k_1}\psi_1 + \int_{k-m}^{k_1} \nabla H(P_{<k_2}\psi_0)P_{k_2}\psi_0 P_{k_1}\psi_1 dk_2.$$

The second term can be estimated as above by $\lesssim_E m\epsilon$. We arrive at

$$P_k(H(\psi_0)\psi_1) = P_k(H(P_{<k-m}\psi_0)P_{[k-m,k+m]}\psi_1) + O_{L^2}(m^2\epsilon + 2^{-m}).$$

This implies (309) via a commutator bound, see (10):

$$\|[P_k, H(P_{<k-m}\psi_0)]P_{[k-m,k+m]}\psi_1\|_{L^2} \lesssim 2^{-k}\|\nabla H(P_{<k-m}\psi_0)\|_{L^\infty}\|P_{[k-m,k+m]}\psi_1\|_{L^2}$$
$$\lesssim_E 2^{-m}.$$

The proof of the lemma is complete. □

This concludes our demonstration of Proposition 11.1. □

References

1. Bahouri, H., Gérard, P.: High frequency approximation of solutions to critical nonlinear wave equations. Amer. J. Math. **121**(1), 131–175 (1999)
2. Bejenaru, I., Ionescu, A., Kenig, C., Tataru, D.: Global Schrödinger maps. http://arXiv.org/abs/0807.0265v1[math.AD], 2008
3. Gromov, M.L.: Isometric imbeddings and immersions. Dokl. Akad. Nauk SSSR **192**, 1206–1209 (1970)
4. Günther, M.: Isometric embeddings of Riemannian manifolds. In: Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990), Tokyo: Math. Soc. Japan, 1991, pp. 1137–1143
5. Hadac, M., Herr, S.S., Koch, H.: Well-posedness and scattering for the kp-ii equation in a critical space. Ann. I. H. Poicavé - AN **26**(3), 917–941 (2009)
6. Kenig, C.E., Merle, F.: Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation. Acta Math. **201**(2), 147–212 (2008)
7. Klainerman, S., Machedon, M.: Space-time estimates for null forms and the local existence theorem. Comm. Pure Appl. Math. **46**(9), 1221–1268 (1993)
8. Klainerman, S., Rodnianski, I.: On the global regularity of wave maps in the critical Sobolev norm. Internat. Math. Res. Notices **13**, 655–677 (2001)
9. Klainerman, S., Selberg, S.: Remark on the optimal regularity for equations of wave maps type. Comm. Part. Diff. Eqs. **22**(5-6), 901–918 (1997)
10. Klainerman, S., Selberg, S.: Bilinear estimates and applications to nonlinear wave equations. Commun. Contemp. Math. **4**(2), 223–295 (2002)
11. Koch, H., Tataru, D.: Dispersive estimates for principally normal pseudodifferential operators. Comm. Pure Appl. Math. **58**(2), 217–284 (2005)
12. Koch, H., Tataru, D.: A priori bounds for the 1D cubic NLS in negative Sobolev spaces. Int. Math. Res. Not. IMRN, **16**:Art. ID rnm053, 36, (2007)
13. Krieger, J., Schlag, W., Tataru, D.: Renormalization and blow up for charge one equivariant critical wave maps. Invent. Math. **171**(3), 543–615 (2008)
14. Krieger, J.: Global regularity of wave maps from $\mathbb{R}^{3+1}$ to surfaces. Commun. Math. Phys. **238**(1-2), 333–366 (2003)
15. Krieger, J.: Global regularity of wave maps from $\mathbb{R}^{2+1}$ to H^2. Small energy. Commun. Math. Phys. **250**(3), 507–580 (2004)
16. Krieger, J., Schlag, W.: Concentration compactness for critical wave-maps. http://arXiv.org/abs/0908.2974v1[math.AP], 2009
17. Nahmod, A., Stefanov, A., Uhlenbeck, K.: On the well-posedness of the wave map problem in high dimensions. Comm. Anal. Geom. **11**(1), 49–83 (2003)
18. Nash, J.: The imbedding problem for Riemannian manifolds. Ann. of Math. (2) **63**, 20–63 (1956)
19. Rodnianski, I., Sterbenz, J.: On the formation of singularities in the critical $O(3)$ sigma-model. http://arXiv.org/abs/math/0605023v3[mth.AP], 2008
20. Shatah, J., Struwe, M.: The Cauchy problem for wave maps. Int. Math. Res. Not. **11**, 555–571 (2002)
21. Sterbenz, J., Tataru, D.: Regularity of Wave-Maps in Dimension 2+1. doi:10.1007/s00220-010-1062-3
22. Tao, T.: Global regularity of wave maps III. Large energy from $\mathbb{R}^{2+1}$ to hyperbolic spaces. http://arXiv.org/abs/0805.4666v3[math.AP], 2009
23. Tao, T.: Global regularity of wave maps IV. Absence of stationary or self-similar solutions in the energy class. http://arXiv.org/abs/0806.3592v2[math.AP], 2009
24. Tao, T.: Global regularity of wave maps V. Large data local well-posedness in the energy class. http://arXiv.org/abs/0808.0368v2[math.AP], 2009
25. Tao, T.: Global regulrity of wave maps VI. Abstract theory of Minimal energy blowup solutions. http://arXiv.org/abs/0906.2833v2[math.AP], 2009
26. Tao, T.: An inverse theorem for the bilinear L^2 Strichartz estimate for the wave equation. http://arXiv.org/abs/0904.2880v1[math.AP], 2009
27. Tao, T.: Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates. Math. Z. **238**(2), 215–268 (2001)

28. Tao, T.: Global regularity of wave maps. I. Small critical Sobolev norm in high dimension. Internat. Math. Res. Notices **6**, 299–328 (2001)
29. Tao, T.: Global regularity of wave maps. II. Small energy in two dimensions. Commun. Math. Phys. **224**(2), 443–544 (2001)
30. Tao, T.: Geometric renormalization of large energy wave maps. In: Journées "Équations aux Dérivées Partielles", pages Exp. No. XI, 32. École Polytech., Palaiseau, 2004
31. Tataru, D.: Local and global results for wave maps. I. Comm. Part. Diff. Eqs. **23**(9-10), 1781–1793 (1998)
32. Tataru, D.: On global existence and scattering for the wave maps equation. Amer. J. Math. **123**(1), 37–77 (2001)
33. Tataru, D.: Rough solutions for the wave maps equation. Amer. J. Math. **127**(2), 293–377 (2005)
34. Wolff, T.: A sharp bilinear cone restriction estimate. Ann. of Math. (2) **153**(3), 661–698 (2001)

Communicated by P. Constantin