CONCENTRATION COMPACTNESS FOR CRITICAL WAVE MAPS

JOACHIM KRIEGER, WILHELM SCHLAG

Abstract. By means of the concentrated compactness method of Bahouri-Gerard [1] and Kenig-Merle [13], we prove global existence and regularity for wave maps with smooth data and large energy from $\mathbb{R}^{2+1} \to \mathbb{H}^2$. The argument yields an a priori bound of the Coulomb gauged derivative components of our wave map relative to a suitable norm $\|\cdot\|_S$ (which holds the solution) in terms of the energy alone. As a by-product of our argument, we obtain a phase-space decomposition of the gauged derivative components analogous to the one of Bahouri-Gerard.

Contents

The authors wish to thank Sergiu Klainerman, Carlos Kenig, and Frank Merle for their interest in this work, as well as their encouragement. The first author is also indebted to Sergiu Klainerman for introducing him to the topic of wave maps in general. Further, they thank Carlos Kenig, Daniel Tataru, and Jacob Sterbenz for helpful discussions. The authors were supported in part by the National Science Foundation, DMS-0757278 (JK) and DMS-0617854 (WS). The first author was supported by a Sloan fellowship, and the second by a Guggenheim fellowship.

1. Introduction and Overview

1.1. **The main result and its history.** Formally speaking, wave maps are the analogue of harmonic maps where the Minkowski metric is imposed on the independent variables. More precisely, for a smooth $\mathbf{u} : \mathbb{R}^{n+1} \to \mathcal{M}$ with $(\mathcal{M}, g)$ Riemannian, define the Lagrangian

$$\mathcal{L}(\mathbf{u}) := \int_{\mathbb{R}^{n+1}} \left(|\partial_t \mathbf{u}|_g^2 - |\nabla \mathbf{u}|_g^2\right) dtdx$$

Then the critical points are defined as $\mathcal{L}'(\mathbf{u}) = 0$ which means that $\Box \mathbf{u} \perp T_u\mathcal{M}$ in case $\mathcal{M}$ is imbedded in some Euclidean space. This is called the *extrinsic* formulation, which can also be written as

$$\Box \mathbf{u} + A(\mathbf{u})(\partial_\alpha \mathbf{u}, \partial^\alpha \mathbf{u}) = 0$$

where $A(\mathbf{u})$ is the second fundamental form. In view of this, it is clear that $\gamma \circ \phi$ is a wave map for any geodesic γ in $\mathcal{M}$ and any free scalar wave ϕ. Moreover, any harmonic map is a stationary wave map. The *intrinsic* formulation is $D^\alpha \partial_\alpha u = 0$, where

$$D_\alpha X^j := \partial_\alpha X^j + \Gamma^j_{ik} \circ \mathbf{u}\, X^i \partial_\alpha u^k$$

is the covariant derivative induced by $\mathbf{u}$ on the pull-back bundle of $T\mathcal{M}$ under $\mathbf{u}$ (with the summation convention in force). Thus, in local coordinates $\mathbf{u} = (u^1, \ldots, u^d)$ one has

$$\Box u^j + \Gamma^j_{ik} \circ \mathbf{u}\, \partial_\alpha u^i\, \partial^\alpha u^k = 0 \tag{1.1}$$

The central problem for wave maps is to answer the following question:

For which $\mathcal{M}$ does the Cauchy problem for the wave map $\mathbf{u} : \mathbb{R}^{n+1} \to \mathcal{M}$ with smooth data $(\mathbf{u}, \dot{\mathbf{u}})|_{t=0} = (\mathbf{u}_0, \mathbf{u}_1)$ have global smooth solutions?

In view of finite propagation speed, one may assume that the data $(\mathbf{u}_0, \mathbf{u}_1)$ are trivial outside of some compact set (i.e., $\mathbf{u}_0$ is constant outside of some compact set, whereas $\mathbf{u}_1$ vanishes outside of that set). Let us briefly describe what is known about this problem.

First, recall that the wave map equation is invariant under the scaling $\mathbf{u} \mapsto \mathbf{u}(\lambda\cdot)$ which is critical relative to $\dot{H}^{\frac{n}{2}}(\mathbb{R}^n)$, whereas the conserved energy

$$\mathcal{E}(\mathbf{u}) = \frac{1}{2} \sum_{\alpha=0}^{n} \int_{\mathbb{R}^n} |\partial_\alpha \mathbf{u}(t, \cdot)|^2\, dx$$

is critical relative to $\dot{H}^1(\mathbb{R}^n)$. In the supercritical case $n \geq 3$ it was observed by Shatah [39] that there are self-similar blowup solutions of finite energy. In the critical case $n = 2$, it is known that there can be no self-similar blowup, see [40]. Moreover, Struwe [48] observed that in the equivariant setting, blowup in this dimension has to result from a strictly slower than self-similar rescaling of a harmonic sphere of finite energy. His arguments were based on the very detailed well-posedness of equivariant wave maps by Christodoulou, Tavildar-Zadeh [4], [5], and Shatah, Tahvildar-Zadeh [42], [43] in the energy class for equivariant wave maps into manifolds that are invariant under the action of $SO(2,\mathbb{R})$. Finally, Rodnianski, Sterbenz [36], as well as the authors together with Daniel Tataru [25] exhibited finite energy wave maps from $\mathbb{R}^{2+1} \to S^2$ that blow up in finite time by suitable rescaling of harmonic maps.

Let us now briefly recall some well-posedness results. The nonlinearity in (1.1) displays a *nullform structure*, which was the essential feature in the subcritical theory of Klainerman-Machedon [17]–[15], and Klainerman-Selberg [19], [20]. These authors proved strong local well-posedness for data in $H^s(\mathbb{R}^n)$ when $s > \frac{n}{2}$. The important critical theory $s = \frac{n}{2}$ was begun by Tataru [62], [61]. These seminal papers proved global well-posedness for smooth data satisfying a smallness condition in $\dot{B}^{\frac{n}{2}}_{2,1}(\mathbb{R}^n) \times \dot{B}^{\frac{n}{2}-1}_{2,1}(\mathbb{R}^n)$. In a breakthrough work, Tao [57], [56] was able to prove well-posedness for data with small $\dot{H}^{\frac{n}{2}} \times \dot{H}^{\frac{n}{2}}$ norm and the sphere as target. For this purpose, he introduced the important *microlocal gauge* in order to remove some "bad" interaction terms from the nonlinearity. Later results by Klainerman, Rodnianski [18], Nahmod, Stephanov, Uhlenbeck [34], Tataru [59], [58], and Krieger [22], [23], [24] considered other cases of targets by using similar methods as in Tao's work.

Recently, Sterbenz and Tataru [45], [46] have given the following very satisfactory answer[1] to the above question: *If the energy of the initial data is smaller than the energy of any nontrivial harmonic map $\mathbb{R}^n \to \mathcal{M}$, then one has global existence and regularity.*
Notice in particular that if there are no harmonic maps other than constants, then one has global existence for all energies. A particular case of this are the hyperbolic spaces $\mathbb{H}^n$ for which Tao [55]–[51] has achieved the same result (with some a priori global norm control).

The purpose of this paper is to apply the method of concentration compactness as in Bahouri, Gerard [1] and Kenig, Merle [13], [14] to the large data wave map problem with the hyperbolic plane $\mathbb{H}^2$ as target. We emphasize that this gives more than global existence and regularity as already in the semilinear case considered by the aforementioned authors. The fact that in the critical case the large data problem should be decided by the geometry of the target is a conjecture going back to Sergiu Klainerman.

Let us now describe our result in more detail. Let $\mathbb{H}^2$ be the upper half-plane model of the hyperbolic plane equipped with the metric $ds^2 = \frac{d\mathbf{x}^2+d\mathbf{y}^2}{\mathbf{y}^2}$. Let $\mathbf{u} : \mathbb{R}^2 \to \mathbb{H}^2$ be a smooth map. Expanding the derivatives $\{\partial_\alpha \mathbf{u}\}_{\alpha=0,1,2}$ (with $\partial_0 := \partial_t$) in the orthonormal frame $\{\mathbf{e}_1, \mathbf{e}_2\} = \{\mathbf{y}\partial_\mathbf{x}, \mathbf{y}\partial_\mathbf{y}\}$ gives rise to smooth coordinate functions $\phi^1_\alpha, \phi^2_\alpha$. In what follows, $\|\partial_\alpha \mathbf{u}\|_X$ will mean $(\sum_{j=1}^2 \|\phi^j_\alpha\|_X^2)^{\frac{1}{2}}$ for any norm $\|\cdot\|_X$ on scalar functions. For example, the energy of $\mathbf{u}$ is

$$E(\mathbf{u}) := \sum_{\alpha=0}^{2} \|\partial_\alpha \mathbf{u}\|_2^2$$

Next, suppose $\pi : \mathbb{H}^2 \to M$ is a covering map with M some hyperbolic Riemann surface with the metric that renders π a local isometry. In other words, $M = \mathbb{H}^2/\Gamma$ for some discrete subgroup $\Gamma \subset PSL(2,\mathbb{R})$ which operates totally discontinuously on $\mathbb{H}^2$. Now suppose $\mathbf{u} : \mathbb{R}^2 \to M$ is a smooth map which is constant outside of some compact set, say. It lifts to a smooth map $\tilde{\mathbf{u}} : \mathbb{R}^2 \to \mathbb{H}^2$ uniquely, up to composition with an element of Γ. We now define $\|\partial_\alpha \mathbf{u}\|_X := \|\partial_\alpha \tilde{\mathbf{u}}\|_X$. In particular, the energy $E(\mathbf{u}) := E(\tilde{\mathbf{u}})$. Note that due to the fact that Γ is a group of isometries of $\mathbb{H}^2$, these definitions are unambiguous. Our main result is as follows.

Theorem 1.1. *There exists a function $K : (0,\infty) \to (0,\infty)$ with the following property: Let M be a hyperbolic Riemann surface. Suppose $(\mathbf{u}_0, \mathbf{u}_1) : \mathbb{R}^2 \to M \times TM$ are smooth and $\mathbf{u}_0 = const$, $\mathbf{u}_1 = 0$ outside of some compact set. Then the wave map evolution $\mathbf{u}$ of these data as a map $\mathbb{R}^{1+2} \to M$ exists*

[1]The conclusions of our work were reached before the appearance of [45], [51]

globally as a smooth function and, moreover, for any $\frac{1}{p}+\frac{1}{2q}\le\frac{1}{4}$ *with* $2\le q<\infty$, $\gamma=1-\frac{1}{p}-\frac{2}{q}$,

$$\sum_{\alpha=0}^{2}\|(-\Delta)^{-\frac{\gamma}{2}}\partial_\alpha\mathbf{u}\|_{L_t^pL_x^q}\le C_q\,K(E) \tag{1.2}$$

Moreover, in the case when $M\hookrightarrow\mathbb{R}^N$ *is a compact Riemann surface, one has scattering:*

$$\max_{\alpha=0,1,2}\|\partial_\alpha\mathbf{u}(t)-\partial_\alpha S(t)(f,g)\|_{L_x^2}\to 0\qquad as\ \ t\to\pm\infty$$

where $S(t)(f,g)=\cos(t|\nabla|)f+\frac{\sin(t|\nabla|)}{|\nabla|}g$ *and suitable* $(f,g)\in(\dot H^1\times L^2)(\mathbb{R}^2;\mathbb{R}^N)$. *Alternatively, if* M *is non-compact, then lifting* $\mathbf{u}$ *to a map* $\mathbb{R}^{1+2}\to\mathbb{H}^2$ *with derivative components* ϕ_α^j *as defined above, one has*

$$\max_{\alpha=0,1,2}\|\phi_\alpha^j(t)-\partial_\alpha S(t)(f^j,g^j)\|_{L_x^2}\to 0\qquad as\ \ t\to\pm\infty$$

where $(f^j,g^j)\in(\dot H^1\times L^2)(\mathbb{R}^2;\mathbb{R})$.

We emphasize that (1.2) can be strengthened considerably in terms of the type of norm applied to the Coulomb gauged derivative components of the wave map:

$$\sum_{\alpha=0}^{2}\|\psi_\alpha\|_S^2\le C\,K(E)^2 \tag{1.3}$$

The meaning ψ_α as well as of the S norm will be explained below. We now turn to describing this result and our methods in more detail. For more background on wave maps see [12], [59], and [40].

1.2. **Wave maps to $\mathbb{H}^2$.** The manifold $\mathbb{H}^2$ is the upper half-plane equipped with the metric $ds^2=\frac{d\mathbf{x}^2+d\mathbf{y}^2}{\mathbf{y}^2}$. Expanding the derivatives $\{\partial_\alpha\mathbf{u}\}_{\alpha=0,1,2}$ (with $\partial_0:=\partial_t$) of a smooth map $\mathbf{u}:\mathbb{R}^{1+2}\to\mathbb{H}^2$ in the orthonormal frame $\{\mathbf{e}_1,\mathbf{e}_2\}=\{\mathbf{y}\partial_\mathbf{x},\mathbf{y}\partial_\mathbf{y}\}$ yields

$$\partial_\alpha\mathbf{u}=(\partial_\alpha\mathbf{x},\partial_\alpha\mathbf{y})=\sum_{j=1}^{2}\phi_\alpha^j\,\mathbf{e}_j$$

whence

$$\mathbf{y}=e^{\sum_{j=1,2}\Delta^{-1}\partial_j\phi_j^2},\quad \mathbf{x}=\sum_{j=1,2}\Delta^{-1}\partial_j(\phi_j^1\mathbf{y}) \tag{1.4}$$

provided we assume the normalization $\lim_{|x|\to\infty}|\ln\mathbf{y}|=\lim_{|x|\to\infty}|\mathbf{x}|=0$. Energy conservation takes the form

$$\int_{\mathbb{R}^2}\sum_{\alpha=0}^{2}\sum_{j=1}^{2}|\phi_\alpha^j(t,x)|^2\,dx=\int_{\mathbb{R}^2}\sum_{\alpha=0}^{2}\sum_{j=1}^{2}|\phi_\alpha^j(0,x)|^2\,dx \tag{1.5}$$

where $x=(x_1,x_2)$ and $\partial_0=\partial_t$. If $\mathbf{u}(t,x)$ is a smooth wave map, then the functions $\{\phi_\alpha^j\}$ for $0\le\alpha\le2$ and $j=1,2$ satisfy the div–curl system

$$\partial_\beta\phi_\alpha^1-\partial_\alpha\phi_\beta^1=\phi_\alpha^1\phi_\beta^2-\phi_\beta^1\phi_\alpha^2 \tag{1.6}$$

$$\partial_\beta\phi_\alpha^2-\partial_\alpha\phi_\beta^2=0 \tag{1.7}$$

$$\partial_\alpha\phi^{1\alpha}=-\phi_\alpha^1\phi^{2\alpha} \tag{1.8}$$

$$\partial_\alpha\phi^{2\alpha}=\phi_\alpha^1\phi^{1\alpha} \tag{1.9}$$

for all $\alpha,\beta=0,1,2$. As usual, repeated indices are being summed over, and lowering or raising is done via the Minkowski metric. Clearly, (1.6) and (1.7) are integrability conditions which are an expression of the curvature of $\mathbb{H}^2$. On the other hand, (1.8) and (1.9) are the actual wave map system. Since the choice of frame was arbitrary, one still has gauge freedom for the system (1.6)–(1.9). We shall exclusively rely on the Coulomb gauge which is given in terms of complex notation by the functions

$$\psi_\alpha:=\psi_\alpha^1+i\psi_\alpha^2=(\phi_\alpha^1+i\phi_\alpha^2)e^{-i\Delta^{-1}\sum_{j=1}^{2}\partial_j\phi_j^1} \tag{1.10}$$

If ϕ_j^1 are Schwartz functions , then $\sum_{j=1}^2 \partial_j \phi_j^1$ has mean zero whence

$$(\Delta^{-1}\sum_{j=1}^2 \partial_j\phi_j^1)(z) = \frac{1}{2\pi}\int_{\mathbb{R}^2} \log|z-\zeta| \sum_{j=1}^2 \partial_j\phi_j^1(\zeta)\, d\zeta\wedge d\bar{\zeta} \tag{1.11}$$

is well-defined and moreover decays like $|z|^{-1}$ (but in general no faster). The gauged components $\{\psi_\alpha\}_{\alpha=0,1,2}$ satisfy the new div–curl system

$$\partial_\alpha\psi_\beta - \partial_\beta\psi_\alpha = i\psi_\beta\Delta^{-1}\sum_{j=1,2}\partial_j(\psi_\alpha^1\psi_j^2 - \psi_\alpha^2\psi_j^1) - i\psi_\alpha\Delta^{-1}\partial_j(\psi_\beta^1\psi_j^2 - \psi_\beta^2\psi_j^1) \tag{1.12}$$

$$\partial_\nu\psi^\nu = i\psi^\nu\Delta^{-1}\sum_{j=1}^2\partial_j(\psi_\nu^1\psi_j^2 - \psi_\nu^2\psi_j^1) \tag{1.13}$$

In particular, one obtains the following system of wave equations for the ψ_α:

$$\begin{aligned}\Box\psi_\alpha &= i\partial^\beta\big[\psi_\alpha\Delta^{-1}\sum_{j=1,2}\partial_j(\psi_\beta^1\psi_j^2 - \psi_\beta^2\psi_j^1)\big] - i\partial^\beta\big[\psi_\beta\Delta^{-1}\partial_j(\psi_\alpha^1\psi_j^2 - \psi_\alpha^2\psi_j^1)\big] \\ &\quad + i\partial_\alpha\big[\psi^\beta\Delta^{-1}\sum_{j=1,2}\partial_j(\psi_\beta^1\psi_j^2 - \psi_\beta^2\psi_j^1)\big]\end{aligned} \tag{1.14}$$

Throughout this paper we shall only consider *admissible* wave maps $\mathbf{u}$. These are characterized as smooth wave maps $\mathbf{u} : I\times\mathbb{R}^2 \to \mathbb{H}^2$ on some time interval I so that the derivative components ϕ_α^j are Schwartz functions on fixed time slices.
By the method of *Hodge decompositions* from[2] [22]–[24] one exhibits the null-structure present in (1.12)–(1.14). Hodge decomposition here refers to writing

$$\psi_\beta = -R_\beta\sum_{k=1}^2 R_k\psi_k + \chi_\beta \tag{1.15}$$

where $R_\beta := \partial_\beta|\nabla|^{-1}$ are the usual Riesz transform. Inserting the hyperbolic terms $R_\beta\sum_{k=1}^2 R_k\psi_k$ into the right-hand sides of (1.12)–(1.14) leads to trilinear nonlinearities with a null structure. As is well-known, such null structures are amenable to better estimates since they annihilate "self-interactions", or more precisely, interactions of waves which propagate along the same characteristics, cf. [17]–[16], as well as [19], [20], [10]. Furthermore, inserting at least one "elliptic term" χ_β from (1.15) leads to a higher order nonlinearity, in fact quintic or higher which are easier to estimate (essentially by means of Strichartz norms). To see this, note that

$$\sum_{j=1}^2\partial_j\chi_j = 0$$
$$\partial_j\chi_\beta - \partial_\beta\chi_j = \partial_j\psi_\beta - \partial_\beta\psi_j$$

whence

$$\chi_\beta = i\sum_{j,k=1}^2\partial_j\Delta^{-1}\big[\psi_\beta\Delta^{-1}\partial_k(\psi_j^1\psi_k^2 - \psi_k^1\psi_j^2) - \psi_j\Delta^{-1}\partial_k(\psi_\beta^1\psi_k^2 - \psi_k^1\psi_\beta^2)\big] \tag{1.16}$$

Since we are only going to obtain a priori bounds on ϕ_α^j, it will suffice to assume throughout that the ϕ_α^j are Schwartz functions, whence the same holds for ψ_α. In what follows, we shall never actually *solve the system* (1.12)–(1.14). To go further, the wave-equation (1.14) by *itself is meaningless without assuming the* ψ_α *to satisfy the compatibility relations* (1.12) *and* (1.13). In fact, it is not even clear that (1.12) and (1.13) will hold for all $t\in(-T,T)$ if they hold at time $t=0$ and (1.14) holds for all $t\in(-T,T)$. Nonetheless, assuming that the ψ_α are defined in terms of the derivative components ϕ_α of a 'sufficiently nice' wave map, it is clear that all three of (1.12) - (1.14) will be satisfied. This being said, we will only use the system (1.14) to derive *a priori estimates* for ψ_α, which will then be shown to lead to suitable

[2]In these papers this decomposition is also referred to as "dynamic decomposition".

bounds on the components ϕ_α^j of derivatives of a wave map $\mathbf{u}$. This is done by means of Tao's device of frequency envelope, see [57] or [22]. This refers to a sequence $\{c_k\}_{k\in\mathbb{Z}}$ of positive reals such that

$$c_k\, 2^{-\sigma|k-\ell|} \le c_\ell \le c_k\, 2^{\sigma|k-\ell|} \tag{1.17}$$

where $\sigma > 0$ is a small number. The most relevant example is given by

$$c_k := \Big(\sum_{\ell\in\mathbb{Z}} 2^{-\sigma|k-\ell|}\|P_\ell\psi(0)\|_2^2\Big)^{\frac12}$$

which controls the initial data. While it is of course clear that (1.6)–(1.9) imply the system (1.12)–(1.14), the reverse implication is not such a simple matter since it involves solving an elliptic system with large solutions. On the other hand, transferring estimates on the ψ_α in $H^s(\mathbb{R}^2)$ spaces to similar bounds on the derivative components ϕ_α^j does not require this full implication. Indeed, assume the bound $\|\psi\|_{L_t^\infty((-T_0,T_1);H^{\delta_1}(\mathbb{R}^2))} < \infty$ for some small $\delta_1 > 0$ (we will obtain such bounds via frequency envelopes with $0 < \delta_1 < \sigma$). For any fixed time $t \in (-T_0, T_1)$ one now has with P_k being the usual Littlewood-Paley projections to frequency 2^k,

$$\begin{aligned}
\|P_\ell\phi_\alpha\|_{H^{\delta_2}} &= \|P_\ell[e^{i\sum_{j=1}^2 \Delta^{-1}\partial_j\phi_j^1}\psi_\alpha]\|_{H^{\delta_2}}\\
&\le \|P_\ell[P_{<\ell-10}(e^{i\sum_{j=1}^2 \Delta^{-1}\partial_j\phi_j^1})P_{[\ell-10,\ell+10]}\psi_\alpha]\|_{H^{\delta_2}}\\
&\quad + \|P_\ell[P_{[\ell-10,\ell+10]}(e^{i\sum_{j=1}^2 \Delta^{-1}\partial_j\phi_j^1})P_{<\ell+15}\psi_\alpha]\|_{H^{\delta_2}}\\
&\quad + \sum_{k>\ell+10} \|P_\ell[P_k(e^{i\sum_{j=1}^2 \Delta^{-1}\partial_j\phi_j^1})P_{k+O(1)}\psi_\alpha]\|_{H^{\delta_2}}\\
&\lesssim \|P_{[\ell-10,\ell+10]}\psi_\alpha\|_{H^{\delta_2}} + \|P_{[\ell-10,\ell+10]}(e^{i\sum_{j=1}^2 \Delta^{-1}\partial_j\phi_j^1})\|_{H^{\delta_2}}\|P_{<\ell+15}\psi_\alpha\|_\infty\\
&\quad + \sum_{k>\ell+10} \|P_k(e^{i\sum_{j=1}^2 \Delta^{-1}\partial_j\phi_j^1})\|_{H^{\delta_2}}\|P_{k+O(1)}\psi_\alpha\|_\infty
\end{aligned}$$

Next, one has the bounds

$$\big\|\nabla_x\, e^{i\Delta^{-1}\sum_{j=1}^2\partial_j\phi_j^1}\big\|_{L_t^\infty L_x^2} \lesssim \|\phi_j^1\|_{L_t^\infty L_x^2}, \qquad \|P_{<\ell+15}\psi_\alpha\|_{L_x^\infty} \lesssim 2^{(1-\delta_1)\ell}\|\psi_\alpha\|_{H^{\delta_1}}$$

where the first one is admissible due to *energy conservation for the derived wave map*, see (1.5). In conclusion,

$$\|P_\ell\phi_\alpha\|_{H^{\delta_2}} \lesssim \|P_{\ell+O(1)}\psi_\alpha\|_{H^{\delta_2}} + 2^{(\delta_2-\delta_1)\ell}\|\phi\|_{L_x^2}\|\psi\|_{H^{\delta_1}}$$

Summing over $\ell \ge 0$ yields

$$\|\phi\|_{L_t^\infty((-T_0,T_1);H^{\delta_2}(\mathbb{R}^2))} < \infty \tag{1.18}$$

By the subcritical existence theory of Klainerman and Machedon, see [17]–[15] as well as [19], [20], the solution can now be extended smoothly beyond this time interval. More precisely, the device of frequency envelopes allows one to place the Schwartz data in $H^s(\mathbb{R}^2)$ for all $s > 0$ initially, and as it turns out, also for all times provided $s > 0$ is sufficiently small. The latter claim is of course the entire objective of this paper. We should also remark that we bring (1.14) into play only because it fits into the framework of the spaces from [57] and [61]. This will allow us to obtain the crucial energy estimate for solutions of (1.14), whereas it is not clear how to do this directly for the system (1.12), (1.13). As already noted in [22], the price one pays for passing to (1.14) lies with the *initial conditions*, or more precisely, the time derivative $\partial_t\psi_\alpha(0,\cdot)$. While $\psi_\alpha(0,\cdot)$ only involves one derivative of the wave map $\mathbf{u}$, this time derivative involves two. This will force us to essentially "randomize" the initial time.

1.3. **The small data theory.** In this section we give a very brief introduction to the spaces which are needed to control the ψ system (1.12), (1.13), and (1.14). A systematic development will be carried out in Section 2 below, largely following [56] (we do need to go beyond both [56] and [22] in some instances such as by adding the sharp Strichartz spaces with the Klainerman-Tataru gain for small scales, and by eventually modifying $\|\cdot\|_{S[k]}$ to the stronger $|\!|\!|\cdot|\!|\!|_{S[k]}$ which allows for a high-high gain in the $S\times S \to L_{tx}^2$ estimate). First note that it is not possible to bound the trilinear nonlinearities in this system in Strichartz spaces due to slow dispersion in dimension two. Moreover, it is not possible to adapt the $X^{s,b}$-space of the subcritical theory to the scaling invariant case as this runs into logarithmic divergences. For this reason, Tataru [61] devised a class of spaces which resolve these logarithmic divergences. His idea was to allow characteristic frames of reference. More precisely, fix $\omega \in S^1$ and define

$$\theta_\omega^\pm := (1, \pm\omega)/\sqrt{2}, \quad t_\omega := (t,x)\cdot\theta_\omega^+, \quad x_\omega := (t,x) - t_\omega\theta_\omega^+$$

which are the coordinates defined by a generator on the light-cone. Now suppose that ψ_i are free waves such that ψ_1 is Fourier supported on $1 \le |\xi| \le 2$, and both ψ_2 and ψ_3 are Fourier supported on $|\xi| \sim 2^k$ where k is large and negative. Finally, we also assume that the three waves are in "generic position", i.e., that their Fourier supports make an angle of about size one. Clearly, $2^{-k}\psi_1\psi_2\psi_3$ is then a representative model for the nonlinearities arising in (1.14). With

$$\psi_3(t,x) = \int_{\mathbb{R}^2} e^{i[t|\xi| + x\cdot\xi]} f(\xi)\, d\xi$$

we perform the *plane-wave* decomposition $\psi_3(t,x) = \int \phi_\omega(\sqrt{2}t_\omega)\, d\omega$ where

$$\phi_\omega(s) := \int e^{irs} f(r\omega)\, rdr$$

By inspection,

$$\int \|\phi_\omega\|_{L^2_{t_\omega} L^\infty_{x_\omega}}\, d\omega \lesssim 2^{\frac{k}{2}} \|\psi_3\|_{L^\infty_t L^2_x} \tag{1.19}$$

Hence,

$$\begin{aligned} 2^{-k} \int \|\phi_\omega\, \psi_2\psi_3\|_{L^1_{t_\omega} L^2_{x_\omega}}\, d\omega &\lesssim 2^{-k} \int \|\phi_\omega\|_{L^2_{t_\omega} L^\infty_{x_\omega}}\, d\omega \|\psi_1\psi_2\|_{L^2_{t_\omega} L^2_{x_\omega}} \\ &\lesssim \|\psi_3\|_{L^\infty_t L^2_x} \|\psi_1\|_{L^\infty_t L^2_x} \|\psi_2\|_{L^\infty_t L^2_x} \end{aligned}$$

which is an example[3] of a *trilinear estimate* which will be studied systematically in Section 5. Here we used both (1.19) and the standard bilinear L^2_{tx} bilinear L^2-bound for waves with angular separation:

$$\|\psi_1\psi_2\|_{L^2_{t_\omega} L^2_{x_\omega}} = \|\psi_1\psi_2\|_{L^2_t L^2_x} \lesssim 2^{\frac{k}{2}} \|\psi_2\|_{L^\infty_t L^2_x} \|\psi_1\|_{L^\infty_t L^2_x}$$

This suggests introducing an *atomic space* with atoms ψ_ω of Fourier support $|\xi| \sim 1$ and satisfying

$$\|\psi_\omega\|_{L^1_{t_\omega} L^2_{x_\omega}} \le 1$$

as part of the space $N[0]$ which holds the nonlinearity (the zero here refers to the Littlewood-Paley projection P_0. Below, we refer to this space as NF). In addition, the space defined by (1.19) is also an atomic space and should be incorporated in the space $S[k]$ holding the solution at frequency 2^k (we refer to this below as the PW space). By duality to $L^1_{t_\omega} L^2_{x_\omega}$ in $N[0]$, we then expect to see $L^\infty_{t_\omega} L^2_{x_\omega}$ as part of $S[0]$. The simple observation here (originating in [61]) is that one can indeed bound the energy along a characteristic frame (t_ω, x_ω) of a free wave as long as its Fourier support makes a positive angle with the direction ω. Indeed, recall the local energy conservation identity $\partial_t e - \operatorname{div}(\partial_t\psi\nabla\psi) = 0$ for a free wave where

$$e = \frac{1}{2}(|\partial_t\psi|^2 + |\nabla\psi|^2)$$

is the energy density, over a region of the form $\{-T \le t \le T\} \cap \{t_\omega > a\}$. From the divergence theorem one obtains that

$$\int_{t_\omega = a} \chi_{[-T\le t\le T]} |\omega^\perp \nabla\psi|^2\, d\mathcal{L}^2 \lesssim \|\psi\|^2_{L^\infty_t L^2_x}$$

where $\mathcal{L}^2$ is the planar Lebesgue measure on $\{t_\omega = a\}$. Sending $T \to \infty$ and letting ρ denote the distance between ω and the direction of the Fourier support of $\psi|_{t=0}$, one concludes that

$$\|\psi\|_{L^\infty_{t_\omega} L^2_{x_\omega}} \lesssim \rho^{-1} \|\psi\|_{L^\infty_t L^2_x}$$

[3]Note that one does not obtain a gain in this case. This fact will be of utmost importance in this paper, forcing us to use a "twisted" wave equation resulting from these high-low-low interactions in the linearized trilinear expressions.

Hence, we should include a piece

$$\sup_{\omega \notin 2\kappa} d(\omega, \kappa) \|\psi\|_{L^\infty_\omega L^2_{x_\omega}}$$

in the norm $S[0]$ holding $P_0\psi$ provided ψ is a wave packet oriented along the cone of dimensions $1 \times 2^k \times 2^{2k}$, projecting onto an angular sector in the ξ-plane associated with the cap $\kappa \subset S^1$, where κ is of size 2^k (this is called NF* below).

Recall that we have made a genericity assumption which guaranteed that the Fourier supports were well separated in the angle. In order to relax this condition, it is essential to invoke the usual device of *nullforms* which cancel out parallel interactions. One of the discoveries of [22] is a genuinely trilinear nullform expansion, see (5.46) and (5.47), which exploit the relative position of all three waves simultaneously. It seems impossible to reduce the trilinear nonlinearities of (1.14) exclusively to the easier bilinear ones.

It is shown in [61] (and then also in [56] which develops much of the functional framework that we use, as well as [22]) that in low dimensions (especially $n = 2$ but these spaces are also needed for $n = 3$), these nullframe spaces are strong enough — in conjunction with more traditional scaling invariant $X^{s,b}$ spaces — to bound the trilinear nonlinearities, as well as weak enough to allow for an energy estimate to hold. This then leads modulo passing to an appropriate gauge to the small energy theory.

The norm $\| \cdot \|_S$ in (1.3) is of the form $\|\psi\|_S := \Big(\sum_{k\in\mathbb{Z}} \|P_k\psi\|_{S[k]}^2 \Big)^{\frac12}$ where $S[k]$ is built from $L^\infty_t L^2_x$, critical $X^{s,b}$, $L^4_t L^\infty_x$ Strichartz norms, as well as the null-frame spaces which we just described.

1.4. **The Bahouri-Gerard concentrated compactness method.** We now come to the core of the argument, namely the Bahouri-Gerard type decomposition and the associated perturbative argument.

In [11] P. Gérard considered defocusing semilinear wave equations in $\mathbb{R}^{3+1}$ of the form $\Box u + f(u) = 0$ with data given by a sequence (ϕ_n, ψ_n) of energy data going weakly to zero. Denote the resulting solutions to the nonlinear problem by u_n, and the free waves with the same data by v_n. Gérard proved that provided $f(u)$ is *subcritical* relative to energy then

$$\|u_n - v_n\|_{L^\infty(I;\mathcal{E})} \to 0 \quad \text{as } n \to \infty$$

where $\mathcal{E}$ is the energy space. In contrast, for this to hold for the energy critical problem he found via the concentrated compactness method of P. L. Lions that it is necessary and sufficient that $\|v_n\|_{L^\infty(I;L^6(\mathbb{R}^3))} \to 0$. In other words, the critical problem experiences a *loss of compactness*.

The origin of this loss of compactness, as well as the meaning of the L^6 condition were later made completely explicit by Bahouri-Gerard [1]. Their result reads as follows: *Let $\{(\phi_n, \psi_n)\}_{n=1}^\infty \subset \dot H^1 \times L^2(\mathbb{R}^3)$ be a bounded sequence, and define v_n to be a free wave with these initial data. Then there exists a subsequence $\{v_n'\}$ of $\{v_n\}$, a finite energy free wave v, as well as free waves $V^{(j)}$ and $(\varepsilon^{(j)}, x^{(j)}) \in (\mathbb{R}^+, \mathbb{R}^3)^{\mathbb{Z}^+}$ for every $j \ge 1$ with the property that for all $\ell \ge 1$,*

$$v_n'(t,x) = v(t,x) + \sum_{j=1}^{\ell} \frac{1}{\sqrt{\varepsilon_n^{(j)}}} V^{(j)}\Big(\frac{t - t_n^{(j)}}{\varepsilon_n^{(j)}}, \frac{x - x_n^{(j)}}{\varepsilon_n^{(j)}} \Big) + w_n^{(\ell)}(t,x) \tag{1.20}$$

where

$$\limsup_{n\to\infty} \|w_n^{(\ell)}\|_{L^5_t(\mathbb{R}, L^{10}_x(\mathbb{R}^3))} \to 0 \qquad \text{as } \ell \to \infty$$

and for any $j \neq k$,

$$\frac{\varepsilon_n^{(j)}}{\varepsilon_n^{(k)}} + \frac{\varepsilon_n^{(k)}}{\varepsilon_n^{(j)}} + \frac{|x_n^{(j)} - x_n^{(k)}| + |t_n^{(j)} - t_n^{(k)}|}{\varepsilon_n^{(j)}} \to \infty \qquad \text{as } n \to \infty$$

Furthermore, the free energy E_0 satisfies the following orthogonality property:

$$E_0(v_n') = E_0(v) + \sum_{j=1}^{\ell} E_0(V^{(j)}) + E_0(w_n^{(\ell)}) + o(1) \qquad \text{as } n \to \infty$$

Note that this result characterized the loss of compactness in terms of the appearance of concentration profiles $V^{(j)}$. Moreover, [1] contains an analogue of this result for so-called Shatah-Struwe solutions of the semi-linear problem $\Box u + |u|^4 u = 0$ which then leads to another proof of the main result in [11]. One of the main applications of their work was to show the existence of a function $A : [0, \infty) \to [0, \infty)$ so that every Shatah-Struwe solution satisfies the bound

$$\|u\|_{L^5_t(\mathbb{R}; L^{10}_x(\mathbb{R}^3))} \leq A(E(u)) \tag{1.21}$$

where $E(u)$ is the energy associated with the semi-linear equation. This is proved by contradiction; indeed, assuming (1.21) fails, one then obtains sequences of bounded energy solutions with uncontrollable Strichartz norm which is then shown to contradict the fact the nonlinear solutions themselves converge weakly to another solution. The decomposition (1.20) compensates for the aforementioned loss of compactness by reducing it precisely to the effect of the *symmetries*, i.e., dilation and scaling. This is completely analogous to the elliptic (in fact, variational) origins of the method of concentrated compactness, see Lions [27] and Struwe [47]. See [1] for more details and other applications.

The importance of [1] in the context of wave maps is made clear by the argument of Kenig, Merle [13], [14]. This method, which will be described in more detail later in this section, represents a general method for attacking global well-posedness problems for energy critical equations such as the wave-map problem. Returning to the Bahouri-Gerard decomposition, we note that any attempt at implementing this technique for wave maps encounters numerous serious difficulties. These are of course all rooted in the difficult nonlinear nature of the system (1.6)–(1.9). Perhaps the most salient feature of our decomposition, performed in detail in section 9.2, as compared to [1] is that the free wave equation no longer captures the correct asymptotic behavior for large times; rather, the atomic components $V^{(j)}$ are defined as solutions of a covariant (or "twisted") wave equation of the form

$$\Box + 2iA_\alpha \partial^\alpha \tag{1.22}$$

where the magnetic potential A_α arises from linearizing the wave map equation in the Coulomb gauge. More precisely, the magnetic term here captures the high-low-low interactions in the trilinear nonlinearities of the wave map system where there is no a priori smallness gain. We shall then obtain the concentration profiles via an inductive procedure over increasing frequency scales; in particular, in (1.22) the Coulomb potential A_α is defined in terms of lower-frequency approximations which are already controlled, see the next subsection for more details.

In keeping with the Kenig-Merle method, the Bahouri-Gerard decomposition is used to show the following: assume that a uniform bound of the form

$$\|\psi\|_S \leq C(E)$$

for some function $C(E)$ fails for some finite energy levels E. In particular, the set

$$\mathcal{A} := \{E \in \mathbf{R}_+ | \sup_{\|\psi\|_{L^2_x} \leq E} \|\psi\|_S = \infty\} \neq \emptyset$$

where we loosely denote the energy by $\|\psi\|_{L^2_x} = (\sum_{\alpha=0}^2 \|\psi_\alpha\|^2_{L^2_x})^{\frac{1}{2}}$, and we can then define a number, denoted throughout the rest of the paper by E_{crit} or also E_C, as follows:

$$E_{crit} = \inf_{E \in \mathcal{A}} E \tag{1.23}$$

Then there must exist a weak wave map $\mathbf{u}_{\text{critical}} : (-T_0, T_1) \to S$ to a compact Riemann surface uniformized by $\mathbb{H}^2$, which enjoys certain compactness properties. In the final part of the argument we then need to rule out the existence of such an object, arriving at an eventual contradiction at the end of the paper.

Starting this grand contradiction argument here, we now assume as above that $\mathcal{A} \neq \emptyset$; this implies that there is a sequence of Schwartz class (on fixed time slices) wave maps $\mathbf{u}^n : (-T_0^n, T_1^n) \times \mathbb{R}^2 \to \mathbb{H}^2$ with the properties that

- $\|\psi^n\|_{L^2_x} \to E_{crit}$
- $\lim_{n\to\infty} \|\psi^n\|_{S((-T_0^n, T_1^n)\times\mathbb{R}^2)} = \infty$

Thus all these wave maps have $t=0$ in their domain of definition. We shall call such a sequence of wave maps *essentially singular*. Roughly speaking, we shall proceed along the following steps. First, recall that the Bahouri-Gerard theorem is a genuine phase-space result in the sense that it identifies the main asymptotic carriers of energy *which are not pure radiation*, which would then sit in $w_n^{(\ell)}$. This refers to the free waves $V^{(j)}$ above, which are "localized" in frequency (namely at scale $(\varepsilon_n^{(j)})^{-1}$) as well as in physical spaces (namely around the space-time points $(t_n^{(j)}, x_n^{(j)})$). The procedure of filtering out the scales $\varepsilon_n^{(j)}$ is due to Metivier-Schochet, see [31].

(1) *Bahouri-Gerard I: filtering out frequency blocks.*

If we apply the frequency localization procedure of Metivier-Schochet to the derivative components $\phi_\alpha^n = (\frac{\partial_\alpha \mathbf{x}^n}{\mathbf{y}^n}, \quad \frac{\partial_\alpha \mathbf{y}^n}{\mathbf{y}^n})$ of an essentially singular sequence at time $t=0$, we run into the problem that the resulting frequency components are not necessarily related to an actual map from $\mathbb{R}^2 \to \mathbb{H}^2$. We introduce a procedure to obtain a frequency decomposition which is "geometric", i.e., the frequency localized pieces are themselves derivative components of maps from $\mathbb{R}^2 \to \mathbb{H}^2$. More specifically, in section 9.2, we start with decompositions

$$\phi_\alpha^n = \sum_{a=1}^{A} \tilde{\phi}_\alpha^{na} + w_\alpha^{nA},\ \alpha = 0,1,2$$

where the $\tilde{\phi}_\alpha^{na}$ are 'frequency atoms' obtained from the first stage of the standard Bahouri-Gerard process, see [1]. Here it may be assumed that the frequency scales in the cases $\alpha = 0,1,2$ are identical. Since the $\tilde{\phi}_\alpha^{na}$ do not necessarily form the derivative components of admissible maps into $\mathbb{H}^2$, one replaces them by components ϕ_α^{na} which are derivative components of admissible maps, subject to the same frequency scales.

(2) *Refining the considerations on frequency localization; frequency localized approximative maps.* In order to deal with the non-atomic (in the frequency sense) derivative components, which may still have large energy, we need to be able to truncate the derivative components arbitrarily in frequency while still retaining the geometric interpretation. Here we shall use arguments just as in the first step to allow us to "build up" the components ψ_α^n from low frequency ones. In the end, we of course need to show that for some subsequence of the ψ_α^n, the frequency support is essentially atomic. If this were to fail, we deduce an a priori bound on $\|\psi_\alpha^n\|_{S((-T_0^n, T_1^n)\times\mathbb{R}^2)}$. Specifically, we show in section 9.3 that judicious choice of an interval J, depending on the position of the Fourier support of the frequency atoms ϕ_α^{na} allows us to truncate the components ϕ_α^n to $P_J\phi_\alpha^n$ while retaining their' geometric significance', i. e. the components $P_J\phi_\alpha^{na}$, $\alpha = 0,1,2$ are also derivative components of a map up to arbitrarily small errors.

(3) *Assuming the presence of a lowest energy non-atomic type component, establish an a priori estimate for its nonlinear evolution.* More precisely, in section 9.4, we replace ϕ_α^n by components $\Phi_\alpha^{nA_0^{(0)}}$, which arise by truncating the frequency support of ϕ_α^n to sufficiently low frequencies such that all frequency atoms with energy above a certain threshold are eliminated. In order to obtain a priori bounds on the evolution of the associated Coulomb components $\Psi_\alpha^{nA_0^{(0)}} = \Phi_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k \Phi_k^{nA_0^{(0)}}}$, we use the previous step to approximate the $\Phi_\alpha^{nA_0^{(0)}}$ by frequency truncated $P_{J_j}\Phi_\alpha^{nA_0^{(0)}}$ for judiciously chosen increasing intervals J_j, whose number only depends on the energy E_{crit}. A finite induction procedure then leads to a priori bounds on the $\Psi_\alpha^{nA_0^{(0)}}$, provided n is chosen large enough (only depending on E_{crit}). Here we already encounter the difficulty that the low frequency components appear to interact strongly with the high-frequency components in the nonlinearity, a stark contrast to the defocussing nonlinear critical wave equation. In particular, in order to 'bootstrap' the bounds on the differences of the Coulomb potentials associated with the $P_{J_j}\Phi_\alpha^{nA_0^{(0)}}$, we have to invoke energy estimates for covariant wave equations of the form $\Box u + 2i\partial^\nu u A_\nu = 0$.

(4) *Bahouri-Gerard II, applied to the first atomic frequency component.* In section 9.6, assuming that we have constructed the first "low frequency approximation" $\Phi_\alpha^{nA_0^{(0)}}$ in the previous step, we need

to filter out the concentration profiles (analogous to the $V^{(j)}$ at the beginning of this subsection) corresponding to the frequency atoms above the minimum energy threshold and at lowest possible frequency. This is where we have to deviate from Bahouri-Gerard: instead of the free wave operator, we need to use the covariant wave operator $\Box_{A^n} = \Box + 2iA^n_\nu \partial^\nu$ to model the asymptotics as $t \to \pm\infty$, where A^n_ν is the Coulomb potential associated with the low frequency approximation $\Phi_\alpha^{nA_0^{(0)}}$. Thus we obtain the concentration profiles as weak limits of the data under the covariant wave evolution. Again a lot of effort needs to be expended on showing that the components we obtain are actually the Coulomb derivative components of Schwartz maps from $\mathbb{R}^2 \to \mathbb{H}^2$, up to arbitrarily small errors in energy. Once we have this, we can then use the result from the stability section in order to construct the time evolution of these pieces and obtain their a priori dispersive behavior.

(5) *Bahouri-Gerard II; completion.* Here we repeat Steps 3 and 4 for the ensuing frequency pieces, to complete the estimate for the ψ^n_α. The conclusion is that upon choosing n large enough, we arrive at a contradiction, unless there is precisely one frequency component and precisely one atomic physical component forming that frequency component. These are the data that then gives rise to the weak wave map with the desired compactness properties.

1.5. **The Kenig-Merle agument.** In [14], [13], Kenig and Merle developed an approach to the global wellposendess for defocusing energy critical semilinear Schrödinger and wave equations; moreover, their argument yields a blowup/global existence dichotomy in the focusing case as well, provided the energy of the wave lies beneath a certain threshold. See [6] for an application of these ideas to wave maps.

Let us give a brief overview of their argument. Consider

$$\Box u + u^5 = 0$$

in $\mathbb{R}^{1+3}$ with data in $\dot{H}^1 \times L^2$. It is standard that this equation is well-posed for small data provided we place the solution in the energy space intersected with suitable Strichartz spaces. Moreover, if I is the maximal interval of existence, then necessarily $\|u\|_{L^8_t(I;L^8_x(\mathbb{R}^3))} = \infty$ and the energy $E(u)$ is conserved.

Now suppose E_{crit} is the maximal energy with the property that all solutions in the above sense with $E(u) < E_{crit}$ exist globally and satisfy $\|u\|_{L^8_t(\mathbb{R};L^8_x(\mathbb{R}^3))} < \infty$. Then by means of the Bahouri-Gerard decomposition, as well as the perturbation theory for this equation one concludes that a critical solution u_C exists on some interval I^* and that $\|u_C\|_{L^8_t(I^*;L^8_x(\mathbb{R}^3))} = \infty$. Moreover, by similar arguments one obtains the crucial property that the set

$$K := \{\big(\lambda^{\frac{1}{2}}(t)u(\lambda(t)(x-y(t)),t), \lambda^{\frac{3}{2}}(t)\partial_t u(\lambda(t)(x-y(t)),t)\big) \,:\, t \in I\}$$

is precompact in $\dot{H}^1 \times L^2(\mathbb{R}^3)$ for a suitable path $\lambda(t), y(t)$. To see this, one applies the Bahouri-Gerard decomposition to a sequence u_n of solutions with energy $E(u_n) \to E_{crit}$ from above. The logic here is that due to the minimality assumption on E_{crit} *only a single limiting profile can arise in* (1.20) up to errors that go to zero in energy as $n \to \infty$. Indeed, if this were not the case then due to fact that the profiles diverge from each other in physical space as $n \to \infty$ one can then apply the perturbation theory to conclude that each of the individual nonlinear evolutions of the limiting profiles (which exist due to the fact that their energies are strictly below E_{crit}) can be superimposed to form a global nonlinear evolution, contradicting the choice of the sequence u_n. The fact that $\ell = 1$ allows one to rescale and re-translate the unique limiting profile to a fixed position in phase space (meaning spatial position and spatial frequency) which then gives the desired nonlinear evolution u_C. The compactness follows by the same logic: assuming that it does not hold, one then obtains a sequence $u_C(\cdot, t_n)$ evaluated at times $t_n \in I^*$ converging to an endpoint of I^* such that for $n \neq n'$, the rescaled and translated versions of $u_C(\cdot, t_n)$ and $u_C(\cdot, t_{n'})$ remain at a minimal positive distance from each other in the energy norm. Again one applies Bahouri-Gerard and finds that $\ell = 1$ by the choice of E_{crit} and perturbation theory. This gives the desired contradiction. The compactness property is of course crucial; indeed, for illustrative purposes suppose that u_C is of the form

$$u_C(t,x) = \lambda(t)^{\frac{1}{2}} U(\lambda(t)(x - x(t)))$$

where $\lambda(t) \to \infty$ as $t \to 1$, say. Then u_C blows up at time $t = 1$ (in the sense that the energy concentrates at the tip of a cone) and

$$\lambda(t)^{-\frac{1}{2}} u_C(\lambda(t)^{-1} x + x(t)) = U(x)$$

is compact for $0 \le t < 1$. Returning to the Kenig-Merle argument, the logic is now to show that u_C acts in some sense like a blow-up solution, at least if I^* is finite in one direction.

The second half of the Kenig-Merle approach then consists of a rigidity argument which shows that a u_C with the stated properties cannot exist. This is done mainly by means of the conservation laws, such as the Morawetz and energy identities. More precisely, the case where I^* is finite at one end is reduced to the self-similar blowup scenario. This, however, is excluded by reducing to the stationary case and an elliptic analysis which proves that the solution would have to vanish. If I^* is infinite, one basically faces the possibility of stationary solutions which are again shown not to exist.

For the case of wave maps, we follow the same strategy. More precisely, our adaptation of the Bahouri-Gerard decomposition to wave maps into $\mathbb{H}^2$ leads to a critical wave map with the desired compactness properties. In the course of our proof, it will be convenient to project the wave map onto a compact Riemann surface $\mathcal{S}$ (so that we can avail ourselves of the *extrinsic formulation* of the wave map equation). However, it will be important to work simultaneously with this object as well as the lifted one which takes its values in $\mathbb{H}^2$ (since it is for the latter that we have a meaningful well-posedness theory for maps with energy data).

The difference from [13] lies mainly with the rigidity part. In fact, in our context the conservation laws are by themselves not sufficient to yield a contradiction. This is natural, since the geometry of the target will need to play a crucial role. As indicated above, the two scenarios that are lead to a contradiction are the self-similar blowup supported inside of a light-cone and the stationary weak wave map, which is of course a weakly harmonic map (which cannot exist since the target $\mathcal{S}$ is compact with negative curvature). The former is handled as follows: in self-similar coordinates, one obtains a harmonic map defined on the disk with the hyperbolic metric and with finite energy (the stationarity is derived as in [13]). Moreover, there is the added twist that one controls the behavior of this map at the boundary in the trace sense (in fact, one shows that this trace is constant). Therefore, one can apply the boundary regularity version of Helein's theorem which was obtained by Qing [35]. Lemaire's theorem [26] then yields the constancy of the harmonic map, whence the contradiction (for a version of this argument under the a priori assumption of regularity all the way to the boundary see Shatah-Struwe [40]).

1.6. **An overview of the paper.** The paper is essentially divided into two parts: the **modified Bahouri-Gerard method** is carried out in its entirety starting with Section 2, and ending with Section 9. Indeed, all that precedes Section 9 leads to this section, which is the core of this paper. The **Kenig-Merle method adapted to Wave Maps** is then performed in the much shorter section 10. We commence by describing in detail the contents of Section 2 to Section 9.

1.6.1. *Preparations for the Bahouri-Gerard process.* As explained above, we describe admissible wave maps $u : R^{2+1} \to \mathbb{H}^2$ mostly in terms of the associated Coulomb derivative components ψ_α. Our goals then are to

- *(1): Develop a suitable functional framework*, in particular a space-time norm $\|\psi\|_{S(\mathbb{R}^{2+1})}$, together with time-localized versions $\|\psi\|_{S([I\times\mathbb{R}^2)}$ for closed time intervals I, which have the property that
$$\limsup_{I\subset\tilde{I}} \|\psi\|_{S(I\times\mathbb{R}^2)} < \infty$$
for some open interval $\tilde{I}$ implies that the underlying wave map u can be extended smoothly and admissibly beyond any endpoint of $\tilde{I}$, provided such exists.
- *(2): Establish an a priori bound* of the form
$$\|\psi\|_{S(I\times\mathbb{R}^2)} \le C(E)$$
for some function $C : \mathbb{R}_+ \to \mathbb{R}_+$ of the energy E. This latter step will be accomplished by the Bahouri Gerard procedure, arguing by contradiction.

We first describe (1) above in more detail: in Section 2, we introduce the norms $\|\cdot\|_{S[k]}$, $\|\cdot\|_{N[k]}$, $k\in\mathbb{Z}$, which are used to control the frequency localized components of ψ and the nonlinear source terms, respectively. The norm $\|\cdot\|_S$ is then obtained by square summation over all frequency blocks. The basic paradigm for establishing estimates on ψ then is to formulate a wave equation

$$\Box\psi = F$$

or more accurately typically in frequency localized form

$$\Box P_0\psi = P_0 F,$$

and to establish bounds for $\|P_0F\|_{N[0]}$ which may then be fed into an energy inequality, see Section 2.3, which establishes the link between the S and N-spaces. In order to be able to estimate the nonlinear source terms F, we need to manipulate the right-hand side of (1.14), making extensive use of (1.15). The precise description of the actual nonlinear source terms that we will use for F is actually rather involved, and given in Section 3. In order to estimate the collection of trilinear as well as higher order terms, we carefully develop the necessary estimates in Sections 4, 5, as well as 6. We note that the estimates in [22], while similar, are not quite strong enough for our purposes, since we need to gain in the largest frequency in case of high-high cascades. This requires us to subtly modify the spaces by comparison to loc. cit. Moreover, the fact that we manage here to build in sharp Strichartz estimates allows us to replace several arguments in [22] by more natural ones, and we opted to make our present account as self-contained as possible.

With the null-form estimates from Sections 4, 5, 6 in hand, we establish the role of $\|\cdot\|_S$ as a "regularity controlling" device in the sense of (1) above in Section 7, see Proposition 7.2. The proof of this reveals a somewhat unfortunate feature of our present setup, namely the fact that working at the level of the differentiated wave map system produces sometimes too many time derivatives, which forces us to use somewhat delicate "randomization" of times arguments. In particular, in the proof of all a priori estimates, we need to distinguish between a "small time" case (typically called Case 1) and a "long time" Case 2, by reference to a fixed frequency scale. In the short time case, one works exclusively in terms of the div-curl system, while in the long-time case, the wave equations start to be essential.

Section 7 furthermore explains the well-posedness theory at the level of the ψ_α, see the most crucial Proposition 7.11. We do not prove this proposition in Section 7, as it follows as a byproduct of the core perturbative Proposition 9.12 in Section 9. Proposition 7.11 and the technically difficult but fundamental Lemma 7.10 allow us to define the "Coulomb wave maps propagation" for a tuple ψ_α, $\alpha=0,1,2$ which are only L^2 functions at time $t=0$, provided the latter are the L^2-limits of the Coulomb components of admissible maps. Indeed, this concept of propagation is independent of the approximating sequence chosen and satisfies the necessary continuity properties.

We also formulate the concept of a "wave map at infinity" at the level of the Coulomb components, see Proposition 7.15 and the following Corollary 7.16. Again the proofs of these results will follow as a byproduct of the fundamental Proposition 9.12 and Proposition 9.30 in the core Section 9.

In Section 8, we develop some auxiliary technical tools from harmonic analysis which will allow us to implement the first stage of the Bahouri Gerard process, namely crystallizing frequency atoms from an "essentially singular" sequence of admissible wave maps. These tools are derived from the imbedding $\dot B^1_{2,\infty}(\mathbb{R}^2)\to \mathrm{BMO}$ as well as weighted (relative to A_p) Coifman-Meyer commutator bounds.

As mentioned before, Section 9 is the core of the present paper. In Section 9.2, starting with an *essentially singular* sequence $\mathbf{u}^n$ of admissible wave maps with deteriorating bounds, i.e., $\|\psi^n_\alpha\|_S\to\infty$ as $n\to\infty$ but with the crucial criticality condition $\lim_{n\to\infty}E(\mathbf{u}^n)=E_{crit}$, we show that the derivative components ϕ^n_α may be decomposed as a sum

$$\phi^n_\alpha = \sum_{a=1}^{A}\phi^{na}_\alpha + w^{nA}_\alpha$$

where the ϕ^{na}_α are derivative components of admissible wave maps which have frequency supports "drifting apart" as $n\to\infty$, while the error w^{nA}_α satisfies

$$\limsup_{n\to\infty}\|w^{nA}_\alpha\|_{\dot B^0_{2,\infty}}<\delta$$

provided $A \geq A_0(\delta)$ is large enough.
In Section 9.3, we then select a number of "principal" frequency atoms ϕ^{na}, $a = 1, 2, \ldots, A_0$, as well as a (potentially very large) collection of "small atoms" ϕ^{na}, $a = A_0 + 1, \ldots, A$. We order these atoms by the frequency scale around which they are supported starting with those of the lowest frequency. The idea now is as follows: under the assumption that there are at least two frequency atoms, or else in case of only one frequency atom that it has energy $< E_{crit}$, we want to obtain a contradiction to the essential criticality of the underlying sequence $\mathbf{u}^n$. To achieve this, we define in Section 9.3 sequence of approximating wave maps, which are essentially obtained by carefully truncating the initial data sequence ϕ^{na} in frequency space.
In Section 9.4, we establish an a priori bound for the lowest frequency approximating map which comprises all the minimum frequency small atoms as well as the component of the small Besov error of smallest frequency, see Proposition 9.9. The proof of this follows again by truncating the data suitably in frequency space, and applying an inductive procedure to a sequence of approximating wave maps. This hinges crucially on the core perturbative result Proposition 9.12, which plays a fundamental role in the paper. The main technical difficulty encountered in the proof of the latter comes from the issue of **divisibility**: let us be given a schematically written expression

$$\partial^\nu \epsilon A_\nu$$

which is *linear* in the perturbation (so that we cannot perform a bootstrap argument based solely on the smallness on ϵ itself), while A_ν denotes some null-form depending on a priori controlled components ψ. "Divisibility" means the property that upon suitably truncating time into finitely many intervals I_j *whose number only depends on* $\|\psi\|_S$, one may bound the expression by

$$\|\partial^\nu \epsilon A_\nu\|_{N(I_j \times \mathbb{R}^2)} \ll \|\epsilon\|_S$$

In other words, by shrinking the time interval, we ensure that we can iterate the term away. While this would be straightforward provided we had an estimate for $\|A_\nu\|_{L_t^1 L_x^\infty}$ (which is possible in space dimensions $n \geq 4$), in our setting, the spaces are much too weak and complicated. Our way out of this impasse is to build those terms for which we have no obvious divisibility into the linear operator, and thereby form a new operator

$$\Box_A \epsilon := \Box \epsilon + 2i\partial^\nu \epsilon A_\nu$$

with a magnetic potential term. Fortunately, it turns out that if A_ν is supported at much lower frequencies than ϵ (which is precisely the case where divisibility fails), one can establish an approximate energy conservation result, which in particular gives a priori control over a certain constituent of $\|\cdot\|_S$. With this in hand, one can complete the bootstrap argument, and obtain full control over $\|\epsilon\|_S$.
Having established control over the lowest-frequency "essentially non-atomic" approximating wave map in Section 9.4, we face the task of "adding the first large atomic component", ϕ^{n1}. It is here that we have to depart crucially from the original method of Bahouri-Gerard: instead of studying the free wave evolution of the data, we extract concentration cores by applying the "twisted" covariant evolution associated with

$$\Box_{A^n} u = 0,$$

which is essentially defined as above. The key property that makes everything work is an almost exact energy conservation property associated with its wave flow. This is a rather delicate point, and uses the Hamiltonian structure of the covariant wave flow. It then requires a fair amount of work to show that the profile decomposition at time $t = 0$ in terms of covariant free waves is "geometric", in the sense that the concentration profiles can indeed by approximated by the Coulomb components of admissible maps, up to a constant phase shift, see Proposition 9.24.
Finally, in Proposition 9.30 we show that we may evolve the data including the first large frequency atom, provided all concentration cores have energy strictly less than E_{crit}.
As most of the work has been done at this point, adding on the remaining frequency atoms in Section 9.9 does not provide any new difficulties, and can be done by the methods of the preceding sections.
In conjunction with the results of Section 7, we can then infer that given an essentially singular sequence of wave maps $\mathbf{u}^n$, we may select a subsequence of them whose Coulomb components ψ_α^n, up to re-scalings and translations, converge to a limiting object $\Psi_\alpha^\infty(t, x)$, which is well-defined on some interval $I \times \mathbb{R}^2$ where

I is either a finite time interval or (semi)-infinite, and the limit of the Coulomb components of admissible maps there. Moreover, most crucially for the sequel, $\Psi_\alpha^\infty(t,x)$ satisfies a remarkable *compactness property*, see Proposition 9.36. This sets the stage for the method of Kenig-Merle, which we adopt to the context of wave maps in section 10.

2. The spaces $S[k]$ and $N[k]$

Sections 2–5 develop the functional framework needed to prove the energy and dispersive estimates required by the wave map system (1.12)–(1.14). The Banach spaces which appear in this context were introduced by Tataru [61], but were specified in this form by Tao [56], and developed further by Krieger [22]. We will largely follow the latter reference although there is much overlap with [56]. We emphasize that this section is self-contained. The spatial dimension is two throughout.

2.1. **Preliminaries.** As usual, P_k denotes a Littlewood-Paley projection[4] to frequencies of size 2^k. More precisely, let m_0 be a nonnegative smooth, even, bump function supported in $|\xi| < 4$ and set $m(\xi) := m_0(\xi) - m_0(2\xi)$. Then

$$\sum_{k\in\mathbb{Z}} m(2^k\xi) = 1 \quad \forall\, \xi \in \mathbb{R}^2 \setminus \{0\}$$

and $\widehat{P_k f}(\xi) := m(2^{-k}\xi)\hat{f}(\xi)$.

In the sequel, we shall call a function f **adapted** to k, provided its Fourier transform is supported at frequency $\sim 2^k$.

The operator Q_j projects to modulation 2^j, i.e.,

$$\widehat{Q_j\phi}(\xi,\tau) := m(2^{-j}(|\tau| - |\xi|))\widehat{\phi}(\tau,\xi)$$

with $\hat{\ }$ referring to the space-time Fourier transform. Similarly,

$$\widehat{Q_j^{\pm}\phi}(\xi,\tau) := m\big(2^{-j}(|\tau| - |\xi|)\big)\chi_{[\pm\tau>0]}\widehat{\phi}(\tau,\xi)$$

Then the relevant $\dot{X}_k^{s,p,q}$ spaces here are defined as

$$\|\phi\|_{\dot{X}_k^{s,p,q}} := 2^{sk}\Big(\sum_j 2^{jpq}\|P_k Q_j\phi\|_{L_t^2 L_x^2}^q\Big)^{\frac{1}{q}}$$

If $P_k\phi = \phi$, then $\|\phi\|_{L_t^\infty L_x^2} \lesssim \|\phi\|_{\dot{X}_k^{0,\frac{1}{2},1}}$ as well as $\|\phi\|_{L_{t,x}^\infty} \lesssim \|\phi\|_{\dot{X}_k^{1,\frac{1}{2},1}}$.

In what follows, $\mathcal{C}_\ell$ is a collection of caps $\kappa \subset S^1$ of size $C^{-1}2^\ell$ and finite overlap (uniformly bounded in ℓ and with C some large absolute constant). There is an associated smooth partition of unity $\sum_{\kappa\in\mathcal{C}_\ell} a_\kappa(\omega) = 1$ for all $\omega \in S^1$, as well as projections $\widehat{P_\kappa f}(\xi) := a_\kappa\big(\widehat{\xi}\big)\hat{f}(\xi)$ where $\widehat{\xi} := \frac{\xi}{|\xi|}$. By construction, $P_{k,\kappa} := P_k \circ P_\kappa$ is a projection to the "rectangle"

$$R_{k,\kappa} := \{|\xi| \sim 2^k,\ \widehat{\xi} \in \kappa\} \tag{2.1}$$

in Fourier space. For space-time functions F we shall follow the convention that

$$P_\kappa F = [a_\kappa(\hat{\xi})\chi_{[\tau>0]}\hat{F}(\xi,\tau)]^\vee + [a_\kappa(-\hat{\xi})\chi_{[\tau<0]}\hat{F}(\xi,\tau)]^\vee$$

We will also encounter other rectangles R which are obtained by dividing $R_{k,\kappa}$ in the radial direction into 2^{-m} many subrectangles of comparable size where $m < 0$ is some integer parameter (it will suffice for us to consider $\ell \le m \le 0$ where $\kappa \in \mathcal{C}_\ell$). The collection of these rectangles will be denoted by $\mathcal{R}_{k,\kappa,m}$, and we introduce projections P_R so that $\sum_{R\in\mathcal{R}_{k,\kappa,m}} P_R = P_{k,\kappa}$. Figure 1 exhibits such a collection of rectangles. The sector $ABCD$ is of length 2^k and width $2^{\ell+k}$, whereas the shorter segments $\overline{AP_1}$, $\overline{P_1P_2}$ etc. are of length 2^{k+m}.

---

[4]Strictly speaking, these are not true projections since $P_k^2 \neq P_k$, but we shall nevertheless follow the customary abuse of language of referring to them as projections. The same applies to smooth localizers to other regions in Fourier space.

FIGURE 1. Rectangles

We shall frequently use Bernstein's inequality: if $\mathrm{supp}(\hat{\phi}) \subset \Omega$, where $\Omega \subset \mathbb{R}^2$ is measurable, then $\|\phi\|_q \lesssim |\Omega|^{\frac{1}{p}-\frac{1}{q}}\|\phi\|_p$ for any choice of $1 \le p \le q \le \infty$. We shall also require the following variant of Bernstein's $L^2 \to L^\infty$ bound, which is obtained by combining the standard form of this bound with the $L_t^4 L_x^\infty(\mathbb{R}^{1+2})$-Strichartz estimate for the wave equation. This type of estimate appears in [57], but the following formulation is from [22], which involves one further localization on the Fourier side. We present the proof for the sake of completeness.

Lemma 2.1. *Let $\mathcal{D}_{k,\ell}$ be a cover of $\{|\xi| \sim 2^k\}$ by disks of radius $2^{k+\ell}$. Then for all $j \le k$,*

$$\Big(\sum_{c\in\mathcal{D}_{k,\ell}} \|P_c Q_j \phi\|_{L_t^2 L_x^\infty}^2\Big)^{\frac{1}{2}} \lesssim 2^{\frac{\ell}{2}} 2^{\frac{3k}{4}} 2^{\frac{j-k}{4}} \|\phi\|_{L_t^2 L_x^2} \tag{2.2}$$

for any ϕ which is adapted to k.

Proof. We follow the argument in [56], but use the small-scale Strichartz estimate of Klainerman-Tataru at a crucial place, see Lemma 2.17 below. First, set $j = 0$, whence $k \gg 1$. Construct a Schwartz function $a(t)$ whose Fourier transform is supported in $|\tau| \ll 1$, and which satisfies

$$1 = \sum_{s\in\mathbb{Z}} a^3(t-s)$$

for all $t \in \mathbb{R}$. Then

$$\begin{aligned}\|P_c Q_0 \psi\|_{L_t^2 L_x^\infty} &\le \|\sum_s a^3(t-s) P_c Q_0 \psi\|_{L_t^2 L_x^\infty} \\ &\lesssim (\sum_s \|a^2(t-s) P_c Q_0 \psi\|_{L_t^2 L_x^\infty}^2)^{\frac{1}{2}} \lesssim (\sum_s \|a(t-s) P_c Q_0 \psi\|_{L_t^4 L_x^\infty}^2)^{\frac{1}{2}}\end{aligned}$$

Now one notes that the function $a(t-s) P_c Q_0 \psi$ satisfies almost the same assumptions about modulation (~ 1) and frequency localization as $P_c Q_0 \psi$. Therefore, we can apply the improved Strichartz estimate of Klainerman-Tataru [21] to estimate

$$\|a(t-s) P_c Q_0 \psi\|_{L_t^4 L_x^\infty} \lesssim 2^{\frac{3k}{4}} 2^{\frac{l}{2}} \|P_c \psi\|_{\dot{X}_k^{0,\frac{1}{2},\infty}}$$

Thus

$$\begin{aligned}\|P_c Q_0 \psi\|_{L_t^2 L_x^\infty} &\lesssim 2^{\frac{3k}{4}} 2^{\frac{l}{2}} (\sum_{s\in\mathbb{Z}} \|a(t-s) Q_0 P_c \psi\|_{\dot{X}_k^{0,\frac{1}{2},\infty}}^2)^{\frac{1}{2}} \\ &\lesssim 2^{\frac{3k}{4}} 2^{\frac{l}{2}} \|P_c Q_0 \psi\|_{L_t^2 L_x^2}\end{aligned}$$

The lemma follows via Plancherel's theorem. □

The previous proof was based on the following small-scale version of the usual $L_t^4 L_x^\infty$-Strichartz estimate. It was obtained by Klainerman and Tataru [21].

Lemma 2.2. *With $\mathcal{D}_{k,\ell}$ as above, one has*

$$\Big(\sum_{c \in \mathcal{D}_{k,\ell}} \|P_c e^{it|\nabla|} f\|^2_{L_t^4 L_x^\infty} \Big)^{\frac12} \lesssim 2^{\frac{\ell}{2}} 2^{\frac{3k}{4}} \|f\|_{L_x^2} \tag{2.3}$$

for any k-adapted f. In particular,

$$\Big(\sum_{c \in \mathcal{D}_{k,\ell}} \|P_c \phi\|^2_{L_t^4 L_x^\infty} \Big)^{\frac12} \lesssim 2^{\frac{\ell}{2}} \|\phi\|_{\dot{X}_k^{\frac34, \frac12, 1}} \tag{2.4}$$

for any Schwartz function ϕ which is adapted to k.

2.2. **The null-frame spaces.** In contrast to sub-critical $\dot{H}^s(\mathbb{R}^2)$ data with $s > 1$, it is well-known that $\dot{X}^{s,b}$ spaces do not suffice in the critical case $s = 1$. Following the aforementioned references, we now develop Tataru's null-frame spaces which will provide sufficient control over the nonlinear interactions in the wave-map system. For fixed[5] $\omega \in S^1$ define

$$\theta_\omega^\pm := (1, \pm\omega)/\sqrt{2}, \quad t_\omega := (t,x) \cdot \theta_\omega^+, \quad x_\omega := (t,x) - t_\omega \theta_\omega^+ \tag{2.5}$$

which are the coordinates defined by a generator on the light-cone. Recall that a plane wave traveling in direction $-\omega \in S^{n-1}$ is a function of the form $h(x \cdot \omega + t)$ (and h sufficiently smooth). We write a free wave ϕ as a superposition of such plane waves: with $\kappa \subset S^1$ and $P_{k,\kappa}$ the projection to $|\xi| \sim 2^k$ and $\widehat{\xi} \in \kappa$ as defined above,

$$\begin{aligned} P_{k,\kappa}\phi(t,x) &= \int_{[|\xi| \sim 2^k]} e^{i(t|\xi| + x\cdot\xi)} \widehat{P_{k,\kappa} f}(\xi)\, d\xi \\ &= \int_\kappa \int_{[r \sim 2^k]} e^{ir(x\cdot\omega + t)} \hat{f}(r\omega)\, r\, dr d\omega \\ &= \int_\kappa \psi_{k,\omega}(t + x\cdot\omega)\, d\omega, \end{aligned} \tag{2.6}$$

where

$$\psi_{k,\omega}(s) := \int_{[r\sim 2^k]} e^{irs} \hat{f}(r\omega) r\, dr$$

The argument of $\psi_{k,\omega}$ in (2.6) is $\sqrt{2}\, t_\omega$, whence

$$\int_\kappa \|\psi_{k,\omega}\|_{L^2_{t_\omega} L^\infty_{x_\omega}}\, d\omega \lesssim |\kappa|^{\frac12} 2^{\frac{k}{2}} \|P_{k,\kappa} f\|_2 \tag{2.7}$$

We now define the following pair of norms[6]

$$\|G\|_{\mathrm{NFA}[\kappa]} := \inf_{\omega \notin 2\kappa} \mathrm{dist}(\omega,\kappa)^{-1} \|G\|_{L^1_{t_\omega} L^2_{x_\omega}} \tag{2.8}$$

$$\|\phi\|_{\mathrm{PWA}[\kappa]} := \inf_{\omega \in \kappa} \|\phi\|_{L^2_{t_\omega} L^\infty_{x_\omega}} \tag{2.9}$$

which are well-defined for general Schwartz functions. The notation here derives from *null-frame* and *plane wave*, respectively. The quantities defined in (2.8) and (2.9) are not norms — in fact, not even pseudo-norms — because they violate the triangle inequality due to the infimum. This indicates that we should be using (2.8) and (2.9) to define *atomic* Banach spaces (which is why we appended "A" in the norms above). First, recall from (2.6) that

$$P_{k_1,\kappa}\phi(t,x) = \int_\kappa \psi_{k,\omega}(\sqrt{2}\, t_\omega)\, d\omega$$

[5]Henceforth, ω will always be a unit vector in the plane.

[6]The $\mathrm{dist}(\omega,\kappa)^{-1}$ factor in the NFA[κ]-norm arises because of a geometric property of the cone, see the proof of Lemma 2.4.

Then (2.7) suggests that we define

$$\|P_{k_1,\kappa}\phi\|_{\mathrm{PW}[\kappa]} := \int_\kappa \|\psi_{k,\omega}\|_{L^2_{t_\omega} L^\infty_{x_\omega}}\, d\omega = \int_\kappa \|\psi_{k,\omega}\|_{\mathrm{PWA}[\kappa]}\, d\omega$$

In other words, PW[κ] is the completion of the space of all functions ϕ which can be written in the form

$$\phi = \sum_j \lambda_j \psi_j, \quad \sum_j |\lambda_j| < \infty, \quad \|\psi_j\|_{\mathrm{PWA}[\kappa]} \le 1 \tag{2.10}$$

where $\lambda_j \in \mathbb{C}$ and ψ_j are Schwartz functions, say. The norm of any such ϕ in PW[κ] is then simply the infimum of $\sum_j |\lambda_j|$ over all representations as in (2.10). By Hölder's inequality we now obtain the simple but crucial estimate

$$\|\phi F\|_{\mathrm{NFA}[\kappa]} \le \mathrm{dist}(\kappa,\kappa')^{-1} \|\phi\|_{\mathrm{PWA}[\kappa']} \|F\|_{L^2_t L^2_x}$$

provided ϕ is a PWA[κ]-atom. This suggests that we also define NF[κ] as the atomic space obtained from NFA[κ] as usual: the atoms of NF[κ] are functions ϕ for which there exists $\omega \notin 2\kappa$ such that $\|\phi\|_{L^1_{t_\omega} L^2_{x_\omega}} \le \mathrm{dist}(\omega,\kappa)$. The previous estimate then implies the bound

$$\|\phi F\|_{\mathrm{NF}[\kappa]} \le \mathrm{dist}(\kappa,\kappa')^{-1} \|\phi\|_{\mathrm{PW}[\kappa']} \|F\|_{L^2_t L^2_x} \tag{2.11}$$

The dual space NF[κ]* is characterized by the norm

$$\|\phi\|_{\mathrm{NF}[\kappa]^*} = \sup_{\omega \notin 2\kappa} \mathrm{dist}(\omega,\kappa) \|\phi\|_{L^\infty_{t_\omega} L^2_{x_\omega}} < \infty$$

We now turn to defining the spaces which hold the wave maps.

Definition 2.3. *Let ϕ be a Schwarz function with* $\mathrm{supp}(\hat\phi) \subset \{\xi \in \mathbb{R}^2 \;:\; |\xi| \sim 2^k\}$. *Henceforth, we shall call such a ϕ* adapted to k. *Define*

$$\|\phi\|_{S[k,\kappa]} := \|\phi\|_{L^\infty_t L^2_x} + |\kappa|^{-\frac12} 2^{-\frac{k}{2}} \|\phi\|_{\mathrm{PW}[\kappa]} + \|\phi\|_{\mathrm{NF}^*[\kappa]} \tag{2.12}$$

$$\|\phi\|_{S[k]} := \|\phi\|_{L^\infty_t L^2_x} + \|Q_{\le k+2}\phi\|_{\dot X^{0,\frac12,\infty}} + \|Q_{\ge k}\phi\|_{\dot X^{-\frac12+\varepsilon,1-\varepsilon,2}} \tag{2.13}$$

$$+ \sup_{j\in\mathbb{Z}} \sup_{\ell\le 0} 2^{-(\frac12-\varepsilon)\ell} 2^{-\frac{3k}{4}} \Big(\sum_{c\in\mathcal{D}_{k,\ell}} \|Q_{<j} P_c \phi\|^2_{L^4_t L^\infty_x} \Big)^{\frac12} \tag{2.14}$$

$$+ \sup_{\pm} \sup_{\ell\le -100} \sup_{\ell\le m\le 0} \Big(\sum_{\kappa\in\mathcal{C}_\ell} \sum_{R\in\mathcal{R}_{k,\pm\kappa,m}} \|P_R Q^{\pm}_{\le k+2\ell}\, \phi\|^2_{S[k,\kappa]} \Big)^{\frac12} \tag{2.15}$$

Here P_κ and P_R are as above, and $\varepsilon > 0$ is a small number ($\varepsilon = \frac{1}{10}$ is sufficient).

The factors $|\kappa|^{-\frac12} 2^{-\frac{k}{2}}$ in (2.12) are from (2.7). By inspection, the norm of $S[k,\kappa]$ is translation invariant, and

$$\|f\phi\|_{S[k,\kappa]} \le \|f\|_{L^\infty_{tx}} \|\phi\|_{S[k,\kappa]} \tag{2.16}$$

One has the following scaling property:

$$\|\phi\|_{S[k]} = \lambda \|\phi(\lambda\cdot)\|_{S[k+m]}, \qquad \lambda = 2^m,\ m\in\mathbb{Z} \tag{2.17}$$

It will be technically convenient to allow noninteger k in Definition 2.3. The only change required for this purpose is to allow $j,\ell,m \in \mathbb{R}$ in (2.14) and (2.15). In that case one has

$$\|\phi\|_{S[k]} = \lambda \|\phi(\lambda\cdot)\|_{S[k+\log_2\lambda]} \qquad \forall \lambda > 0 \tag{2.18}$$

Later we will need to address the question whether $\|P_k\phi\|_{S[k+h]}$ is continuous in h near $h=0$ for a fixed Schwartz function ϕ. Henceforth, we shall use the operator $I := \sum_{k\in\mathbb{Z}} P_k Q_{\le k}$ and $I^c := 1 - I$ (we will also use $Q_{\le k+C}$ instead of $Q_{\le k}$). Moreover, we refer to functions which belong to the range of I as "hyperbolic" and to those in the range of I^c as "elliptic". Since

$$\|Q_{\ge k} P_k \phi\|_{\dot X_k^{0,\frac12,1}} \lesssim \|Q_{\ge k} P_k \phi\|_{\dot X_k^{-\frac12+\varepsilon,1-\varepsilon,2}}$$

one concludes that the energy norm $L^\infty_t L^2_x$ in (2.13) as well as the Strichartz norm of (2.14) are controlled by the final norm of (2.13) for the case of elliptic functions (for the Strichartz norm use Lemma 2.2).

We first verify that temporally truncated free waves lie in these spaces (with an imbedding constant that does not depend on the length of the truncation interval).

Lemma 2.4. *Let $\kappa \subset S^1$ be arbitrary. Then*

$$\|\phi\|_{S[k,\kappa]} \lesssim \|P_{k,\kappa}\phi\|_{\dot X^{0,\frac12,1}} \tag{2.19}$$

as well as

$$\|P_k Q_{\le k}\phi\|_{S[k]} \lesssim \|P_k Q_{\le k}\phi\|_{\dot X^{0,\frac12,1}} \tag{2.20}$$

In particular, if f is adapted to k, then

$$\|\chi(t/T)e^{it\sqrt{-\Delta}}f\|_{S[k]} \le C\|f\|_{L^2} \tag{2.21}$$

with a constant that depends on the Schwartz function χ but not on $T \ge 2^{-k}$.

Proof. We assume that ϕ is an $\dot X^{0,\frac12,1}$-atom with $P_{0,\kappa}\phi = \phi$. Then from Plancherel's theorem and Minkowski's and Hölder's inequalities,

$$\begin{aligned}
\|\phi\|_{L^\infty_t L^2_x} &\lesssim \|\hat\phi\|_{L^2_\xi L^1_\tau} \lesssim 2^{\frac{j}{2}}\|\phi\|_{L^2_t L^2_x}\\
\|\phi\|_{L^2_{t_\omega} L^\infty_{x_\omega}} &\lesssim (2^j|\kappa|)^{\frac12}\|\hat\phi\|_{L^2_{\xi_\omega} L^2_{\tau_\omega}} = (2^j|\kappa|)^{\frac12}\|\phi\|_{L^2_t L^2_x}\\
\|\phi\|_{L^\infty_{t_\omega} L^2_{x_\omega}} &\lesssim \frac{2^{\frac{j}{2}}}{\mathrm{dist}(\kappa,\omega)}\|\phi\|_{L^2_{\xi_\omega} L^2_{\tau_\omega}} \qquad (2.22)
\end{aligned}$$

In the final estimate (2.22) we used that $\sphericalangle(\ell_\omega, T_{\omega'}) \sim \sphericalangle(\omega,\omega')^2$ where ℓ_ω is the line oriented along the generator parallel to $(1,\omega)$ and $T_{\omega'}$ is the tangent plane to the cone which touches the cone along the generator $\ell_{\omega'}$. To establish (2.20) we begin with

$$\sup_{\ell \le -100}\Big(\sum_{\kappa\in\mathcal C_\ell}\sum_{R\in\mathcal R_{k,\pm\kappa,\lambda}}\|P_R Q^{\pm}_{\le 2\ell}\phi\|^2_{\dot X_0^{0,\frac12,1}}\Big)^{\frac12} \lesssim \|\phi\|_{\dot X_0^{0,\frac12,1}}$$

which is obvious from orthogonality of the $P_{0,\pm\kappa}$. In view of (2.19), this bound yields the square function in (2.15). The energy is controlled via the imbedding $\|\phi\|_{L^\infty L^2} \lesssim \|\phi\|_{\dot X^{0,\frac12,1}}$, whereas the Strichartz component of $S[k]$ is controlled by Lemma 2.2.

Finally, the statement concerning the free wave reduces to the case $k=0$ for which we need to verify the bound

$$\begin{aligned}
&\sum_{j\in\mathbb Z} 2^{\frac{j}{2}}\|T\hat\chi(T|\tau\pm|\xi||)m(2^{-j}|\tau\pm|\xi||)\hat f(\xi)\|_{L^2_\tau L^2_\xi} +\\
&+\Big(\sum_{j\in\mathbb Z}2^{2j}\|T\hat\chi(T|\tau\pm|\xi||)m(2^{-j}|\tau\pm|\xi||)\hat f(\xi)\|^2_{L^2_\tau L^2_\xi}\Big)^{\frac12} \lesssim \|f\|_2
\end{aligned}$$

which are both clear provided $T \ge 1$ due to the rapid decay of $\hat\chi$. □

Naturally, $S[k]$ contains more general functions than just free waves. One way of obtaining such functions is to take $\phi = \Box^{-1}F$, in other words from the Duhamel formula. We will study this in much greater generality in the context of the energy estimate below, but for now we take F to be a Schwartz function.

Remark 2.5. *The bounded function ϕ defined via its Fourier transform*

$$\widehat\phi(\tau,\xi) = \chi_1(\xi)\chi_2(|\xi|-|\tau|)(|\xi|-|\tau|)^{-1}$$

belongs to $S[0]$ but is not a truncated free wave. Here χ_1 is a smooth cut-off to $|\xi|\sim 1$, and $\chi_2(u)$ is a smooth cut-off to $|u| < 1/10$. We leave it to the reader to construct other functions which lie in $S[0]$ and which are not (truncated) free waves.

The following basic estimates will be used repeatedly:

- if ϕ is adapted to k, then

$$\|Q_j\phi\|_{L^2L^2} \lesssim \min(2^{-(j-k)(\frac{1}{2}-\varepsilon)}, 1)2^{-\frac{j}{2}}\|\phi\|_{S[k]} \tag{2.23}$$

$$\|Q_j\phi\|_{L^2L^\infty} \lesssim 2^k 2^{\frac{j-k}{4}\wedge 0}\min(2^{-(j-k)(\frac{1}{2}-\varepsilon)}, 1)2^{-\frac{j}{2}}\|\phi\|_{S[k]} \tag{2.24}$$

  This follows from the $\dot{X}^{s,b,q}$ components of the $S[k]$-norms, as well as the improved Bernstein's inequality of Lemma 2.1.
- The duality between $\mathrm{NF}[\kappa]$ and $\mathrm{NF}^*[\kappa]$ implies

$$|\langle\phi, F\rangle| \lesssim \|\phi\|_{S[k,\kappa]}\|F\|_{\mathrm{NF}[\kappa]} \tag{2.25}$$

In what follows, $\Theta := \mathrm{sign}(\tau)\widehat{\xi}$, and for any $\omega \in S^1$, Π_ω denotes the orthogonal projection onto $\mathrm{NP}(\omega) := \theta_\omega^\perp$ (the *null-plane* of ω).

Lemma 2.6. *The projection Π_ω satisfies the following properties:*

- *Let $\mathcal{F} \subset \mathcal{C}_\ell$ be a collection of disjoint caps. Suppose that $\omega \in S^1$ satisfies $\mathrm{dist}(\omega,\kappa) \in [\alpha, 2\alpha]$ for any $\kappa \in \mathcal{F}$ where $\alpha > 2^\ell$ is arbitrary but fixed. Define*[7]

$$\mathcal{T}_{\kappa,\alpha} := \big\{(\tau,\xi) \;:\; |\xi| \sim 1,\ \Theta \in \kappa,\ ||\xi| - |\tau|| \lesssim \alpha 2^\ell\big\} \tag{2.26}$$

  Then $\{\Pi_\omega(\mathcal{T}_{\kappa,\alpha})\}_{\kappa\in\mathcal{F}} \subset \mathrm{NP}(\omega)$ have finite overlap, i.e.,

$$\sum_{\kappa\in\mathcal{F}} \chi_{\Pi_\omega(\mathcal{T}_{\kappa,\alpha})} \le C$$

  where C is some absolute constant.
- *Let*

$$\mathcal{S} := \Big\{(\pm|\xi|, \xi) \;:\; \xi \in \mathbb{R}^2,\ \widehat{\xi} \in \pm\kappa\Big\}$$

  be a sector on the light-cone where $\kappa \subset S^1$ is any cap. Furthermore, let $\omega \notin 2\kappa$ and $\widetilde{\mathcal{S}} := \Pi_\omega(\mathcal{S})$. Then on $\widetilde{\mathcal{S}}$ the Jacobian $\frac{\partial\xi}{\partial\xi_\omega}$ satisfies

$$\Big|\frac{\partial\xi}{\partial\xi_\omega}\Big| \sim d(\omega,\kappa)^{-2} \tag{2.27}$$

  The same holds on $\Pi_\omega(\mathcal{S}_a)$ where

$$\mathcal{S}_a := \Big\{(\pm|\xi| + a, \xi) \;:\; \xi \in \mathbb{R}^2,\ |\xi| \sim 1,\ \widehat{\xi} \in \pm\kappa\Big\}$$

  provided a is fixed with $|a| \le |\kappa| d(\omega,\kappa)$.

Proof. Denote

$$\mathcal{S}_\kappa := \{s(1,\omega') + \rho(1,-\omega') \;:\; \omega' \in \kappa,\ 1 \le s \le 2,\ |\rho| < h\}$$

where h will be determined. Then

$$\{\Pi_\omega(\mathcal{S}_\kappa)\}_{\kappa\in\mathcal{F}} = \{s\vec{v} + \rho\vec{w} \;:\; \omega' \in \kappa,\ 1 \le s \le 2,\ |\rho| < h\}$$

where

$$\vec{v} := \vec{v}(\omega,\omega') = (1,\omega') - \lambda(1,\omega), \quad \vec{w} := (1,-\omega') - \mu(1,\omega)$$

with $\lambda = \frac{1}{2}(1+\omega\cdot\omega')$, $\mu = \frac{1}{2}(1-\omega\cdot\omega')$. Recall that $\mathrm{dist}(\omega,\kappa) \sim \mathrm{dist}(\omega,\kappa') =: \alpha$ where $\kappa \in \mathcal{F}$ is arbitrary. Moreover, $\mathrm{diam}(\kappa) \sim 2^\ell =: \beta$. One checks that

$$|\vec{v}| = \sqrt{2(1-\lambda^2)} \sim \sqrt{\mu} \sim \alpha$$

Furthermore, $\partial\vec{v} := (0,\omega'^\perp) - \frac{1}{2}\omega\cdot\omega'^\perp(1,\omega)$ denotes the derivative $\partial_{\theta'}\vec{v}$ where we have written $\omega' = e^{i\theta'}$. Then $|\partial\vec{v}| \sim 1$ and

$$\vec{v}\wedge\partial\vec{v} = (\mu, \omega' - \omega + \mu\omega) \wedge (0,\omega'^\perp) - \frac{1}{2}\omega\cdot\omega'^\perp(1,\omega')\wedge(1,\omega)$$

[7] An important detail here is that these dimensions deviate from the usual wave-packets of dimension $1\times 2^\ell \times 2^{2\ell}$.

FIGURE 2. The projected sectors

satisfies $|\vec{v} \wedge \partial \vec{v}| \sim \alpha^2$. In conjunction with $|\vec{v}| \sim \alpha$ this implies that $|\sphericalangle(\vec{v}, \partial\vec{v})| \sim \alpha$. Since

$$|\vec{v}(\omega, \omega') - \vec{v}(\omega, \omega'')| \gtrsim |\omega' - \omega''| \quad \forall\, \omega', \omega'' \in \kappa'$$

it follows that

$$\mathrm{dist}(\sigma(\omega, \omega'), \sigma(\omega, \omega'')) \gtrsim \alpha|\omega' - \omega''| \tag{2.28}$$

where

$$\sigma(\omega, \omega') := \{s\vec{v}(\omega, \omega') \;:\; 1 \le s \le 2\}$$

Therefore, one needs to take $h = \alpha\beta$ to insure the property of finite overlap of the projections. This is optimal, since one can check that $\vec{v}$ and $\vec{w}$ always satisfy $|\cos(\sphericalangle(\vec{v}, \vec{w})| \le \frac{1}{2}$. In Figure 2 the left-hand side depicts four sectors as they would appear on the light-cone, whereas the right-hand side is the projected configuration in $\mathrm{NP}(\omega)$ with $A' := \Pi_\omega(A)$ etc. Note that the segments $\overline{A'B'}$ as well as $\overline{A'P'}$, $\overline{P'Q'}$, $\overline{Q'R'}$, $\overline{R'B'}$ have lengths comparable to the corresponding ones on the left, i.e., $\overline{AB}$ etc., whereas the lengths of $\overline{A'D'}$, $\overline{B'C'}$ are those of $\overline{AD}$ and $\overline{BC}$ contracted by the factor α. Finally, we have shown that $\sphericalangle(A'B'C') \sim \alpha$ (and similarly for the angles at the points P', Q', R') so that the height of the parallelogram $A'P'X'D'$ is proportional to α times the length of $\overline{A'P'}$, see (2.28).
The second statement of the lemma follows from the consideration of the preceding paragraph. □

As a consequence of Lemma 2.6, we now show that the square-function in (2.15) can always be refined in terms of the angle.

Lemma 2.7. *Let $\mathcal{F} \subset \mathcal{C}_\ell$ be a collection of disjoint caps and let $\kappa' \in \mathcal{C}_{\ell'}$ be a cap with $\bigcup_{\mathcal{F}} \kappa \subset \kappa'$. Suppose further that for every $\kappa \in \mathcal{F}$ there is a Schwartz function ϕ_κ adapted to $k \in \mathbb{Z}$ and which is supported on*

$$\mathcal{T}_{\kappa,k} := \big\{\Theta := \mathrm{sign}(\tau)\widehat{\xi} \in \kappa,\ ||\xi| - |\tau|| \lesssim 2^{\ell+\ell'+k}\big\}$$

with some $k \in \mathbb{Z}$. Then

$$\Big\| \sum_{\kappa\in\mathcal{F}} \phi_\kappa \Big\|_{S[k,\kappa']} \le C\Big(\sum_{\kappa\in\mathcal{F}} \|\phi_\kappa\|^2_{S[k,\kappa]}\Big)^{\frac{1}{2}}$$

with some absolute constant C.

Proof. First, one may take $k = 0$ and $\tau > 0$ (the latter by conjugation symmetry). The $L^\infty L^2$-component of (2.12) satisfies the required property due to orthogonality, whereas the $\mathrm{PW}[\kappa]$-component is reduced to Cauchy-Schwarz (via the $|\kappa|^{-\frac{1}{2}}$-factor). For the final $\mathrm{NF}^*[\kappa]$ (i.e., $L^\infty_{t_\omega} L^2_{x_\omega}$)-component one exploits orthogonality relative to x_ω via the preceding lemma. Here $\omega \in S^1 \setminus (2\kappa')$ is arbitrary but fixed. □

Later we will prove bi- and trilinear estimates involving S and N space. The following bilinear bounds will be a basic ingredient in that context.

Lemma 2.8. *One has the estimates*

$$\|\phi F\|_{\mathrm{NF}[\kappa]} \lesssim \frac{|\kappa'|^{\frac12}2^{\frac{k'}{2}}}{\mathrm{dist}(\kappa,\kappa')}\|\phi\|_{S[k',\kappa']}\|F\|_{L^2_tL^2_x} \tag{2.29}$$

$$\|\phi\psi\|_{L^2_tL^2_x} \lesssim \frac{|\kappa|^{\frac12}2^{\frac{k}{2}}}{\mathrm{dist}(\kappa,\kappa')}\|\phi\|_{S[k,\kappa]}\|\psi\|_{S[k',\kappa']} \tag{2.30}$$

For the final two bounds we require that $2\kappa\cap 2\kappa'=\emptyset$.

Proof. The second one follows from the definition of the spaces, whereas (2.30) follows from (2.29) and the duality bound (2.25). □

Note that both of these estimates have a dispersive character, as they involve space-time integrals. By applying ideas from the energy estimate, we will improve on (2.30) in the high-high case, see Lemma 4.5. Next, we define the spaces which will hold the nonlinearities. These spaces differ from those used for example in [22] as far as the 'elliptic norm' $\|\cdot\|_{\dot X_k^{-\frac12+\varepsilon,-1-\varepsilon,2}}$ is concerned. Here the extra ε ensures that we get exponential gains in the maximal frequencies for certain high-high-low interactions.

Definition 2.9. *$N[k]$ is generated by the following four types of atoms: with F being k-admissible, either*

- $\|F\|_{L^1_tL^2_x}\le 2^k$
- *$\hat F$ is supported on $||\xi|-|\tau||\sim 2^j\le 2^k$ and $\|F\|_{\dot X_k^{-1,-\frac12,1}}\le 1$*
- *$F=Q_{\ge k}F$, $\|F\|_{\dot X_k^{-\frac12+\varepsilon,-1-\varepsilon,2}}\le 1$ where $\varepsilon>0$ is as in the $S[k]$ spaces*
- *F is the sum of wave packets F_κ: there exists $\ell\le -100$ such that $F=\sum_{\kappa\in\mathcal{C}_\ell}F_\kappa$ with all $\mathrm{supp}(\widehat{F_\kappa})$ supported on either $\tau>0$ or $\tau<0$, with $\widehat{F_\kappa}$ supported on $|\xi|\sim 2^k$, $||\xi|-|\tau||\le C^{-1}2^{k+2\ell}$, $\Theta:=\mathrm{sign}(\tau)\widehat{\xi}\in\kappa$ and to that the bound*

$$\Big(\sum_\kappa \|F_\kappa\|^2_{\mathrm{NF}[\kappa]}\Big)^{\frac12}\le 2^k$$

  holds.

We refer to these types as energy, $\dot X^{s,b,q}$, and wave-packet atoms*, respectively.*

In what follows, we refer to functions ϕ adapted to some $k\in\mathbb{Z}$ as "elliptic" iff $P_kQ_{\ge k}\phi=\phi$, whereas those satisfying $P_kQ_{\le k}\phi=\phi$ as "hyperbolic". This terminology has to do with the behavior of the wave operator $\Box$ in these respective regimes. We now record a fundamental duality property of $N[k]$.

Lemma 2.10. *For any $\phi\in S[k]$ and $F\in N[k]$ with $F=P_kQ_{\le k}F$*

$$|\langle\phi,F\rangle|\lesssim 2^k\|\phi\|_{S[k]}\|F\|_{N[k]} \tag{2.31}$$

$$\|F\|_{\dot X_k^{-1,-\frac12,\infty}}\lesssim\|F\|_{N[k]}\lesssim\|F\|_{\dot X_k^{-1,-\frac12,1}} \tag{2.32}$$

Proof. The duality relation (2.31) is proved by taking F to be an atom; for the wave-packet atom use (2.25). By definition of $N[k]$, one has $\|F\|_{N[k]}\lesssim\|F\|_{\dot X_k^{-1,-\frac12,1}}$. For the left-hand bound in (2.32) use (2.19) and (2.31). □

As an application of the geometric considerations of Lemma 2.6 we now show that *refining* a wave-packet atom yields another wave-packet atom.

Lemma 2.11. *Let $F=\sum_{\kappa\in\mathcal{C}_\ell}F_\kappa$ be a wave-packet atom as in Definition 2.9. Then*

$$\sup_{\ell'\le\ell}\sup_{\ell'\le j\le 0}\Big(\sum_\kappa\sum_{\substack{\kappa'\in\mathcal{C}_{\ell'}\\ \kappa'\subset\kappa}}\sum_{R\in\mathcal{R}_{k,\kappa',j}}\|P_RP_{\kappa'}Q_{<\ell+\ell'+k}F_\kappa\|^2_{\mathrm{NF}[\kappa']}\Big)^{\frac12}\le C\,2^k \tag{2.33}$$

with some absolute constant C.

Proof. By scaling invariance, we can set $k=0$. Moreover, fix $\ell' \le \ell$ and $\ell' \le j \le 0$. Choose $\omega' = \omega(\kappa') \in S^1 \setminus (2\kappa')$ for each κ' which attain the respective NF$[\kappa']$ norm. Then one has

$$\sum_{\kappa} \sum_{\substack{\kappa' \in \mathcal{C}_{\ell'} \\ \kappa' \subset \kappa}} \sum_{R \in \mathcal{R}_{0,\kappa',j}} \|P_R P_{\kappa'} Q_{<\ell+\ell'} F_\kappa\|^2_{\mathrm{NF}[\kappa']}$$

$$\lesssim \sum_{\kappa} \sum_{\substack{\kappa' \in \mathcal{C}_{\ell'} \\ \kappa' \subset \kappa}} \sum_{R \in \mathcal{R}_{0,\kappa',j}} d(\omega',\kappa')^{-2} \|P_R P_{\kappa'} Q_{<\ell+\ell'} F_\kappa\|^2_{L^1_{t_\omega} L^2_{x_\omega}}$$

$$\lesssim \sum_{\kappa} \inf_{\omega \in S^1 \setminus (2\kappa)} d(\omega,\kappa)^{-2} \Big\| \Big(\sum_{\substack{\kappa' \in \mathcal{C}_{\ell'} \\ \kappa' \subset \kappa}} \sum_{R \in \mathcal{R}_{0,\kappa',j}} \|P_R P_{\kappa'} Q_{<\ell+\ell'} F_\kappa\|^2_{L^2_{x_\omega}} \Big)^{\frac12} \Big\|^2_{L^1_{t_\omega}} \tag{2.34}$$

$$\lesssim \sum_{\kappa} \inf_{\omega \in S^1 \setminus (2\kappa)} d(\omega,\kappa)^{-2} \|F_\kappa\|^2_{L^1_{t_\omega} L^2_{x_\omega}} \tag{2.35}$$

To pass to (2.34) we used the inclusion $\ell^2(L^1_{t_\omega}) \supset L^1_{t_\omega}(\ell^2)$, whereas orthogonality implies (2.35). Indeed, first note that

$$\bigcup_{t_\omega \in \mathbb{R}} \mathrm{supp}\big([P_R P_{\kappa'} Q_{<\ell+\ell'} F_\kappa)(t_\omega,\cdot)]^\wedge\big) \subset \Pi_\omega\big(\mathrm{supp}(\mathcal{F}[P_R P_{\kappa'} Q_{<\ell+\ell'} F_\kappa])\big)$$

where the Fourier transform on the left-hand side is in x_ω and on the right-hand side in (t_ω, x_ω). Second, the sets on the right-hand side enjoy a finite overlap property by Lemma 2.6. □

In what follows, we will often need to split a wave ϕ into $\phi^+ + \phi^-$ where

$$\phi^+ := \big(\chi_{[\tau \ge 0]} \hat{\phi}(\cdot,\tau)\big)^\vee, \qquad \phi^- := \big(\chi_{[\tau < 0]} \hat{\phi}(\cdot,\tau)\big)^\vee$$

The question arises whether the spaces $S[k]$ and $N[k]$ are preserved under these operations.

Lemma 2.12. *For any Schwartz function ϕ which is adapted to k,*

$$\|\phi^\pm\|_{S[k]} \le C\|\phi\|_{S[k]}, \qquad \|F^\pm\|_{N[k]} \le C\|F\|_{N[k]}$$

with some absolute constant C.

Proof. We set $k=0$ and assume that ϕ is adapted to $k=0$. Let χ_0 be a bump function on the line with $\chi_0(\tau)=1$ on $\tau \ge -C^{-1}$ and $\chi_0(\tau)=0$ if $\tau \le -2C^{-1}$ where $C>1$ is some large constant. Then

$$\widehat{\phi^+}(\tau,\xi) = \chi_0(\tau-|\xi|)\chi_{[\tau\ge 0]}\widehat{\phi}(\tau,\xi) + (1-\chi_0)(\tau-|\xi|)\chi_{[\tau\ge 0]}\widehat{\phi}(\tau,\xi)$$

Denote the two functions on the right-hand side by $\phi^{(+,1)}$ and $\phi^{(+,2)}$, respectively. Then

$$\phi^{(+,1)} = \phi * \mu \tag{2.36}$$

where μ is a measure of bounded mass. Therefore,

$$\|\phi^{(+,1)}\|_{S[0]} \le C\|\phi\|_{S[0]}$$

Next,

$$\|\phi^{(+,2)}\|_{S[0]} \le C\|\phi^{(+,2)}\|_{\dot{X}_0^{0,\frac12,1}} \le C\|\phi\|_{S[0]}$$

where we used the Plancherel theorem in the final step.

For $N[k]$ it will suffice to check the case of $L^1_t L^2_x$-atoms. For these, we write

$$F^+ = F^{(+,1)} + F^{(+,2)}$$

as above. The first term here is fine from (2.36), whereas the second is placed in $L^2_t L^2_x$ and bounded by means of (2.32). □

Another piece of terminology used by Tao is the following:

Definition 2.13. *We shall say that a family* $\{m_\alpha\}_\alpha$ *is* disposable, *if* $T_\alpha f := (m_\alpha \hat{f})^\vee = f * \mu_\alpha$ *where* μ_α *are measures with uniformly bounded mass:*

$$\sup_\alpha \|\mu_\alpha\| \le C < \infty$$

with some universal constant C.

Clearly, disposable multipliers give rise to bounded operators on any translation invariant Banach space. Thus, if X is a Banach space of functions on $\mathbb{R}^{n+1}$ with the property that for all $f \in X$ one has

$$\|f(\cdot - y)\|_X = \|f\|_X \qquad \forall\, y \in \mathbb{R}^{n+1}$$

then $\sup_\alpha \|T_\alpha f\|_X \le C\|f\|_X$. The following observation will be a useful device for removing frequency cut-offs.

Lemma 2.14. *The families*

$$\big\{P_{k,\kappa}\big\}_{k,\kappa}, \quad \big\{P_k Q_j\big\}_{j\ge k}, \quad \big\{P_k Q_{<j}\big\}_{j\ge k}, \quad \big\{P_k Q^{\pm}_{<k-C}\big\}_k$$

are disposable. In the first family κ *is any cap, whereas in the last family* $C > 0$ *has to be chosen such that the support of the multiplier associated with* $P_k Q_{<k-C}$ *does not intersect* $\tau = 0$. *In addition,*

$$\big\{P_{k,\kappa} Q_{<k+2\ell}\big\}$$

is disposable where $k \in \mathbb{Z}$ *and* κ *is any cap with* $\mathrm{diam}(\kappa) \sim 2^\ell$ *with* $\ell \le -100$ *arbitrary.*

Proof. Without loss of generality one may take $k = 0$. Then these statements reduce to simple exercises in harmonic analysis. □

The following fact will serve as a substitute for the previous problem in a non-disposable context.

Lemma 2.15. Q_j, $Q_{<j}$ *are bounded on* $L^p L^2$ *for every* $1 \le p \le \infty$ *with a constant independent of* $j \in \mathbb{Z}$.

Proof. The inverse Fourier transform of $Q_{<j}$ with respect to time alone is

$$\begin{aligned}&\int e^{it\tau} m_0(2^{-j}(\tau - |\xi|))\widehat{F}(\tau,\xi)\,d\tau \\ &= 2^j \int \widehat{m_0}(2^j(t-s))e^{i|\xi|(t-s)} F(s,\hat{\xi})\,ds\end{aligned}$$

where $F(s,\hat{\xi})$ in the second line denotes the Fourier transform with respect to the second variable. Consequently,

$$\|Q_{<j}F\|_{L^1_t L^2_x} \le \|\widehat{m_0}\|_1 \|F\|_{L^1_t L^2_x}$$

as claimed. □

The previous result, combined with Lemma 2.7, implies the following square-function bound.

Corollary 2.16. *For all* $j,k \in \mathbb{Z}$ *and all* k*–adapted Schwartz functions* ϕ *one has* $\|Q_{<j}\phi\|_{S[k]} \le C\|\phi\|_{S[k]}$ *with some absolute constant* C.

Proof. We may again take $k = 0$. The $L^\infty_t L^2_x$-component of the $S[0]$-norm is covered by Lemma 2.15. The $\dot{X}^{s,b,q}$-components are obvious , the Strichartz norms as well by construction, and the square-function is a consequence of Lemma 2.7. □

We remark that the analogous statement for $N[k]$ holds as well, see Corollary 2.23 below. Next, for the sake of completeness we state the full range of Strichartz estimates that follow from (2.14).

Lemma 2.17. *For any* $4 \le p \le \infty$ *and* $2 \le q \le \infty$ *which satisfy* $\frac{1}{p} + \frac{1}{2q} \le \frac{1}{4}$,

$$\Big(\sum_{c\in\mathcal{D}_{k,\ell}} \|P_c\phi\|^2_{L^p_t L^q_x}\Big)^{\frac12} \le C 2^{\ell(1-\frac{2}{p}-\frac{2}{q}-\frac{4\varepsilon}{p})} 2^{k(1-\frac{1}{p}-\frac{2}{q})} \|\psi\|_{S[k]} \tag{2.37}$$

for any $k \in \mathbb{Z}$, $\ell \le 0$, *and with an absolute constant* C.

Proof. Assume first that $\frac{1}{p}+\frac{1}{2q}=\frac{1}{4}$. By interpolation, and with $\theta=\frac{4}{p}$,

$$\begin{aligned}\Big(\sum_{c\in\mathcal{D}_{0,\ell}}\|P_c\phi\|_{L_t^pL_x^q}^2\Big)^{\frac12}&\le\Big(\sum_{c\in\mathcal{D}_{0,\ell}}\|P_c\psi\|_{L_t^4L_x^\infty}^{2\theta}\|P_c\phi\|_{L_t^\infty L_x^2}^{2(1-\theta)}\Big)^{\frac12}\\&\le\Big(\sum_{c\in\mathcal{D}_{0,\ell}}\|P_c\psi\|_{L_t^4L_x^\infty}^2\Big)^{\frac{\theta}{2}}\Big(\sum_{c\in\mathcal{D}_{0,\ell}}\|P_c\psi\|_{L_t^\infty L_x^2}^2\Big)^{\frac{1-\theta}{2}}\qquad(2.38)\\&\lesssim 2^{\theta(\frac12-\varepsilon)\ell}\|\psi\|_{S[0]}\end{aligned}$$

To pass from (2.38) to the last line, one uses (2.14) as well as the energy component of (2.15). For larger q, one gains a factor $2^{2\ell(\frac12-\frac2p-\frac1q)}$ by Bernstein's inequality, and rescaling to frequency 2^k yields a factor of $2^{k(1-\frac1p-\frac2q)}$ as claimed. □

Finally, we conclude this section with the following useful fact.

Lemma 2.18. *Let ϕ be adapted to* 0*. Then for any $m_0\le -10$,*

$$\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}\phi\|_{L_t^\infty L_x^2}^2\Big)^{\frac12}\lesssim|m_0|\|\phi\|_{S[0]}$$

Proof. First,

$$\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}Q_{\le 2m_0}\phi\|_{L_t^\infty L_x^2}^2\Big)^{\frac12}\lesssim\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}Q_{\le 2m_0}\phi\|_{S[0,\kappa]}^2\Big)^{\frac12}\lesssim\|\phi\|_{S[0]}$$

by (2.15). Second,

$$\begin{aligned}\sum_{2m_0\le\ell\le0}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}Q_\ell\phi\|_{L_t^\infty L_x^2}^2\Big)^{\frac12}&\lesssim\sum_{2m_0\le\ell\le0}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}Q_\ell\phi\|_{\dot X_0^{0,\frac12,\infty}}^2\Big)^{\frac12}\\&\lesssim|m_0|\|\phi\|_{S[0]}\end{aligned}$$

And third,

$$\|P_{0,\kappa}Q_{\ge0}\phi\|_{L_t^\infty L_x^2}\lesssim\|P_{0,\kappa}Q_{\ge0}\phi\|_{\dot X_0^{0,\frac12,1}}\lesssim\|P_{0,\kappa}Q_{\ge0}\phi\|_{\dot X_0^{0,1-\varepsilon,2}}$$

whence

$$\begin{aligned}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}Q_{\ge0}\phi\|_{L_t^\infty L_x^2}^2\Big)^{\frac12}&\lesssim\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_{0,\kappa}Q_{\ge0}\phi\|_{\dot X_0^{0,1-\varepsilon,2}}^2\Big)^{\frac12}\\&\lesssim\|P_0Q_{\ge0}\phi\|_{\dot X_0^{0,1-\varepsilon,2}}\lesssim\|\phi\|_{S[0]}\end{aligned}$$

as claimed. □

The central problems concerning the $S[k]$ and $N[k]$ spaces are how to obtain an energy estimate and how to control the trilinear nonlinearities appearing in the gauged wave-map system. We begin with the energy estimate, and then develop bilinear bounds which are preliminary to the central trilinear bounds.

2.3. **The energy estimate.** The purpose of this section is to prove the energy estimate in the context of the $S[k]$ and $N[k]$ spaces, see Proposition 2.26 below. First, we require some technical lemmas. The first two of these lemmas will arise in the Duhamel integral.

Lemma 2.19. *For any F which is k-adapted and satisfies $F=Q_{\le k+C}F$,*

$$\|\chi_{\mathbb{R}^+}F\|_{N[k]}\lesssim\|F\|_{N[k]}\qquad(2.39)$$

where $\chi_{\mathbb{R}^+}$ acts only in time.

Proof. We may assume that $k=0$. This is clear if F is an energy atom. Next, we consider the $\dot X^{0,-\frac12,1}$-atoms. First, let $\hat F$ be supported on $|\xi|\ \sim 1, ||\xi|-|\tau||\sim 2^j$ with $\|F\|_{L^2_tL^2_x}\lesssim 2^{j/2}$. Then

$$\begin{aligned}\|\chi_{\mathbb{R}^+}F\|_{N[0]} &\lesssim \|P_{<j}(\chi_{\mathbb{R}^+})F\|_{N[0]}+\|P_{\geq j}(\chi_{\mathbb{R}^+})F\|_{N[0]}\\ &\lesssim 2^{-j/2}\|P_{<j}(\chi_{\mathbb{R}^+})F\|_{L^2_tL^2_x}+\|P_{\geq j}(\chi_{\mathbb{R}^+})F\|_{L^1L^2}\\ &\lesssim 2^{-j/2}\|P_{<j}(\chi_{\mathbb{R}^+})\|_{L^\infty_t}\|F\|_{L^2_tL^2_x}+\|P_{\geq j}(\chi_{\mathbb{R}^+})\|_{L^2_t}\|F\|_{L^2L^2}\\ &\lesssim 2^{-j/2}\|F\|_{L^2_tL^2_x}\lesssim 1\end{aligned}$$

Now let F be a wave-packet atom, i.e., for some $\ell\leq -100$,

$$F=\sum_{\kappa\in\mathcal{C}_\ell}F_\kappa,\quad \mathrm{supp}(\widehat{F_\kappa})\subset\big\{\tau>0,\ |\xi|\sim 1,\ ||\xi|-\tau|\sim 2^{2\ell},\ \Theta\in\kappa\big\}$$

and $\sum_\kappa\|F_\kappa\|^2_{\mathrm{NF}[\kappa]}\leq 1$. We write, with $j=2\ell$,

$$\chi_{\mathbb{R}^+}=P_{<j}(\chi_{\mathbb{R}^+})+P_{\geq j}(\chi_{\mathbb{R}^+})$$

as before. Then $P_{<j}(\chi_{\mathbb{R}^+})$ does not significantly change the support properties of F_κ. Moreover, since $\|P_{<j}(\chi_{\mathbb{R}^+})\|_\infty\lesssim 1$, we see that $P_{<j}(\chi_{\mathbb{R}^+})F$ is essentially a wave-packet atom. On the other hand, since $\|F\|_{L^2_tL^2_x}\lesssim 2^{j/2}$ from (2.32) we conclude that

$$\|P_{\geq j}(\chi_{\mathbb{R}^+})F\|_{L^1L^2}\lesssim 2^{-j/2}\|F\|_{L^2_tL^2_x}\lesssim 1 \tag{2.40}$$

which proves (2.39). □

It is important to note that the previous lemma *fails* for functions in $N[0]$ which are "elliptic" since the $\dot X_k^{-\frac12+\varepsilon,-1-\varepsilon,2}$-norm is finite on functions which are too singular. But in the elliptic regime, there will be no need for the Duhamel formula and thus for Lemma 2.19.

A technical variant of the preceding lemma will be needed in the proof of Proposition 9.14:

Lemma 2.20. *Let F be as in the preceding lemma, $\kappa\subset S^1$, and $\omega\notin\pm 2\kappa$. Then for any $c\in\mathbb{R}$ we have*

$$\|P_\kappa Q^\pm_{<k+C}\chi_{t_\omega\geq c}F\|_{N[k]}\lesssim\|F\|_{N[k]},$$

where the implied constant only depends on C and $|\kappa|$.

Proof. It is essentially identical to the preceding one: one replaces $\chi_{\mathbb{R}^+}$ by $\chi_{t_\omega\geq c}$ and $L^1_tL^2_x$ by $L^1_{t_\omega}L^2_{x_\omega}$. □

The Duhamel formula (in other words, $\square^{-1}$) introduces a Hilbert transform in the normal direction to the light-cone. The following lemma is of this type.

Lemma 2.21. *Let η be a smooth function on $\mathbb{R}$ such that $0\leq\eta\leq 1$, $\eta(u)=1$ on $-1\leq u\leq 1$,* $\mathrm{supp}(\eta)\subset[-2,2]$*, and $\eta'(u)\geq 0$ on $u\leq 0$, and $\eta'(u)\leq 0$ on $u\geq 0$. Define $\eta_T^+(t):=\chi_{[0,\infty)}\eta(t/T)$ for each $T\geq 1$. Then, with $\chi:=\chi_{[0,\infty)}\eta'$,*

$$\widehat{\eta_T^+}(\tau)=-\frac{1}{i\tau}(\widehat{\chi}(T\tau)+1) \tag{2.41}$$

In particular, $\widehat{\eta_T^+}(\tau)=a\widehat{\eta^+_{\frac{T}{a}}}(a\tau)$ for all $0<a<1$ and

$$\big|\widehat{\eta_T^+}(\tau)\big|\lesssim|\tau|^{-1},\quad \big|\frac{d}{d\tau}\widehat{\eta_T^+}(\tau)\big|\lesssim|\tau|^{-2}$$

Moreover, let $\mu=\mu(\tau)$ be a smooth function on $[-1,1]$ with $\mu(0)=1$ and $\mu\geq 1$ on $[-1,1]$. Then

$$\sup_{|\tau|\leq 1}\big|\widehat{\eta_T^+}(\tau)-\widehat{\eta_T^+}(\mu(\tau)\tau)\big|\leq C\|\mu'\|_\infty$$

with an absolute constant C. Finally, if $T'\in[T/2,2T]$, then

$$\big|\widehat{\eta_T^+}(\tau)-\widehat{\eta_{T'}^+}(\tau)\big|\leq CT\min\big(1,(T|\tau|)^{-100}\big)$$

with a constant C that only depends on χ.

Proof. Integrating by parts in
$$\widehat{\eta_T^+}(\tau) = \int e^{-i\tau u}\eta_T^+(u)\,du$$
yields (2.41). In particular,
$$\big|\widehat{\eta_T^+}(\tau)\big| \lesssim |\tau|^{-1}\min(T|\tau|, (T|\tau|)^{-100})$$
and similarly for the derivatives. Next, write
$$\widehat{\eta_T^+}(\tau) - \widehat{\eta_T^+}(\mu(\tau)\tau) = -\frac{1}{i\tau}(\widehat{\chi}(T\tau)+1) + \frac{1}{i\mu(\tau)\tau}(\widehat{\chi}(T\mu(\tau)\tau)+1)$$
In view of our assumptions on μ,
$$\big|\tau^{-1}(1-\mu(\tau)^{-1})\big| \lesssim \|\mu'\|_\infty$$
and similarly for the terms involving $\widehat{\chi}(T\tau)$. The final statement is an immediate consequence of (2.41). □

The following representation of waves $\Box^{-1}F$ with F a null-frame atom will be useful in several instances. Hence, we state it as a separate fact.

Lemma 2.22. *Assume that $F \in N[0]$ is a wave-packet atom, i.e., $F = F^+ = \sum_{\kappa\in\mathcal{C}_\ell} F_\kappa$ with*
$$\sum_{\kappa\in\mathcal{C}_\ell} \|F_\kappa\|^2_{\mathrm{NF}[\kappa]} \le 1$$
for some $\ell \le -100$, see Definition 2.9. Then
$$\phi(t) := \Box^{-1}F(t) = \int_0^t \frac{\sin((t-s)|\nabla|}{|\nabla|}F(s)\,ds$$
admits a decomposition of the form
$$\phi = \Box^{-1}F_1 + \sum_{\kappa\in\mathcal{C}_\ell}\int_{\mathbb{R}} \big(\Psi^1_{\kappa,a} + B_{\kappa,a}\,\Psi^2_{\kappa,a}\big)\,da \tag{2.42}$$
where $\|F_1\|_{L^1_tL^2_x} \lesssim 1$ and
$$\sup_{\kappa,a}\|B_{\kappa,a}\|_{L^\infty_{t,x}} \le C, \quad \sum_{j=1}^2 \int \|\Psi^j_{\kappa,a}\|_{\dot X_0^{0,\frac12,1}}\,da \le C\|F_\kappa\|_{\mathrm{NF}[\kappa]}$$
with an absolute constant C whence
$$\sup_{j=1,2}\sum_\kappa \Big(\int \|\Psi^j_{\kappa,a}\|_{\dot X_0^{0,\frac12,1}}\,da\Big)^2 \lesssim 1$$
Finally[8]*, for $j=1,2$*
$$\mathrm{supp}\big(\widehat{\Psi^j_{\kappa,a}}\big) \subset C\mathrm{supp}\big(\widehat{F_\kappa}\big), \quad \mathrm{supp}\big(\widehat{B_{\kappa,a}\,\Psi^2_{\kappa,a}}\big) \subset C\mathrm{supp}\big(\widehat{F_\kappa}\big) \tag{2.43}$$
for all a and κ and some absolute constant C.

Proof. As in the proof of Lemma 2.19, we first write
$$\chi_{\mathbb{R}^+} = P_{\ge 2\ell}(\chi_{\mathbb{R}^+}) + P_{<2\ell}(\chi_{\mathbb{R}^+}) =: \chi_1 + \chi_2$$
Then $F_1 := \chi_1 F$ satisfies $\|F_1\|_{L^1_tL^2_x} \lesssim 1$, see (2.40). On the other hand, $F_2 := \chi_2 F$ is again a wave-packet atom at essentially the same scale as F, i.e., $F_2 = \sum_{\kappa\in\mathcal{C}_\ell}\widetilde{F}_\kappa$ with
$$\sum_{\kappa\in\mathcal{C}_\ell}\|\widetilde{F}_\kappa\|^2_{\mathrm{NF}[\kappa]} \le 1$$
Define $\Phi := \Box^{-1}F_2$. Then $\Phi = \sum_\kappa \lim_{T\to\infty}\Phi_{T,\kappa}$ with
$$\Phi_{T,\kappa}(t) := \int_{-\infty}^{\infty}\frac{\sin((t-s)|\nabla|)}{|\nabla|}\eta_T^+(t-s)\widetilde{F}_\kappa(s)\,ds$$

[8] Here cE denotes the dilation of the convex set E about its center of mass by the constant c.

It suffices to prove that

$$\Phi_{T,\kappa} = \int \big(\Psi^1_{T,\kappa,a} + B_{T,\kappa,a}\Psi^2_{T,\kappa,a}\big)\, da$$

where

$$\sup_{T\ge 1}\sup_a \|B_{T,\kappa,a}\|_\infty \lesssim 1$$

and

$$\sup_{j=1,2}\sup_{T\ge 1}\int \|\Psi^j_{T,\kappa,a}\|_{\dot X_0^{0,\frac12,1}}\, da \lesssim \|\tilde F_\kappa\|_{\mathrm{NF}[\kappa]} \tag{2.44}$$

both uniformly in κ.
Fix $\kappa\in\mathcal{C}_\ell$ and $\omega=\omega(\kappa)\in S^1\setminus(2\kappa)$ so that

$$d(\omega,\kappa)^{-1}\|\widetilde F_\kappa\|_{L^1_{t_\omega}L^2_{x_\omega}} \le 2\|\widetilde F_\kappa\|_{\mathrm{NF}[\kappa]}$$

As usual, we foliate relative to t_ω. More precisely, define

$$f_a(x_\omega) = \widetilde F_\kappa((t,x)(a,x_\omega))$$

where $t_\omega=a$ means that

$$(t,x)(a,x_\omega) := a\theta^+_\omega + x_\omega$$

By Lemma 2.6

$$\mathrm{supp}(\hat f_a) \subset \Pi_\omega\big(\big\{(\tau,\xi)\ :\ |\xi|\sim 1,\ \widehat\xi\in\kappa,\ |\tau-|\xi||\lesssim 2^{2\ell}\big\}\big) =: R_{\kappa,\omega} \tag{2.45}$$

Let (t_ω,x_ω) denote the null-frame coordinates. Then

$$\tau-|\xi| = \frac{2\xi^1_\omega}{\tau+|\xi|}(\tau_\omega - h(\xi_\omega)) \tag{2.46}$$

where $\xi^1_\omega := \xi_\omega\cdot\theta^-_\omega$ and, with $|\xi_\omega|^2 = (\xi^1_\omega)^2+(\xi^2_\omega)^2$, one has $h(\xi_\omega) := \frac{(\xi^2_\omega)^2}{2\xi^1_\omega}$. Moreover, $|\xi^1_\omega|\sim d(\omega,\kappa)^2$ and $|\xi^2_\omega|\lesssim d(\omega,\kappa)$ by elementary geometry (cf. Lemma 2.6). We define

$$P_{\kappa,\omega} f := \mathcal{F}^{-1}[\chi_{R_{\kappa,\omega}}(\xi_\omega) f(\tau_\omega,\xi_\omega)]$$

where $\chi_{R_{\kappa,\omega}}$ is a smooth cut-off adapted to the rectangle $\chi_{R_{\kappa,\omega}}$ in the ξ_ω-plane. Furthermore, we set

$$Q^+_{\le j,\omega} f := \mathcal{F}^{-1}\big[m_0\big(2^{-j-C}d^2(\kappa,\omega)(\tau_\omega-h(\xi_\omega))\big) f(\tau_\omega,\xi_\omega)\big]$$

By construction, $P_{\kappa,\omega}Q^+_{\le 2\ell,\omega}$ is essentially the same as $P_{0,\kappa}Q_{\le 2\ell}$, see Lemma 2.6. In fact, one has

$$\widetilde F_\kappa = P_{\kappa,\omega}Q^+_{\le 2\ell,\omega}\widetilde F_\kappa$$

and $P_{\kappa,\omega}Q^+_{\le 2\ell,\omega}$ is disposable. Clearly,

$$\Phi_{T,\kappa} = \int \Phi_{T,\kappa,a}\, da$$

where

$$\Phi_{T,\kappa,a}(t) := P_{\kappa,\omega}Q^+_{\le 2\ell,\omega}\int_{-\infty}^{\infty}\frac{\sin((t-s)|\nabla|)}{|\nabla|}\eta^+_T(t-s)\delta(s_\omega-a) f_a\, ds$$

Then $\Phi_{T,\kappa,a} = P_{0,C\kappa}Q^+_{\le 2\ell+C}\Phi_{T,\kappa,a}$ and

$$\Phi_{T,\kappa,a} = P_{\kappa,\omega}Q^+_{\le 2\ell,\omega}\ \mathcal{F}^{-1}[(\widehat{\eta^+_T}(\tau-|\xi|) - \widehat{\eta^+_T}(|\xi|+\tau))e^{-i\tau_\omega a}\hat f_a(\xi_\omega)] \tag{2.47}$$

We claim that the contribution of $|\widehat{\eta^+_T}(|\xi|+\tau)|\lesssim 1$ to (2.47) can be added to $\Psi^1_{T,\kappa,a}$. In fact,

$$\begin{aligned}
&\big\|Q^+_{\le 2\ell+C}\ \mathcal{F}^{-1}[O(1)e^{-i\tau_\omega a}\chi_{R_{\kappa,\omega}}(\xi_\omega)\hat f_a(\xi_\omega)]\big\|_{\dot X_0^{0,\frac12,1}}\\
&\lesssim 2^\ell\big\|m_0(2^{-2\ell-C}(\tau-|\xi|))\chi_{R_{\kappa,\omega}}(\xi_\omega)\hat f_a(\xi_\omega)\big\|_{L^2_\tau L^2_\xi}\\
&\lesssim 2^{2\ell}d(\omega,\kappa)^{-1}\|f_a\|_{L^2}
\end{aligned} \tag{2.48}$$

which is better than needed. To pass to the final estimate here we used Lemma 2.6, especially (2.27); the latter estimate can be applied for fixed τ, since then $\xi_\omega = \xi_\omega(\xi, \tau)$. Next, we split the contribution of $\widehat{\eta_T^+}(|\xi| - \tau)$ to (2.47) into several pieces. Since $\tau_\omega - h(\xi_\omega) = 0$ implies that

$$2|\xi| = \tau + |\xi| = 2[h(\xi_\omega)/\sqrt{2} + \xi_\omega \cdot e_1] =: g(\xi_\omega)$$

where $e_1 = (1, 0, 0)$, one has by Lemma 2.21

$$\widehat{\eta_T^+}(\tau - |\xi|) - \widehat{\eta_T^+}\big(\frac{2\xi_\omega^1}{g(\xi_\omega)}(\tau_\omega - h(\xi_\omega)\big) = O\big(\frac{1}{\xi_\omega^1}\big) = O\big(d(\omega, \kappa)^{-2}\big) \tag{2.49}$$

In view of (2.48) (which *gains* a factor of $2^{2\ell} \lesssim d(\omega, \kappa)^2$), the contribution of (2.49) to (2.47) can again be added to $\Psi^1_{T,\kappa,a}$. Set $b = b(\xi_\omega) = \frac{2\xi_\omega^1}{g(\xi_\omega)}$. Furthermore, set $b_0 := b(\xi_\omega^{(0)})$ where $\xi_\omega^{(0)} \in R_{\kappa,\omega}$ is fixed, cf. (2.45). In view of Lemma 2.21,

$$\begin{aligned}\widehat{\eta_T^+}\big(\frac{2\xi_\omega^1}{g(\xi_\omega)}(\tau_\omega - h(\xi_\omega)\big) &= b^{-1}\widehat{\eta_{bT}^+}(\tau_\omega - h(\xi_\omega)) \\ &= b^{-1}\widehat{\eta_{b_0T}^+}(\tau_\omega - h(\xi_\omega)) + O\big[T\min(1, (T|\tau - |\xi||)^{-100})\big] \qquad (2.50)\end{aligned}$$

where we used that $b \sim b_0$ on $R_{\kappa,\omega}$. The computation from (2.48) above now shows that the $O(\cdot)$ term in (2.50) can be added to $\Psi^1_{T,\kappa,a}$. It therefore remains to analyze the contribution of the first term in (2.50) to (2.47). Define

$$B_{T,a,\kappa}(t, x) := \int \eta_{b_0T}^+(t_\omega - s_\omega)e^{-ia(t_\omega - s_\omega)}\lambda\widehat{m_0}(\lambda s_\omega)\, ds_\omega$$

where $2^{j+C}d^{-2}(\kappa, \omega) =: \lambda$ (recall that m_0 is even). On the one hand, $\|B_{T,a,\kappa}\| \le \|\widehat{m_0}\|_1$ and on the other hand,

$$\begin{aligned}&P_{\kappa,\omega}Q^+_{\le 2\ell,\omega}\ \mathcal{F}^{-1}[b^{-1}\widehat{\eta_{b_0T}^+}(\tau_\omega - h(\xi_\omega))e^{-i\tau_\omega a}\hat{f}_a(\xi_\omega)] \\ &= B_{T,a,\kappa}\ \mathcal{F}^{-1}\Big[\delta(\tau_\omega - h(\xi_\omega))\,\chi_{R_{\kappa,\omega}}(\xi_\omega)\frac{g(\xi_\omega)}{2\xi_\omega^1}e^{-ih(\xi_\omega)a}\hat{f}_a(\xi_\omega)\Big] \\ &=: B_{T,a,\kappa}\Psi^2_{T,\kappa,a}\end{aligned}$$

By inspection, the Fourier support of $\Psi^2_{T,\kappa,a}$ as well as that of $B_{T,a,\kappa}\Psi^2_{T,\kappa,a}$ are no larger than that of the original wave-packet F_κ (up to a dilation by a constant). Finally, by a calculation similar to (2.48),

$$\begin{aligned}&\|\Psi^2_{T,\kappa,a}\|_{\dot{X}_0^{0,\frac12,1}} \lesssim \big\|\mathcal{F}^{-1}\big[\delta(\tau_\omega - h(\xi_\omega))\,\frac{g(\xi_\omega)}{2\xi_\omega^1}e^{-ih(\xi_\omega)a}\chi_{R_{\kappa,\omega}}(\xi_\omega)\hat{f}_a(\xi_\omega)\big]\big\|_{\dot{X}_0^{0,\frac12,1}} \\ &\lesssim d(\omega, \kappa)^2 \limsup_{M\to\infty} \big\|\mathcal{F}^{-1}[M\eta(M(\tau - |\xi|))\,\frac{g(\xi_\omega)}{2\xi_\omega^1}e^{-ih(\xi_\omega)a}\chi_{R_{\kappa,\omega}}(\xi_\omega)\hat{f}_a(\xi_\omega)]\big\|_{\dot{X}_0^{0,\frac12,1}} \\ &\lesssim d(\omega, \kappa)^{-1}\|f_a\|_{L^2}\end{aligned}$$

This concludes the proof of the lemma. □

In passing, we now prove the analogue of Lemma 2.15 for null-frame coordinates, which then gives Corollary 2.16 for the $N[k]$ spaces.

Corollary 2.23. *For all $F \in N[k]$ and all $j \in \mathbb{Z}$ one has $\|Q_{\le j}F\|_{N[k]} \le C\|F\|_{N[k]}$ with some absolute constant C.*

Proof. This is clear if F is either an energy or a $\dot{X}^{s,b}$-atom. Therefore, suppose that $F = \sum_\kappa F_\kappa$ is a wave-packet atom with $k = 0$. It suffices to prove that

$$\|Q_{\le j}F_\kappa\|_{\mathrm{NF}[\kappa]} \le C\|F_\kappa\|_{\mathrm{NF}[\kappa]}$$

This in turn follows from

$$\|Q_{\le j}F_\kappa\|_{L^1_{t_\omega}L^2_{x_\omega}} \le C\|F_\kappa\|_{L^1_{t_\omega}L^2_{x_\omega}} \tag{2.51}$$

which holds uniformly in $\omega \in S^1 \setminus (2\kappa)$. Fix such an ω and apply Plancherel's theorem in x_ω. By (2.46),

$$\mathcal{F}_2\, Q_{<j} F_\kappa(t_\omega, \xi_\omega) = \int m_0\Big(2^{-j}\frac{2\xi^1_\omega}{\tau+|\xi|}(\tau_\omega - h(\xi_\omega))\Big) e^{i\tau_\omega(t_\omega - s_\omega)}\, d\tau_\omega\; \mathcal{F}_2 F_\kappa(s_\omega, \xi_\omega)\, ds_\omega$$

where for our purposes here $\mathcal{F}_2$ refers to a partial Fourier transform relative to the second variable x_ω. In view of $|\xi^1_\omega| \sim d(\omega,\kappa)^2$,

$$\begin{aligned}\big|\mathcal{F}_2\, Q_{<j} F_\kappa(t_\omega, \xi_\omega)\big| &\le \Big|\int m_0\Big(2^{-j}\frac{2\xi^1_\omega}{\tau+|\xi|}(\tau_\omega - h(\xi_\omega))\Big) e^{i\tau_\omega(t_\omega - s_\omega)}\, d\tau_\omega\Big| \big|\mathcal{F}_2 F_\kappa(s_\omega, \xi_\omega)\big|\, ds_\omega \\ &\lesssim_N 2^j d(\omega,\kappa)^{-2} \int \langle 2^j d(\omega,\kappa)^{-2}(t_\omega - s_\omega)\rangle^{-N} \big|\mathcal{F}_2 F_\kappa(s_\omega, \xi_\omega)\big|\, ds_\omega\end{aligned}$$

Performing an $L^2_{\xi_\omega}$ estimate followed by an $L^1_{t_\omega}$ bound yields (2.51). □

Finally, there is the following simple fact that will play a role in the proof of the Strichartz component of $\|\cdot\|_{S[k]}$.

Lemma 2.24. *Let $a_{km} \ge 0$ for all $1 \le m \le M$ and $1 \le k \le K$. Suppose $\sum_{k=1}^K a_{km} \le \sigma$ for all m where $\sigma \ge 0$ is arbitrary. Then*

$$\sum_{k=1}^K \Big(\sum_{m=1}^M a^2_{km}\Big)^{\frac12} \le \sigma M^{\frac{1+\theta}{2}} K^{\frac{1-\theta}{2}} \tag{2.52}$$

for all $0 \le \theta \le 1$.

Proof. Denote the sum in (2.52) by S. On the one hand,

$$S \le \sum_{k=1}^K \sum_{m=1}^M a_{km} \le \sigma\, M$$

On the other hand,

$$S \le \sqrt{K}\Big(\sum_{k=1}^K \sum_{m=1}^M a^2_{km}\Big)^{\frac12} \le \sigma\sqrt{KM}$$

and the lemma is proved. □

Now we can state the main energy bound. We begin with the easier elliptic regime.

Lemma 2.25. *Let F be a space-time Schwartz function which is adapted to $k \in \mathbb{Z}$. Assume furthermore that $F = I^c F$ and set $\phi := \Box^{-1} F$, which is defined via division by $\tau^2 - |\xi|^2$ on the Fourier side. Then*

$$\|\phi\|_{S[k]} \lesssim \|F\|_{N[k]}$$

with an absolute implicit constant.

Proof. We may again assume that $k = 0$. We then need to prove that

$$\|\phi\|_{\dot X_k^{0,1-\varepsilon,2}} \lesssim \min\big(\|F\|_{L^1_t L^2_x}, \|F\|_{\dot X_0^{0,-1-\varepsilon,2}}\big) \tag{2.53}$$

since, as we observed after Definition 2.3, the norm on the left-hand side dominates the other norms which make up $\|\cdot\|_{S[k]}$. If we select $\|F\|_{\dot X_0^{0,-1-\varepsilon,2}}$ on the right-hand side of (2.53), then this inequality is obvious. On the other hand, if we select $\|F\|_{L^1_t L^2_x}$, then one concludes via Bernstein's inequality in time. □

Next, we deal with the hyperbolic regime.

Proposition 2.26. *Let $k \in \mathbb{Z}$ and suppose ϕ_0, ϕ_1 are Schwartz functions in $\mathbb{R}^2$ which are adapted to k. Further, suppose F is a space-time Schwartz function which is adapted to k, and which is moreover hyperbolic, i.e., $F = IF$. Then the unique smooth solution of*

$$\Box\phi = F, \quad (\phi(0), \partial_t\phi(0)) = (\phi_0, \phi_1)$$

satisfies

$$\|\phi\|_{S[k]} \lesssim \|(\phi_0,\phi_1)\|_{L^2\times \dot H^{-1}} + \|F\|_{N[k]} \tag{2.54}$$

with an absolute implicit constant.

Proof. By scaling we may assume that $k = 0$. We first assume that $F = 0$. Then

$$\widehat{\phi(t)}(\xi) = \cos(t|\xi|)\widehat{\phi_0}(\xi) + \frac{\sin(t|\xi|)}{|\xi|}\widehat{\phi_1}(\xi)$$

Consequently, (2.54) follows from (2.21) upon sending $T \to \infty$.

Next, we assume that $\phi_0 = \phi_1 = 0$. By the Duhamel formula,

$$\widehat{\phi(t)} = \int_0^t \frac{\sin((t-s)|\xi|)}{|\xi|}\widehat{F(s)}(\xi)\,ds$$

In other words, we need to show that

$$\Big\| \int_{-\infty}^t \frac{\sin((t-s)|\nabla|)}{|\nabla|}\chi_{[0,\infty)}(s)F(s)\,ds \Big\|_{S[0]} \lesssim \|F\|_{N[0]}$$

In view of Lemma 2.19, we may remove the indicator function $\chi_{[0,\infty)}(t) = \chi_{\mathbb{R}^+}(t)$ on the left-hand side. This is where we use that $F = IF$, but after this point we may no longer assume that $F = IF$ since $\chi_{\mathbb{R}^+}F$ loses this property.

The goal is now to prove uniformly in $T \geq 1$

$$\Big\| \int_{-\infty}^\infty \sin((t-s)|\nabla|)\eta_T^+(t-s)F(s)\,ds \Big\|_{S[0]} \lesssim \|F\|_{N[0]} \tag{2.55}$$

where $\eta_T^+(u) := \eta(u/T)\chi_{\mathbb{R}^+}(u)$ is a bump function as specified in Lemma 2.21. Denote

$$\phi(t) = \int_{-\infty}^\infty \sin((t-s)|\nabla|)\eta_T^+(t-s)F(s)\,ds \tag{2.56}$$

Then the space-time Fourier transform of ϕ equals (up to a multiplicative constant)

$$\widehat{\phi}(\tau,\xi) = \big(\widehat{\eta_T^+}(\tau - |\xi|) - \widehat{\eta_T^+}(\tau + |\xi|)\big)\widehat{F}(\tau,\xi) \tag{2.57}$$

whence, by Lemma 2.21,

$$|\widehat{\phi}(\tau,\xi)| \lesssim \Big(\big||\tau| - |\xi|\big|^{-1}\chi_{[|\tau|<10]} + \tau^{-2}\chi_{[|\tau|\geq 10]} \Big)|\widehat{F}(\tau,\xi)| \tag{2.58}$$

and thus also

$$\|Q_{\leq 0}\phi\|_{\dot{X}_0^{0,\frac12,\infty}} + \|Q_{>0}\phi\|_{\dot{X}_0^{0,1-\varepsilon,2}} \lesssim \|F\|_{N[0]} \tag{2.59}$$

from (2.32) and Lemma 2.25. By Lemma 2.4 it suffices to assume that F is either an energy or a wave-packet atom. Moreover, in each of these cases the $\dot{X}^{0,\frac12,\infty}$ and $\dot{X}^{0,1-\varepsilon,2}$-components of the $S[k]$ norm of ϕ can be ignored due to (2.59). Moreover, since the $\dot{X}^{0,1-\varepsilon,2}$-norm controls the entire $S[0]$-norm in the elliptic regime, it suffices to consider only $Q_{\leq 0}\phi$.

In case F is an energy atom, i.e., $\|F\|_{L^1L^2} \leq 1$ standard $X^{s,b}$ and Strichartz norms for the wave equation bound the norms in (2.13) and (2.14), see Lemma 2.2. We are therefore reduced to bounding (2.15), for which it suffices to verify that

$$\sup_{\ell \leq -100} \sup_{T \geq 1} \sup_{\kappa \in \mathcal{C}_\ell} \big\| P_{0,\kappa} Q^+_{<2\ell} \sin(t|\nabla|)\eta_T^+(t) f \big\|_{S[0,\kappa]} \lesssim \|f\|_{L^2_x}$$

for any f which is 0-adapted (the case of Q^- being analogous). We can ignore the further localization to the rectangle R due to orthogonality, cf. (2.15). The Fourier transform of the function inside the norms on the left-hand side is

$$\chi_\kappa(\hat{\xi})m_0(2^{-2\ell}(|\xi| - \tau))(\widehat{\eta_T^+}(\tau - |\xi|) - \widehat{\eta_T^+}(|\xi| + \tau))\hat{f}(\xi)$$

where χ_κ is a cut-off adapted to the cap κ. The contribution by $\widehat{\eta_T^+}(|\xi| + \tau)$ is controlled by (2.19). As for $\widehat{\eta_T^+}(|\xi| - \tau)$, one needs to show that

$$\big\| [\eta_T^+ * 2^{2\ell}\widehat{m_0}(2^{2\ell}\cdot)] P_{0,\kappa} e^{it|\nabla|} f \big\|_{S[0,\kappa]} \lesssim \|f\|_{L^2_x}$$

However, since the term in brackets is a bounded function uniformly in ℓ, one can again apply Lemma 2.4.

Now assume that F is a wave-packet atom, i.e., $F = \sum_{\kappa\in\mathcal{C}_\ell} F_\kappa$ with

$$\sum_{\kappa\in\mathcal{C}_\ell} \|F_\kappa\|^2_{\mathrm{NF}[\kappa]} \le 1 \tag{2.60}$$

where the F_κ have the wave-packet form as specified in Definition 2.9. We need to show that

$$\sup_{\pm}\ \sup_{\ell'\le -100}\ \sup_{\ell'\le m\le 0} \Big(\sum_{\kappa'\in\mathcal{C}_{\ell'}} \sum_{R\in\mathcal{R}_{0,\pm\kappa',m}} \|P_R Q^{\pm}_{\le 2\ell'-C}\, \phi\|^2_{S[0,\kappa']}\Big)^{\frac12} \lesssim 1 \tag{2.61}$$

We first consider the case $\ell' \le \ell$. Lemma 2.11 implies that it suffices to assume that $\ell' = \ell$ and to show that, uniformly in $\kappa\in\mathcal{C}_\ell$,

$$\|\phi_\kappa\|_{S[0,\kappa]} \le C\|F_\kappa\|_{\mathrm{NF}[\kappa]}$$

with an absolute constant C where

$$\phi_\kappa := \int_{-\infty}^{\infty} \sin((t-s)|\nabla|)\eta_T^+(t-s)F_\kappa(s)\,ds$$

However, this follows immediately from Lemma 2.22 applied to ϕ_κ, the stability property (2.16), and the imbedding (2.19); note that the term $\Box^{-1}F_1$ in (2.42) can be ignored as it was dealt with in the beginning of this proof. Finally, the case $\ell' \ge \ell$ is reduced the to $\ell' = \ell$ by means of Lemma 2.7 (note that the Fourier-support of ϕ_κ equals that of F_κ).

It remains to control the Strichartz norms (2.14). Due to Corollary 2.23, we may ignore the projection $Q_{<j}$. We split the argument into two parts: First, we will prove the estimate

$$\Big(\sum_{c\in\mathcal{D}_{0,\ell}} \Big\| P_c \int_{-\infty}^{\infty} e^{\pm i(t-s)|\nabla|} F(s)\,ds\Big\|^2_{L^4_t L^\infty_x}\Big)^{\frac12} \lesssim 2^{\frac{\ell}{2}}\|F\|_{N[0]} \tag{2.62}$$

for any F as in (2.60), cf. Lemma 2.1. Second, we take the η_T^+ cut-off as in (2.56) into account which then yields the full result. This second step is done by an adaptation of the Christ-Kiselev argument and will result in the loss of a power $2^{\ell\delta}$ where $\delta > 0$ can be made arbitrarily small. Lemma 2.2 reduces the proof of (2.62) to the bound

$$\Big\| \int_{-\infty}^{\infty} e^{\mp is|\nabla|} F(s)\,ds \Big\|_{L^2_x} \lesssim \|F\|_{N[0]}$$

By orthogonality, it suffices to show that uniformly in κ

$$\Big\| \int_{-\infty}^{\infty} e^{\mp is|\nabla|} F_\kappa(s)\,ds \Big\|_{L^2_x} \lesssim \inf_{\omega\notin 2\kappa} d(\omega,\kappa)^{-1}\|F_\kappa\|_{L^1_{t_\omega} L^2_{x_\omega}}$$

with F_κ as in (2.60). By Plancherel, this is the same as

$$\big\|\widehat{F}_\kappa(\pm|\xi|,\xi)\big\|_{L^2_\xi} \lesssim d(\omega,\kappa)^{-1}\|F_\kappa\|_{L^1_{t_\omega} L^2_{x_\omega}}$$

where we choose an arbitrary $\omega \notin 2\kappa$. As above, we may set $F_\kappa = \delta(t_\omega - t_\omega^{(0)})f_\kappa(x_\omega)$ where $t_\omega^{(0)} \in \mathbb{R}$ is an arbitrary number and $f_\kappa \in L^2_{x_\omega}$ is an arbitrary function whose Fourier support is contained in the projection of the Fourier support of F_κ onto the ξ_ω-plane. This reduces us further to the bound

$$\|\hat f_\kappa(\xi_\omega)\|_{L^2_\xi} \lesssim d(\omega,\kappa)^{-1}\|f_\kappa\|_{L^2_{\xi_\omega}} \tag{2.63}$$

where on the left-hand side we regard ξ_ω as a function of ξ. By Lemma 2.6 the Jacobian obeys $\left|\frac{\partial\xi}{\partial\xi_\omega}\right| \sim d(\omega,\kappa)^{-2}$ which implies (2.63). This concludes the first step, i.e., the proof of (2.62). Note that our proof of (2.62) applies to any F which can be written in the form $F = \sum_\kappa F_\kappa$ provided F_κ satisfy

$$\mathrm{supp}(\widehat{F}_\kappa) \subset \{\xi\in\mathbb{R}^2 \,:\, |\xi|\sim 1,\ \widehat{\xi}\in\kappa\}$$

In other words, one does not need any condition on the modulations of F_κ. This fact will be most important for the remainder of the proof (since we will need to multiply F by cutoff functions in time). Our next goal is to establish the estimate, with $\delta > 0$ arbitrarily small,

$$\Big(\sum_{c\in\mathcal{D}_{0,\ell}} \Big\| P_c \int_{-\infty}^{\infty} e^{\pm i(t-s)|\nabla|} \eta_T^+(t-s) F(s)\, ds \Big\|_{L_t^4 L_x^\infty}^2 \Big)^2 \lesssim 2^{(\frac12-\delta)\ell} \|F\|_{N[0]} \tag{2.64}$$

for any F as in (2.60) but without any restriction on the modulations of each F_κ. For the remainder of the proof we will *fix* such a Schwartz function F. Moreover, $\|\cdot\|^2$ without any subscripts will mean the sum in (2.60). As mentioned before, we prove (2.64) by an adaptation of the Christ-Kiselev lemma. The latter does not apply directly since the null-frame norm in (2.60) is not of pure Lebesgue type. We make the following preliminary observation. Let $\chi_E = \chi_E(t)$ act only in the time variable and define the map $\mu(E) := \|\chi_E F\|^2$ as a set function on the Borel sets of $\mathbb{R}$. Then one has the following σ-subadditivity property with $\{E_j\} \subset \mathbb{R}$ an arbitrary collection of pairwise disjoint Borel sets:

$$\begin{aligned}
\sum_j \mu(E_j) &= \sum_j \|\chi_{E_j} F\|^2 = \sum_\kappa \sum_j \inf_{\omega\notin 2\kappa} d(\omega,\kappa)^{-2} \|\chi_{E_j} F_\kappa\|_{L^1_{t_\omega} L^2_{x_\omega}}^2 \\
&\le \sum_\kappa \inf_{\omega\notin 2\kappa} d(\omega,\kappa)^{-2} \sum_j \|\chi_{E_j} F_\kappa\|_{L^1_{t_\omega} L^2_{x_\omega}}^2 \\
&\le \sum_\kappa \inf_{\omega\notin 2\kappa} d(\omega,\kappa)^{-2} \Big\| \Big(\sum_j \|\chi_{E_j} F_\kappa\|_{L^2_{x_\omega}}^2 \Big)^{\frac12} \Big\|_{L^1_{t_\omega}}^2 \\
&= \sum_\kappa \inf_{\omega\notin 2\kappa} d(\omega,\kappa)^{-2} \|F_\kappa\|_{L^1_{t_\omega} L^2_{x_\omega}}^2 \\
&= \sum_\kappa \|F_\kappa\|_{\mathrm{NF}[\kappa]}^2 \le 1
\end{aligned}$$

In view of this property it suffices to prove (2.64) for F which are supported on intervals[9] of size T in time and we may also replace η_T^+ by the indicator $\chi_{[s<t]}$. We now perform a Whitney decomposition of the triangle

$$\Delta_T := \{(t,s) \,:\, 0 \le s \le t \le T\}$$

by means of squares (we have shifted the support of F to be contained in $[0,T]$). This yields finitely many disjoint squares of the form

$$\mathcal{Q} := \Big\{ I_{m,n} \times J_{m,n} \big\}_{n\ge 0,\ 1\le m\le M_n}$$

with intervals $I_{m,n}, J_{m,n}$ such that $M_n \le 2^n$ and

$$\Delta = \bigcup_{n\ge0} \bigcup_{1\le m\le 2^n} I_{m,n} \times J_{m,n}$$

$$|I_{m,n}| = |J_{m,n}| = T2^{-n} \quad \forall\, 1 \le m \le M_n,\ n\ge 0$$

$$\mathrm{dist}(I_{m,n}\times J_{m,n}, \{s=t\}) \in (T2^{-n}/10, 10T2^{-n}) \quad \forall\, 1\le m\le M_n,\ n\ge 0$$

We call any two intervals I, J of length $T2^{-n}$ *related* provided $I \times J \in \mathcal{Q}$. Note that any I can be related to at most 20 of the J intervals. To each $n \ge 0$ we now also associate 2^n pairwise disjoint intervals $\{\widetilde{J}_{m,n}\}_{1\le m\le 2^n}$ which partition $[0,T]$ and with the property that

$$\mu(\widetilde{J}_{m,n}) = \mu(\widetilde{J}_{m',n}) \quad \forall\, 1 \le m, m' \le 2^n$$

The subadditivity of μ implies that $\mu(\widetilde{J}_{m,n}) \le 2^{-n}$. Finally, we introduce an auxiliary function Φ which is piece-wise linear, strictly increasing on $[0,T]$ and which has the property that $\Phi(J_{m,n}) = \widetilde{J}_{m,n}$. In view

[9] Strictly speaking, one would need to choose something like $10T$ here to accommodate the support of η_T^+, but we ignore this issue.

of all these properties

$$\sum_{c\in\mathcal{D}_{0,\ell}} \Big\| P_c \int_{-\infty}^t e^{\pm i(t-s)|\nabla|} F(s)\,ds \Big\|_{L_t^4 L_x^\infty}^2$$

$$\lesssim \sum_{c\in\mathcal{D}_{0,\ell}} \Big(\sum_{n=0}^\infty \Big\| \sum_{m=1}^{M_n} \chi_{\Phi(I_{m,n})}(t) P_c \int_{-\infty}^t e^{\pm i(t-s)|\nabla|} \chi_{\Phi(J_{m,n})}(s) F(s)\,ds \Big\|_{L_t^4 L_x^\infty} \Big)^2$$

Applying Cauchy-Schwarz to the sum over n allows one to bound this further as

$$\lesssim \sum_{c\in\mathcal{D}_{0,\ell}} \sum_{n=0}^\infty (1+n)^2 \Big\| \sum_{m=1}^{M_n} \chi_{\Phi(I_{m,n})}(t) P_c \int_{-\infty}^t e^{\pm i(t-s)|\nabla|} \chi_{\Phi(J_{m,n})}(s) F(s)\,ds \Big\|_{L_t^4 L_x^\infty}^2$$

$$\lesssim \sum_{n=0}^\infty (1+n)^2 \sum_{c\in\mathcal{D}_{0,\ell}} \Big(\sum_{m=1}^{2^n} \Big\| P_c \int_{-\infty}^\infty e^{\pm i(t-s)|\nabla|} \chi_{\widetilde{J}_{m,n}}(s) F(s)\,ds \Big\|_{L_t^4 L_x^\infty}^4 \Big)^{\frac12} \tag{2.65}$$

Label the disks $c\in\mathcal{D}_{0,\ell}$ by $\{c_k\}_{k=1}^K$, $K\sim 2^{-2\ell}$, and denote for fixed n,

$$a_{km,n} := \Big\| P_{c_k} \int_{-\infty}^\infty e^{\pm i(t-s)|\nabla|} \chi_{\widetilde{J}_{m,n}}(s) F(s)\,ds \Big\|_{L_t^4 L_x^\infty}^2$$

The previous bound now takes the form

$$\sum_{c\in\mathcal{D}_{0,\ell}} \Big\| P_c \int_{-\infty}^t e^{\pm i(t-s)|\nabla|} F(s)\,ds \Big\|_{L_t^4 L_x^\infty}^2 \lesssim \sum_{n=0}^\infty (1+n)^2 \sum_{k=1}^K \Big(\sum_{m=1}^{2^n} a_{km,n}^2 \Big)^{\frac12}$$

In view of (2.62) (and the remark at the end of its proof concerning time cutoffs)

$$\sum_{k=1}^K a_{km,n} \lesssim 2^\ell \big\| \chi_{\widetilde{J}_{m,n}} F \big\|^2 = 2^\ell \mu(\widetilde{J}_{m,n}) \le 2^\ell 2^{-n}$$

By Lemma 2.24 with $\sigma = 2^{\ell-n}$, $M=2^n$, $K=2^{-2\ell}$,

$$\sum_{k=1}^K \Big(\sum_{m=1}^{2^n} a_{km,n}^2 \Big)^{\frac12} \lesssim 2^{(1-2\delta)\ell} 2^{-\delta n}$$

for any $0\le\delta\le 1$. In view of (2.65), one obtains (2.64). □

As a simple corollary, we now obtain the following continuity result. Recall that the norm of $S[k]$ can also be defined for non-integer k, cf. (2.18). The continuity in k is not obvious due to the various Fourier multipliers in (2.14) and (2.15) over infinitely many scales.

Corollary 2.27. *Let ϕ be a Schwartz function in $\mathbb{R}^{1+2}$ which is adapted to $k\in\mathbb{R}$. Then*

$$\lim_{h\to 0} \|\phi\|_{S[k+h]} = \|\phi\|_{S[k]}$$

Proof. By (2.18),

$$\lambda^{-1} \|\phi(\lambda^{-1}\cdot)\|_{S[k]} = \|\phi\|_{S[k+\log_2\lambda]}$$

It therefore suffices to note that by the energy estimate

$$\begin{aligned} \big| \lambda^{-1}\|\phi(\lambda^{-1}\cdot)\|_{S[k]} - \|\phi\|_{S[k]} \big| &\le \|\lambda^{-1}\phi(\lambda^{-1}\cdot) - \phi\|_{S[k]} \\ &\le \|(\lambda^{-1}\phi(\lambda^{-1}\cdot)-\phi)[0]\|_{L^2\times\dot H^{-1}} + \|\Box(\lambda^{-1}\phi(\lambda^{-1}\cdot)-\phi)\|_{N[k]} \\ &\lesssim \|(\lambda^{-1}\phi(\lambda^{-1}\cdot)-\phi)[0]\|_{L^2\times\dot H^{-1}} + \|\Box(\lambda^{-1}\phi(\lambda^{-1}\cdot)-\phi)\|_{L_t^1 \dot H^{-1}} \to 0 \end{aligned}$$

as $\lambda\to 1$. □

2.4. **A stronger $S[k]$-norm, and time localizations.** The energy estimate of Proposition 2.26 and Lemma 2.25 can be summarized as the statement that $\|\phi\|_{S[k]} \lesssim |||\phi|||_{S[k]}$ where

$$|||\phi|||_{S[k]} := \|\phi\|_{L_t^\infty L_x^2} + \|\Box\phi\|_{N[k]} \tag{2.66}$$

for any space-time Schwartz function ϕ which is adapted to $k \in \mathbb{Z}$. To see this, one estimates

$$\begin{aligned} \|\phi\|_{S[k]} &\lesssim \|I\phi\|_{S[k]} + \|I^c\phi\|_{S[k]} \\ &\lesssim \|((I\phi)(0), (\partial_t I\phi)(0))\|_{L^2\times \dot H^{-1}} + \|I\Box\phi\|_{N[k]} + \|I^c\Box\phi\|_{N[k]} \\ &\lesssim \|\phi\|_{L_t^\infty L_x^2} + \|\Box\phi\|_{N[k]} = |||\phi|||_{S[k]} \end{aligned}$$

To remove I from the right-hand side here one uses Corollary 2.23.

We shall henceforth use this stronger norm and the resulting smaller $S[k]$-space. We introduce this norm because it leads to an improvement over the bilinear bound (2.30) in the case of high-high interactions, see Lemma 4.5 below. This improvement reflects a smoothing effect of convolutions of measures supported on the light cone. It thus *cannot* be obtained using the $S[k,\kappa]$ norms alone, since (2.30) is based on Hölder's inequality

$$L^2_{t_\omega} L^\infty_{x_\omega} \cdot L^\infty_{t_\omega} L^2_{x_\omega} \hookrightarrow L^2_t L^2_x$$

which does not allow for any gain in regularity. It will be essential to note that Corollary 2.16 still applies to the stronger norm $|||\cdot|||$:

Lemma 2.28. *For all ϕ which are adapted to $k \in \mathbb{Z}$ and all $j \in \mathbb{Z}$ one has $|||Q_{\le j}\phi|||_{S[k]} \le C|||\phi|||_{S[k]}$ with some absolute constant C.*

Proof. This follows immediately from Lemma 2.15 and Corollary 2.23. □

Another property which the stronger norm inherits is that it is finite on free wave, cf. Lemma 2.4. More precisely, for any ϕ which is adapted to k and satisfies $\phi = Q_{\le k}\phi$,

$$\begin{aligned} |||\phi|||_{S[k]} &= \|\phi\|_{L_t^\infty L_x^2} + \|\Box\phi\|_{N[k]} \\ &\le \|\phi\|_{L_t^\infty L_x^2} + \|\Box\phi\|_{\dot X_k^{0,-\frac12,1}} \lesssim \|\phi\|_{\dot X_k^{0,\frac12,1}} \end{aligned}$$

As in [22], one needs to allow for time-localized versions of $S[k]$, both relative to the original $\|\cdot\|_{S[k]}$, as well as the stronger $|||\cdot|||$-norm. This has to do with the fact that the we need to derive a priori bounds in these spaces for Schwartz functions ψ_α which satisfy (1.12)–(1.14) on some time interval $[-T,T]$. Since the norms of the $S[k]$ and $N[k]$ spaces are defined in phase space, one cannot simply define these norms by time truncations. Rather, one proceeds as in [57] and [22] by means of Schwartz extensions: with ψ and $\tilde\psi$ both Schwarz functions, and $T \ge 0$,

$$\begin{aligned} \|\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)} &:= \inf_{\tilde\psi|_{[-T,T]}=\psi|_{[-T,T]}} \|P_k\tilde\psi\|_{S[k]} \\ |||\psi|||_{S[k]([-T,T]\times\mathbb{R}^2)} &:= \inf_{\tilde\psi|_{[-T,T]}=\psi|_{[-T,T]}} |||P_k\tilde\psi|||_{S[k]} \end{aligned} \tag{2.67}$$

It is easy to see that the triangle inequality holds for these expressions and that they are actually norms. Moreover, it is clear that these norms are nondecreasing in T. Following [22], we now verify that these norms are continuous in T.

Lemma 2.29. *Let ψ be the restriction of some Schwartz function ψ_0 in $\mathbb{R}^{1+2}$ to the time interval $[-T_0,T_0]$ where $T_0 > 0$. Then*

$$\|\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)} \quad \textit{and} \quad |||\psi|||_{S[k]([-T,T]\times\mathbb{R}^2)}$$

are nondecreasing and continuous in $0 \le T < T_0$.

Proof. The definition of $S[k]$ with respect to either norm can be extended to non-integer k. Given $T > 0$, let $|\varepsilon|$ be very small and set $\lambda := \frac{T+\varepsilon}{T}$. Then

$$\|P_k\psi\|_{S[k]([-T-\varepsilon,T+\varepsilon]\times\mathbb{R}^2)} = \|P_{k+\mu}\psi_\lambda\|_{S[k+\mu]([-T,T]\times\mathbb{R}^2)}$$

where $\mu := \log_2 \lambda$ and $\psi_\lambda(t,x) := \lambda\psi(\lambda t, \lambda x)$, and similarly for $\|\!| \cdot |\!\|$. Clearly, for $\varepsilon > 0$,

$$\begin{aligned}
&\big| \|P_k\psi\|_{S[k]([-T-\varepsilon,T+\varepsilon]\times\mathbb{R}^2)} - \|P_k\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)} \big| \\
&= \big| \|P_{k+\mu}\psi_\lambda\|_{S[k+\mu]([-T,T]\times\mathbb{R}^2)} - \|P_k\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)} \big| \\
&\le \big| \|P_{k+\mu}\psi\|_{S[k+\mu]([-T,T]\times\mathbb{R}^2)} - \|P_k\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)} \big| + \|P_{k+\mu}(\psi-\psi_\lambda)\|_{S[k+\mu]([-T,T]\times\mathbb{R}^2)}
\end{aligned}$$

By the energy estimate,

$$\begin{aligned}
\|P_{k+\mu}(\psi-\psi_\lambda)\|_{S[k+\mu]([-T,T]\times\mathbb{R}^2)} &\lesssim \|P_{k+\mu}(\psi-\psi_\lambda)\|_{S[k+\mu]} \\
&\lesssim \|(\psi-\psi_\lambda)[0]\|_{L^2\times\dot H^{-1}} + \|\Box P_{k+\mu}(\psi-\psi_\lambda)\|_{N[k+\mu]} \\
&\lesssim \|(\psi-\psi_\lambda)[0]\|_{L^2\times\dot H^{-1}} + \|\Box P_{k+\mu}(\psi-\psi_\lambda)\|_{L^1_t\dot H^{-1}(\mathbb{R}^{1+2})} \to 0
\end{aligned}$$

as $\lambda \to 1$. By Corollary 2.27,

$$\lim_{\lambda\to 1} \|P_{k+\mu}\psi_\lambda\|_{S[k+\mu]([-T,T]\times\mathbb{R}^2)} = \|P_k\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)}$$

which implies that

$$\lim_{\varepsilon\to 0+} \|P_k\psi\|_{S[k]([-T-\varepsilon,T+\varepsilon]\times\mathbb{R}^2)} = \|P_k\psi\|_{S[k]([-T,T]\times\mathbb{R}^2)}$$

as claimed. The case of $T = 0$ follows directly from the energy estimate. The case of $\|\!| \cdot |\!\|$ is essentially the same. $\square$

We define localized $N[k]$-norms similarly, i.e.,

$$\|\psi\|_{N[k]([-T,T]\times\mathbb{R}^2)} := \inf_{\tilde\psi|_{[-T,T]}=\psi|_{[-T,T]}} \|P_k\tilde\psi\|_{N[k]}$$

for Schwartz functions. In particular, one has a localized version of (2.66)

$$\|\!|\phi|\!\|_{S[k]([-T,T]\times\mathbb{R}^2)} := \|\phi\|_{L^2(I;L^2(\mathbb{R}^2))} + \|\Box\phi\|_{N[k]([-T,T]\times\mathbb{R}^2)}$$

Furthermore, later we will also need localized norms on asymmetric time intervals $[-T', T]$ for which the results here of course continue to hold.

Finally, in the perturbative steps to follow, we will need to piece together solutions of time-localized wave equations to solutions on larger time intervals. To justify this procedure we rely on the following lemma.

Lemma 2.30. *Let $I \subset \mathbb{R}$ be a closed interval, with a covering $I = \cup_{j=1}^N I_j$ by closed intervals; assume that the I_j overlap at most two at a time, and that consecutive intervals have intersection with non-empty interior. Then if we are given k-adapted ψ_j with*

$$\|\psi_j\|_{S[k](I_j\times\mathbb{R}^2)} \le c_j, \quad j = 1, 2, \ldots, N,$$

such that $\psi_j|_{I_j\cap I_l} = \psi_\ell|_{I_j\cap I_\ell}$, then defining ψ via $\psi|_{I_j} := \psi_j$, we have

$$\|\psi\|_{S[k](I\times\mathbb{R}^2)} \lesssim \sum_{j=1}^N \|\psi_j\|_{S[k](I_j\times\mathbb{R}^2)}$$

where the implied constant is universal (independent of the decomposition of I or N). The same applies to the norms $\|\!| \cdot |\!\|_{S[k](I_j\times\mathbb{R}^2)}$.

Proof. Chose a partition of unity $\{\chi_j\}$ subordinate to the cover $\{I_j\}$, such that $\mathrm{supp}\chi_j \subset I_j$. We shall select the χ_j in such fashion that $|\mathrm{supp}\chi_j'| \ll |I_j \cap I_k|$, provided the latter is non-zero (which happens only for at most two other k). We first deal with the $\|\cdot\|_{S[k]}$-norms. By assumption, we can find Schwartz extensions $\tilde\psi_j$ of ψ_j, $\forall j$, such that $\|\tilde\psi_j\|_{S[k](\mathbb{R}^{2+1})} \le 2c_j$. We now define

$$\tilde\psi := \sum_{j=1}^N \chi_j\tilde\psi_j$$

and verify the desired bound $\|\tilde{\psi}\|_{S[k](\mathbb{R}^{2+1})} \lesssim \sum c_j$. For simplicity, consider a single interval half-infinite I_1 with neighboring half-infinite I_2, and the corresponding expression

$$\chi_1\tilde{\psi}_1 + \chi_2\tilde{\psi}_2$$

Note that $\chi_1' + \chi_2' = 0$ on the overlap of the intervals. It is easy to see that the only potential difficulty in controlling $\|\chi_1\tilde{\psi}_1 + \chi_2\tilde{\psi}_2\|_{S[k]}$ comes from the "elliptic portion" of $\|\cdot\|_{S[k]}$, as we have introduced the cutoffs whose derivatives we do not a priori control. By scaling invariance, it suffices to consider $k = 0$. Hence consider now

$$P_0Q_j[\chi_1\tilde{\psi}_1 + \chi_2\tilde{\psi}_2]$$

for some $j \gg 1$. We decompose this by applying a frequency trichotmoy

$$\begin{aligned} P_0Q_j[\chi_1\tilde{\psi}_1 + \chi_2\tilde{\psi}_2] =& P_0Q_j[Q_{[j-10,j+10]}(\chi_1)\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\tilde{\psi}_2] \\ &+ P_0Q_j[Q_{<j-10}(\chi_1)\tilde{\psi}_1 + Q_{<j-10}(\chi_2)\tilde{\psi}_2] \\ &+ P_0Q_j[Q_{>j+10}(\chi_1)\tilde{\psi}_1 + Q_{>j+10}(\chi_2)\tilde{\psi}_2] \end{aligned} \tag{2.68}$$

We start by estimating the last line: we have

$$\begin{aligned} \|P_0Q_j[Q_{>j+10}(\chi_1)\tilde{\psi}_1 + Q_{>j+10}(\chi_2)\tilde{\psi}_2]\|_{L^2_{t,x}} &\leq \sum_{r>j+10} \|P_0Q_j[Q_r(\chi_1)Q_{[r-5,r+5]}\tilde{\psi}_1 + Q_r(\chi_2)Q_{[r-5,r+5]}\tilde{\psi}_2]\|_{L^2_{t,x}} \\ &\lesssim \sum_{\ell=1,2}\sum_{r>j+10} 2^{-(1-\epsilon)r}\|Q_{[r-5,r+5]}\tilde{\psi}_\ell\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,1-\epsilon,2}} \end{aligned}$$

¿From here we easily obtain

$$\|\sum_{j>O(1)} P_0Q_j[Q_{>j+10}(\chi_1)\tilde{\psi}_1 + Q_{>j+10}(\chi_2)\tilde{\psi}_2]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,1-\epsilon,2}} \lesssim \sum_{\ell=1,2}\|\tilde{\psi}_\ell\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,1-\epsilon,2}}$$

The second line in (2.68) is estimated similarly, and so we reduce to estimating

$$P_0Q_j[Q_{[j-10,j+10]}(\chi_1)\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\tilde{\psi}_2]$$

We may assume that $\mathcal{F}(\chi_{1,2})$ decay rapidly away from frequency scale $2^R \gg 1$, say. Write

$$\begin{aligned} &\partial_t P_0Q_j[Q_{[j-10,j+10]}(\chi_1)\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\tilde{\psi}_2] \\ &= P_0Q_j[Q_{[j-10,j+10]}\partial_t(\chi_1)\tilde{\psi}_1 + Q_{[j-10,j+10]}\partial_t(\chi_2)\tilde{\psi}_2] \\ &+ P_0Q_j[Q_{[j-10,j+10]}(\chi_1)\partial_t\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\partial_t\tilde{\psi}_2] \end{aligned} \tag{2.69}$$

We start by estimating the second row: we will consider the case $j = R + O(1)$, since in the other cases one obtains additional exponential gains from the frequency localization of $\chi_{1,2}$. But then we can write

$$Q_{[j-10,j+10]}\partial_t(\chi_1) = \tilde{\chi}_1 Q_{[j-10,j+10]}\partial_t(\chi_1) + O_{L^2_t}(R^{-N})$$

where $\tilde{\chi}$ localizes to an interval around $\mathrm{supp}\chi_1'$ of length $R^{-\frac{1}{2}}$, say, and similarly for χ_2. By picking R large enough, we may assume that

$$\tilde{\chi}_1\tilde{\psi}_1 = \tilde{\chi}_2\tilde{\psi}_2$$

Thus we obtain

$$P_0Q_j[Q_{[j-10,j+10]}\partial_t(\chi_1)\tilde{\psi}_1 + Q_{[j-10,j+10]}\partial_t(\chi_2)\tilde{\psi}_2] = \sum_{\ell=1,2} O_{L^2_t}(R^{-N})\tilde{\psi}_\ell$$

and from here we infer

$$\|P_0Q_j[Q_{[j-10,j+10]}\partial_t(\chi_1)\tilde{\psi}_1 + Q_{[j-10,j+10]}\partial_t(\chi_2)\tilde{\psi}_2]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,1-\epsilon,2}} = O(R^{-N'})\sum_{\ell=1,2}\|\tilde{\psi}_\ell\|_{S[0](\mathbb{R}^{2+1})}$$

where we recall the assumption $j = R + O(1)$. The remaining cases $j \leq R$, $j \geq R$ are lead to a similar bound. Next, consider the last line of (2.69); here we write

$$\begin{aligned} &P_0Q_j[Q_{[j-10,j+10]}(\chi_1)\partial_t\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\partial_t\tilde{\psi}_2] \\ &= P_0Q_j[Q_{[j-10,j+10]}(\chi_1)\partial_t Q_{<j+20}\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\partial_t Q_{<j+20}\tilde{\psi}_2] \end{aligned} \tag{2.70}$$

But this we can estimate by

$$
\begin{aligned}
&\|P_0 Q_j[Q_{[j-10,j+10]}(\chi_1)\partial_t Q_{<j+20}\tilde{\psi}_1 + Q_{[j-10,j+10]}(\chi_2)\partial_t Q_{<j+20}\tilde{\psi}_2]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,1-\epsilon,2}} \\
&\lesssim \sum_{\ell=1,2} 2^{-(1-\epsilon)j}\|\partial_t Q_{<j+20}\tilde{\psi}_\ell\|_{L^2_{t,x}}
\end{aligned}
\tag{2.71}
$$

One can now perform the square summation over $j > O(1)$, and gets the upper bound $\lesssim \sum_{l=1,2}\|\tilde{\psi}_l\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,1-\epsilon,2}}$, where the implied constant is universal.
The argument for controlling the $\|\!|\cdot|\!\|_{S[k]}$-norm is similar. One uses

$$\Box\sum_{j=1}^N \chi_j\tilde{\psi}_j = \sum_{j=1}^N \chi_j\Box\tilde{\psi}_j,$$

as well as Lemma 2.19. □

Remark 2.31. In the sequel, we shall use the preceding lemma freely without an explicit reference.

2.5. **Solving the inhomogeneous wave equation in the Coulomb gauge.** Consider the wave equation (1.14), i.e., $\Box\psi_\alpha = F_\alpha$. Here F_α is a nonlinear expression in ψ, but we will not pay attention to this now. In the sequel, we shall require a priori bounds on ψ_α in the $S[k]$-space. To do so, we reduce matters to the energy estimates of Section 2.3 as follows: writing (suppressing α for simplicity)

$$\Box\psi = IF + I^cF$$

one concludes (with both ψ and F global space-time Schwartz functions adapted to frequency 1),

$$\psi(t) = S(t-t_0)(I\psi)[t_0] + \int_{t_0}^t U(t-s)IF(s)\,ds + \Box^{-1}I^cF \tag{2.72}$$

where the final term is obtained by division by the symbol[10] of $\Box$, and the first two terms represent the free wave and the Duhamel integral, respectively. Note that the first term here implicitly depends on all of ψ, not just $\psi[t_0]$, and so in order to actually obtain a bound on $\|\psi\|_S$, one needs to implement a bootstrap argument. Specifically, assume that we a priori have a bound on

$$\|\psi|_{[-T_0,T_0]}\|_S$$

for some $T_0 > 0$. Also, assume that we define $I = \sum_{k\in\mathbb{Z}} P_k Q_{<k+C}$ where $2^C \gg T_0^{-1}$. Then, using the energy estimate from Section 2.3, we claim that

$$\|\psi\|_S \lesssim T_0^{-1}\|\psi|_{[-T_0,T_0]}\|_S + \|F\|_N \tag{2.73}$$

where the implied constant is absolute (the T_0^{-1} here comes from the time-derivative in the initial data). Indeed, this follows from

$$(I\psi)[t_0] = (I(\chi_{[-T_0,T_0]}\psi))[t_0] + (I([1-\chi_{[-T_0,T_0]}]\psi))[t_0]$$

and

$$\min_{t_0\in[-T_0,T_0]}\|\big((I([1-\chi_{[-T_0,T_0]}]\psi))\big)[t_0]\|_{L^2_x\times\dot{H}^{-1}} \ll \|\psi\|_S$$

as well as

$$\|(I(\chi_{[-T_0,T_0]}\psi))[t_0]\|_{L^2_x} = \|\nabla_{t,x}I(\chi_{[-T_0,T_0]}\psi)\|_{L^2_x} \lesssim T_0^{-1}\|\psi|_{[-T_0,T_0]}\|_S$$

due to our choice of I. The above energy inequality then follows immediately.
It is apparent that in order to use this energy inequality, one needs to establish an a priori bound for ψ on a small time interval $[-T_0,T_0]$. In fact, in later applications we will always split the estimates for $P_k\psi$ into the small-time case $|t-t_0| \leq \varepsilon_1 2^{-k}$ and the large time case $|t-t_0| \geq \varepsilon_1 2^{-k}$ (with a small ε_1 that is determined by the specific context - this then requires the constant C in the definition of I to be large). In the small time case, the necessary a priori bound is derived from the div-curl system (1.12), (1.13) for the gauged components. This information is then fed into the large-time case as described above.

[10]In the sequel, we shall understand the operator $\Box^{-1}$ to be division by the symbol unless otherwise stated.

3. Hodge decomposition and null-structures

Here we introduce the actual system of wave equations for which our S and N-spaces allow us to deduce a priori estimates. From the discussion at the very beginning, we recall that the Coulomb components ψ_α satisfy the system (1.14), which has the schematic form

$$\Box\psi_\alpha = i\partial^\beta[\psi_\alpha A_\beta] - i\partial^\beta[\psi_\beta A_\alpha] + i\partial_\alpha[\psi^\beta A_\beta] \tag{3.1}$$

where A_β denotes the Coulomb gauge potential

$$A_\beta = \sum_{j=1,2} \triangle^{-1}\partial_j[\psi^1_\beta\psi^2_j - \psi^2_\beta\psi^1_j]$$

This system in and of itself does not appear to lend itself to good estimates, and to overcome this we have to use a key additional feature, namely the fact that *the flow of* (1.14) *preserves the div-curl system* (1.12), (1.13) in the obvious sense: if the ψ_α at time $t = 0$ are the Coulomb derivative components of an actual map, whence (1.12), (1.13) holds at time $t = 0$, then the corresponding solution of (1.14) satisfies this system on its entire time interval of existence. The div-curl system allows us to decompose the components ψ_α as the sum of a gradient term and an error term solving an elliptic equation, see (1.15). Thus we have schematic identities of the form

$$\psi_\alpha = R_\alpha\psi + \chi_\alpha$$

Substituting the gradient terms introduces the desired null-structure. The present section serves to make this decomposition of the nonlinear source terms precise. We now describe this procedure for each of the three terms on the right-hand side of (3.1). First, define $\partial_j^{-1} := \Delta^{-1}\partial_j$ and

$$\begin{aligned}
\mathcal{Q}_{\beta j}(\psi,\psi) &= R_\beta\psi^1 R_j\psi^2 - R_j\psi^1 R_\beta\psi^2\\
\mathcal{Q}_{\beta j}(\psi,\chi) &= R_\beta\psi^1 \chi^2_j - R_j\psi^1 \chi^2_\beta\\
\mathcal{Q}_{\beta j}(\chi,\psi) &= \chi^1_\beta R_j\psi^2 - \chi^1_j R_\beta\psi^2\\
\mathcal{Q}_{\beta j}(\chi,\chi) &= \chi^1_\beta\chi^2_j - \chi^1_j\chi^2_\beta
\end{aligned}$$

Then, adopting the Einstein summation convention,

$$\begin{aligned}
i\partial^\beta[\psi_\alpha A_\beta] &= i\partial^\beta[\psi_\alpha\, I^c\partial_j^{-1}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\partial^\beta[\psi_\alpha\, I\partial_j^{-1}\mathcal{Q}_{\beta j}(\psi,\psi)]\\
&\quad + i\partial^\beta[\psi_\alpha\,\partial_j^{-1}\mathcal{Q}_{\beta j}(\psi,\chi)] + i\partial^\beta[\psi_\alpha\,\partial_j^{-1}\mathcal{Q}_{\beta j}(\chi,\psi)] + i\partial^\beta[\psi_\alpha\,\partial_j^{-1}\mathcal{Q}_{\beta j}(\chi,\chi)]
\end{aligned} \tag{3.2}$$

The two main terms here are the trilinear ones in ψ. We introduced the modulation cutoff I in front of $\mathcal{Q}_{\beta j}$ since the two resulting expressions are estimated differently: for the second, one uses a trilinear null-form structure, see (5.46) below, whereas for the first the bilinear null-form $\mathcal{Q}_{\beta j}$ suffices. Note that the other three terms involving χ are quintilinear and septilinear in ψ, respectively, due to (1.16). These are discussed in greater detail below, under the heading "higher order errors".
Next,

$$\begin{aligned}
-i\partial^\beta[\psi_\beta A_\alpha] &= -i\partial^\beta[\psi_\beta\,\partial_j^{-1}\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\partial^\beta[\psi_\beta\,\partial_j^{-1}\mathcal{Q}_{\alpha j}(\psi,\chi)] - i\partial^\beta[\psi_\beta\,\partial_j^{-1}\mathcal{Q}_{\alpha j}(\chi,\psi)]\\
&\quad - i\partial^\beta[\psi_\beta\,\partial_j^{-1}\mathcal{Q}_{\alpha j}(\chi,\chi)]
\end{aligned}$$

The χ-terms need to be decomposed further, whereas the main term here is again the trilinear one in ψ, which we now rewrite as follows:

$$-i\partial^\beta[\psi_\beta\,\partial_j^{-1}\mathcal{Q}_{\alpha j}(\psi,\psi)] = -i\partial^\beta[\psi_\beta\,\partial_j^{-1}I^c\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\partial^\beta[\psi_\beta\,\partial_j^{-1}I\mathcal{Q}_{\alpha j}(\psi,\psi)] \tag{3.3}$$

The first term on the right-hand side will be estimated as is, whereas the second term now needs to be rewritten according to the Littlewood-Paley trichotomy, in order to make it amenable to our estimates:

$$
\begin{aligned}
&-i\partial^\beta[\psi_\beta\,\partial_j^{-1}I\mathcal{Q}_{\alpha j}(\psi,\psi)] = \\
&= -i\sum_k P_k[\partial^\beta\psi_\beta\,\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\sum_k P_k[R_\beta\psi\;\partial_j^{-1}IP_{<k-5}\partial^\beta\mathcal{Q}_{\alpha j}(\psi,\psi)] \\
&\qquad\qquad - i\sum_k P_k[\chi_\beta\;\partial_j^{-1}IP_{<k-5}\partial^\beta\mathcal{Q}_{\alpha j}(\psi,\psi)]
\end{aligned} \tag{3.4}
$$

$$
-i\sum_k \partial^\beta P_k[P_{>k}R_\beta\psi\,\partial_j^{-1}IP_{>k+5}\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\sum_k \partial^\beta P_k[P_{>k}\chi_\beta\,\partial_j^{-1}IP_{>k+5}\mathcal{Q}_{\alpha j}(\psi,\psi)] \tag{3.5}
$$

$$
-i\sum_k \partial^\beta[P_{<k+10}R_\beta\psi\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\sum_k \partial^\beta[P_{<k+10}\chi_\beta\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\alpha j}(\psi,\psi)] \tag{3.6}
$$

The terms involving χ are expanded further as explained below. For the first term on the right-hand side of (3.4) one replaces $\partial^\beta\psi_\beta$ by the right-hand side of (1.13) which leads to a quintilinear term. The second term can be estimated since the ∂^β-term falls on the small frequencies.
Finally, the third term in (3.1) is treated as follows:

$$
i\partial_\alpha[\psi^\beta A_\beta] = i\partial_\alpha[\psi^\beta I^c A_\beta] + i\partial_\alpha[\psi^\beta IA_\beta] \tag{3.7}
$$

The first term on the right-hand side of (3.7) is estimated as is; in fact, it is essential that one does not perform the Hodge decomposition in the first slot since otherwise $\beta = 0$ would create problems if ψ has large modulation. For the second term, one needs to distinguish frequency interactions as before:

$$
i\partial_\alpha[\psi^\beta IA_\beta] = i\sum_k P_k[\partial_\alpha\psi^\beta\,\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\sum_k P_k[\psi^\beta\;\partial_j^{-1}IP_{<k-5}\partial_\alpha\mathcal{Q}_{\beta j}(\psi,\psi)] \tag{3.8}
$$

$$
+\,i\sum_k \partial_\alpha P_k[P_{>k+5}R^\beta\psi\,\partial_j^{-1}IP_{>k}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\sum_k \partial_\alpha P_k[P_{>k+5}\chi^\beta\,\partial_j^{-1}IP_{>k}\mathcal{Q}_{\beta j}(\psi,\psi)] \tag{3.9}
$$

$$
+\,i\sum_k \partial_\alpha[P_{<k+10}R^\beta\psi\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\sum_k \partial_\alpha[P_{<k+10}\chi^\beta\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\beta j}(\psi,\psi)] \tag{3.10}
$$

The χ-terms need to be expanded further, see below, whereas the ψ-terms in (3.9) and (3.10) are estimated as they are. The second term on right-hand side of (3.8) is expanded by means of the Hodge decomposition:

$$
i\sum_k P_k[\psi^\beta\;\partial_j^{-1}IP_{<k-5}\partial_\alpha\mathcal{Q}_{\beta j}(\psi,\psi)] = i\sum_k P_k[R^\beta\psi\;\partial_j^{-1}IP_{<k-5}\partial_\alpha\mathcal{Q}_{\beta j}(\psi,\psi)] \tag{3.11}
$$

$$
+\,i\sum_k P_k[\chi^\beta\;\partial_j^{-1}IP_{<k-5}\partial_\alpha\mathcal{Q}_{\beta j}(\psi,\psi)] \tag{3.12}
$$

The trilinear estimates of Section 5 cover (3.11), and (3.12) is handled below, under 'higher order errors'.
Finally, the first term on the right-hand side of (3.8) is rewritten by means of (1.12):

$$
i\sum_k P_k\partial_\alpha\psi^\beta\,\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\beta j}(\psi,\psi) = i\sum_k P_k\partial^\beta\psi_\alpha\,\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\beta j}(\psi,\psi) + \texttt{quintilinear terms} \tag{3.13}
$$

where the quintilinear terms arise by using the curl identity for $\partial_\alpha\psi^\beta - \partial^\beta\psi_\alpha$ into this expression. Note that we have switched the derivatives ∂_α and ∂_β.
We still have to explain how to deal with the higher order terms involving at least one factor of χ.

Higher order errors.
Note that these arise in two ways: first, we generate errors by replacing the Gauge potential A_β in

$$
i\partial^\beta[\psi_\alpha A_\beta]
$$

by a $\mathcal{Q}_{\beta j}(\psi,\psi)$ null-form, and similarly for the remaining types of terms

$$
i\partial^\beta[\psi_\beta A_\alpha], \quad i\partial_\alpha[\psi^\beta A_\beta]
$$

We shall call the higher order terms generated by this process (and later further Hodge decompositions applied to them) *of the first type or kind.*
Second, we generate errors of the schematic form

$$\chi\nabla^{-1}IQ_{\beta j}(\psi,\psi),$$

and we call these together with all the terms generated by them upon applying further Hodge decompositions *of the second type or kind.* For simplicity, we omit frequency localizations in the ensuing discussion. Considering the errors of the first kind, these are of the schematic form

$$\nabla_{x,t}[\psi\nabla^{-1}[\chi\psi]], \quad \nabla_{x,t}[\psi\nabla^{-1}[\chi\chi]],$$

where we recall from the very beginning, section 1, that

$$\chi = \nabla^{-1}[\psi\nabla^{-1}(\psi^2)],$$

whence the above terms may be thought of as quintilinear and septilinear. Now as they are written, we cannot yet quite estimate these expressions, and we need to introduce more null-structure, by expanding the $\nabla^{-1}(\psi^2)$ in

$$\chi = \nabla^{-1}[\psi\nabla^{-1}(\psi^2)],$$

into a $Q_{\nu j}$-null-form as well as even higher order error terms. To keep track of things we associate an expansion graph, i.e., a simple binary tree with the expressions generated: represent the original terms

$$\nabla_{x,t}[\psi A_\beta]$$

by a simple node, and whenever we replace one of the factors in the (schematically written)

$$A_\beta = \nabla^{-1}(\psi^2)$$

by the corresponding χ, we draw a downward edge pointing left or right corresponding to which factor we replace. We can now exactly specify the full expansion of the higher order errors of first type:

FIGURE 3. An example of an expansion graph

Precise description of expansion for errors of first type:
keep applying Hodge decompositions to the inner $\nabla^{-1}(\psi^2)$ *in all factors*

$$\chi = \nabla^{-1}[\psi\nabla^{-1}(\psi^2)],$$

generated until the associated expansion graph has a directed subgraph of length four. Then the process stops. Note that formally, the terms with a directed subgraph of length four thereby generated are up to at least the 11th degree in ψ.

Next, we apply a similar process to the errors of the second type. We represent the first such error, schematically given by

$$\nabla_{x,t}[\chi\nabla^{-1}IQ_{\nu j}(\psi,\psi)]$$

by a simple node, and whenever we apply a Hodge decomposition to one of the factors of $\nabla^{-1}(\psi^2)$ in

$$\chi = \nabla^{-1}(\psi\nabla^{-1}(\psi^2))$$

we draw a downward edge pointing left or right, thereby generating an associated expansion graph. Then we have

Precise description of expansion for errors of second type:
Keep applying Hodge decompositions as above until the associated expansion graph has a directed subgraph of length three. Then the process stops. Again we generated a list of errors of degree of multilinearity up to order 11 and more in ψ.

To summarize this discussion, we have now recast our system of equations in the form

$$\Box\psi_\alpha = \sum_{i=1}^{5} F_\alpha^{2i+1}$$

where the superscript indicates the minimum degree of multilinearity of the corresponding terms in ψ ($i-1$ indicates the length of a directed subgraph in the corresponding graph representation), and the *leading cubic terms* F_α^3 can be expressed as

(3.14)
$$\begin{aligned}
F_\alpha^3 =& i\partial^\beta[\psi_\alpha\, I^c\partial_j^{-1}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\partial^\beta[\psi_\alpha\, I\partial_j^{-1}\mathcal{Q}_{\beta j}(\psi,\psi)] - i\partial^\beta[\psi_\beta\,\partial_j^{-1}I^c\mathcal{Q}_{\alpha j}(\psi,\psi)]\\
&- i\sum_k P_k[R_\beta\psi\ \partial_j^{-1}IP_{<k-5}\partial^\beta\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\sum_k \partial^\beta P_k[P_{>k}R_\beta\psi\,\partial_j^{-1}IP_{>k+5}\mathcal{Q}_{\alpha j}(\psi,\psi)]\\
&- i\sum_k \partial^\beta P_k[P_{<k+10}R_\beta\psi\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\alpha j}(\psi,\psi)] + i\partial_\alpha[\psi^\beta I^c\partial_j^{-1}\mathcal{Q}_{\beta j}(\psi,\psi)]\\
&+ i\sum_k P_k[\partial^\beta\psi_\alpha\,\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\sum_k P_k[R^\beta\psi\ \partial_j^{-1}IP_{<k-5}\partial_\alpha\mathcal{Q}_{\beta j}(\psi,\psi)]\\
&+ i\sum_k \partial_\alpha P_k[P_{>k}R^\beta\psi\,\partial_j^{-1}IP_{>k+5}\mathcal{Q}_{\beta j}(\psi,\psi)] + i\sum_k \partial_\alpha[P_{<k+10}R^\beta\psi\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\beta j}(\psi,\psi)]
\end{aligned}$$

Here it is very important to note that the second as well as the eighth term on the right contribute a *magnetic potential interaction term* of the form

$$2i\sum_k P_k\partial^\beta\psi_\alpha\,\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\beta j}(\psi,\psi), \tag{3.15}$$

the idea being that we interpret the low-frequency term $\partial_j^{-1}IP_{<k-5}\mathcal{Q}_{\beta j}(\psi,\psi)$ as a magnetic gauge potential. The main issue here is that these high-low interactions cannot be made small in general which creates problems for a bootstrap argument. Hence, in order to prove the core perturbative results in Section 9 we shall have to move these interaction terms to the left-hand side, i.e., build them into the linear operator. For later reference, we shall denote by F_α^{3k}, $k=1,2,3$, those trilinear terms contributed by the first, second or third term in (3.1); thus for example, we write

(3.16)
$$\begin{aligned}
F_\alpha^{32} =& i\partial^\beta[\psi_\beta\,\partial_j^{-1}I^c\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\sum_k \partial^\beta P_k[P_{>k}R_\beta\psi\,\partial_j^{-1}IP_{>k+5}\mathcal{Q}_{\alpha j}(\psi,\psi)]\\
&- i\sum_k P_k[R_\beta\psi\ \partial_j^{-1}IP_{<k-5}\partial^\beta\mathcal{Q}_{\alpha j}(\psi,\psi)] - i\sum_k \partial^\beta[P_{<k+10}R_\beta\psi\,\partial_j^{-1}IP_{[k-5,k+5]}\mathcal{Q}_{\alpha j}(\psi,\psi)]
\end{aligned}$$

Furthermore, we denote by

$$F_\alpha^{3k}(\psi_1,\psi_2,\psi_3)$$

the corresponding multilinear expressions. We also introduce frequency localized versions

$$F_\alpha^{3k}(\psi_1;P_{<\ell};\psi_2,\psi_3)$$

in which one includes a cutoff $P_{<\ell}$ in front of all instances of $\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)$, and similarly for other multipliers $P_{\leq\ell}$ etc.

4. Bilinear estimates involving S and N spaces

In this section we develop some of the required bilinear bounds. First, we present some bounds from $S \times S$ into L^2_{tx}, in particular one which involves a gain in the high-high case and which does not appear in [57] or [22], see Lemma 4.7 below. This result allows for better control on products $\phi_1 \phi_2$ of S-waves and will be most useful in the trilinear case. In addition, as in the aforementioned references we consider the case of $\phi_1 \in S$ and $\phi_2 \in N$. This section concludes with bilinear estimates for null-forms.

4.1. **Basic L^2-bounds.** To begin with, we present the following geometric lemma for cones, see [56] for a similar result. It will be used repeatedly.

Lemma 4.1. *Suppose ϕ_1, ϕ_2 are such that*

$$\mathrm{supp}(\widehat{\phi_j}) \subset \big\{(\xi,\tau) \mid |\xi| \sim 2^{k_j},\ \big||\xi| - |\tau|\big| \sim 2^{\ell_j}\big\}$$

for $j = 1, 2$. Let $\ell_0, k_0 \in \mathbb{Z}$ and assume that there exists $j_0 \in \{0,1,2\}$ so that

$$\ell_{j_0} > \ell_j + C \qquad \forall\, j \in \{0,1,2\} \setminus \{j_0\} \tag{4.1}$$

Then there is the following dichotomy:

(A) If $k_0 = k_{\max} + O(1)$, then

$$P_{k_0} Q_{\ell_0}(\phi_1 \phi_2) = P_{k_0} Q_{\ell_0}\Big(\sum_{\kappa_1,\kappa_2} P_{k_1,\kappa_1}\phi_1 \cdot P_{k_2,\kappa_2}\phi_2\Big) \tag{4.2}$$

where κ_1, κ_2 are caps of size $C^{-1}r$ and separation $\mathrm{dist}(\kappa_1,\kappa_2) \sim r$ *with*

$$r := 2^{(\ell_{\max} - k_{\min})/2}$$

In particular, $\ell_{\max} \le k_{\min} + O(1)$.

(B) If $k_0 < k_{\max} - C$, then

$$P_{k_0} Q_{\ell_0}(\phi_1 \phi_2) = \sum_{\varepsilon = \pm} P_{k_0} Q_{\ell_0}\Big(\sum_{\kappa} P_{k_1,\kappa}\phi_1^{(\varepsilon)} \cdot P_{k_2,-\kappa}\phi_2^{(\varepsilon)}\Big) \tag{4.3}$$

$$+ \sum_{\varepsilon=\pm} P_{k_0} Q_{\ell_0}\Big(\sum_{\kappa_1,\kappa_2} P_{k_1,\kappa_1}\phi_1^{(\varepsilon)} \cdot P_{k_2,-\kappa_2}\phi_2^{(-\varepsilon)}\Big) \tag{4.4}$$

the sum in (4.4) *runs over caps of size $C^{-1}r$ with*

$$r := 2^{k_0 - k_{\max}} 2^{(\ell_{\max} - k_{\min})/2}$$

and with separation $\mathrm{dist}(\kappa_1,\kappa_2) \sim r$*, whereas the sum in* (4.3) *runs over caps of size r' where $2^{k_0 - k_{\max}} \le r' \le 1$ is arbitrary but fixed. The sum* (4.3) *is empty if $\ell_{\max} < k_{\max} - C$ and* (4.4) *is nonzero only if $\ell_{\max} \le k_{\min} + O(1)$. Finally, if* (4.1) *fails, then the same representations hold provided $r \lesssim 1$ and one replaces* $\mathrm{dist}(\kappa_1,\kappa_2) \sim r$ *with* $\mathrm{dist}(\kappa_1,\kappa_2) \lesssim r$*.*

Proof. We consider first the $(++)$ and $(--)$ cases, i.e., when τ_1, τ_2 have the same sign. Then

$$|\xi_1| + |\xi_2| - |\xi_1 + \xi_2| \sim 2^{\ell_{\max}} \tag{4.5}$$

whence

$$(|\xi_1| + |\xi_2|)^2 - |\xi_1 + \xi_2|^2 \sim 2^{\ell_{\max} + k_{\max}}$$

and thus

$$\sphericalangle(\xi_1, \xi_2) \sim 2^{(\ell_{\max} + k_{\max} - k_1 - k_2)/2}$$

Now assume further that $k_0 = k_{\max} + O(1)$. Then it follows that

$$\sphericalangle(\xi_1, \xi_2) \sim 2^{(\ell_{\max} - k_{\min})/2}$$

If on the other hand $k_0 < k_{\max} - C$, then $k_1 = k_2 + O(1) = k_{\max} + O(1)$ and from (4.5), $\ell_{\max} = k_{\max} + O(1)$. Furthermore, $\xi_2 = -\xi_1 + O(2^{k_0})$ implies that

$$|\sphericalangle(\xi_1, -\xi_2)| \sim \frac{|\xi_1 \wedge \xi_2|}{|\xi_1||\xi_2|} = O(2^{k_0 - k_{\max}})$$

Next, consider the $(+-)$ or $(-+)$ cases. Then

FIGURE 4. Opposing $(++)$ waves

$$|\xi_1+\xi_2| - ||\xi_1|-|\xi_2|| \sim 2^{\ell_{\max}} \tag{4.6}$$

which implies that

$$|\xi_1+\xi_2|^2 - ||\xi_1|-|\xi_2||^2 \sim 2^{\ell_{\max}}(|\xi_1+\xi_2| + ||\xi_1|-|\xi_2||)$$

or equivalently,

$$2^{k_1+k_2}\sphericalangle^2(\xi_1,-\xi_2) \sim 2^{\ell_{\max}+k_0} \tag{4.7}$$

If $k_0 = k_{\max}+O(1)$, then

$$\sphericalangle(\xi_1,-\xi_2) \sim 2^{(\ell_{\max}-k_{\min})/2}$$

If, on the other hand, $k_0 \le k_{\max}-C$, then

$$\sphericalangle(\xi_1,-\xi_2) \sim 2^{k_0-k_{\max}}2^{(\ell_{\max}-k_{\min})/2}$$

and we are done. While it is clear that $\ell_{\max} \le k_{\min}+O(1)$ if $k_0 = k_{\max}+O(1)$, some proof is needed in case $k_0 < k_{\max}-C$. Thus, suppose $|\xi_1| \ge |\xi_2|$ whence

$$|\xi_1+\xi_2| - |\xi_1| + |\xi_2| \sim 2^{\ell_{\max}}$$

which implies that

$$2^{k_0+k_1}\sphericalangle^2(\xi_1+\xi_2,-\xi_2) \sim 2^{\ell_{\max}+k_{\max}}$$

since $2^{k_0+k_1} \sim 2^{k_{\min}+k_{\max}}$, the claim follows.

Finally, if (4.1) fails, then (4.5) turns into

$$|\xi_1|+|\xi_2| - |\xi_1+\xi_2| \lesssim 2^{\ell_{\max}}$$

which then leads to the claimed loss of separation between the sectors. However, their maximal distances are controlled by the same quantities as before. $\square$

The special appearance of (4.3) derives from the contributions of waves which lie on opposing sides of the light-cone. In fact, Figure 3 shows two vectors on the same half (i.e., $\tau > 0$) but opposing sides of the light cone. They add up to produce a wave of small frequency but large modulation, as described by (4.3). This is the mechanism by which nonlinearities can turn free waves into "elliptic objects". This phrase refers to functions whose Fourier support has large separation from the characteristic variety of $\Box$. Also, following Tao we refer to (4.1) as the *modulation imbalanced case*, whereas its opposite is the *modulation balanced case*.

Remark 4.2. *Lemma 4.1 is optimal in the following sense:*

- *Given* $\ell_0 \le k_0 \le -10$ *there exist* $\xi_1, \xi_2 \in \mathbb{R}^n$ *with* $1 \le |\xi_1|, |\xi_2| \le 2$, $\sphericalangle(\xi_1,\xi_2) \sim 2^{(\ell_0+k_0)/2}$ *and such that*
$$\begin{aligned} |\xi_1+\xi_2| - ||\xi_1|-|\xi_2|| &\sim 2^{\ell_0} \\ |\xi_1+\xi_2| &\sim 2^{k_0} \end{aligned}$$
- *Given* $\ell_0 \le k_1 \le -10$ *there exist* $\xi_1, \xi_2 \in \mathbb{R}^n$ *with* $2^{k_1-1} \le |\xi_1| \le 2^{k_1}$, $1 \le |\xi_1|, |\xi_2| \le 2$ *and* $\sphericalangle(\xi_1,\xi_2) \sim 2^{(\ell_0-k_1)/2}$ *and so that*
$$|\xi_1|+|\xi_2| - |\xi_1+\xi_2| \sim 2^{\ell_0}.$$

Our immediate goal now is the proof of Lemma 4.5. It is important to note that the improvement of $2^{\frac{k}{2}}$ over (2.30) which is obtained in Lemma 4.5 coincides with the gain for the case of free waves. In order to accomplish this, we require three preparatory lemmas, all of which are well-known. The first is Mockenhaupt's "square function estimate" (more precisely, its geometric content), see [32], [33]. Recall that $\Theta = \mathrm{sign}(\tau)\hat{\xi}$.

Lemma 4.3. *Let* $\kappa, \tilde\kappa \in \mathcal{C}_\ell$ *with* $\mathrm{dist}(\kappa,\tilde\kappa) \sim |\kappa| \ll 1$ *and suppose that* $\mathcal{F}_i \subset \mathcal{C}_{\ell_i}$ *for* $i=1,2$ *are partitions of* κ *and* $\tilde\kappa$*, respectively, by pairwise disjoint caps. Further, let* $r \in (0,1)$, $\mu \in (1,2)$*, and define for any cap* $\kappa' \subset S^1$
$$\mathcal{T}_{\kappa',\mu,r} := \{(\tau,\xi) \;:\; ||\xi|-\mu| \le r,\ \Theta \in \kappa',\ ||\tau|-|\xi|| \le |\kappa'|^2\}$$
Set $M_i := \#\mathcal{F}_i$*. Then*
$$\sup_{\mu_1,\mu_2\sim 1} \Big\| \sum_{\kappa_1\in\mathcal{F}_1} \sum_{\kappa_2\in\mathcal{F}_2} \chi_{\mathcal{T}_{\kappa_1,\mu_1,r}+\mathcal{T}_{\kappa_2,\mu_2,r}} \Big\|_{L^\infty(\mathbb{R}^3)} \le C\max(1, r(M_1+M_2)) \tag{4.8}$$
where C *is some absolute constant.*

Proof. Fix $r\in(0,1)$, and $\mu_1,\mu_2\sim 1$. Applying a Lorentz transform, one may assume that $\ell=-10$, say. Also, suppose without loss of generality that $\ell_1 \le \ell_2$ whence $M_1 \ge M_2$. We first consider the case where $rM_1 \ge 1$. Fix $(\tau,\xi)\in\mathbb{R}^3$ such that[11]
$$\sum_{\kappa_1\in\mathcal{F}_1} \sum_{\kappa_2\in\mathcal{F}_2} \chi_{\frac12(\mathcal{T}_{\kappa_1,\mu_1,r}+\mathcal{T}_{\kappa_2,\mu_2,r})}(\tau,\xi) \ge 1$$
Suppose $\mathcal{T}_{\kappa_1,\mu_1,r}$ with $\kappa_1\in\mathcal{F}_1$ contributes to the sum on the left-hand side. Define a mirror-image $\mathcal{T}^*_{\kappa_1,\mu_1,r}$ of $\mathcal{T}_{\kappa_1,\mu_1,r}$ by reflecting $\mathcal{T}_{\kappa_1,\mu_1,r}$ about the point (τ,ξ). Due to $\ell=-10$ and the dimensions of the tubes $\mathcal{T}$, the mirror images of all $\{\mathcal{T}_{\kappa_1,\mu_1,r}\}_{\kappa_1\in\mathcal{F}_1}$ have uniformly bounded overlap. The same applies with the role of $\mathcal{F}_1$ and $\mathcal{F}_2$ reversed. In conclusion, each $\mathcal{T}_{\kappa_1,\mu_1,r}$ can pair up with at most $O(1)$-many $\mathcal{T}_{\kappa_2,\mu_2,r}$ so as to give a contribution to (4.8), whence the bound of M_1 for (4.8). To obtain the factor r improvement, we further note that due to fixed μ_1 and μ_2, only those contributions to (4.8) need to be counted which derive from pairs $(\mathcal{T}_{\kappa_1,\mu_1,r}, \mathcal{T}_{\kappa_2,\mu_2,r})$ which lie in fixed cylinders $||\xi_i|-\mu_i|<r$, $i=1,2$. In terms of equations, we are given $(\sigma,\zeta)\in\mathbb{R}^3$ and we need to consider the sets of $(\tau_i, r_i\omega_i)$, $i=1,2$ with $\omega_i\in S^1$ satisfying the transversality condition $\sphericalangle(\omega_1,\omega_2)\in[\frac{1}{100},\frac{1}{50}]$, say, and such that
$$\begin{gathered} \tau_1+\tau_2=\sigma, \quad r_1\omega_1+r_2\omega_2=\zeta \\ |r_1-\mu_1|<r, \quad |r_2-\mu_2|<r \\ ||\tau_1|-r_1|<2^{2\ell_1}, \quad ||\tau_2|-r_2|<2^{2\ell_2} \end{gathered}$$
It follows from the second, third, and fourth conditions that
$$\mu_1\omega_1+\mu_2\omega_2 = \zeta + O(r)$$
and since the circular arcs containing ω_1 and ω_2 are transverse to each other, they must be of lengths $\lesssim r$. Consequently, we can only count tubes which correspond to an $r\times r$ disk on the light-cone and of those

[11] The $\frac12$- factor is a convenient modification that can be made due to scaling.

there are at most rM_1-many. In case $rM_1 \le 1$, then the number of the allowed pairs is $\lesssim 1$ in light of this construction and we are done. $\square$

Next, we present a standard bilinear L^2 bound for free waves.

Lemma 4.4. *Let $\kappa, \tilde\kappa \in \mathcal{C}_\ell$ with $\mathrm{dist}(\kappa,\tilde\kappa) \sim |\kappa| := \beta$ and suppose $\kappa_1 \subset \kappa$, $\kappa_2 \subset \tilde\kappa$ are arbitrary caps. Let $r \in (0,1)$ and $\mu_1, \mu_2 \sim 1$. Then*

$$\|e^{it|\nabla|} f_1\, e^{\pm it|\nabla|} f_2\|_{L^2_t L^2_x} \lesssim \beta^{-1} \sqrt{\min\big(r\beta, |\kappa_1|, |\kappa_2|\big)}\; \|f_1\|_2 \|f_2\|_2 \tag{4.9}$$

provided

$$\begin{aligned}
\mathrm{supp}(\widehat{f_1}) &\subset \{\xi \in \mathbb{R}^2 \;:\; \hat\xi \in \kappa_1,\; ||\xi| - \mu_1| \lesssim r\} \\
\mathrm{supp}(\widehat{f_2}) &\subset \{\xi \in \mathbb{R}^2 \;:\; \hat\xi \in \pm\kappa_2,\; ||\xi| - \mu_2| \lesssim r\}
\end{aligned}$$

and the sign in the last sign is chosen to be the same as in (4.9)*.*

Proof. The proof reduces to the following well-known property of convolutions: suppose

$$\begin{aligned}
\Gamma_1 &:= \{(|\xi|, \xi) \in \mathbb{R}^3 \;:\; \hat\xi \in \kappa_1,\; ||\xi| - \mu_1| \lesssim r\} \\
\Gamma_2 &:= \{(\pm|\xi|, \xi) \in \mathbb{R}^3 \;:\; \hat\xi \in \pm\kappa_2,\; ||\xi| - \mu_2| \lesssim r\}
\end{aligned}$$

Note that $\sphericalangle(\xi, \pm\eta) \gtrsim \beta$ for any $(|\xi|, \xi) \in \Gamma_1$ and $(\pm|\eta|, \eta) \in \Gamma_2$. Then

$$\|f\sigma_{\Gamma_1} * g\sigma_{\Gamma_2}\|_{L^2(\mathbb{R}^3)} \lesssim \beta^{-1} \sqrt{\min\big(r\beta, |\kappa_1|, |\kappa_2|\big)}\; \|f\|_{L^2(d\sigma_{\Gamma_1})} \|g\|_{L^2(d\sigma_{\Gamma_2})} \tag{4.10}$$

where σ_{Γ_1} and σ_{Γ_2} are the lifts of the measure in $\mathbb{R}^2$ to the sectors Γ_1, Γ_2 on the light-cones. To prove (4.10), interpolate between L^1 and L^∞. On L^1 we have the standard fact that $\|\mu * \nu\| \le \|\mu\| \|\nu\|$ for measures and their total variation norms. This fact does not use the angular separation of the supports nor their sizes. On L^∞, however, this separation and size are crucial and yield

$$\|f\sigma_{\Gamma_1} * g\sigma_{\Gamma_2}\|_{L^\infty(\mathbb{R}^3)} \lesssim \beta^{-1} \min(r, |\kappa_1|\beta^{-1}) \|f\|_{L^\infty(d\sigma_{\Gamma_1})} \|g\|_{L^\infty(d\sigma_{\Gamma_2})}$$

assuming as we may that $|\kappa_1| \le |\kappa_2|$. To obtain this bound, consider δ-neighborhoods of Γ_1 and Γ_2, respectively. In other words, replace $d\sigma_1$ by

$$d\widetilde{\sigma}_j^{(\delta)} := \delta^{-1} \chi_{[\mathrm{dist}((\xi,\tau),\Gamma_j) < \delta]} d\xi d\tau$$

for small $\delta > 0$ and observe that

$$\limsup_{\delta \to 0+} \|d\widetilde{\sigma}_1^{(\delta)} * d\widetilde{\sigma}_2^{(\delta)}\|_{L^\infty_{\xi,\tau}} \lesssim \beta^{-1} \min(r, |\kappa_1|\beta^{-1}) \tag{4.11}$$

by elementary geometry. To pass from (4.10) to estimates for the wave equation use Plancherel's theorem. $\square$

We can now state the aforementioned improved bilinear L^2 bound. The norm $|\!|\!| \cdot |\!|\!|$ is the one from (2.66).

Lemma 4.5. *Let ϕ_i be adapted to k_i for $i = 1, 2$. Assume further that we are in the high-high case $k_1 = k_2 + O(1)$ and that $\phi_i = Q_{\le j+k-2k_1-C}\phi_i$ for $i = 1, 2$. Then*

$$\|P_k Q_j(\phi_1\, \phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{k_1}{2}} 2^{\frac{k-j}{4}} |\!|\!|\phi_1|\!|\!|_{S[k_1]} |\!|\!|\phi_2|\!|\!|_{S[k_2]} \tag{4.12}$$

for any $j \le k \le k_1 + O(1)$. Moreover, in the same range of j,

$$\|P_k Q_j(R_\alpha\phi_1 R_\beta\phi_2 - R_\beta\phi_1 R_\alpha\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{-\frac{k_1}{2}} 2^{\frac{3k+j}{4}} |\!|\!|\phi_1|\!|\!|_{S[k_1]} |\!|\!|\phi_2|\!|\!|_{S[k_2]} \tag{4.13}$$

for any $\alpha, \beta = 0, 1, 2$.

Proof. We assume that $k_1 = k_2 + O(1) = 0$. At first, we also assume that $k \le -C$ so as to exclude the opposing $(++)$ and $(--)$ waves in Lemma 4.1. We need to prove that

$$\|P_k(\phi_1\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{k-j}{4}} \big(\|(f_1,g_1)\|_{L^2\times \dot H^{-1}} + \|F_1\|_{N[k_1]}\big)\cdot \big(\|(f_2,g_2)\|_{L^2\times \dot H^{-1}} + \|F_2\|_{N[k_2]}\big) \tag{4.14}$$

for any k_i-adapted Schwartz functions f_i, g_i, F_i, $i = 1,2$ and

$$\phi_i(t) = \cos(t|\nabla|) f_i + \frac{\sin(t|\nabla|)}{|\nabla|} g_i + \int_0^t \frac{\sin((t-s)|\nabla|)}{|\nabla|} F_i(s)\, ds \tag{4.15}$$

We reduce this to three cases:

$$\|P_k Q_j(\phi_1\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{k-j}{4}} \|(f_1,g_1)\|_{L^2\times \dot H^{-1}} \|(f_2,g_2)\|_{L^2\times \dot H^{-1}} \tag{4.16}$$

$$\|P_k Q_j(\phi_1\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{k-j}{4}} \|(f_1,g_1)\|_{L^2\times \dot H^{-1}} \|F_2\|_{N[k_2]} \tag{4.17}$$

$$\|P_k Q_j(\phi_1\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{k-j}{4}} \|F_1\|_{N[k_1]} \|F_2\|_{N[k_2]} \tag{4.18}$$

where the absence of terms on the right-hand side implies that the corresponding functions are zero (thus, $F_1 = F_2 = 0$ in (4.16) etc.) We begin with (4.16) which follows easily from Lemma 4.4. To see this, we decompose ϕ_i into caps of size $\ell = (j+k)/2$ as in Lemma 4.1. Adopting the convention that $\kappa_1 \sim \kappa_2$ means that $\mathrm{dist}(\kappa_1,\kappa_2) \sim 2^\ell$, and setting $g_1 = g_2 = 0$ for simplicity, one has[12]

$$\begin{aligned}
\|P_k Q_j(\phi_1\phi_2)\|_{L^2_t L^2_x} &\lesssim \sum_{\kappa_1\sim\kappa_2\in\mathcal{C}_\ell} \|P_k Q_j(P_{k_1,\kappa_1}\phi_1\, P_{k_2,\kappa_2}\phi_2)\|_{L^2_t L^2_x} \\
&\lesssim \sum_{c\in\mathcal{D}_{0,k}} \sum_{\kappa_1\sim\kappa_2\in\mathcal{C}_\ell} \|P_{k_1,\kappa_1} P_c\phi_1\, P_{k_2,\kappa_2} P_{-c}\phi_2\|_{L^2_t L^2_x} \\
&\lesssim \sum_{c\in\mathcal{D}_{0,k}} \sum_{\kappa_1\sim\kappa_2\in\mathcal{C}_\ell} 2^{\frac{k-\ell}{2}} \|P_{k_1,\kappa_1} P_c f_1\|_2 \|P_{k_2,\kappa_2} P_{-c} f_2\|_2 \\
&\lesssim 2^{\frac{k-\ell}{2}} \|f_1\|_2 \|f_2\|_2
\end{aligned} \tag{4.19}$$

as needed. The estimate in (4.19) follows from (4.8) since $k \ge \ell$.

To prove (4.17) and (4.18) it will suffice as usual to assume that F_i are $N[k_i]$-atoms for $i = 1,2$. In fact, if F_2 in (4.17) is either and energy or an $\dot X^{s,b}$-atom, then one again reduces matters to the free case. Consequently, we may restrict ourselves to (4.18) when both F_1 and F_2 are null-frame atoms. Using Lemma 2.11 to refine these null-frame atoms one can thus assume that

$$F_1 = \sum_{\kappa'\in\mathcal{C}_{\ell'}} F_{\kappa'}, \quad F_1 = \sum_{\kappa''\in\mathcal{C}_{\ell''}} \widetilde F_{\kappa''} \tag{4.20}$$

where $\ell', \ell'' \le \ell$. Again by Lemma 2.11, we can further assume that there exists a fixed $c \in \mathcal{D}_{0,k}$ so that $P_c F_1 = F_1$ and $P_{-c} F_2 = F_2$. Applying the same decomposition as in (4.19), fix $\kappa_1 \sim \kappa_2$. In view of Lemma 2.22,

$$P_{k_1,\kappa_1}\phi_1 = \Box^{-1} G_{\kappa_1} + \sum_{\substack{\kappa'\in\mathcal{C}_{\ell'}\\ \kappa'\subset\kappa_1}} \int_{\mathbb{R}} \big(\Psi^1_{\kappa',a} + B_{\kappa',a}\,\Psi^2_{\kappa',a}\big)\, da \tag{4.21}$$

$$P_{k_2,\kappa_2}\phi_2 = \Box^{-1} \widetilde G_{\kappa_2} + \sum_{\substack{\kappa''\in\mathcal{C}_{\ell''}\\ \kappa''\subset\kappa_2}} \int_{\mathbb{R}} \big(\widetilde\Psi^1_{\kappa'',a} + \widetilde B_{\kappa'',a}\,\widetilde\Psi^2_{\kappa'',a}\big)\, da \tag{4.22}$$

where the functions on the right-hand side satisfy the bounds specified in that lemma. Moreover, the Fourier supports of the functions appearing inside the integral in (4.21) and (4.22) satisfy (2.43), and they also retain the P_c and P_{-c} localization property, respectively, due to the fact that $k \ge \ell$. We can

[12] Recall our convention about $P_{k_i,\kappa}$ which takes the sign of τ into account.

ignore the terms involving G_{κ_1} and $\widetilde{G}_{\kappa_2}$ as they are reducible to free waves. For simplicity, we also set $\Psi^1_{\kappa',a} = \widetilde{\Psi}^1_{\kappa'',a} = 0$. By Plancherel's theorem and Lemma 4.3,

$$
\begin{aligned}
&\|P_k[P_{k_1,\kappa_1}\phi_1 P_{k_2,\kappa_2}\phi_2]\|_{L^2_t L^2_x} \\
&\lesssim \sqrt{1+2^k(M_1+M_2)} \sum_{c\in\mathcal{D}_{0,k}} \Big(\sum_{\substack{\kappa'\in\mathcal{C}_{\ell'}\\ \kappa'\subset\kappa_1}} \sum_{\substack{\kappa''\in\mathcal{C}_{\ell''}\\ \kappa''\subset\kappa_2}} \|P_{k_1,\kappa'}P_c\phi_1 P_{k_1,\kappa''}P_{-c}\phi_2\|^2_{L^2_t L^2_x} \Big)^{\frac12}
\end{aligned}
$$

where $M_1 = 2^{\ell-\ell'}, M_2 = 2^{\ell-\ell''}$. On the other hand, applying Lemma 2.22 to $P_c\phi_1$, $P_{-c}\phi_2$ and using Lemma 4.4 implies that

$$
\begin{aligned}
&\|P_{k_1,\kappa'}P_c\phi_1 P_{k_1,\kappa''}P_{-c}\phi_2\|_{L^2_t L^2_x} \\
&\lesssim \int_{\mathbb{R}^2} \|B^c_{\kappa',a}\,\Psi^{2c}_{\kappa',a}\widetilde{B}^{-c}_{\kappa'',b}\,\widetilde{\Psi}^{2-c}_{\kappa'',b}\|_{L^2_t L^2_x}\,dadb \\
&\lesssim \int_{\mathbb{R}^2} \|\Psi^{2c}_{\kappa',a}\,\widetilde{\Psi}^{2-c}_{\kappa'',b}\|_{L^2_t L^2_x}\,dadb \\
&\lesssim 2^{-\ell}\sqrt{\min(2^{k+\ell},2^{\ell'},2^{\ell''})}\int_{\mathbb{R}^2} \|\Psi^{2c}_{\kappa',a}\|_{\dot X_0^{0,\frac12,1}}\,\|\widetilde{\Psi}^{2-c}_{\kappa'',b}\|_{\dot X_0^{0,\frac12,1}}\,dadb \\
&\lesssim 2^{-\ell}\sqrt{\min(2^{k+\ell},2^{\ell'},2^{\ell''})}\|P_cF_{\kappa'}\|_{\mathrm{NF}[\kappa']}\,\|\widetilde{P}_{-c}F_{\kappa''}\|_{\mathrm{NF}[\kappa'']}
\end{aligned}
$$

One checks that

$$
\sqrt{1+2^k(M_1+M_2)}\;2^{-\ell}\sqrt{\min(2^{k+\ell},2^{\ell'},2^{\ell''})} \lesssim 2^{\frac{k-\ell}{2}}
$$

whence

$$
\sum_{c\in\mathcal{D}_{0,k}} \|P_{k_1,\kappa_1}P_c\phi_1 P_{k_2,\kappa_2}P_{-c}\phi_2\|_{L^2_t L^2_x} \lesssim 2^{\frac{k-\ell}{2}} \sum_{c\in\mathcal{D}_{0,k}} \Big(\sum_{\substack{\kappa'\in\mathcal{C}_{\ell'}\\ \kappa'\subset\kappa_1}} \sum_{\substack{\kappa''\in\mathcal{C}_{\ell''}\\ \kappa''\subset\kappa_2}} \|P_cF_{\kappa'}\|^2_{\mathrm{NF}[\kappa']}\,\|\widetilde{P}_{-c}F_{\kappa''}\|^2_{\mathrm{NF}[\kappa'']} \Big)^{\frac12}
$$

In conclusion,

$$
\begin{aligned}
\|P_kQ_j(\phi_1\phi_2)\|_{L^2_t L^2_x} &\lesssim \sum_{c\in\mathcal{D}_{0,k}} \sum_{\kappa_1\sim\kappa_2\in\mathcal{C}_\ell} \|P_{k_1,\kappa_1}P_c\phi_1 P_{k_2,\kappa_2}P_{-c}\phi_2\|_{L^2_t L^2_x} \\
&\lesssim 2^{\frac{k-\ell}{2}} \sum_{c\in\mathcal{D}_{0,k}} \sum_{\kappa_1\sim\kappa_2\in\mathcal{C}_\ell} \Big(\sum_{\substack{\kappa'\in\mathcal{C}_{\ell'}\\ \kappa'\subset\kappa_1}} \sum_{\substack{\kappa''\in\mathcal{C}_{\ell''}\\ \kappa''\subset\kappa_2}} \|P_cF_{\kappa'}\|^2_{\mathrm{NF}[\kappa']}\,\|\widetilde{P}_{-c}F_{\kappa''}\|^2_{\mathrm{NF}[\kappa'']} \Big)^{\frac12} \\
&\lesssim 2^{\frac{k-\ell}{2+}} \Big(\sum_{\kappa'\in\mathcal{C}_{\ell'}} \|F_{\kappa'}\|^2_{\mathrm{NF}[\kappa']} \Big)^{\frac12} \Big(\sum_{\kappa''\in\mathcal{C}_{\ell''}} \|\widetilde{F}_{\kappa''}\|^2_{\mathrm{NF}[\kappa'']} \Big)^{\frac12}
\end{aligned}
$$

as desired; in the last step we have also used Lemma 2.11. This concludes the proof of (4.18) for the case of null-frame atoms F_1, F_2. As indicated, the other cases are easier since they can be reduced to free waves. Finally, if $k = O(1)$, then the proof is easier. In fact, it follows via a cap-decomposition from the basic bilinear bound (2.30). We leave those details to the reader.

The second bound (4.13) follows by the same argument. The only difference from (4.12) lies with an additional gain of 2^ℓ which is precisely the size of the angle in the above decompositions into caps. $\square$

Later, we shall require the following technical variant of the previous bound.

Corollary 4.6. *Under the assumptions of Lemma 4.5, for any $j \le k \le k_1 + O(1)$ and any $m_0 \le -10$,*

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}} \|P_kQ_j(P_{k_1,\kappa_1}\phi_1\, P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x} \lesssim 2^{\frac{k_1}{2}}2^{\frac{k-j}{4}}\,\|\!|\phi_1|\!\|_{S[k_1]}\|\!|\phi_2|\!\|_{S[k_2]} \tag{4.23}$$

$$\Big(\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(P_{k_1,\kappa}\phi_1\,\phi_2)\|^2_{L^2_tL^2_x}\Big)^{\frac12} \lesssim |m_0|\,2^{\frac{k_1}{2}}2^{\frac{k-j}{4}}\,\|\!|\phi_1|\!\|_{S[k_1]}\|\!|\phi_2|\!\|_{S[k_2]} \tag{4.24}$$

Moreover, analogous bounds hold for the null form in (4.13) *with an extra gain of $2^{\frac{j+k}{2}}$. Finally, the left-hand side in* (4.23) *vanishes unless $j+k\le 2m_0\le -100$.*

Proof. The final statement here is due to Lemma 4.1. Note that one cannot simply square sum the bounds of Lemma 4.5 applied to $P_{k_1,\kappa_1}\phi_1$ and $P_{k_2,\kappa_2}\phi_2$ due to the fact that $\sum_\kappa \|\!|P_{k,\kappa}\phi|\!\|^2_{S[k]}$ (or $\sum_\kappa \|P_{k,\kappa}\phi\|^2_{S[k]}$ for that matter) cannot be controlled. However, since we may assume that $\frac{j+k}{2}\le m_0$, the angular decomposition induced by the frequency and modulation cutoffs P_kQ_j is *finer* than the one superimposed by κ_1 and κ_2. Inspection of the proof now reveals that either by orthogonality or by organizing the finer caps into subsets of the $\kappa_1,\kappa_2\in\mathcal{C}_{m_0}$, and applying the Cauchy-Schwarz inequality yields the stated bound. For (4.24) one needs to distinguish two cases: either $m_0\ge\frac{j+k}{2}$ or not. In the former case, the decomposition into caps in $\mathcal{C}_{m_0}$ is coarser than the one coming from Lemma 4.1 and one can again argue by means of Cauchy-Schwarz as before. In the latter case, however, we split the modulation of the first input as follows:

$$Q_{<j+k-C} = Q_{<2m_0-C} + Q_{2m_0-C\le\cdot<j+k-C}$$

The contribution of $Q_{<2m_0-C}\phi_1$ is handled exactly as in the Lemma 4.5 since one may always refine the null-frame representation, cf. (4.20). On the other hand, $Q_{2m_0-C\le\cdot<j+k-C}\phi_1$ is controlled by means of Lemma 2.4. More precisely, for any $2m_0-C\le\ell<j+k-C$ one has $Q_\ell\phi_1 = Q_\ell\Box^{-1}F_1$, see (4.15). Since (2.32) implies that

$$\|Q_\ell\phi_1\|_{\dot X_0^{0,\frac12,\infty}} = \|Q_\ell\Box^{-1}F_1\|_{\dot X_0^{0,\frac12,\infty}} \lesssim \|Q_\ell F_1\|_{\dot X_0^{0,-\frac12,\infty}} \lesssim \|F_1\|_{N[0]}$$

one can reduce the contribution of $Q_\ell\phi_1$ to the case of free waves as in the proof of Lemma 4.5. Summing over all ℓ in this range loses a factor of at most $|m_0|$, as claimed. Finally, the claim concerning the null-forms is immediate. ☐

Removing the modulation restrictions on the inputs in Lemma 4.5 results in the following estimates.

Lemma 4.7. *If ϕ_1 and ϕ_2 are adapted to k_1 and k_2, respectively, then for $j\le k\le k_1+O(1)=k_2+O(1)$,*

$$\|P_kQ_j(\phi_1\phi_2)\|_{L^2_tL^2_x} \lesssim 2^{\frac{k-j}{4}}2^{\frac{k_1}{2}}\,\|\!|\phi_1|\!\|_{S[k_1]}\|\!|\phi_2|\!\|_{S[k_2]} \tag{4.25}$$

whereas for $j\le k_2\le k=k_1+O(1)$,

$$\|P_kQ_j(\phi_1\phi_2)\|_{L^2_tL^2_x} \lesssim 2^{\frac{3k_2}{4}}2^{-\frac{j}{4}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]} \tag{4.26}$$

Proof. Consider the high-high case $j\le k\le k_1+O(1)=k_2+O(1)=0$. On the one hand, there is the bound

$$\|P_kQ_j(Q_{\le j+k-C}\phi_1 Q_{\le j+k-C}\phi_2)\|_{L^2_tL^2_x} \lesssim 2^{\frac{k-j}{4}}\,\|\!|\phi_1|\!\|_{S[k_1]}\,\|\!|\phi_2|\!\|_{S[k_2]} \tag{4.27}$$

which is given by Lemma 4.5. On the other hand, by the improved Bernstein bound of Lemma 2.1,

$$\begin{aligned}\|P_kQ_j(Q_{>j+k-C}\phi_1\cdot\phi_2)\|_{L^2_tL^2_x} &\lesssim 2^{\frac{j-k}{4}\wedge 0}2^k\|Q_{>j+k-C}\phi_1\cdot\phi_2\|_{L^2_tL^1_x}\\ &\lesssim 2^{\frac{j-k}{4}}2^k\|Q_{>j+k-C}\phi_1\|_{L^2_tL^2_x}\|\phi_2\|_{L^\infty_tL^2_x}\\ &\lesssim 2^{\frac{k-j}{4}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned} \tag{4.28}$$

In the high-low case $j\le k_2\le k=k_1+O(1)=0$ consider the following three subcases. First,

$$\|P_kQ_j(Q_{<j-C}\phi_1Q_{<j-C}\phi_2)\|_{L^2_tL^2_x} \lesssim 2^{-\frac{j-k_2}{4}}2^{\frac{k_2}{2}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}$$

by a decomposition into caps of size $2^{\frac{j-k_2}{2}}$ and the L^2-bilinear bound (2.30). Next, by the improved Bernstein estimate Lemma 2.1,

$$\begin{aligned}\|P_kQ_j(\phi_1 Q_{\geq j-C}\phi_2)\|_{L_t^2L_x^2} &\lesssim \|\phi_1\|_{L^\infty L^2}\|Q_{\geq j-C}\phi_2\|_{L_t^2L_x^\infty}\\ &\lesssim 2^{\frac{j-k_2}{4}}2^{k_2}2^{-\frac{j}{2}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

And third,

$$\begin{aligned}\|P_kQ_j(Q_{\geq j-C}\phi_1 Q_{<j-C}\phi_2)\|_{L_t^2L_x^2} &\lesssim \sum_{m\geq j+O(1)}\|P_kQ_j(Q_m\phi_1 Q_{<j-C}\phi_2)\|_{L_t^2L_x^2}\\ &\lesssim \sum_{m\geq j+O(1)}\sum_{\kappa_1,\kappa_2}\|P_{k_1,\kappa_1}Q_m\phi_1 P_{k_2,\kappa_2}Q_{<j-C}\phi_2\|_{L_t^2L_x^2}\\ &\lesssim \sum_{m\geq j+O(1)}\sum_{\kappa_1,\kappa_2}\|P_{k_1,\kappa_1}Q_m\phi_1\|_{L_t^2L_x^2}\|P_{k_2,\kappa_2}Q_{<j-C}\phi_2\|_{L_{t,x}^\infty}\\ &\lesssim \sum_{m\geq j+O(1)}2^{-\frac{m}{2}}2^{k_2}2^{\frac{m-k_2}{4}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\\ &\lesssim 2^{\frac{3k_2}{4}}2^{-\frac{j}{4}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

as claimed. The inner sums run over $\kappa_1,\kappa_2\in\mathcal{C}_{\frac{m-k_2}{2}}$ with $\mathrm{dist}(\kappa_1,\kappa_2)\lesssim 2^{\frac{m-k_2}{2}}$. □

Later we shall also need the following technical variants, both of which are in the same spirit as Corollary 4.6.

Corollary 4.8. *Let ϕ be adapted to k_1 and suppose for every $\kappa\in\mathcal{C}_{m_0}$ with $m_0\leq -100$ there is a Schwarz function ψ_κ which is adapted to k_2. Then, provided $j\leq k_2\leq k=k_1+O(1)$,*

$$\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_kQ_j(P_{k_1,\kappa}\phi\,\psi_\kappa)\|_{L_t^2L_x^2}\lesssim |m_0|2^{\frac{3k_2}{4}}2^{-\frac{j}{4}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|^2_{S[k_2]}\Big)^{\frac12} \tag{4.29}$$

Proof. One uses the argument for the high-low case of Lemma 4.7. In particular, $k=k_1+O(1)=0$. First, with $m=\frac{j-k_2}{2}$,

$$\sum_{\kappa\in\mathcal{C}_{m_0}}P_kQ_j(Q_{<j-C}P_{k_1,\kappa}\phi\,Q_{<j-C}\psi_\kappa)=\sum_{\kappa\in\mathcal{C}_{m_0}}\sum_{\kappa_1,\kappa_2\in\mathcal{C}_m}P_kQ_j(Q_{<j-C}P_{k_1,\kappa_1}P_{k_1,\kappa}\phi\,Q_{<j-C}P_{k_2,\kappa_2}\psi_\kappa)$$

If $m\leq m_0$, then by the L^2-bilinear bound (2.30)

$$\begin{aligned}&\sum_{\kappa\in\mathcal{C}_{m_0}}\|P_kQ_j(Q_{<j-C}P_{k_1,\kappa}\phi\,Q_{<j-C}\psi_\kappa)\|_{L_t^2L_x^2}\\ &\lesssim \sum_{\kappa\in\mathcal{C}_{m_0}}\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_m\\ \kappa_1\subset\kappa}}\|P_kQ_j(Q_{<j-C}P_{k_1,\kappa_1}\phi\,Q_{<j-C}P_{k_2,\kappa_2}\psi_\kappa)\|_{L_t^2L_x^2}\\ &\lesssim \sum_{\kappa\in\mathcal{C}_{m_0}}\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_m\\ \kappa_1\subset\kappa}}2^{-\frac{j-k_2}{4}}2^{\frac{k_2}{2}}\|Q_{<j-C}P_{k_1,\kappa_1}\phi\|_{S[k_1,\kappa_1]}\|Q_{<j-C}P_{k_2,\kappa_2}\psi_\kappa\|_{S[k_2,\kappa_2]}\\ &\lesssim 2^{-\frac{j-k_2}{4}}2^{\frac{k_2}{2}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|^2_{S[k_2]}\Big)^{\frac12}\end{aligned}$$

where we applied Cauchy-Schwarz twice to pass to the last line. If, on the other hand, $m > m_0$, then we first consider smaller modulations of ϕ. In fact, dropping the $Q_{<j-C}$ on ϕ as we may one has

$$\begin{aligned}
&\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(Q_{<2m_0-C}P_{k_1,\kappa}\phi\, Q_{<j-C}\psi_\kappa)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_m\\ \kappa\subset\kappa_1}} \sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(Q_{<2m_0-C}P_{k_1,\kappa_1}\phi\, Q_{<j-C}P_{k_2,\kappa_2}\psi_\kappa)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_m\\ \kappa\subset\kappa_1}} \sum_{\kappa\in\mathcal{C}_{m_0}} 2^{-\frac{j-k_2}{4}}2^{\frac{k_2}{2}}\|Q_{<2m_0-C}P_{k_1,\kappa}\phi\|_{S[k_1,\kappa]}\|Q_{<j-C}P_{k_2,\kappa_2}\psi_\kappa\|_{S[k_2,\kappa_2]}\\
&\lesssim 2^{-\frac{j-k_2}{4}}2^{\frac{k_2}{2}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|_{S[k_2]}^2\Big)^{\frac12}
\end{aligned}$$

where we again applied Cauchy-Schwarz twice to pass to the last line. Finally, we need to account for $Q_{2m_0-C\le\cdot<j-C}\phi$. Fix ℓ with $2m_0-C\le\ell<j-C$ and repeat the previous estimate. This yields

$$\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(Q_\ell P_{k_1,\kappa}\phi\, Q_{<j-C}\psi_\kappa)\|_{L_t^2L_x^2} \lesssim 2^{-\frac{j-k_2}{4}}2^{\frac{k_2}{2}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|_{S[k_2]}^2\Big)^{\frac12}$$

which, upon summing in ℓ yields the same bound with the loss of a factor of $(j-2m_0)_+$. Replacing this by the larger $|m_0|$ then implies the bound of the corollary. Next, by the improved Bernstein estimate of Lemma 2.1, and Lemma 2.18,

$$\begin{aligned}
\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(P_{k_1,\kappa}\phi\, Q_{\ge j-C}\psi_\kappa)\|_{L_t^2L_x^2} &\lesssim \sum_{\kappa\in\mathcal{C}_{m_0}} \|P_{k_1,\kappa}\phi\|_{L^\infty L^2}\|Q_{\ge j-C}\psi_\kappa\|_{L_t^2L_x^\infty}\\
&\lesssim |m_0|\,2^{\frac{j-k_2}{4}}2^{k_2}2^{-\frac{j}{2}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|_{S[k_2]}^2\Big)^{\frac12}
\end{aligned}$$

And third,

$$\begin{aligned}
&\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(Q_{\ge j-C}P_{k_1,\kappa}\phi\, Q_{<j-C}\psi_\kappa)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{m\ge j+O(1)}\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(Q_mP_{k_1,\kappa}\phi\, Q_{<j-C}\psi_\kappa)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{m\ge j+O(1)}\sum_{\kappa\in\mathcal{C}_{m_0}}\sum_{\kappa_1\sim\kappa_2} \|P_{k_1,\kappa_1}Q_mP_{k_1,\kappa}\phi\, P_{k_2,\kappa_2}Q_{<j-C}\psi_\kappa\|_{L_t^2L_x^2}\\
&\lesssim \sum_{m\ge j+O(1)}\sum_{\kappa\in\mathcal{C}_{m_0}}\sum_{\kappa_1\sim\kappa_2} \|P_{k_1,\kappa_1}Q_mP_{k_1,\kappa}\phi\|_{L_t^2L_x^2}\,2^{k_2}2^{\frac{m-k_2}{4}}\|P_{k_2,\kappa_2}Q_{<j-C}\psi_\kappa\|_{L_t^\infty L_x^2}\\
&\lesssim 2^{k_2}2^{\frac{j-k_2}{4}}\sum_{m=j+O(1)}\Big(\sum_{\kappa_1,\kappa}\|P_{k_1,\kappa_1}Q_mP_{k_1,\kappa}\phi\|_{L_t^2L_x^2}^2\Big)^{\frac12}\Big(\sum_{\kappa_2,\kappa'}\|P_{k_2,\kappa_2}Q_{<j-C}\psi_{\kappa'}\|_{L_t^\infty L_x^2}^2\Big)^{\frac12}\\
&\lesssim 2^{\frac{3k_2}{4}}2^{-\frac{j}{4}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|_{S[k_2]}^2\Big)^{\frac12}
\end{aligned}$$

as claimed. The inner sums run over $\kappa_1,\kappa_2\in\mathcal{C}_{\frac{m-k_2}{2}}$ and $\kappa_1\sim\kappa_2$ denotes $\mathrm{dist}(\kappa_1,\kappa_2)\lesssim 2^{\frac{m-k_2}{2}}$. □

Remark 4.9. We note here that one may also gain in terms of m_0 at the expense of some losses in terms of the frequencies/modulations; specifically, using similar reasoning and under the same assumptions as above, one gets

$$\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_kQ_j(P_{k_1,\kappa}\phi\,\psi_\kappa)\|_{L_t^2L_x^2} \lesssim 2^{\delta_1m_0}2^{\delta_2(k_1-j)}2^{\frac{3k_2}{4}}2^{-\frac{j}{4}}\|\phi\|_{S[k_1]}\Big(\sum_{\kappa\in\mathcal{C}_{m_0}}\|\psi_\kappa\|_{S[k_2]}^2\Big)^{\frac12} \tag{4.30}$$

for suitable $\delta_{1,2} > 0$.

We shall also require the following estimates which gain something in terms of the small angle.

Corollary 4.10. *Given $\delta > 0$ small and $L \gg 1$, there exists $m_0(\delta, L) \ll -1$ with the following property: let $k, k_1, k_2 \in \mathbb{Z}$ so that $\max_{i=1,2} |k - k_i| \leq L$. For any ϕ_1 and ϕ_2 which are adapted to k_1, k_2, respectively, and $j \leq k + C$,*

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\leq 2^{m_0}}} \|P_kQ_j(P_{k_1,\kappa_1}\phi_1\, P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x} \leq \delta 2^{\frac{k-j}{3}} 2^{\frac{k_1}{2}} \|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]} \tag{4.31}$$

In the high-low case $k = k_1 + O(1)$, $k_2 \leq k_1 - C$,

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\leq 2^{m_0}}} \|P_kQ_j(P_{k_1,\kappa_1}\phi_1\, P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x} \leq \delta 2^{\frac{k_2-j}{3}} 2^{\frac{k_2}{2}} \|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]} \tag{4.32}$$

as well as

$$\Big(\sum_{\kappa_2\in\mathcal{C}_{m_0}} \|P_kQ_j(\phi_1\, P_{k_2,\kappa}\phi_2)\|^2_{L^2_tL^2_x}\Big)^{\frac12} \leq \delta 2^{\frac{k_2-j}{3}} 2^{\frac{k_2}{2}} \|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]} \tag{4.33}$$

Proof. Let $k_1 = 0$ whence $|k| \leq L$ and $|k_2| \leq 2L$. Implicit constants will be allowed to depend on L. By Corollary 4.6 and (4.27),

$$\begin{aligned}&\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\leq 2^{m_0}}} \|P_kQ_j(Q_{\leq j+k-C}P_{k_1,\kappa_1}\phi_1\, Q_{\leq j+k-C}P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x}\\ &\lesssim 2^{\frac{k+j}{4}} 2^{-\frac{j}{2}} \|\phi_1\|_{S[k_1]}\, \|\phi_2\|_{S[k_2]} \leq \delta 2^{-\frac{j}{4-}} \|\phi_1\|_{S[k_1]}\, \|\phi_2\|_{S[k_2]}\end{aligned}$$

which is sufficient. Note that we used interpolation and $2^{\frac{k+j}{4}} \leq 2^{\frac{m_0}{2}}$ which gives the desired gain of δ provided m_0 is small enough relative to δ and L. For the remaining cases we use a variant of (4.28): with $2 > r > 1$, $\theta = \frac{2}{r} - 1$, and $\frac1p = \frac1r - \frac12$,

$$\begin{aligned}&\|P_kQ_j(Q_{>j+k-C}P_{k_1,\kappa_1}\phi_1 \cdot P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x} &(4.34)\\ &\lesssim 2^{\frac{j}{4}\theta}\|Q_{>j+k-C}\phi_1 \cdot P_{k_2,\kappa_2}\phi_2\|_{L^2_tL^r_x}\\ &\lesssim 2^{\frac{j}{4}\theta}\|Q_{>j+k-C}P_{k_1,\kappa_1}\phi_1\|_{L^2_tL^2_x}\|P_{k_2,\kappa_2}\phi_2\|_{L^\infty_tL^p_x}\\ &\lesssim 2^{\frac{j}{4}(\theta-2)}2^{m_0(\frac12-\frac1p)}\|P_{k_1,\kappa_1}\phi_1\|_{\dot X_0^{0,\frac12,\infty}}\|P_{k_2,\kappa_2}\phi_2\|_{L^\infty_tL^2_x} &(4.35)\end{aligned}$$

Taking θ close to 1, one can make this $\leq \delta 2^{-\frac{j}{4-}}$ as desired. This bound can be summed over κ_1, κ_2 by Cauchy-Schwarz and the definition of the $S[k]$-norm; see also Lemma 2.18.

In the high-low case $j \leq k_2 \leq k = k_1 + O(1) = 0$ we proceed as follows. First,

$$\begin{aligned}&\|P_kQ_j(Q_{<j-C}P_{k_1,\kappa_1}\phi_1Q_{<j-C}P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x}\\ &\lesssim 2^{-\frac{j-k_2}{2}} \min\big(2^{\frac{m_0}{2}}, 2^{\frac{j-k_2}{4}}\big)2^{\frac{k_2}{2}}\|P_{k_1,\kappa_1}Q_{<j-C}\phi_1\|_{S[k_1]}\|P_{k_2,\kappa_2}Q_{<j-C}\phi_2\|_{S[k_2]}\\ &\lesssim \delta 2^{-\frac{j-k_2}{3}}2^{\frac{k_2}{2}}\|P_{k_1,\kappa_1}Q_{<j-C}\phi_1\|_{S[k_1]}\|P_{k_2,\kappa_2}Q_{<j-C}\phi_2\|_{S[k_2]}\end{aligned}$$

by a decomposition into caps of size $2^{\frac{j-k_2}{2}}$ and the L^2-bilinear bound (2.30). The summation over κ_1 and κ_2 can be carried out since it leads to the square function (2.15). Next, by the improved Bernstein estimate, see Lemma 2.1,

$$\begin{aligned}
&\|P_k Q_j(P_{k_1,\kappa_1}\phi_1 Q_{\geq j-C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \|P_{k_1,\kappa_1}\phi_1\|_{L^\infty L^2}\|Q_{\geq j-C}P_{k_2,\kappa_2}\phi_2\|_{L_t^2 L_x^\infty} \\
&\lesssim \min(2^{\frac{j-k_2}{4}}, 2^{\frac{m_0}{2}})2^{k_2}2^{-\frac{j}{2}}\|P_{k_1,\kappa_1}\phi_1\|_{L_t^\infty L_x^2}\|P_{k_2,\kappa_2}\phi_2\|_{\dot X_{k_2}^{0,\frac12,\infty}} \\
&\leq \delta 2^{-\frac{j-k_2}{3}}2^{\frac{k_2}{2}}\|P_{k_1,\kappa_1}\phi_1\|_{L_t^\infty L_x^2}\|P_{k_2,\kappa_2}\phi_2\|_{\dot X_{k_2}^{0,\frac12,\infty}}
\end{aligned}$$

and summation over κ_1, κ_2 is again admissible. Finally,

$$\begin{aligned}
&\|P_k Q_j(Q_{\geq j-C}P_{k_1,\kappa_1}\phi_1 Q_{<j-C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m\geq j+O(1)} \|P_k Q_j(Q_m P_{k_1,\kappa_1}\phi_1 Q_{<j-C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m\geq j+O(1)}\sum_{\kappa_1',\kappa_2'} \|P_{k_1,\kappa_1}Q_m P_{k_1,\kappa_1'}\phi_1 P_{k_2,\kappa_2}Q_{<j-C}P_{k_2,\kappa_2'}\phi_2\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m\geq j+O(1)}\sum_{\kappa_1',\kappa_2'} \|Q_m P_{k_1,\kappa_1}P_{k_1,\kappa_1'}\phi_1\|_{L_t^2 L_x^2}\|Q_{<j-C}P_{k_2,\kappa_2}P_{k_2,\kappa_2'}\phi_2\|_{L_{t,x}^\infty} \\
&\lesssim \sum_{m\geq j+O(1)} 2^{-\frac{m}{2}}2^{k_2}\min(2^{\frac{m-k_2}{4}}, 2^{\frac{m_0}{2}})\|P_{k_1,\kappa_1}\phi_1\|_{S[k_1]}\|P_{k_2,\kappa_2}\phi_2\|_{S[k_2]} \\
&\lesssim \delta 2^{\frac{k_2}{2}}2^{-\frac{j-k_2}{3}}\|Q_{\geq j+O(1)}P_{k_1,\kappa_1}\phi_1\|_{\dot X_{k_1}^{0,\frac12,\infty}}\|P_{k_2,\kappa_2}\phi_2\|_{L_t^\infty L_x^2} \qquad (4.36)
\end{aligned}$$

as claimed. The inner sums run over $\kappa_1', \kappa_2' \in \mathcal{C}_{\frac{j-k_2}{2}}$ with $\mathrm{dist}(\kappa_1', \kappa_2') \lesssim 2^{\frac{j-k_2}{2}}$. The bound in (4.36) can be summed over the caps κ_1, κ_2 by definition of the $S[k]$ norm.

Finally, (4.33) follows from the preceding since the gain of δ was obtained only from the low-frequency function ϕ_2. We can therefore square-sum the final estimate to obtain the desired conclusion. □

4.2. **An algebra estimate for $S[k]$.** The following bilinear bound expresses something close to an algebra property of the $S[k]$ spaces. It is obtained by removing the restriction on the modulation of the output in Lemma 4.7.

Lemma 4.11. *For any $j, k \in \mathbb{Z}$,*

$$\|P_k Q_j(\phi\psi)\|_{\dot X^{0,\frac12,\infty}} \lesssim 2^{k_1\wedge k_2}2^{\frac{k-k_1\vee k_2}{4}}2^{\frac{j-k_1\wedge k_2}{4}\wedge 0}\,2^{(k_1\vee k_2-j)(\frac12-\varepsilon)\wedge 0}\,\|\phi\|_{S[k_1]}\|\psi\|_{S[k_2]} \qquad (4.37)$$

provided ϕ, ψ are Schwartz functions which are adapted to k_1 and k_2, respectively.

Proof. We commence with the high-high case $k_1 = k_2 + O(1) = 0$ and $k \leq O(1)$. We need to prove that

$$\|P_k Q_j(\phi\psi)\|_{\dot X^{0,\frac12,\infty}} \lesssim 2^{\frac{k}{4}}\min\left(2^{\frac{j}{4}}, 2^{-j(\frac12-\varepsilon)}\right)\|\phi\|_{S[k_1]}\|\psi\|_{S[k_2]}$$

To begin with, one has

$$\begin{aligned}
2^{j/2}\|P_k Q_j(Q_{>j-C}\phi\cdot\psi)\|_{L_t^2 L_x^2} &\lesssim 2^k 2^{j/2}2^{\frac{j-k}{4}\wedge 0}\|P_k Q_j(Q_{>j-C}\phi\cdot\psi)\|_{L_t^2 L_x^1} \\
&\lesssim 2^k 2^{j/2}2^{\frac{j-k}{4}\wedge 0}\|Q_{>j-C}\phi\|_{L^2L^2}\|\psi\|_{L^\infty L^2} \\
&\lesssim 2^k 2^{\frac{j-k}{4}\wedge 0}\min(1, 2^{-(\frac12-\varepsilon)j})\|\phi\|_{S[k_1]}\|\psi\|_{S[k_2]}
\end{aligned}$$

which is admissible. So it suffices to estimate $P_k Q_j(Q_{\leq j-C}\phi\cdot Q_{\leq j-C}\psi)$. As usual, we perform a wave-packet decomposition by means of Lemma 4.1. Note that (4.1) holds here. We begin with (4.3) where we choose $r' := 2^k$. Thus, $k < -C$ and $j = O(1)$, and in view of (2.30)

$$\begin{aligned}\|P_k Q_j(Q_{\le j-C}\phi^+ \cdot Q_{\le j-C}\psi^+)\|_{L^2_t L^2_x} &\lesssim \sum_{\kappa\in\mathcal{C}_k} \|P_\kappa Q_{\le j-C}\phi^+ \cdot P_{-\kappa} Q_{\le j-C}\psi^+\|_{L^2_t L^2_x} \\ &\lesssim \sum_{\kappa\in\mathcal{C}_k} |\kappa|^{\frac12} \|P_\kappa Q_{\le j-C}\phi^+\|_{S[k_1,\kappa]} \|P_{-\kappa} Q_{\le j-C}\psi^+\|_{S[k_2,-\kappa]} \\ &\lesssim 2^{k/2+} \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]}\end{aligned}$$

where we invoked Lemma 2.12 in the final step. The same estimate applies to ϕ^- and ψ^-. It therefore suffices to assume that $j \le k + O(1)$; but then Lemma 4.7 applies.
Next, we consider the low-high case $k = k_2 + O(1) = 0$, $k_1 < -C$. We need to prove that

$$2^{\frac{j}{2}} \|P_0 Q_j(\phi\psi)\|_{L^2_t L^2_x} \lesssim 2^{k_1} 2^{\frac{j-k_1}{4}\wedge 0} \min(1, 2^{-j(\frac12-\varepsilon)}) \|\phi\|_{S[k_1]} \|\psi\|_{S[0]}$$

In view of Lemma 4.7 we can assume that $j \ge k_1$. From the $\dot X^{s,b,q}$ components of the $S[k]$ norm,

$$\begin{aligned}&2^{j/2} \|P_0 Q_j(Q_{\ge j-C}\phi \cdot \psi)\|_{L^2_t L^2_x} \lesssim 2^{j/2} \|Q_{\ge j-C}\phi\|_{L^2 L^\infty} \|\psi\|_{L^\infty L^2} \\ &\lesssim 2^{k_1} 2^{\frac{j-k_1}{4}\wedge 0} \min(1, 2^{-(\frac12-\varepsilon)j}) \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]}\end{aligned}$$

Finally, it remains to bound

$$2^{j/2} \|P_0 Q_j(Q_{<j-C}\phi \cdot Q_{\ge j-C}\psi)\|_{L^2_t L^2_x}$$

which will be done using the usual angular decomposition. In fact, from Lemma 4.1, and provided $j \le C$, and with $\ell = \frac{m-k_1}{2} \wedge 0$,

(4.38)

$$\begin{aligned}2^{j/2} \|P_0 Q_j(Q_{<j-C}\phi \cdot Q_{\ge j-C}\psi)\|_{L^2_t L^2_x} &\lesssim \sum_{m\ge j-C} 2^{j/2} \sum_{\kappa,\kappa'\in\mathcal{C}_\ell, \kappa\sim\kappa'} \|P_0 Q_j(P_{k_1,\kappa} Q_{<j-C}\phi \cdot P_{k_2,\kappa'} Q_m\psi)\|_{L^2_t L^2_x} \\ &\lesssim \sum_{m\ge j-C} 2^{j/2} \sum_{\kappa,\kappa'\in\mathcal{C}_\ell, \kappa\sim\kappa'} \|P_{k_1,\kappa} Q_{<j-C}\phi\|_{L^\infty L^\infty} \|P_{k_2,\kappa'} Q_m\psi\|_{L^2_t L^2_x} \\ &\lesssim \sum_{m\ge j-C} 2^{j/2} \sum_{\kappa,\kappa'\in\mathcal{C}_\ell, \kappa\sim\kappa'} 2^{k_1} 2^{\frac{m-k_1}{4}\wedge 0} \|P_{k_1,\kappa}\phi\|_{L^\infty L^2} \|P_{k_2,\kappa'} Q_m\psi\|_{L^2_t L^2_x} \\ &\lesssim 2^{k_1} 2^{\frac{j-k_1}{4}\wedge 0} \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]}\end{aligned}$$

where we used Corollary 2.16 in the final inequality. If $j \ge C$, then only $m = j + O(1)$ contributes to the sum in (4.38). The $\dot X^{0,1-\varepsilon,2}$ component of the $S[k]$-norm then leads to a gain of $\min(1, 2^{-(\frac12-\varepsilon)j})$ and we are done. □

Corollary 4.12. *Under the same conditions as in the previous lemma and provided $k_1 \ll k_2$ one has*

$$\|P_k(\phi Q_{<a}\psi)\|_{\dot X^{0,\frac12,1}} \lesssim 2^{k_1} \big(1 + (k_2 \wedge a - k_1)_+\big) \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]} \tag{4.39}$$

where $k = O(1) + k_2$.

Proof. Summing (4.37) over j yields (4.39) with $a \geq k_2$. It thus suffices to consider $a \leq k_2$. If $a \leq k_1$ we use $Q_{<a} = Q_{<k_1} Q_{<a}$ to reduce matters to $a = k_1$ (see Corollary 2.16). If $a = k_1$, then

$$\begin{aligned}
&\sum_j 2^{j/2} \|P_k Q_j(\phi Q_{<a}\psi)\|_{L^2_t L^2_x} \\
&\leq \sum_{j \leq a+10} 2^{j/2} \|P_k Q_j(Q_{<a}\phi Q_{<a}\psi)\|_{L^2_t L^2_x} \\
&\quad + \sum_{j \geq a+10} 2^{j/2} \|P_k Q_j(Q_{j+O(1)}\phi Q_{<a}\psi)\|_{L^2_t L^2_x} \\
&\lesssim 2^{k_1} \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]} + \sum_{j \geq a+10} 2^{j/2} \|Q_{j+O(1)}\phi\|_{L^2 L^\infty} \|Q_{<a}\psi\|_{L^\infty L^2} \\
&\lesssim 2^{k_1} \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]} + \sum_{j \geq a+10} 2^{j/2} 2^{(\frac{3}{2}-\varepsilon)k_1} 2^{-j(1-\varepsilon)} \|\phi\|_{S[k_1]} \|Q_{<a}\psi\|_{S[k_2]} \\
&\lesssim 2^{k_1} \|\phi\|_{S[k_1]} \|\psi\|_{S[k_2]}
\end{aligned}$$

as desired. The sum over $j \leq a + 10$ was estimated via Lemma 4.11. If $k_1 < a \leq k_2$, one proceeds similarly. □

4.3. **Bilinear estimates involving both $S[k_1]$ and $N[k_2]$ waves.** The following lemma is a crucial tool. In essence, it expresses the property $N \times S \hookrightarrow N$.

Lemma 4.13. *For ϕ and F which are k_1 and k_2-adapted, respectively, one has*

$$\|P_k(\phi F)\|_{N[k]} \lesssim 2^{k_1 \wedge k_2}\, 2^{\frac{j-k\wedge k_1 \wedge k_2}{4} \wedge 0} \|\phi\|_{S[k_1]} \|F\|_{N[k_2]} \tag{4.40}$$

provided $P_{k_2} Q_j F = F$ and under the following condition

$$P_k Q_{\leq j-C}(Q_{<j-C}\phi \cdot F) = P_k Q_{\leq j-C}(Q_{<j+k-k_1-C}\phi \cdot F) \tag{4.41}$$

in the case $k_1 = k_2 + O(1) \geq k + O(1) \geq j$. If (4.41) fails, then one loses a factor of $1 + (k_1 - k)_+$ on the right-hand side of (4.40); alternatively, one has the following weaker version of (4.40)

$$\|P_k(\phi F)\|_{N[k]} \lesssim 2^{k_1 \wedge k_2}\, 2^{\frac{j-k\wedge k_1 \wedge k_2}{4} \wedge 0} \|\phi\|_{\dot{X}^{0,\frac{1}{2},1}_{k_1}} \|F\|_{N[k_2]} \tag{4.42}$$

Proof. We remark beforehand that this proof will only use the $\dot{X}^{0,\frac{1}{2},\infty}_{k_1}$-norm for the elliptic regime $\phi = P_k Q_{\geq k}\phi$ of the $S[k]$ norm. In particular, the imbedding $\|\phi\|_{S[k_1]} \lesssim \|\phi\|_{\dot{X}^{0,\frac{1}{2},\infty}_{k_1}}$ holds without any restrictions on the modulation, cf. (2.20). We start with the high-high case $k_1 = k_2 + O(1) = 0$. Throughout this proof, we shall freely use Lemma 2.15 in order to remove $Q_{<j-C}$ from various estimates. First, we consider the case where $\phi = Q_{\geq j-C}\phi$. If $j \geq k$, then by Bernstein's and Hölder's inequalities

$$\begin{aligned}
\|P_k(\phi F)\|_{N[k]} &\lesssim 2^{-k} \|P_k(\phi F)\|_{L^1 L^2} \lesssim \|\phi F\|_{L^1 L^1} \\
&\lesssim \|\phi\|_{L^2_t L^2_x} \|F\|_{L^2_t L^2_x} \lesssim \|\phi\|_{S[k_1]} 2^{-\frac{j}{2}} \|F\|_{L^2_t L^2_x} \\
&\lesssim \|\phi\|_{S[k_1]} \|F\|_{N[k_2]}
\end{aligned}$$

which is admissible. If on the other hand $j \leq k$, then we again have to consider several subcases. If $\phi = Q_{\geq k}\phi$, then

$$\begin{aligned}
2^{-k} \|Q_{\geq k}\phi \cdot F\|_{L^1_t L^2_x} &\lesssim \|Q_{\geq k}\phi \cdot F\|_{L^1_t L^1_x} \lesssim \|Q_{\geq k}\phi\|_{L^2_t L^2_x} \|F\|_{L^2_t L^2_x} \\
&\lesssim 2^{\frac{j-k}{2}} \|\phi\|_{S[k_1]} \|F\|_{N[k_2]}
\end{aligned}$$

which is admissible. Hence it suffice to assume that $\phi = Q_{j-C\leq\cdot\leq k}\phi$. Furthermore, we can assume that the output is at modulation $\leq j$. In fact, by the improved Bernstein's inequality,

$$\begin{aligned}\|P_kQ_{>j}(\phi F)\|_{N[k]} &\lesssim 2^{-k}\sum_{\ell>j}2^{-\frac{\ell}{2}}\|P_kQ_\ell(\phi F)\|_{L^2_tL^2_x}\\ &\lesssim \sum_{\ell>j}2^{-\frac{\ell}{2}}2^{\frac{\ell-k}{4}\wedge 0}\|\phi F\|_{L^2L^1}\\ &\lesssim \sum_{\ell>j}2^{\frac{j-\ell}{2}}2^{\frac{\ell-k}{4}\wedge 0}\|\phi\|_{L^\infty_tL^2_x}2^{-\frac{j}{2}}\|F\|_{L^2_tL^2_x}\\ &\lesssim 2^{\frac{j-k}{4}}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

as desired. Now consider the output of modulation at most j. We also first restrict ourselves to the contributions by $Q_{j-C<\cdot<j+C}\phi$. Thus, by Lemma 4.1 and Lemma 2.15,

$$\begin{aligned}&\|P_kQ_{\leq j}(Q_{j-C<\cdot<j+C}\phi\cdot F)\|_{N[k]}\\ &\lesssim 2^{-k}\sum_{\ell=j+O(1)}\|P_kQ_{\leq j}(Q_\ell\phi\cdot F)\|_{L^1_tL^2_x}\\ &\lesssim 2^{-k}\sum_{\ell=j+O(1)}\sum_{D\in\mathcal{D}_k}\sum_{\kappa\sim\kappa'\in\mathcal{C}_{(\ell+k)/2}}\|P_kQ_{\leq j}(P_DP_\kappa Q_\ell\,\phi\cdot P_{\kappa'}P_{-D}F)\|_{L^1_tL^2_x}\\ &\lesssim 2^{-k}\sum_{\ell=j+O(1)}\sum_{D\in\mathcal{D}_k}\sum_{\kappa\sim\kappa'\in\mathcal{C}_{(\ell+k)/2}}\|P_DP_\kappa Q_\ell\,\phi\cdot P_{\kappa'}P_{-D}F\|_{L^1_tL^2_x}\end{aligned}$$

where $\kappa\sim\kappa'$ means $\mathrm{dist}(\kappa,\kappa')\lesssim\mathrm{diam}(\kappa)$. Moreover, $\mathcal{D}_k$ is a cover of $\{|\xi|\sim 1\}$ by disks of diameter 2^k and with overlap uniformly bounded in k; the associated projections are P_D. Hence, one can further estimate (recall $j\leq k$)

$$\begin{aligned}&\|P_kQ_{\leq j}(\phi\cdot F)\|_{N[k]}\\ &\lesssim 2^{-k}\sum_{\ell=j+O(1)}\sum_{D\in\mathcal{D}_k}\sum_{\kappa\sim\kappa'\in\mathcal{C}_{(\ell+k)/2}}\|P_DP_\kappa Q_\ell\,\phi\|_{L^2_tL^\infty_x}\|P_{\kappa'}P_{-D}F\|_{L^2_tL^2_x}\\ &\lesssim 2^{-k}\sum_{\ell=j+O(1)}\sum_{D\in\mathcal{D}_k}\sum_{\kappa\sim\kappa'\in\mathcal{C}_{(\ell+k)/2}}2^{\frac{\ell+3k}{4}}\|P_DP_\kappa Q_\ell\,\phi\|_{L^2_tL^2_x}\|P_{\kappa'}P_{-D}F\|_{L^2_tL^2_x}\\ &\lesssim 2^{(j-k)/4}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

Next, we consider the output of modulation at most j and $\phi = Q_{j+C\leq\cdot\leq k}\phi$. Then we are in the "imbalanced case" of Lemma 4.1 whence

$$P_kQ_{\leq j}(\phi F) = \sum_{k\geq\ell\geq j+C}\sum_{\kappa,\kappa',\kappa''}P_{k,\kappa}Q_{\leq j}(P_{k_1,\kappa'}Q_l\phi\cdot P_{k_2,\kappa''}F)$$

where $\kappa\in\mathcal{C}_{\frac{\ell-k}{2}}$, $\kappa',\kappa''\in\mathcal{C}_{(\ell+k)/2}$ and $\mathrm{dist}(\kappa,\kappa')\sim 2^{\frac{\ell-k}{2}}$, $\mathrm{dist}(\kappa',\kappa'')\sim 2^{\frac{\ell+k}{2}}$. Using (2.29) one obtains

$$\begin{aligned}&\|P_kQ_{\leq j}(\phi F)\|_{N[k]}\\ &\leq 2^{-k}\sum_{k\geq\ell\geq j+C}\sum_{\kappa,\kappa',\kappa''}\|P_{k,\kappa}Q_{\leq j}(P_{k_1,\kappa'}Q_\ell\phi\cdot P_{k_2,\kappa''}F)\|_{\mathrm{NF}[\kappa]}\\ &\lesssim \sum_{k\geq\ell\geq j+C}2^{-k}\frac{2^{\frac{\ell+k}{4}}}{2^{\frac{\ell-k}{2}\wedge 0}}\sum_{\kappa',\kappa''}\|P_{k_1,\kappa'}Q_\ell\phi\|_{S[k_1,\kappa]}\|P_{k_2,\kappa''}F\|_{L^2_tL^2_x}\\ &\lesssim \sum_{k\geq\ell\geq j+C}2^{\frac{j-k}{4}}2^{\frac{j-\ell}{4}}\|Q_\ell\phi\|_{S[k_1]}2^{-\frac{j}{2}}\|F\|_{L^2_tL^2_x}\lesssim 2^{\frac{j-k}{4}}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

as desired. To pass to the final inequality one uses (2.20) as well as $\|Q_\ell\phi\|_{S[k_1]}\sim\|Q_\ell\phi\|_{\dot X^{0,\frac12,\infty}_{k_1}}$.

Now assume $Q_{<j-C}\phi = \phi$. We first dispose of outputs of modulation exceeding $j - C$. If $j \geq k$, then

$$\begin{aligned}\|P_k Q_{>j-C}(\phi F)\|_{N[k]} &\lesssim 2^{-\frac{j}{2}} \|P_k Q_{>j-C}(\phi F)\|_{L^2_t L^1_x} \lesssim 2^{-\frac{j}{2}} \|\phi\|_{L^\infty L^2} \|F\|_{L^2 L^2} \\ &\lesssim 2^{-\frac{j}{2}} \|\phi\|_{S[k_1]} \|F\|_{L^2_t L^2_x} \lesssim \|\phi\|_{S[k_1]} \|F\|_{N[k_2]}\end{aligned}$$

which is admissible. On the other hand, if $j \leq k$, then

$$\begin{aligned}\|P_k Q_{>j-C}(\phi F)\|_{N[k]} &\lesssim 2^{-k} \sum_{\ell > j-C} 2^{-\frac{\ell}{2}} \|P_k Q_\ell(\phi F)\|_{L^2_t L^2_x} \\ &\lesssim \sum_{k \geq \ell > j-C} 2^{-\frac{\ell}{2}} 2^{\frac{\ell-k}{4}} \|P_k Q_\ell(\phi F)\|_{L^2_t L^1_x} + 2^{-\frac{3k}{2}} \|P_k Q_{\geq k}(\phi F)\|_{L^2_t L^2_x} \lesssim 2^{\frac{j-k}{4}} \|\phi\|_{S[k_1]} \|F\|_{N[k_2]}\end{aligned}$$

as desired. It therefore remains to consider

$$P_k Q_{\leq j-C}(Q_{<j-C}\phi \cdot F) = \sum_{\pm} P_k Q_{\leq j-C}(Q_{<j-C}\phi^{\pm} \cdot F^{\pm})$$

where all four possibilities $(++), (+-), (-+), (--)$ are allowed on the right-hand side. We first dispose of the contributions "opposing waves" as described by (4.3). This occurs only if $k < -C$ and $j = O(1)$, in fact,

$$P_k Q_{\leq j-C}(Q_{<j-C}\phi^+ \cdot F^+) = P_k Q_{-C<\cdot<C}(Q_{<j-C}\phi^+ \cdot F^+)$$

whence

$$\begin{aligned}&\|P_k Q_{\leq j-C}(Q_{<j-C}\phi^+ \cdot F^+)\|_{N[k]} \\ &\lesssim \|\phi^+ F^+\|_{L^2 L^1} \lesssim \|\phi^+\|_{L^\infty L^2} \|F^+\|_{L^2_t L^2_x} \\ &\lesssim \|\phi\|_{S[k_1]} \|F\|_{N[k_2]}\end{aligned}$$

which is admissible. Therefore, we can now ignore the contribution of (4.3). Let us now also assume without loss of generality that $\phi = \phi^+$, see Lemma 2.12. Using duality and Lemma 4.1, one obtains in view of (4.41)

$$P_k Q_{\leq j-C}(Q_{<j-C}\phi \cdot F) = \sum_{\pm} \sum_{\kappa,\kappa',\kappa''} P_{k,\pm\kappa} Q^{\pm}_{\leq j-C}(P_{k_1,\kappa'} Q_{<j+k-C}\phi \cdot P_{k_2,\kappa''}F) \tag{4.43}$$

with caps $\kappa \in \mathcal{C}_\ell$, $\kappa', \kappa'' \in \mathcal{C}_\ell$ satisfying $\mathrm{dist}(\kappa', \kappa'') \sim 2^m$, $\mathrm{dist}(\kappa, \kappa') \sim 2^\ell$ where $\ell = (j-k)/2$, $m = (j+k)/2$. Note that Lemma 4.1 also implies that $j \leq k + O(1)$. Since

$$P_{k,\pm\kappa} Q^{\pm}_{\leq j-C} = P_{k,\pm\kappa} Q^{\pm}_{\leq k+2\ell-C} \tag{4.44}$$

the right-hand side of (4.43) represents a wave-packet decomposition in the sense of Definition 2.9. Moreover, the operators in (4.44) are disposable in the sense of Lemma 2.14. Therefore,

$$\|P_k Q_{\leq j-C}(Q_{<j-C}\phi \cdot F)\|_{N[k]} \lesssim 2^{-k} \max_{\pm} \sum_{\kappa',\kappa''} \left\|P_{k_1,\kappa'} Q_{<j+k-C}\phi \cdot P_{k_2,\kappa''}F\right\|_{\mathrm{NF}[\kappa]}$$

We could discard κ here since the choice of κ' leaves only a finite number of choice of κ. Invoking (2.29), this can be further estimated by

$$\begin{aligned}&\lesssim 2^{-k} \max_{\pm} \sum_{\kappa',\kappa''} \frac{2^{\frac{j+k}{4}}}{2^{\frac{j-k}{2}}} \|P_{k_1,\kappa'} Q_{<j+k-C}\phi\|_{S[k_1,\kappa']} \|P_{k_2,\kappa''}F\|_{L^2_t L^2_x} \\ &\lesssim 2^{(j-k)/4} \|\phi\|_{S[k_1]} \|F\|_{N[k_2]}\end{aligned}$$

To pass to the final inequality here it was essential that (4.41) reduced the modulation of ϕ from $< j - C$ to $< j + k - C$. Indeed, if (4.41) fails, then we need to write $Q_{<j-C}\phi = Q_{<j+k-C}\phi + Q_{j+k-C\leq\cdot<j-C}\phi$. For the first summand here one applies the argument we just gave, whereas for the second summand the best one can do is to invoke (2.20) which results in the loss of of factor of k a claimed. This concludes the high-high case.

Let us now consider the low-high case $k_1 < -C$, $k_2 = k = O(1)$. Since (2.24) implies that

$$\begin{aligned}\|Q_{\geq j-C}\phi \cdot F\|_{L^1L^2} &\lesssim \|Q_{\geq j-C}\phi\|_{L^2L^\infty}\|F\|_{L^2_tL^2_x}\\ &\lesssim 2^{-\frac{j}{2}}2^{\frac{j-k_1}{4}\wedge 0}2^{k_1}\|\phi\|_{S[k_1]}\|F\|_{L^2_tL^2_x}\\ &\lesssim 2^{\frac{j-k_1}{4}\wedge 0}2^{k_1}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

it will suffice to bound $\|Q_{\leq j-C}\phi \cdot F\|_{N[k]}$. Moreover, if $j \geq k_1 + C$, then the modulation of $Q_{\leq j-C}\phi \cdot F$ is on the order of j whence

$$\begin{aligned}\|Q_{\leq j-C}\phi \cdot F\|_{N[k]} &\lesssim 2^{-\frac{j}{2}}\|Q_{\leq j-C}\phi \cdot F\|_{L^2_tL^2_x}\\ &\lesssim 2^{-\frac{j}{2}}\|Q_{\leq j-C}\phi\|_{L^\infty L^\infty}\|F\|_{L^2_tL^2_x}\\ &\lesssim 2^{-\frac{j}{2}}2^{k_1}\|\phi\|_{S[k_1]}\|F\|_{L^2_tL^2_x} \lesssim 2^{k_1}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

as desired. We may therefore assume that $j \leq k_1 + C$. We first consider the case where the output has modulation $\geq j - C$. More precisely, let $j - C \leq m \leq k_1 + C$, $\ell = (m - k_1)/2$, as well as without loss of generality $F = F^+$. Then by the balanced modulation case of Lemma 4.1,

$$Q_m(Q_{\leq j-C}\phi \cdot F) = \sum_{\kappa,\kappa'} P_{k,\kappa}Q_m^+(P_{k_1,\kappa'}Q_{\leq j-C}\phi \cdot F)$$

where κ, κ' are caps of size $C^{-1}2^\ell$ and with $\mathrm{dist}(\kappa,\kappa') \lesssim 2^\ell$. Therefore,

$$\begin{aligned}\|Q_m(Q_{\leq j-C}\phi \cdot F)\|_{N[k]} &\lesssim 2^{-m/2}\|\sum_{\kappa,\kappa'} P_{k,\kappa}Q_m^+(P_{k_1,\kappa'}Q_{\leq j-C}\phi \cdot F)\|_{L^2_tL^2_x}\\ &\lesssim 2^{-m/2}\Big(\sum_{\kappa,\kappa'}\|P_{k_1,\kappa'}Q_{\leq j-C}\phi\|^2_{L^\infty L^\infty}\|F\|^2_{L^2L^2}\Big)^{\frac{1}{2}}\\ &\lesssim 2^{-m/2}\Big(\sum_{\kappa,\kappa'}2^{2k_1}2^\ell\|P_{k_1,\kappa}Q_{\leq j-C}\phi\|^2_{L^\infty L^2}\ \|F\|^2_{L^2L^2}\Big)^{\frac{1}{2}}\\ &\lesssim 2^{-m/2}2^{-(k_1-m)/4}2^{k_1}\|Q_{\leq j-C}\phi\|_{L^\infty L^2}\|F\|_{L^2_tL^2_x}\\ &\lesssim 2^{-m/2}2^{-(k_1-m)/4}2^{k_1}\|\phi\|_{S[k_1]}\|F\|_{L^2_tL^2_x}\end{aligned}$$

where we used Lemma 2.15 in the final step. Summing over $m \geq j - C$ implies that

$$\begin{aligned}\|Q_{>j-C}(Q_{\leq j-C}\phi \cdot F)\|_{N[k]} &\lesssim 2^{k_1}2^{-j/2}2^{-(k_1-j)/4}\|\phi\|_{S[k_1]}\|F\|_{L^2_tL^2_x}\\ &\lesssim 2^{k_1}2^{-(k_1-j)/4}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

which is admissible.

It therefore remains to bound $\|Q_{<j-C}(Q_{\leq j-C}\phi \cdot F)\|_{N[k]}$ for which we shall again apply a wave-packet decomposition as in Lemma 4.1. Since $j \leq k_1 + C$ and $k_1 \ll -1$, we can assume that $j \leq -C$ in applying Lemma 4.1 (which allows us to ignore the opposing $(++)$ or $(--)$ contributions in (4.3)). Without loss of generality, we assume further that $\phi = \phi^+$ (see Lemma 2.12). Then with caps κ, κ' of size $C^{-1}2^m$ and separation $\sim 2^m$ where $m := (j - k_1)/2$,

$$\begin{aligned}\|Q_{<j-C}(Q_{\leq j-C}\phi \cdot F)\|_{N[k]} &\lesssim \Big\|\sum_{\kappa\sim\kappa'} P_{k,\kappa}Q_{<j-C}(P_{k_1,\kappa'}Q_{\leq j-C}\phi \cdot F)\Big\|_{N[k]}\\ &\lesssim \Big(\sum_{\kappa\sim\kappa'}\Big\|P_{k,\kappa}Q_{<j-k_1}(P_{k_1,\kappa'}Q_{\leq j-C}\phi \cdot F)\Big\|^2_{N[k]}\Big)^{\frac{1}{2}}\end{aligned} \tag{4.45}$$

where we used Corollary 2.23 to dispose of $Q_{\leq j-C}$. In view of (2.29) this is further bounded by

$$\begin{aligned}&\lesssim 2^{\frac{k_1}{2}}2^{-\frac{j-k_1}{4}}\Big(\sum_{\kappa'}\|P_{k_1,\kappa'}Q_{\leq j-C}\phi\|^2_{S[k_1,\kappa']}\|F\|^2_{L^2_tL^2_x}\Big)^{\frac{1}{2}}\\ &\lesssim 2^{(j-k_1)/4}2^{k_1}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

as desired.

It remains to consider the high-low case $k = k_1 + O(1) = 0$, $k_2 < -C$. First,

$$\begin{aligned}\|Q_{>j+k_2}\phi \cdot F\|_{N[k]} &\lesssim \|Q_{>j+k_2}\phi \cdot F\|_{L^1L^2} \\ &\lesssim \|Q_{>j+k_2}\phi\|_{L^2L^2}\|F\|_{L^2L^\infty} \\ &\lesssim 2^{-(j+k_2)/2}\|\phi\|_{S[k_1]}2^{k_2}2^{\frac{j-k_2}{4}\wedge 0}\|F\|_{L^2_tL^2_x} \\ &\lesssim 2^{k_2}2^{\frac{j-k_2}{4}\wedge 0}\|\phi\|_{S[k_1]}2^{-\frac{j}{2}}2^{-k_2}\|F\|_{L^2_tL^2_x}\end{aligned}$$

which is acceptable with a factor of $2^{\frac{k_2}{2}}$ to spare. The reason for using $Q_{>j+k_2}$ rather than $Q_{>j}$ will become clear momentarily. Next,

$$\|Q_{\le j+k_2}\phi \cdot F\|_{N[k]} \lesssim \|Q_{\ge j+k_2-C}[Q_{\le j+k_2}\phi \cdot F]\|_{N[k]} \tag{4.46}$$

$$+ \|Q_{<j+k_2-C}[Q_{\le j+k_2}\phi \cdot F]\|_{N[k]} \tag{4.47}$$

As usual, (4.46) is controlled in the $\dot{X}^{-1,-\frac{1}{2},1}$ norm whence

$$\begin{aligned}(4.46) &\lesssim 2^{-(j+k_2)/2}\|Q_{\le j+k_2}\phi \cdot F\|_{L^2_tL^2_x} \\ &\lesssim 2^{-(j+k_2)/2}\|Q_{\le j+k_2}\phi\|_{L^\infty L^2}\|F\|_{L^2L^\infty} \\ &\lesssim 2^{-(j+k_2)/2}\|Q_{\le j+k_2}\phi\|_{L^\infty L^2}2^{k_2}2^{\frac{j-k_2}{4}\wedge 0}\|F\|_{L^2_tL^2_x} \\ &\lesssim 2^{k_2}2^{\frac{j-k_2}{4}\wedge 0}\|\phi\|_{S[k_1]}2^{-\frac{j}{2}}2^{-k_2}\|F\|_{L^2_tL^2_x}\end{aligned}$$

which is again acceptable. Finally, we perform a wave-packet decomposition on (4.47) via Lemma 4.1 in the imbalanced case and duality. Thus, one has

$$Q_{<j+k_2-C}(Q_{\le j+k_2}\phi \cdot F) = \sum_{\kappa,\kappa'} P_{k,\kappa}Q^+_{<j+k_2-C}(P_{k_1,\kappa'}Q_{\le j+k_2}\phi \cdot F)$$

where the sum runs over pairs of caps κ, κ' of size $C^{-1}2^\ell$ with $\ell := (j+k_2)/2$ and $\mathrm{dist}(\kappa,\kappa') \sim 2^\ell$. Moreover, $j \le k_2 + O(1)$ since the only other possibility $j = O(1)$ allowed by (4.3) contributes a vanishing term (as does $Q^-_{<j+k_2-C}$). Therefore, with $\kappa' \sim \kappa$ denoting the admissible pairs,

$$\begin{aligned}&\lesssim \Big(\sum_\kappa \Big\| \sum_{\kappa'\sim\kappa} P_{k,\kappa}Q^+_{<j+k_2-C}(P_{k_1,\kappa'}Q_{\le j+k_2}\phi \cdot F)\Big\|^2_{N[\kappa]}\Big)^{\frac{1}{2}} \\ &\lesssim 2^{-\frac{\ell}{2}}2^{j/2}2^{k_2}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]} \\ &\lesssim 2^{k_2}2^{(j-k_2)/4}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]}\end{aligned}$$

as desired. □

There is the following general estimate that does not require (4.41) since we restrict ourselves to $k \ge k_1 + O(1)$.

Corollary 4.14. *For ϕ and F which are k_1 and k_2-adapted, respectively, one has*

$$\|P_k(\phi F)\|_{N[k]} \lesssim 2^{k_1\wedge k_2}\, 2^{\frac{j-k\wedge k_1\wedge k_2}{4}\wedge 0}\|\phi\|_{S[k_1]}\|F\|_{N[k_2]} \tag{4.48}$$

provided $P_{k_2}Q_jF = F$ and $k = k_1 \vee k_2 + O(1)$.

Proof. This is an immediate consequence of Lemma 4.13. □

Another important technical variant of Lemma 4.13 has to do with an additional angular localization of the inputs. This will be important later in the trilinear section. Its statement is somewhat technically cumbersome, but this is precisely the form in which we shall use it later.

Corollary 4.15. *Let ϕ be k_1-adapted, and assume that for some $m_0 \le -100$, for every $\kappa \in \mathcal{C}_{m_0}$ there is a Schwarz function F_κ which is adapted to k_2 and so that $P_{k_2} Q_j F_\kappa = F_\kappa$. Then*

$$\sum_{\kappa \in \mathcal{C}_{m_0}} \|P_k(P_{k_1,\kappa}\phi\, F_\kappa)\|_{N[k]} \lesssim |m_0|\, 2^{k_1}\, 2^{\frac{j-k_1}{4}\wedge 0} \|\phi\|_{S[k_1]} \Big(\sum_{\kappa \in \mathcal{C}_{m_0}} \|F_\kappa\|^2_{N[k_2]} \Big)^{\frac12} \tag{4.49}$$

provided we are in the low-high case $k = k_2 + O(1) \ge k_1$. The sum here runs over caps with $\mathrm{dist}(\kappa_1, \kappa_2) \lesssim 2^{m_0}$.

Proof. For this, one simply repeats the proof of the low-high case of Lemma 4.13 with one additional twist: since $\sum_\kappa \|P_{k_1,\kappa}\phi\|^2_{S[k]}$ cannot be controlled by $\|\phi\|_{S[k_1]}$, one has to check carefully that the square summation — which (4.49) leads to after Cauchy-Schwarz — is compatible with the estimates we are making (the norm for F is always $L^2_t L^2_x$). This is the case if we place $P_{k_1,\kappa}\phi$ in $L^\infty_t L^2_x$ or an $\dot{X}^{s,b}$-norm. In the latter case one does not incur any loss due to orthogonality, whereas in the former case there is a loss of $|m_0|$, see Lemma 2.18. The only place where one cannot use either of these norms is (4.45). Indeed, if $k_1 + 2m_0 \le j - C$, then the caps of sizes 2^{m_0} are smaller than those of size $2^\ell = 2^{\frac{j-k_1}{2}}$ in the wave-packet decomposition of (4.45). In this case, however, one considers a wave-packet decomposition induced by the projections $P_{k_1,\kappa} Q_{<k_1+2m_0}$ with $\kappa \in \mathcal{C}_{m_0}$ which leads to the desired bound; the remaining projection $P_{k_1,\kappa} Q_{k_1+2m_0 \le \cdot \le j-C}$ is then controlled by means of Lemma 2.7 leading to a loos of $|m_0|$ as claimed. If, on the other hand, $k_1 + 2m_0 > j - C$, then this issue does not arise at all and the estimate (4.45) is performed essentially as in Lemma 4.13 — the only difference being that the caps in the wave-packet decomposition are grouped together inside the larger $\mathcal{C}_{m_0}$–caps. □

4.4. **Nullform bounds in the high-high case.** Henceforth, $\|\cdot\|_{S[k]}$ will mean the stronger norm $|\!|\!|\cdot|\!|\!|_{S[k]}$. The following definition introduces the basic nullforms as well as the method of "pulling out a derivative".

Definition 4.16. *The nullforms $\mathcal{Q}_{\alpha\beta}$ for $0 \le \alpha, \beta \le 2$, $\alpha \ne \beta$, are defined as*

$$\mathcal{Q}_{\alpha\beta}(\phi,\psi) := R_\alpha\phi R_\beta\psi - R_\beta\phi R_\alpha\psi$$

whereas

$$\mathcal{Q}_0(\phi,\psi) := R_\alpha\phi R^\alpha\psi$$

By "pulling out a derivative from" from $\mathcal{Q}_{\alpha\beta}$ we mean writing

$$\mathcal{Q}_{\alpha\beta}(\phi,\psi) = \partial_\alpha(|\nabla|^{-1}\phi R_\beta\psi) - \partial_\beta(|\nabla|^{-1}\phi R_\alpha\psi)$$

or the analogous expression with ϕ and ψ interchanged.

Recall the L^2-bound (4.13) of Lemma 4.5 for $\mathcal{Q}_{\alpha\beta}$-nullforms. We separate the nullform bounds according to high-high vs. high-low and low-high interactions. The high-high case is slightly more involved due to the possibility of opposing $(++)$ or $(--)$ waves with comparable frequencies and very small modulations which produce a wave of small frequency but very large modulation.

Lemma 4.17. *For any $\ell \le k + O(1)$, and ϕ_j adapted to k_j with $k_1 = k_2 + O(1)$,*

$$\|P_k Q_\ell \mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{\ell-k}{4+}} 2^{\frac{k}{2}} 2^{\frac{k-k_1}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]} \tag{4.50}$$

In particular,

$$\|P_k Q_{\le k+C} \mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{k-\frac{k_1}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]} \tag{4.51}$$

Finally, for any $m_0 \le -10$,

$$\Big(\sum_{\kappa \in \mathcal{C}_{m_0}} \|P_k Q_{\le k+C} \mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa}\phi_1,\phi_2)\|^2_{L^2_t L^2_x} \Big)^{\frac12} \lesssim |m_0|\, 2^{\frac{k_1}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]} \tag{4.52}$$

Proof. We can take $k_1 = k_2 + O(1) = 0$. First, by (4.13),

$$\|P_k Q_\ell \mathcal{Q}_{\alpha\beta}(Q_{\le k+\ell-C}\phi_1, Q_{\le k+\ell-C}\phi_2)\|_{L^2_t L^2_x} \lesssim 2^{\frac{\ell+k}{4}} 2^{\frac{k}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}$$

Second, by an angular decomposition into caps of size $2^{\frac{\ell+k}{2}}$,

$$\begin{aligned}&\sum_{\ell+k-C\le m\le \ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_m\phi_1,Q_{\le m}\phi_2)\|_{L^2_tL^2_x}\\ &\lesssim \sum_{\ell+k-C\le m\le\ell}2^{\frac{\ell+k}{2}}2^{\frac{\ell-k}{4}}2^k\|Q_m\phi_1\|_{L^2_tL^2_x}\|Q_{\le m}\phi_2\|_{L^\infty_tL^2_x}\\ &\lesssim 2^{\frac{\ell+3k}{4}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}\tag{4.53}$$

To pass to (4.53) one uses the improved Bernstein inequality, which yields a factor of $2^k2^{\frac{\ell-k}{4}}$, whereas the $2^{\frac{\ell+k}{2}}$ corresponds to the angular gain from the nullform (note that the error coming from the modulation is at most $2^m\le 2^\ell$ which is less than this gain). And third, by the improved Bernstein inequality and a decomposition into caps of size $2^{\frac{m+k}{2}}$,

$$\begin{aligned}\sum_{\ell\le m\le C}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_m\phi_1,Q_{\le m}\phi_2)\|_{L^2_tL^2_x}&\lesssim\sum_{\ell\le m\le C}2^{\frac{\ell-k}{4}}2^k\big(2^{\frac{m+k}{2}}+2^m\big)\|Q_m\phi_1\|_{L^2_tL^2_x}\|Q_{\le m}\phi_2\|_{L^\infty_tL^2_x}\\ &\lesssim\sum_{\ell\le m\le C}2^{\frac{\ell-k}{4}}2^k\big(2^{\frac{m+k}{2}}+2^m\big)2^{-\frac{m}{2}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\\ &\lesssim 2^{\frac{\ell-k}{4+}}2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

The factor $2^{\frac{m+k}{2}}+2^m$ here is made up out of the angular gain $2^{\frac{m+k}{2}}$ and the loss of 2^m in modulation (in case $\beta=0$). And finally, due to $\varepsilon<\frac12$,

$$\begin{aligned}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge C}\phi_1,\phi_2)\|_{L^2_tL^2_x}&\lesssim 2^k2^{\frac{\ell}{2}}\|\mathcal{Q}_{\alpha\beta}(Q_{\ge C}\phi_1,\phi_2)\|_{L^1_tL^1_x}\\ &\lesssim\sum_{m\ge C}2^k2^{\frac{\ell}{2}}2^m\|Q_m\phi_1\|_{L^2_tL^2_x}\|\tilde{Q}_m\phi_2\|_{L^2_tL^2_x}\\ &\lesssim 2^k2^{\frac{\ell}{2}}\sum_{m\ge C}2^m2^{-2m(1-\varepsilon)}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\\ &\lesssim 2^{\frac{\ell-k}{4}}2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

as desired.

Next, we consider (4.52). Here one essentially repeats the proof of (4.51) verbatim. The only difference being that instead of Lemma 4.5 one uses Corollary 4.6, in fact the null-form version of (4.24). Note that this loses a factor of $|m_0|$. To sum over the caps one also needs to invoke Lemma 2.18 in case of a $L^\infty_tL^2_x$-norm, which incurs the same loss. □

We shall also require the following technical variant of the estimate of Lemma 4.17. It obtains an improvement for the case of angular alignment in the Fourier supports of the inputs.

Lemma 4.18. *Let $\delta>0$ be small and $L>1$ be large. Then there exists $m_0=m_0(\delta,L)<0$ large and negative such that for any ϕ_j adapted to k_j for $j=1,2$,*

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\|P_kQ_{\le k+C}\mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa_1}\phi_1,P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x}\le\delta\,2^{\frac{k_1}{2}}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\tag{4.54}$$

provided $\max_{j=1,2}|k-k_j|\le L$. *The constant C is an absolute constant which does not depend on L or δ.*

Proof. Set $k=0$. We first note that summing (4.50) over $\ell\le -B$ already yields an improvement over (4.51) provided B is large enough (in relation to δ and L). Hence it suffices to consider the contribution of $P_0Q_\ell\mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa_1}\phi_1,P_{k_2,\kappa_2}\phi_2)$ with $-B\le\ell\le O(1)$ fixed. First, if we choose m_0 to be a sufficiently large negative integer, then

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}P_0Q_\ell\mathcal{Q}_{\alpha\beta}(Q_{\le\ell-C}P_{k_1,\kappa_1}\phi_1,\ Q_{\le\ell-C}P_{k_2,\kappa_2}\phi_2)=0$$

by Lemma 4.1. Second, by an angular decomposition into caps of size $2^{\frac{\ell}{2}}$,

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\ \sum_{\ell-C\le m\le C}\|P_0Q_\ell\mathcal{Q}_{\alpha\beta}(Q_mP_{k_1,\kappa_1}\phi_1,Q_{\le m}P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x}\\
&\le C(L,\delta)\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\ \sum_{\ell-C\le m\le C}\|Q_mP_{k_1,\kappa_1}\phi_1\|_{L^2_tL^2_x}\|Q_{\le m}P_{k_2,\kappa_2}\phi_2\|_{L^\infty_tL^\infty_x}\\
&\le C(L,\delta)\,|m_0|\,2^{\frac{m_0}{2}}\,\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\le\delta\,\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}
\end{aligned}$$

To pass to the last line we applied Cauchy-Schwarz to the sum over the caps as well as Lemma 2.18. The case dealing with $Q_{\le m}P_{k_1,\kappa_1}\phi_1$ and $Q_mP_{k_2,\kappa_2}\phi_2$ is analogous. And finally, due to $\varepsilon<\frac12$,

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\|P_0Q_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge C}P_{k_1,\kappa_1}\phi_1,P_{k_2,\kappa_2}\phi_2)\|_{L^2_tL^2_x}\\
&\lesssim\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\|P_0Q_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge C}P_{k_1,\kappa_1}\phi_1,P_{k_2,\kappa_2}\phi_2)\|_{L^1_tL^2_x}\\
&\le C(L,\delta)\sum_{m\ge C}2^m\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\|Q_mP_{k_1,\kappa_1}\phi_1\|_{L^2_tL^2_x}\|\tilde{Q}_mP_{k_2,\kappa_2}\phi_2\|_{L^2_tL^\infty_x}\\
&\le C(L,\delta)2^{\frac{m_0}{2}}\sum_{m\ge C}2^m2^{-2m(1-\varepsilon)}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\le\delta\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}
\end{aligned}$$

as desired. □

In case the output has "elliptic" rather than hyperbolic character, there is the following bound.

Lemma 4.19. *For any ϕ_j adapted to k_j with $k_1=k_2+O(1)$,*

$$\sum_{\ell\ge k+C}2^{-\varepsilon\ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_tL^2_x}\lesssim 2^{\frac{k}{2}}2^{-\varepsilon k_1}\langle k_1-k\rangle^2\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}$$

Furthermore,

$$\sum_{\ell\ge k+C}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\le k_1+C}\phi_1,Q_{\le k_2+C}\phi_2)\|_{L^2_tL^2_x}\lesssim 2^{\frac{k}{2}}\langle k_1-k\rangle^2\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\tag{4.55}$$

Proof. We set $k_1=k_2+O(1)=0$. One has the decomposition

$$\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_tL^2_x}\lesssim\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge\ell-C}\phi_1,Q_{\le k_1+C}\phi_2)\|_{L^2_tL^2_x}\tag{4.56}$$
$$+\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge\ell-C}\phi_1,Q_{>k_1+C}\phi_2)\|_{L^2_tL^2_x}\tag{4.57}$$
$$+\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{<\ell-C}\phi_1,Q_{\ge\ell-C}\phi_2)\|_{L^2_tL^2_x}\tag{4.58}$$
$$+\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{<\ell-C}\phi_1,Q_{<\ell-C}\phi_2)\|_{L^2_tL^2_x}\tag{4.59}$$

We begin with the estimate

$$\sum_{\substack{\ell\ge k+C\\ \ell\ne k_1+O(1)}}2^{-\varepsilon\ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_tL^2_x}\lesssim 2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}$$

which is stronger than what we claim – this is due to the fact that the the case of opposing $(++)$ and $(--)$ waves is excluded in this sum. We first consider the case $\ell\le k_1-C'$ where C' is large but still smaller than the constant C in (4.56)–(4.59). Then the term in (4.59) vanishes. On the one hand,

$$\begin{aligned}
(4.56)\lesssim{}&2^k\|P_kQ_\ell[\partial_\beta(Q_{\ge\ell-C}|\nabla|^{-1}\phi_1\cdot Q_{\le k_2+C}\partial_\alpha|\nabla|^{-1}\phi_2)\\
&-\partial_\alpha(Q_{\ge\ell-C}|\nabla|^{-1}\phi_1\cdot Q_{\le k_2+C}\partial_\beta|\nabla|^{-1}\phi_2)]\|_{L^2_tL^1_x}\\
\lesssim{}&2^{k+\ell}\|Q_{\ge\ell-C}\phi_1\|_{L^2_tL^2_x}\|\phi_2\|_{S[k_2]}
\end{aligned}$$

Here we used that

$$\|Q_{\leq k_2+C}\partial_\beta|\nabla|^{-1}\phi_2\|_{L_t^\infty L_x^2} \lesssim \|\phi_2\|_{S[k_2]}$$

Furthermore,

$$\begin{aligned}\sum_{k+C\leq\ell\leq k_1-C}\ \sum_{m\geq\ell-C} 2^{(1-\varepsilon)\ell}2^{k-k_1}\|Q_m\phi_1\|_{L_t^2L_x^2}\|\phi_2\|_{S[k_2]} &\lesssim 2^k\|\phi_1\|_{\dot X_{k_1}^{0,1-\varepsilon,2}}\|\phi_2\|_{S[k_2]}\\ &\lesssim 2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

as desired. The term (4.58) satisfies the same bound, more precisely, it can be reduced to (4.56), (4.57). Next, note that due to $\ell \leq k_1 - C'$ it suffices to consider $\phi_1 = Q_{\geq k_1+C}\phi_1$ in (4.57). Consequently,

$$\begin{aligned}\sum_{k+C\leq\ell\leq k_1-C} 2^{-\varepsilon\ell}(4.57) &\lesssim \sum_{k+C\leq\ell\leq k_1-C} 2^{-\varepsilon\ell}\sum_{m\geq k_1+C}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_m\phi_1,\tilde Q_m\phi_2)\|_{L_t^2L_x^2}\\ &\lesssim \sum_{k+C\leq\ell\leq k_1-C} 2^{-\varepsilon\ell}\sum_{m\geq k_1+C} 2^k2^{\frac{\ell}{2}}\|\mathcal{Q}_{\alpha\beta}(Q_m\phi_1,\tilde Q_m\phi_2)\|_{L_t^1L_x^1}\\ &\lesssim \sum_{k+C\leq\ell\leq k_1-C} 2^{-\varepsilon\ell}\sum_{m\geq k_1+C} 2^k2^{\frac{\ell}{2}}2^m2^{-2m(1-\varepsilon)}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\\ &\lesssim 2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

where we used that $\varepsilon < \frac{1}{4}$ in the final step. Second, suppose that $\ell \geq k_1 + C'$. Then

$$\begin{aligned}&\sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L_t^2L_x^2}\\ &\lesssim \sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(\tilde Q_\ell\phi_1,Q_{\leq\ell-5}\phi_2)\|_{L_t^2L_x^2} && (4.60)\\ &\quad+\sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\leq\ell-5}\phi_1,\tilde Q_\ell\phi_2)\|_{L_t^2L_x^2} && (4.61)\\ &\quad+\sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}\sum_{m\geq\ell-5}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_m\phi_1,\tilde Q_m\phi_2)\|_{L_t^2L_x^2} && (4.62)\end{aligned}$$

which are in turn estimated as follows:

$$\begin{aligned}(4.60) &\lesssim \sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}2^k\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(\tilde Q_\ell\phi_1,Q_{\leq\ell-5}\phi_2)\|_{L_t^2L_x^1}\\ &\lesssim \sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}2^k2^\ell\|\tilde Q_\ell\phi_1\|_{L_t^2L_x^2}\|\phi_2\|_{L_t^\infty L_x^2}\\ &\lesssim 2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

and similarly for (4.61), whereas (4.62) is bounded by

$$\begin{aligned}&\lesssim \sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}\sum_{m\geq\ell-5} 2^{\frac{\ell}{2}}2^k\|\mathcal{Q}_{\alpha\beta}(Q_m\phi_1,\tilde Q_m\phi_2)\|_{L_t^1L_x^1}\\ &\lesssim \sum_{\ell\geq k_1+C'} 2^{-\varepsilon\ell}\sum_{m\geq\ell-5} 2^{\frac{\ell}{2}}2^k2^m2^{-2m(1-\varepsilon)}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\\ &\lesssim 2^k\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]}\end{aligned}$$

as desired.

It remains to consider the case $|\ell - k_1| = |\ell| \leq C'$, which gives us the weaker bound stated in the lemma. We use the decomposition (4.56)–(4.59). The terms (4.56)–(4.58) give a bound of 2^k as before. The main difference lies with (4.59) which is nonzero only due to the contribution to opposing $(++)$ or $(--)$ waves, see Lemma 4.1. In fact, one has with $\ell = k_1 + O(1) = O(1)$,

$$\begin{aligned}
&\|P_k Q_{O(1)} \mathcal{Q}_{\alpha\beta}(Q_{<-C}\phi_1, Q_{<-C}\phi_2)\|_{L^2_t L^2_x} \\
&\lesssim \sum_{\pm} \sum_{\kappa\in\mathcal{C}_k} \|P_k Q_{O(1)} \mathcal{Q}_{\alpha\beta}(Q_{<-C}P_\kappa\phi_1^\pm, Q_{<-C}P_{-\kappa}\phi_2^\pm)\|_{L^2_t L^2_x} \\
&\lesssim \sum_{\pm} \sum_{\kappa\in\mathcal{C}_k} 2^{\frac{k}{2}} \|Q_{<-C}P_\kappa\phi_1^\pm\|_{S[k_1,\kappa]} \|Q_{<-C}P_{-\kappa}\phi_2^\pm\|_{S[k_2,-\kappa]} \\
&\lesssim 2^{\frac{k}{2}} \sum_{\pm} \Big(\sum_{\kappa\in\mathcal{C}_k} \|Q_{<-C}P_\kappa\phi_1^\pm\|^2_{S[k_1,\kappa]}\Big)^{\frac12} \Big(\sum_{\kappa\in\mathcal{C}_k} \|Q_{<-C}P_{-\kappa}\phi_2^\pm\|^2_{S[k_2,-\kappa]}\Big)^{\frac12} \\
&\lesssim 2^{\frac{k}{2}} k^2 \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned} \tag{4.63}$$

To pass to the last line, we wrote

$$\begin{aligned}
\sum_{\kappa\in\mathcal{C}_k} \|Q_{<-C}P_\kappa\phi_1^\pm\|^2_{S[k_1,\kappa]} &\lesssim \sum_{\kappa\in\mathcal{C}_k} \|Q_{<2k}P_\kappa\phi_1^\pm\|^2_{S[k_1,\kappa]} + \sum_{\kappa\in\mathcal{C}_k}\Big(\sum_{2k\le j\le -C} \|Q_j P_\kappa \phi_1^\pm\|_{S[k_1,\kappa]}\Big)^2 \\
&\lesssim \|\phi_1^\pm\|^2_{S[k_1]} + |k| \sum_{2k\le j\le -C} \sum_{\kappa\in\mathcal{C}_k} \|Q_j P_\kappa\phi_1^\pm\|^2_{\dot X_{k_1}^{0,\frac12,\infty}} \\
&\lesssim |k|^2 \|\phi_1^\pm\|^2_{S[k_1]}
\end{aligned} \tag{4.64}$$

and the result follows.
The second statement (4.55) follows by essentially the same proof. □

Remark 4.20. It is important to note that the logarithmic loss of $\langle k_1 - k\rangle^2$ in (4.55) only results from the case of opposing waves in the high-high case. Later we will use (4.55) without this loss in those cases where these interactions are excluded.

Later, we shall also require the following technical refinement of Lemma 4.19 dealing with a further angular restriction of the first input.

Corollary 4.21. *Under the assumptions of Lemma 4.19 and for any $m_0 \le -10$,*

$$\Big(\sum_{\kappa\in\mathcal{C}_{m_0}} \Big(\sum_{\ell\ge k+C} 2^{-\varepsilon\ell} \|P_k Q_\ell \mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa}\phi_1, \phi_2)\|_{L^2_t L^2_x}\Big)^2\Big)^{\frac12} \lesssim |m_0|\, 2^{\frac{k}{2}} 2^{-\varepsilon k_1} \langle k_1 - k\rangle^2 \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}$$

with an absolute implicit constant.

Proof. This can be seen by reviewing the proof[13] of Lemma 4.19. Specifically, up until (4.63), one places $P_{k_1,\kappa}\phi_1$ either in the $\dot X^{s,b}$ or $L^\infty_t L^2_x$ norms. The norms are amenable to square summation, in the latter case at the expense of a factor $|m_0|$, see Lemma 2.18. However, as far as (4.63) is concerned, we distinguish two cases: $k \le m_0$ and $k > m_0$. In the former case, the caps in $\mathcal{C}_k$ are smaller than those in $\mathcal{C}_{m_0}$ and (4.63) applies directly (one organizes the caps in $\mathcal{C}_k$ into subsets of the larger $\mathcal{C}_{m_0}$–caps). In the latter case, however, the $\mathcal{C}_{m_0}$–caps are smaller which forces us to write

$$Q_{<-C}\phi_1 = Q_{<2m_0}\phi_1 + Q_{2m_0<\cdot<-C}\phi_1$$

The former is subsumed in a square-function bound as in (4.63), whereas the latter leads to a loss of $|m_0|$ as in (4.64) and the corollary is proved. □

Next, we obtain an improvement in case of angular alignment of the inputs. This is analogous the case of low modulations, see Lemma 4.18.

[13]It is important to observe that one *cannot* square sum the bound of Lemma 4.19 directly due to the fact that $\sum_{\kappa\in\mathcal{C}_{m_0}} \|P_{k_1,\kappa}\phi_1\|^2_{S[k]}$ cannot be controlled.

Lemma 4.22. *Let $\delta > 0$ be small and $L > 1$ be large. Then there exists $m_0 = m_0(\delta, L) < 0$ large and negative such that for any ϕ_j adapted to k_j for $j = 1, 2$,*

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}} \sum_{\ell\ge k+C} 2^{-\varepsilon\ell}\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa_1}\phi_1, P_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2} \le \delta\, 2^{(\frac12-\varepsilon)k_1}\|\phi_1\|_{S[k_1]}\|\phi_2\|_{S[k_2]} \tag{4.65}$$

provided $\max_{j=1,2}|k-k_j| \le L$*. The constant C in* (4.65) *is an absolute constant which does not depend on L or δ.*

Proof. The proof consists of checking that one can glean a gain from angular alignment by following the proof of Lemma 4.19. In effect, this will always be done by means of Bernstein's inequality. The only case where this is not possible is (4.63), but that case is excluded by the angular alignment assumption.

We set $k_1 = 0$ whence $|k| \le L$ and $|k_2| \le 2L$. Implicit constants here will be allowed to depend on L, but not the constants C appearing in modulation cutoffs. As before, one has the decomposition

$$\|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa_1}\phi_1, P_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2} \lesssim \|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge\ell-C}P_{k_1,\kappa_1}\phi_1, Q_{\le k_1+C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2} \tag{4.66}$$
$$+ \|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{\ge\ell-C}P_{k_1,\kappa_1}\phi_1, Q_{>k_1+C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2} \tag{4.67}$$
$$+ \|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{<\ell-C}P_{k_1,\kappa_1}\phi_1, Q_{\ge\ell-C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2} \tag{4.68}$$
$$+ \|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_{<\ell-C}P_{k_1,\kappa_1}\phi_1, Q_{<\ell-C}P_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2} \tag{4.69}$$

We first consider the case $\ell \le k_1 - C'$ where C' is large but still smaller than the constant C in (4.66)–(4.69). Then the term in (4.69) vanishes by Lemma 4.1. By Bernstein's inequality,

$$\begin{aligned}\sum_{k+C\le\ell\le k_1-C} 2^{-\varepsilon\ell}(4.66) &\lesssim \sum_{k+C\le\ell\le k_1-C} \|P_kQ_\ell[\partial_\beta(Q_{\ge\ell-C}|\nabla|^{-1}P_{k_1,\kappa_1}\phi_1\cdot Q_{\le k_2+C}\partial_\alpha|\nabla|^{-1}P_{k_2,\kappa_2}\phi_2)\\
&\qquad - \partial_\alpha(Q_{\ge\ell-C}|\nabla|^{-1}P_{k_1,\kappa_1}\phi_1\cdot Q_{\le k_2+C}\partial_\beta|\nabla|^{-1}P_{k_2,\kappa_2}\phi_2)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k+C\le\ell\le k_1-C} \|Q_{\ge\ell-C}P_{k_1,\kappa_1}\phi_1\|_{L_t^2L_x^2}\|P_{k_2,\kappa_2}\phi_2\|_{L_t^\infty L_x^\infty}\\
&\lesssim 2^{\frac{m_0}{2}} \sum_{k+C\le\ell\le k_1-C} \|Q_{\ge\ell-C}P_{k_1,\kappa_1}\phi_1\|_{L_t^2L_x^2}\|P_{k_2,\kappa_2}\phi_2\|_{L_t^\infty L_x^2}\\
&\le \delta\|P_{k_1,\kappa_1}\phi_1\|_{\dot X_0^{0,\frac12,\infty}}\|P_{k_2,\kappa_2}\phi_2\|_{L_t^\infty L_x^2}\end{aligned}$$

Summing over κ_1, κ_2 now yields the desired bound by Cauchy-Schwarz (see also Lemma 2.18). The term (4.68) satisfies the same bound. Next, note that due to $\ell \le k_1 - C'$ it suffices to consider $\phi_1 = Q_{\ge k_1+C}\phi_1$ in (4.67). Consequently,

$$\begin{aligned}\sum_{k+C\le\ell\le k_1-C} 2^{-\varepsilon\ell}(4.67) &\lesssim \sum_{k+C\le\ell\le k_1-C}\sum_{m\ge k_1+C} \|P_kQ_\ell\mathcal{Q}_{\alpha\beta}(Q_mP_{k_1,\kappa_1}\phi_1, \tilde Q_mP_{k_2,\kappa_2}\phi_2)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k+C\le\ell\le k_1-C}\sum_{m\ge k_1+C} \|\mathcal{Q}_{\alpha\beta}(Q_mP_{k_1,\kappa_1}\phi_1, \tilde Q_mP_{k_2,\kappa_2}\phi_2)\|_{L_t^1L_x^2}\\
&\lesssim \sum_{k+C\le\ell\le k_1-C}\sum_{m\ge k_1+C} 2^m2^{-2m(1-\varepsilon)}\|P_{k_1,\kappa_1}Q_m\phi_1\|_{L_t^2L_x^2}\|P_{k_2,\kappa_2}\tilde Q_m\phi_2\|_{L_t^2L_x^\infty}\\
&\lesssim \delta\|P_{k_1,\kappa_1}\phi_1\|_{\dot X_0^{0,1-\varepsilon,2}}\|P_{k_2,\kappa_2}\phi_2\|_{\dot X_0^{0,1-\varepsilon,2}}\end{aligned}$$

Summing over κ_1, κ_2 again leads to the desired bound. Second, suppose that $\ell \geq k_1 + C'$. Then

$$\sum_{\ell \geq k_1 + C'} 2^{-\varepsilon\ell} \| P_k Q_\ell \mathcal{Q}_{\alpha\beta}(P_{k_1,\kappa_1}\phi_1, P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2}$$

$$\lesssim \sum_{\ell \geq k_1 + C'} 2^{-\varepsilon\ell} \| P_k Q_\ell \mathcal{Q}_{\alpha\beta}(\tilde{Q}_\ell P_{k_1,\kappa_1}\phi_1, Q_{\leq \ell-5} P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \tag{4.70}$$

$$+ \sum_{\ell \geq k_1 + C'} 2^{-\varepsilon\ell} \| P_k Q_\ell \mathcal{Q}_{\alpha\beta}(Q_{\leq \ell-5} P_{k_1,\kappa_1}\phi_1, \tilde{Q}_\ell P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \tag{4.71}$$

$$+ \sum_{\ell \geq k_1 + C'} 2^{-\varepsilon\ell} \sum_{m \geq \ell - 5} \| P_k Q_\ell \mathcal{Q}_{\alpha\beta}(Q_m P_{k_1,\kappa_1}\phi_1, \tilde{Q}_m P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \tag{4.72}$$

which are in turn estimated as follows:

$$\begin{aligned}
\sum_{\substack{\kappa_1,\kappa_2 \in \mathcal{C}_{m_0} \\ \mathrm{dist}(\kappa_1,\kappa_2) \leq 2^{m_0}}} (4.70) &\lesssim \sum_{\substack{\kappa_1,\kappa_2 \in \mathcal{C}_{m_0} \\ \mathrm{dist}(\kappa_1,\kappa_2) \leq 2^{m_0}}} \sum_{\ell \geq k_1 + C'} 2^{-\varepsilon\ell} \| P_k Q_\ell \mathcal{Q}_{\alpha\beta}(\tilde{Q}_\ell P_{k_1,\kappa_1}\phi_1, Q_{\leq \ell-5} P_{k_2,\kappa_2}\phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{\substack{\kappa_1,\kappa_2 \in \mathcal{C}_{m_0} \\ \mathrm{dist}(\kappa_1,\kappa_2) \leq 2^{m_0}}} \sum_{\ell \geq k_1 + C'} 2^{(1-\varepsilon)\ell} \| \tilde{Q}_\ell P_{k_1,\kappa_1}\phi_1\|_{L_t^2 L_x^2} \| P_{k_2,\kappa_2}\phi_2\|_{L_t^\infty L_x^2} \\
&\leq \delta \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}$$

and similarly for (4.71), whereas

$$\begin{aligned}
\sum_{\substack{\kappa_1,\kappa_2 \in \mathcal{C}_{m_0} \\ \mathrm{dist}(\kappa_1,\kappa_2) \leq 2^{m_0}}} (4.72) &\lesssim \sum_{\substack{\kappa_1,\kappa_2 \in \mathcal{C}_{m_0} \\ \mathrm{dist}(\kappa_1,\kappa_2) \leq 2^{m_0}}} \sum_{\ell \geq k_1 + C'} 2^{-\varepsilon\ell} \sum_{m \geq \ell - 5} 2^{\frac{\ell}{2}} \| \mathcal{Q}_{\alpha\beta}(Q_m P_{k_1,\kappa_1}\phi_1, \tilde{Q}_m P_{k_2,\kappa_2}\phi_2)\|_{L_t^1 L_x^2} \\
&\lesssim 2^{\frac{m_0}{2}} \sum_{\substack{\kappa_1,\kappa_2 \in \mathcal{C}_{m_0} \\ \mathrm{dist}(\kappa_1,\kappa_2) \leq 2^{m_0}}} \sum_{m+5 \geq \ell \geq k_1 + C'} 2^{(\frac{1}{2}-\varepsilon)\ell} 2^m 2^{-2m(1-\varepsilon)} \| P_{k_1,\kappa_1}\phi_1\|_{S[k_1]} \| P_{k_2,\kappa_2}\phi_2\|_{S[k_2]} \\
&\leq \delta \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}$$

as desired.
It the remaining case $|\ell - k_1| = |\ell| \leq C'$ we use the decomposition (4.66)–(4.69). The terms (4.66)–(4.68) give a bound of δ as before. The main difference lies with (4.69) which is nonzero only due to the contribution to opposing $(++)$ or $(--)$ waves, see Lemma 4.1. However, this case is excluded due to the angular alignment assumption. □

4.5. **Nullform bounds in the low-high and high-low cases.** We now derive analogues of the previous two lemmas in the high-low case, with the low-high case being completely analogous.

Lemma 4.23. *For any ϕ_j adapted to k_j with $k_2 \leq k_1 + O(1) = k$ one has*

$$\| P_k Q_{\leq k+C} \mathcal{Q}_{\alpha\beta}(\phi_1, \phi_2)\|_{L_t^2 L_x^2} \lesssim 2^{(\frac{1}{2}-\varepsilon)k_2} 2^{\varepsilon k_1} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}$$

Proof. We may take $k = k_1 + O(1) = 0$ and $k_2 \leq -C$. Assume first that $Q_{<k_2}\phi_i = \phi_i$ for $i = 1, 2$. Then the modulation of the output does not exceed 2^{k_2}, and we are reduced to bounding the following three expressions:

$$\sum_{j \leq k_2 + O(1)} \| P_0 Q_j (R_\alpha Q_{<j-C}\phi_1 R_\beta Q_{<j-C}\phi_2 - R_\beta Q_{<j-C}\phi_1 R_\alpha Q_{<j-C}\phi_2)\|_{L_t^2 L_x^2} \tag{4.73}$$

$$+ \sum_{j \leq k_2 + O(1)} \| P_0 Q_{<j+C} (R_\alpha Q_j\phi_1 R_\beta Q_{\leq j}\phi_2 - R_\beta Q_j\phi_1 R_\alpha Q_{\leq j}\phi_2)\|_{L_t^2 L_x^2} \tag{4.74}$$

$$+ \sum_{j \leq k_2 + O(1)} \| P_0 Q_{<j+C} (R_\alpha Q_{<j}\phi_1 R_\beta Q_j\phi_2 - R_\beta Q_{<j}\phi_1 R_\alpha Q_j\phi_2)\|_{L_t^2 L_x^2} \tag{4.75}$$

Each of the summands here is bounded by $2^{(j+k_2)/4}$. For the first, one decomposes into caps of size $2^{\frac{j-k_2}{2}}$:

$$
\begin{aligned}
(4.73) \lesssim & \sum_{j\le k_2+O(1)} \sum_{\kappa\sim\kappa'\in\mathcal{C}_{\frac{j-k_2}{2}}} \|P_0 Q_j(R_\alpha Q_{<j-C}P_\kappa\phi_1 R_\beta Q_{<j-C}P_{\kappa'}\phi_2 \\
& - R_\beta Q_{<j-C}P_\kappa\phi_1 R_\alpha Q_{<j-C}P_{\kappa'}\phi_2)\|_{L^2_tL^2_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j-k_2}{4}} 2^{\frac{k_2}{2}} \|Q_{<j-C}P_\kappa\phi_1\|_{S[k_1,\kappa]} \|Q_{<j-C}P_{\kappa'}\phi_2\|_{S[k_2,\kappa']} \\
\lesssim & \, 2^{\frac{k_2}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}
$$

where we applied (2.29) in the last step. Note that the nullform gains a factor of the angle in this bound. As for (4.74), one performs a similar cap decomposition but without the separation between the caps:

$$
\begin{aligned}
(4.74) \lesssim & \sum_{j\le k_2+O(1)} \sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{j-k_2}{2}}} \|P_0 Q_{<j+C}(R_\alpha Q_j P_\kappa\phi_1 R_\beta Q_{\le j}P_{\kappa'}\phi_2 \\
& - R_\beta Q_j P_\kappa\phi_1 R_\alpha Q_{\le j}P_{\kappa'}\phi_2)\|_{L^2_tL^2_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j-k_2}{2}} \|Q_jP_\kappa\phi_1\|_{L^2_tL^2_x} \|Q_{\le j}P_{\kappa'}\phi_2\|_{L^\infty_tL^\infty_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j-k_2}{2}} 2^{-\frac{j}{2}} 2^{\frac{j-k_2}{4}} 2^{k_2} \|Q_jP_\kappa\phi_1\|_{S[k_1,\kappa]} \|P_{\kappa'}\phi_2\|_{L^\infty_tL^2_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j+k_2}{4}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]} \lesssim 2^{\frac{k_2}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}
$$

Finally,

$$
\begin{aligned}
(4.75) \lesssim & \sum_{j\le k_2+O(1)} \sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{j-k_2}{2}}} \|P_0 Q_{<j+C}(R_\alpha Q_{<j} P_\kappa\phi_1 R_\beta Q_j P_{\kappa'}\phi_2 \\
& - R_\beta Q_{<j} P_\kappa\phi_1 R_\alpha Q_j P_{\kappa'}\phi_2)\|_{L^2_tL^2_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j-k_2}{2}} \|Q_{<j}P_\kappa\phi_1\|_{L^\infty_tL^2_x} \|Q_jP_{\kappa'}\phi_2\|_{L^2_tL^\infty_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j-k_2}{2}} 2^{\frac{j-k_2}{4}} 2^{k_2} \|Q_{<j}P_\kappa\phi_1\|_{S[k_1,\kappa]} \|Q_jP_{\kappa'}\phi_2\|_{L^2_tL^2_x} \\
\lesssim & \sum_{j\le k_2+O(1)} 2^{\frac{j+k_2}{4}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]} \lesssim 2^{\frac{k_2}{2}} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}
$$

If $Q_{k_2\le\cdot\le C}\phi_2 = \phi_2$, then we may take $\phi_1 = Q_{\le C}\phi_1$ whence

$$
\begin{aligned}
\|P_0 Q_{\le O(1)}\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_tL^2_x} & \lesssim \|\phi_1\|_{L^\infty_tL^2_x} \|R_0\phi_2\|_{L^2_tL^\infty_x} \\
& \lesssim \|\phi_1\|_{L^\infty_tL^2_x} \sum_{k_2\le j\le C} 2^j \|Q_j\phi_2\|_{L^2_tL^2_x} \\
& \lesssim 2^{(\frac12-\varepsilon)k_2} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}
$$

On the other hand, if $\phi_2 = Q_{\ge C}\phi_2$, then necessarily also $\phi_1 = Q_{\ge C}\phi_1$ so that

$$
\begin{aligned}
\|P_0 Q_{\le O(1)}\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^2_tL^2_x} & \lesssim \|P_0 Q_{\le O(1)}\mathcal{Q}_{\alpha\beta}(\phi_1,\phi_2)\|_{L^1_tL^1_x} \\
& \lesssim \sum_{m\ge C} \|Q_m\phi_1\|_{L^2_tL^2_x} 2^m \|\tilde{Q}_m\phi_2\|_{L^2_tL^2_x} \\
& \lesssim \|\phi_1\|_{S[k_1]} \sum_{m\ge C} 2^m 2^{-2m(1-\varepsilon)} 2^{k_2(\frac12-\varepsilon)} \|\phi_2\|_{S[k_2]} \\
& \lesssim 2^{(\frac12-\varepsilon)k_2} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}
$$

and the lemma is proved. □

Next, we deal with the case of outputs with large modulation.

Lemma 4.24. *For any ϕ_j adapted to k_j with $k_2 \le k_1 + O(1) = k$ one has*

$$\sum_{\ell \ge k+C} 2^{-\varepsilon\ell} \|P_k Q_\ell \mathcal{Q}_{\alpha\beta}(\phi_1, \phi_2)\|_{L_t^2 L_x^2} \lesssim 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}$$

Proof. Set $k = k_1 + O(1) = 0$ and $k_2 \le -C$. Then

$$\begin{aligned}
&\sum_{\ell \ge C} 2^{-\varepsilon\ell} \|P_0 Q_\ell \mathcal{Q}_{\alpha\beta}(\phi_1, \phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{\ell \ge C} 2^{-\varepsilon\ell} \|P_0 Q_\ell \mathcal{Q}_{\alpha\beta}(\tilde{Q}_\ell \phi_1, Q_{<\ell-C}\phi_2)\|_{L_t^2 L_x^2} \qquad (4.76)\\
&\quad + \sum_{\ell \ge C} 2^{-\varepsilon\ell} \|P_0 Q_\ell \mathcal{Q}_{\alpha\beta}(Q_{\ge \ell-C}\phi_1, Q_{\ge\ell-C}\phi_2)\|_{L_t^2 L_x^2} \qquad (4.77)
\end{aligned}$$

First, taking $\alpha = 0$ and $\beta = 1$,

$$\begin{aligned}
(4.76) &\lesssim \sum_{\ell \ge C} 2^{-\varepsilon\ell} \|P_0 Q_\ell (R_0 \tilde{Q}_\ell \phi_1 R_1 Q_{<\ell-C}\phi_2 - R_1 \tilde{Q}_\ell \phi_1 R_0 Q_{<\ell-C}\phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{\ell \ge C} 2^{-\varepsilon\ell} (2^\ell \|Q_\ell \phi_1\|_{L_t^2 L_x^2} \|\phi_2\|_{L_t^\infty L_x^\infty} + \|\tilde{Q}_\ell \phi_1\|_{L_t^2 L_x^2} \|R_0 Q_{<\ell-C}\phi_2\|_{L_t^\infty L_x^\infty}) \\
&\lesssim \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]} 2^{k_2} + \sum_{\ell \ge C} \|\tilde{Q}_\ell \phi_1\|_{L_t^2 L_x^2} 2^{(\frac{1}{2}+\varepsilon)\ell} 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_2\|_{S[k_2]} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}$$

To pass to the second to last line we used the estimate

$$\begin{aligned}
\|R_0 Q_{<\ell-C}\phi_2\|_{L_t^\infty L_x^\infty} &\lesssim 2^{k_2} \|Q_{\le k_2}\phi_2\|_{L_t^\infty L_x^2} + \sum_{k_2 < j < \ell - C} 2^{\frac{3j}{2}} \|Q_j \phi_2\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_2} \|Q_{\le k_2}\phi_2\|_{L_t^\infty L_x^2} + \sum_{k_2 < j < \ell - C} 2^{\frac{3j}{2}} 2^{-(1-\varepsilon)j} 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_2\|_{S[k_2]} \\
&\lesssim 2^{k_2} \|Q_{\le k_2}\phi_2\|_{L_t^\infty L_x^2} + \sum_{k_2 < j < \ell - C} 2^{(\frac{1}{2}+\varepsilon)j} 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_2\|_{S[k_2]} \\
&\lesssim 2^{k_2} \|Q_{\le k_2}\phi_2\|_{L_t^\infty L_x^2} + 2^{(\frac{1}{2}+\varepsilon)\ell} 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_2\|_{S[k_2]}
\end{aligned}$$

On (4.77) one has the bound (again for $\alpha = 0$ and $\beta = 1$)

$$\begin{aligned}
(4.77) &\lesssim \sum_{m \ge \ell \ge C} 2^{-\varepsilon\ell} \|P_0 Q_\ell (R_0 Q_m \phi_1 R_1 \tilde{Q}_m \phi_2 - R_1 Q_m \phi_1 R_0 \tilde{Q}_m \phi_2)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m \ge \ell \ge C} 2^{(\frac{1}{2}-\varepsilon)\ell} 2^m \|Q_m \phi_1\|_{L_t^2 L_x^2} \|\tilde{Q}_m \phi_2\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m \ge C} 2^{-(\frac{1}{2}-\varepsilon)m} \|\phi_1\|_{S[k_1]} \, 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_2\|_{S[k_2]} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)k_2} \|\phi_1\|_{S[k_1]} \|\phi_2\|_{S[k_2]}
\end{aligned}$$

as claimed. □

5. Trilinear estimates

The purpose of this section is to derive the estimates on the trilinear nonlinearities which govern the wave map system. To clarify the role that these estimates play, we include here a table which explains

TABLE 1. Relation of trilinear estimates to null-forms

Estimate	Trilinear expression
Lemma 5.1	$1, 2$
Corollary 5.2	$3, 7$
Lemma 5.3	4-6, 8-11
Corollary 5.4	$1, 3, 7$
Lemma 5.5	$2,\ 5, 6,$ 8-11
Remark 5.6	2
Lemma 5.7	4, 9
Lemma 5.9	$1, 3, 7$
Corollary 5.10	$2,\ 5, 6,$ 8-11
Lemma 5.11	$2,\ 5, 6,$ 8-11

how these estimates relate to the trilinear terms arising in section 3. Recall that the trilinear null-forms are summarized in the expression (3.14). In the table, we number the terms in this expression 1-11.

The gist of the estimates to follow is that whenever one is given a trilinear null-form of the schematic form

$$\nabla_{x,t} P_0 \big[\psi_1 \nabla^{-1}[\psi_2 \psi_3]\big]$$

with each ψ_i localized to spatial frequency $\sim 2^{k_i}$, $i = 1, 2, 3$, then one gains exponentially in the difference of the largest to the smallest logarithmic frequency present, unless one is in the situation where $k_1 = O(1)$ and $k_{2,3} \ll -1$, i. e. the high-low-low case. This latter feature forces us to modify the procedure of Bahouri-Gerard in section 9, and it also informs our choice of the weight $w(k_1, k_2, k_3)$ below.
In addition to the bilinear estimates of the previous sections, we will also heavily use the Strichartz component of the $S[k]$-norm, see (2.14). As already in [22], we will partially rely on Tao's trilinear estimate from [57] which states that (relative to our norms in the $S[k]$-spaces)

$$\|P_0[\psi_1\, R^\beta \psi_2\, R_\beta \psi_3]\|_{N[0]} \lesssim 2^{\sigma_1 (k_2 \wedge k_3 - k_1)\wedge 0} 2^{k_1} \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]} \tag{5.1}$$

for some $\sigma_1 > 0$. To obtain (5.1) from [57], one observes that $\|\nabla\psi\|_{S[k]}$ (strictly) dominates the $S[k]$-norms of [57], whereas $\|P_0 \cdot\|_{N[0]}$ is dominated by the respective $N[0]$-norm used in [57]. Because of this property, the trilinear bound from [57] can be adapted to this setting provided the correct scaling is taken into account. Moreover, throughout this section we define, with $k_{\max} := k_1 \vee k_2 \vee k_3$, $k_{\min} := k_1 \wedge k_2 \wedge k_3$, and k_{med} the median of k_1, k_2, k_3,

$$w(k_1, k_2, k_3) := \begin{cases} 2^{-\sigma_0 k_{\max}}\, 2^{\sigma_0 k_{\min} \wedge 0} & \text{if } \ k_{\max} \geq C \\ 2^{-\sigma_0 (k_{\text{med}} - k_{\min})} & \text{if } \ k_1 = k_{\max} = O(1) \\ 2^{\sigma_0 (k_1 + k_2 \wedge k_3)} & \text{if } \ k_1 < k_{\max} = O(1) \end{cases}$$

where $\sigma_0 > 0$ is some fixed small constant.
We split our argument into two cases, depending on whether all inputs are "hyperbolic" or not. This distinction is based on modulation vs. frequency.

5.1. **Reduction to the hyperbolic case.** The following lemma deals with the case where at least one of the inputs or the interior null-form have "elliptic" character. Recall that $I := \sum_{k\in\mathbb{Z}} P_k Q_{\leq k+C}$ and $I^c := 1 - I$ (here C is an absolute constant, $C = 10$ will suffice). Throughout this section, we will write $\tilde{P}_k$ to denote a projection $\sum_{k'=k+O(1)} P_{k'}$, and similarly with $\tilde{Q}_k$.

Lemma 5.1. *Let ψ_i be Schwarz functions adapted to k_i for $i = 0, 1, 2$. Then for any $\alpha = 0, 1, 2$, and $j = 1, 2$,*

$$\|P_0 \partial^\beta A_0[A_1 R_\alpha \psi_1 \Delta^{-1} \partial_j \tilde{A}_1 \mathcal{Q}_{\beta j}(A_2 \psi_2, A_3 \psi_3)]\|_{N[0]} \lesssim w(k_1, k_2, k_3) \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]}$$

where A_i and $\tilde{A}_1$ are either I or I^c, with at least one being I^c. Moreover, we impose the condition that $A_1 = \tilde{A}_1 = I^c$ implies $\alpha \neq 0$.

Proof. Case 1: $0 \le k_1 \le k_2 + O(1) = k_3 + O(1)$. We begin with $A_0 = I^c$ and $A_1 = I$. Then we can drop IR_α from ψ_1 and estimate[14]

$$\|P_0 Q_{\geq 0}\partial^\beta[\psi_1 \Delta^{-1}\partial_j \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \lesssim \|P_0 Q_{\geq 0}\partial^\beta[\psi_1 \Delta^{-1}\partial_j Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.2}$$

$$+ \|P_0 Q_{\geq 0}\partial^\beta[\psi_1 \Delta^{-1}\partial_j Q_{\geq k_1+C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.3}$$

By Lemma 4.17, placing (5.2) into $\dot{X}_0^{0,-1-\varepsilon,2}$ implies

$$\begin{aligned}\|\psi_1 \Delta^{-1}\partial_j Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^1} &\lesssim \|\psi_1\|_{L_t^\infty L_x^2} 2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \\ &\lesssim 2^{-\frac{k_2}{2}}\|\psi_1\|_{S[k_1]}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\end{aligned}$$

whereas

$$(5.3) \lesssim \sum_{m\geq k_1+C}\|P_0 Q_m \partial^\beta[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.4}$$

$$+ \sum_{m\geq k_1+C}\|P_0 Q_{\geq 0}\partial^\beta[Q_{\geq m-C}\psi_1 \Delta^{-1}\partial_j Q_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.5}$$

Lemma 4.19 yields the following bound on (5.4):

$$\begin{aligned}&\sum_{m\geq k_1+C}\|P_0 Q_m \partial^\beta[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\ &\lesssim \sum_{m\geq k_1+C}\|P_0 Q_m \partial^\beta[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\ &\lesssim \sum_{m\geq k_1+C} 2^{-m\varepsilon}\|\psi_1\|_{L_t^\infty L_x^2} 2^{-k_1}\|\tilde{P}_{k_1}\tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2 L_x^2} \\ &\lesssim 2^{-\varepsilon k_2} 2^{-\frac{k_1}{2}}\langle k_2-k_1\rangle^2 \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}\end{aligned}$$

The bound on (5.5) proceeds similarly:

$$\begin{aligned}(5.5) &\lesssim \sum_{m\geq k_1+C}\|P_0 Q_{\geq 0}\partial^\beta[Q_{>m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\ &\lesssim \sum_{m\geq k_1+C}\sum_{0\leq \ell\leq m+C}\|P_0 Q_\ell \partial^\beta[Q_{>m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\ &\quad + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C}\|P_0 Q_\ell \partial^\beta[\tilde{Q}_\ell \psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\ &\lesssim \sum_{m\geq k_1+C}\sum_{0\leq \ell\leq m+C} 2^{(\frac12-\varepsilon)\ell}\|Q_{>m-C}\psi_1\|_{L_t^2 L_x^2}\|\tilde{P}_{k_1}\Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\ &\quad + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} 2^{-\varepsilon\ell}\|\tilde{Q}_\ell \psi_1\|_{L_t^2 L_x^2}\|\tilde{P}_{k_1}\Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^2}\end{aligned}$$

In the second to last line we applied Bernstein's inequality in the time variable to switch from L_t^2 to L_t^1. We now replace the L_t^∞ on the right-hand side of the last line by an L_t^2 at the expense of a factor of $2^{\frac{m}{2}}$.

[14]It is convenient to prove the somewhat stronger bound with $A_0 = Q_{\geq 0}$ here.

Together with Lemma 4.19 this yields

$$
\begin{aligned}
(5.5) \lesssim & \sum_{m\geq k_1+C}\sum_{0\leq \ell\leq m+C} 2^{-k_1+(\frac12-\varepsilon)\ell}\|Q_{>m-C}\psi_1\|_{L_t^2L_x^2}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
& + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} 2^{-k_1-\varepsilon\ell}2^{\frac{m}{2}}\|\tilde{Q}_\ell\psi_1\|_{L_t^2L_x^2}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
\lesssim & \sum_{m\geq k_1+C} 2^{-(\frac12+\varepsilon)k_1}2^{-\frac12 m}\|\psi_1\|_{S[k_1]}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
& + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} 2^{-k_1-\varepsilon\ell}2^{\frac{m}{2}}2^{-(1-\varepsilon)\ell}2^{(\frac12-\varepsilon)k_1}\|\psi_1\|_{S[k_1]}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
\lesssim & \, 2^{-\frac{k_1}{2}}2^{-\varepsilon k_2}\langle k_2-k_1\rangle^2\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Next, we consider the case where both $A_0 = I^c$ and $A_1 = I^c$. If $\alpha \neq 0$, then one can drop R_α altogether so that the previous analysis applies. Otherwise, if $\alpha = 0$, then by assumption $\tilde{A}_1 = I$ and

$$
\begin{aligned}
& \|P_0Q_{\geq 0}\partial^\beta[Q_{\geq k_1+C}R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{N[0]}\\
(5.6)\qquad & \leq \sum_{m\geq k_1+10C}\|P_0Q_m\partial^\beta[\tilde{Q}_mR_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{N[0]}\\
(5.7)\qquad & + \|P_0Q_{0\leq\cdot\leq k_1+10C}\partial^\beta[Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{N[0]}
\end{aligned}
$$

By Lemma 4.17, (5.7) is bounded by

$$
\begin{aligned}
& \|P_0Q_{0\leq\cdot\leq k_1+10C}\partial^\beta[Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}}\\
\lesssim & \|Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\ \Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)\|_{L_t^2L_x^1}\\
\lesssim & \|Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\|_{L_t^\infty L_x^2}\,\|\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)\|_{L_t^2L_x^2}\\
\lesssim & \, 2^{-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

On the other hand, (5.6) is estimated as follows:

$$
\begin{aligned}
& \sum_{m\geq k_1+10C}\|P_0Q_m\partial^\beta[\tilde{Q}_mR_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}}\\
\lesssim & \sum_{m\geq k_1} 2^{-m\varepsilon}\|\tilde{Q}_mR_\alpha\psi_1\|_{L_t^2L_x^2}\|\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{L_t^\infty L_x^2}\\
\lesssim & \, 2^{-(\frac12+\varepsilon)k_1}\|\psi_1\|_{S[k_1]}\,2^{-\frac{k_1}{2}}\,\|\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{L_t^2L_x^2}\\
\lesssim & \, 2^{-\varepsilon k_1-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

where we applied Bernstein's inequality relative to t as well as Lemma 4.17 to pass to the last line.

Now suppose $A_0 = I$ (in fact, $A_0 = Q_{\leq 0}$), but at least one of A_1 or $\tilde{A}_1$ equals I^c. But then the modulations of ψ_1 and $\mathcal{Q}_{\beta j}$ essentially agree, whence $\alpha \neq 0$ and

$$\begin{aligned}
&\sum_{m \geq k_1 + C} \|P_0 Q_{\leq 0} \partial^\beta [Q_m R_\alpha \psi_1 \Delta^{-1} \partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m \geq k_1 + C} \|P_0 Q_{\leq 0} \partial^\beta [Q_m R_\alpha \psi_1 \Delta^{-1} \partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^1_{tx}} \\
&\lesssim \sum_{m \geq k_1 + C} \|Q_m \psi_1\|_{L^2_t L^2_x} 2^{-k_1} \|\tilde{P}_{k_1} \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)\|_{L^2_t L^2_x} \\
&\lesssim \sum_{m \geq k_1 + C} 2^{(\frac{1}{2} - \varepsilon) k_1} 2^{-m(1-2\varepsilon)} \|\psi_1\|_{S[k_1]} 2^{-k_1} 2^{-m\varepsilon} \|\tilde{P}_{k_1} \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)\|_{L^2_t L^2_x} \\
&\lesssim 2^{-\frac{k_1}{2} - \varepsilon k_2} \langle k_2 - k_1 \rangle^2 \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

The final estimate here uses Lemma 4.24. The last case which we need to consider is $A_0 = A_1 = \tilde{A}_1 = I$ and either one of A_2, A_3 equal to I^c. But then necessarily $A_2 = A_3 = I^c$ whence

$$\begin{aligned}
&\|P_0 \partial^\beta I [I R_\alpha \psi_1 \Delta^{-1} \partial_j I \mathcal{Q}_{\beta j}(Q_{\geq k_2 + C} \psi_2, Q_{\geq k_2 + C} \psi_3)]\|_{N[0]} \\
&\lesssim \|I R_\alpha \psi_1 \Delta^{-1} \partial_j I \mathcal{Q}_{\beta j}(Q_{\geq k_2 + C} \psi_2, Q_{\geq k_2 + C} \psi_3)\|_{L^1_t L^1_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{m \geq k_2 + C} \|\tilde{P}_{k_1} Q_{\leq k_1 + C} \mathcal{Q}_{\beta j}(Q_m \psi_2, \tilde{Q}_m \psi_3)]\|_{L^1_t L^1_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{m \geq k_2 + C} 2^{m - k_2} 2^{-2m(1-\varepsilon)} 2^{(1-2\varepsilon)k_2} \|\psi_2\|_{S[k_2]} \|\psi_3\|_{S[k_3]} \lesssim 2^{-k_2} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which concludes Case 1.

Case 2: $0 \leq k_1 = k_2 + O(1), k_3 \leq k_2 - C$. We again begin with $A_0 = I^c$, $A_1 = I$ and the representation (5.2) and (5.3) (dropping IR_α from ψ_1 as before). By Lemma 4.23, (5.2) is bounded by

$$\|\psi_1 \Delta^{-1} \partial_j Q_{\leq k_1 + C} \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^2_t L^1_x} \lesssim \|\psi_1\|_{L^\infty_t L^2_x} 2^{(\frac{1}{2} - \varepsilon) k_3} 2^{-(1-\varepsilon) k_1} \|\psi_2\|_{S[k_2]} \|\psi_3\|_{S[k_3]}$$

whereas

$$(5.3) \lesssim \sum_{m \geq k_1 + C} \|P_0 Q_m \partial^\beta [Q_{\leq m - C} \psi_1 \Delta^{-1} \partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \tag{5.8}$$

$$+ \sum_{m \geq k_1 + C} \|P_0 Q_{\geq 0} \partial^\beta [Q_{\geq m - C} \psi_1 \Delta^{-1} \partial_j Q_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \tag{5.9}$$

Lemma 4.24 yields the following bound on (5.8):

$$\begin{aligned}
&\sum_{k_1 + C \leq m} \|P_0 Q_m \partial^\beta [Q_{\leq m - C} \psi_1 \Delta^{-1} \partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \\
&\lesssim \sum_{k_1 + C \leq m} \|P_0 Q_m \partial^\beta [Q_{\leq m - C} \psi_1 \Delta^{-1} \partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{\dot{X}_0^{0, -1-\varepsilon, 2}} \\
&\lesssim \sum_{k_1 + C \leq m} 2^{-m\varepsilon} \|\psi_1\|_{L^\infty_t L^2_x} 2^{-k_1} \|\tilde{P}_{k_1} \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2, \psi_3)\|_{L^2_t L^2_x} \\
&\lesssim 2^{(\frac{1}{2} - \varepsilon) k_3} 2^{-k_1} \prod_{i=1}^3 \|\psi\|_{S[k_i]}
\end{aligned}$$

The bound on (5.9) proceeds similarly:

$$
\begin{aligned}
(5.9) &\lesssim \sum_{m\geq k_1+C} \|P_0 Q_{\geq 0}\partial^\beta[Q_{>m-C}\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\geq k_1+C}\sum_{0\leq \ell\leq m+C} \|P_0 Q_\ell\partial^\beta[Q_{>m-C}\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\quad + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} \|P_0 Q_\ell\partial^\beta[\tilde{Q}_\ell\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \sum_{m\geq k_1+C}\sum_{0\leq \ell\leq m+C} 2^{(\frac{1}{2}-\varepsilon)\ell}\|Q_{>m-C}\psi_1\|_{L_t^2L_x^2}\|\tilde{P}_{k_1}\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2} \quad (5.10)\\
&\quad + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} 2^{-\varepsilon\ell}\|\tilde{Q}_\ell\psi_1\|_{L_t^2L_x^2}\|\tilde{P}_{k_1}\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^2} \quad (5.11)
\end{aligned}
$$

To pass to (5.10) we used Bernstein's inequality to switch from L_t^2 to L_t^1, which costs $2^{\frac{\ell}{2}}$. We now replace the L_t^∞ on the right-hand side of the last line by an L_t^2 at the expense of a factor of $2^{\frac{m}{2}}$. In view of Lemma 4.24 one concludes that

$$
\begin{aligned}
(5.9) &\lesssim \sum_{m\geq k_1+C}\sum_{0\leq \ell\leq m+C} 2^{-k_1+(\frac{1}{2}-\varepsilon)\ell}\|Q_{>m-C}\psi_1\|_{L_t^2L_x^2}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2} \\
&\quad + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} 2^{-k_1-\varepsilon\ell}2^{\frac{m}{2}}\|\tilde{Q}_\ell\psi_1\|_{L_t^2L_x^2}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2} \\
&\lesssim \sum_{m\geq k_1+C} 2^{-(\frac{1}{2}+\varepsilon)k_1}2^{-\frac{1}{2}m}\|\psi_1\|_{S[k_1]}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2} \\
&\quad + \sum_{m\geq k_1+C}\sum_{\ell\geq m+C} 2^{-k_1-\varepsilon\ell}2^{\frac{m}{2}}2^{-(1-\varepsilon)\ell}2^{(\frac{1}{2}-\varepsilon)k_1}\|\psi_1\|_{S[k_1]}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2} \\
&\lesssim 2^{-k_1}2^{(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Next, we consider the case where both $A_0 = I^c$ and $A_1 = I^c$. If $\alpha \neq 0$, then one can drop R_α altogether so that the previous analysis applies. Otherwise, if $\alpha = 0$, then by assumption $\tilde{A}_1 = I$ and as in Case 1 one obtains (5.6) and (5.7). By Lemma 4.23, (5.7) is bounded by

$$
\begin{aligned}
&\|P_0 Q_{0\leq\cdot\leq k_1+10C}\partial^\beta[Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \|Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\ \Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)\|_{L_t^2L_x^1} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)k_1}\|Q_{0\leq\cdot\leq k_1+10C}R_\alpha\psi_1\|_{L_t^2L_x^2}\ \|\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)\|_{L_t^2L_x^2} \\
&\lesssim 2^{-\frac{k_1}{2}}2^{(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

On the other hand, (5.6) is estimated as follows:

$$
\begin{aligned}
&\sum_{m\geq k_1+10C}\|P_0Q_m\partial^\beta[\tilde{Q}_mR_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \sum_{m\geq k_1} 2^{-m\varepsilon}\|\tilde{Q}_m\nabla_{t,x}|\nabla|^{-1}\psi_1\|_{L_t^2L_x^2}\ 2^{-k_1}\|\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{L_t^\infty L_x^2} \\
&\lesssim 2^{-(\frac{1}{2}+\varepsilon)k_1}\|\psi_1\|_{S[k_1]}\,2^{-\frac{k_1}{2}}\|\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2,A_3\psi_3)]\|_{L_t^2L_x^2} \\
&\lesssim 2^{-k_1+(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

where we applied Bernstein's inequality relative to t as well as Lemma 4.23.

We now turn to the case where $A_0 = I$, but at least one of A_1 or $\tilde{A}_1$ equals I^c. But then the modulations of ψ_1 and $\mathcal{Q}_{\beta j}$ essentially agree whence $\alpha \neq 0$. Bounding $N[0]$ by $L_t^1 L_x^2$ and invoking Lemma 4.24 yields

$$
\begin{aligned}
&\sum_{m\geq k_1+C} \|P_0 Q_{\leq 0}\partial^\beta[Q_m\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{m\geq k_1+C} \|Q_m\psi_1\|_{L_t^2L_x^2}\, 2^{-k_1}\|\tilde{P}_{k_1}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim 2^{-(\frac{3}{2}-\varepsilon)k_1+(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \lesssim 2^{-k_1}2^{(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

The last case which we need to consider is $A_0 = A_1 = \tilde{A}_1 = I$ and either one of A_2, A_3 equal to I^c. We begin with $A_2 = I^c$. But then necessarily $A_2 = A_3 = I^c$ whence

$$
\begin{aligned}
&\|P_0\partial^\beta I[I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(Q_{\geq k_2+C}\psi_2, Q_{\geq k_2+C}\psi_3)]\|_{N[0]}\\
&\lesssim \|I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(Q_{\geq k_2+C}\psi_2, Q_{\geq k_2+C}\psi_3)\|_{L_t^1L_x^1}\\
&\lesssim \|\psi_1\|_{L_t^\infty L_x^2}2^{-k_1}\sum_{m\geq k_2+C}\|\tilde{P}_{k_1}Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(Q_m\psi_2,\tilde{Q}_m\psi_3)]\|_{L_t^1L_x^2}\\
&\lesssim \|\psi_1\|_{L_t^\infty L_x^2}\sum_{m\geq k_2+C}2^{m-k_2}2^{-2(1-\varepsilon)m}2^{(\frac{1}{2}-\varepsilon)k_2}2^{(\frac{1}{2}-\varepsilon)k_3}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{-(\frac{3}{2}-\varepsilon)k_1+(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \lesssim 2^{-k_1}2^{(\frac{1}{2}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

It remains to consider the case $A_2 = I$ and $A_3 = I^c$. We begin by reducing the modulation of the entire output. Indeed, by Lemma 4.23,

$$
\begin{aligned}
&\|P_0\partial^\beta Q_{(1-3\varepsilon)k_3\leq\cdot\leq C}[Q_{\leq k_1+C}\psi_1\Delta^{-1}\partial_j Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{N[0]}\\
&\lesssim 2^{-\frac{1}{2}(1-3\varepsilon)k_3}\|\psi_1\|_{L_t^\infty L_x^2}\,2^{-k_1}\|I\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{-k_1}2^{-\frac{1}{2}(1-3\varepsilon)k_3}\|\psi_1\|_{L_t^\infty L_x^2}\,2^{(\frac{1}{2}-\varepsilon)k_3}2^{\varepsilon k_2}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{-(1-\varepsilon)k_1}2^{\frac{\varepsilon}{2}k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Next, we reduce the modulation of ψ_1:

$$
\begin{aligned}
&\|P_0\partial^\beta Q_{\leq(1-3\varepsilon)k_3}[Q_{\geq(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{N[0]}\\
&\lesssim \|P_0\partial^\beta Q_{\leq(1-3\varepsilon)k_3}[Q_{\geq(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\leq k_1+C}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{L_t^1L_x^1}\\
&\lesssim \|Q_{\geq(1-3\varepsilon)k_3-k_1}\psi_1\|_{L_t^2L_x^2}\,2^{-k_1}\|I\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{-\frac{k_1}{2}}2^{-\frac{1}{2}(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{(\frac{1}{2}-\varepsilon)k_3}2^{\varepsilon k_2}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{-(\frac{1}{2}-\varepsilon)k_1}2^{\frac{\varepsilon}{2}k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Finally, we reduce the modulation of the interior null-form using Lemma 4.13:

$$
\begin{aligned}
&\|P_0\partial^\beta Q_{\leq(1-3\varepsilon)k_3}[Q_{\leq(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{k_3\leq\cdot\leq k_1+C}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{N[0]}\\
&\lesssim \|\psi_1\|_{S[k_1]}\sum_{k_3\leq\ell\leq k_1+C}2^{-\frac{\ell}{4}}2^{-k_1}\langle k_1\rangle\|P_{k_1}Q_\ell\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{-(1-2\varepsilon)k_1}2^{(\frac{1}{4}-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is again admissible. After these preparations, we are faced with the following decomposition:

$$\begin{aligned}&P_0\partial^\beta Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\\&= P_0\partial^\beta Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(Q_{k_3\le\cdot\le k_2+C}\psi_2, Q_{k_3+C\le\cdot\le k_2+C}\psi_3)]\\&= \sum_{\kappa,\kappa'\in\mathcal{C}_\ell} P_{0,\kappa}\partial^\beta Q_{\le(1-3\varepsilon)k_3}[P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(Q_{k_3\le\cdot\le k_2+C}\psi_2, Q_{k_3+C\le\cdot\le k_2+C}\psi_3)]\end{aligned}$$

where $\ell = \frac{1}{2}(1-3\varepsilon)k_3$ and $\mathrm{dist}(\kappa,\kappa')\lesssim 2^\ell$. Placing the entire expression in $L^1_t L^2_x$ and using Bernstein's inequality results in the following estimate: with $J := Q_{k_3\le\cdot\le k_2+C}$,

$$\begin{aligned}&\|P_0\partial^\beta Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{L^1_tL^2_x}\\&\le \Big\|\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_\ell}\|P_{0,\kappa}[P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2, J\psi_3)]\|^2_{L^2_x}\Big)^{\frac12}\Big\|_{L^1_t}\\&\le 2^{\frac{\ell}{2}}\Big\|\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_\ell}\|P_{0,\kappa}[P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2, J\psi_3)]\|^2_{L^1_x}\Big)^{\frac12}\Big\|_{L^1_t}\\&\le 2^{\frac{\ell}{2}}\Big\|\Big(\sum_{\kappa'\in\mathcal{C}_\ell}\|P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\|^2_{L^2_x}\|\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2, J\psi_3)]\|^2_{L^2_x}\Big)^{\frac12}\Big\|_{L^1_t}\\&\le 2^{\frac{\ell}{2}}\|Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\|_{L^\infty_tL^2_x}\|\Delta^{-1}\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2, J\psi_3)]\|_{L^1_tL^2_x}\\&\lesssim 2^{\frac14(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{-k_1}\|\nabla_{t,x}|\nabla|^{-1}J\psi_2\|_{L^2_tL^2_x}\|\nabla_{t,x}|\nabla|^{-1}J\psi_3\|_{L^2_tL^\infty_x}\\&\lesssim 2^{\frac14(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{-k_1}2^{-\frac{k_3}{2}}\|\psi_2\|_{S[k_2]}\,2^{(\frac12-\varepsilon)k_3}2^{\varepsilon k_2}\|\psi_3\|_{S[k_3]}\end{aligned}$$

which is again admissible for small $\varepsilon>0$.

Case 3: $0\le k_1 = k_3+O(1), k_2\le k_3-C$. This case is symmetric to the previous one.

Case 4: $O(1)\le k_2 = k_3+O(1), k_1\le -C$. This case proceeds similarly to Case 1. We again begin with $A_0 = I^c$ and $A_1 = I$. Then we can drop IR_α from ψ_1 and estimate

$$\|P_0Q_{\ge0}\partial^\beta[\psi_1\Delta^{-1}\partial_j\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \lesssim \|P_0Q_{\ge0}\partial^\beta[\psi_1\Delta^{-1}\partial_j\widetilde{P}_0Q_{<C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.12}$$

$$+\|P_0Q_{\ge0}\partial^\beta[\psi_1\Delta^{-1}\partial_j\widetilde{P}_0Q_{\ge C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.13}$$

where we write $\widetilde{P}_0 = P_{[-C,C]}$ for simplicity. By Lemma 4.17, placing (5.12) into $\dot X_0^{0,-1-\varepsilon,2}$ implies

$$\begin{aligned}\|\psi_1\Delta^{-1}\partial_j\widetilde{P}_0Q_{<C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_tL^2_x} &\lesssim \|\psi_1\|_{L^\infty_tL^\infty_x}\,2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\&\lesssim 2^{k_1}2^{-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

whereas

$$(5.13)\lesssim\sum_{m\ge C}\|P_0Q_m\partial^\beta[Q_{\le m-C}\psi_1\Delta^{-1}\partial_j\widetilde{P}_0\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.14}$$

$$+\sum_{m\ge C}\|P_0Q_{\ge0}\partial^\beta[Q_{\ge m-C}\psi_1\Delta^{-1}\partial_j\widetilde{P}_0Q_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \tag{5.15}$$

Lemma 4.19 yields the following bound on (5.14):

$$\begin{aligned}
&\sum_{m\geq C} \|P_0 Q_m \partial^\beta [Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\geq C} \|P_0 Q_m \partial^\beta [Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \sum_{m\geq C} 2^{-m\varepsilon} \|\psi_1\|_{L_t^\infty L_x^\infty} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_1} 2^{-\varepsilon k_2} \langle k_2 - k_1\rangle^2 \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible. The bound on (5.15) proceeds similarly:

$$\begin{aligned}
(5.15) &\lesssim \sum_{m\geq C} \|P_0 Q_{\geq 0} \partial^\beta [Q_{>m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\geq C} \sum_{0\leq \ell\leq m+C} \|P_0 Q_\ell \partial^\beta [Q_{>m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\quad + \sum_{m\geq C} \sum_{\ell\geq m+C} \|P_0 Q_\ell \partial^\beta [\tilde{Q}_\ell\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \sum_{m\geq C} \sum_{0\leq \ell\leq m+C} 2^{(\frac12-\varepsilon)\ell} \|Q_{>m-C}\psi_1\|_{L_t^2 L_x^\infty} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\quad + \sum_{m\geq C} \sum_{\ell\geq m+C} 2^{-\varepsilon\ell} \|\tilde{Q}_\ell\psi_1\|_{L_t^2 L_x^\infty} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^2}
\end{aligned}$$

In the second to last line we applied Bernstein's inequality in the time variable to switch from L_t^2 to L_t^1. We now replace the L_t^∞ on the right-hand side of the last line by an L_t^2 at the expense of a factor of $2^{\frac{m}{2}}$. Together with Lemma 4.19 this yields

$$\begin{aligned}
(5.15) &\lesssim \sum_{m\geq C} \sum_{0\leq \ell\leq m+C} 2^{(\frac12-\varepsilon)\ell} 2^{k_1} \|Q_{>m-C}\psi_1\|_{L_t^2 L_x^2} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\quad + \sum_{m\geq C} \sum_{\ell\geq m+C} 2^{-\varepsilon\ell} 2^{k_1} 2^{\frac{m}{2}} \|\tilde{Q}_\ell\psi_1\|_{L_t^2 L_x^2} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_1} \sum_{m\geq C} 2^{-\frac12 m} \|\psi_1\|_{S[k_1]} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\quad + 2^{k_1} \sum_{m\geq C} \sum_{\ell\geq m+C} 2^{-\varepsilon\ell} 2^{\frac{m}{2}} 2^{-(1-\varepsilon)\ell} \|\psi_1\|_{S[k_1]} \|\widetilde{P}_0 \tilde{Q}_m \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{(\frac32-\varepsilon)k_1} 2^{-\varepsilon k_2} \langle k_2\rangle^2 \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible. Next, we consider the case where both $A_0 = I^c$ and $A_1 = I^c$. If $\alpha \neq 0$, then one can drop R_α altogether so that the previous analysis applies. Otherwise, if $\alpha = 0$, then by assumption $\tilde{A}_1 = I$ and

$$\|P_0 Q_{\geq 0} \partial^\beta [Q_{\geq C} R_\alpha \psi_1 \Delta^{-1}\partial_j \widetilde{P}_0 Q_{\leq C} \mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{N[0]}$$

$$\leq \sum_{m\geq 10C} \|P_0 Q_m \partial^\beta [\tilde{Q}_m R_\alpha \psi_1 \Delta^{-1}\partial_j \widetilde{P}_0 Q_{\leq C} \mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{N[0]} \tag{5.16}$$

$$+ \|P_0 Q_{0\leq\cdot\leq 10C} \partial^\beta [Q_{0\leq\cdot\leq 10C} R_\alpha \psi_1 \Delta^{-1}\partial_j \widetilde{P}_0 Q_{\leq C} \mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{N[0]} \tag{5.17}$$

By Lemma 4.17, (5.17) is bounded by

$$\begin{aligned}
&\|P_0 Q_{0\le\cdot\le 10C}\partial^\beta[Q_{0\le\cdot\le 10C}R_\alpha\psi_1\Delta^{-1}\partial_j\widetilde{P}_0 Q_{\le C}\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \|Q_{0\le\cdot\le 10C}R_\alpha\psi_1\ \Delta^{-1}\partial_j\widetilde{P}_0 Q_{\le C}\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)\|_{L_t^2 L_x^2} \\
&\lesssim \|Q_{0\le\cdot\le k_1+10C}R_\alpha\psi_1\|_{L_t^\infty L_x^\infty}\ \|\Delta^{-1}\partial_j\widetilde{P}_0 Q_{\le C}\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_1}2^{-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

On the other hand, (5.16) is estimated as follows:

$$\begin{aligned}
&\sum_{m\ge 10C}\|P_0 Q_m\partial^\beta[\tilde{Q}_m R_\alpha\psi_1\Delta^{-1}\partial_j\widetilde{P}_0 Q_{\le C}\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{\dot{X}_0^{0,-1-\varepsilon,2}} \\
&\lesssim \sum_{m\ge 0}2^{-m\varepsilon}\|\tilde{Q}_m\nabla_{t,x}|\nabla|^{-1}\psi_1\|_{L_t^2 L_x^\infty}\|\widetilde{P}_0 Q_{\le k_1+C}\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{L_t^\infty L_x^2} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)k_1}\|\psi_1\|_{S[k_1]}\|\widetilde{P}_0 Q_{\le C}\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)k_1-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

where we applied Bernstein's inequality relative to t as well as Lemma 4.17. Now suppose $A_0 = I$ (in fact, $A_0 = Q_{\le 0}$), but at least one of A_1 or $\tilde{A}_1$ equals I^c. If $\tilde{A}_1 = I^c$, then the modulations of ψ_1 and $\mathcal{Q}_{\beta j}$ essentially agree, whence $\alpha \ne 0$ and

$$\begin{aligned}
&\sum_{m\ge C}\|P_0 Q_{\le 0}\partial^\beta[Q_m R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\ge k_1+C}\|P_0 Q_{\le 0}\partial^\beta[Q_m R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^1 L_x^2} \\
&\lesssim \sum_{m\ge C}\|Q_m\psi_1\|_{L_t^2 L_x^\infty}\|\widetilde{P}_0\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m\ge C}2^{(\frac{3}{2}-\varepsilon)k_1}2^{-m(1-2\varepsilon)}\|\psi_1\|_{S[k_1]}\,2^{-m\varepsilon}\|\widetilde{P}_0\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2 L_x^2} \\
&\lesssim 2^{(\frac{3}{2}-\varepsilon)k_1-\varepsilon k_2}\langle k_2\rangle^2\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

The final estimate here uses Lemma 4.24. Now suppose that $\tilde{A}_1 = I$ and $A_1 = I^c$. Then

$$\begin{aligned}
&\|P_0 Q_{\le 0}\partial^\beta[I^c R_\alpha\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \|I^c R_\alpha\psi_1\Delta^{-1}\partial_j\ \widetilde{P}_0 I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^1 L_x^2} \\
&\lesssim \|I^c\nabla_{t,x}|\nabla|^{-1}\psi_1\|_{L_t^2 L_x^\infty}\|\widetilde{P}_0 Q_{\le C}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2 L_x^2} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)k_1}\|\psi_1\|_{L_t^2 L_x^2}2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}
\end{aligned}$$

The last case which we need to consider is $A_0 = A_1 = \tilde{A}_1 = I$ and either one of A_2, A_3 equal to I^c. But then necessarily $A_2 = A_3 = I^c$ whence

$$\begin{aligned}
&\|P_0\partial^\beta I[I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(Q_{\geq k_2+C}\psi_2, Q_{\geq k_2+C}\psi_3)]\|_{N[0]}\\
&\lesssim \|I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(Q_{\geq k_2+C}\psi_2, Q_{\geq k_2+C}\psi_3)\|_{L^1_t L^1_x}\\
&\lesssim \|\psi_1\|_{L^\infty_t L^\infty_x} \sum_{m\geq k_2+C} \|\widetilde{P}_0 Q_{\leq C}\mathcal{Q}_{\beta j}(Q_m\psi_2, \tilde{Q}_m\psi_3)\|_{L^1_t L^1_x}\\
&\lesssim 2^{k_1}\|\psi_1\|_{L^\infty_t L^2_x} \sum_{m\geq k_2+C} 2^{m-k_2} 2^{-2m(1-\varepsilon)} 2^{(1-2\varepsilon)k_2}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{k_1-k_2}\prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which concludes Case 4.

Case 5: $O(1) = k_1$, $k_2 = k_3 + O(1)$. We begin with $A_0 = I^c$ and $\tilde{A}_1 = I$ (in fact, $A_0 = Q_{\geq 0}$ suffices here as usual). Moreover, we will drop R_α from ψ_1 which amounts to excluding the case $A_1 = I^c$ and $\alpha = 0$ but nothing else. Then, from Lemma 4.17,

$$\begin{aligned}
&\|P_0 I^c\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)} \|\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_t L^2_x}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)} \|\psi_1\|_{L^\infty_t L^2_x}\, 2^{-k}\|IP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_t L^\infty_x}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)} \|\psi_1\|_{L^\infty_t L^2_x} 2^{k-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{-\frac{|k_2|}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is better than needed. Now suppose $\alpha = 0$ and $A_0 = A_1 = I^c$, which implies that $\tilde{A}_1 = I$. Then

$$\begin{aligned}
&\|P_0 I^c\partial^\beta[I^c R_0\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)}\sum_{m\geq 0} 2^{-\varepsilon m}\|P_0 Q_m[\tilde{Q}_m R_0\psi_1\,\Delta^{-1}\partial_j P_k I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_t L^2_x}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)}\sum_{m\geq 0} 2^{(1-\varepsilon)m}\|\tilde{Q}_m\psi_1\|_{L^2_t L^2_x} 2^{-k}\|P_k I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^\infty_t L^\infty_x}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)} \|\psi_1\|_{S[k_1]}\, 2^{\frac{k}{2}}\|IP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_t L^2_x}\\
&\lesssim \sum_{k\leq k_2\wedge 0+O(1)} 2^{\frac{k}{2}}\|\psi_1\|_{S[k_1]} 2^{k-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{-\frac{|k_2|}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Next, consider the case $A_0 = I^c$, and $\tilde{A}_1 = I^c$. Since $I^c R_0\psi_1$ is now excluded, we may drop $A_1 R_\alpha$ altogether. Then

$$\|P_0 I^c\partial^\beta[\psi_1\Delta^{-1}\partial_j I^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}$$

$$\lesssim \sum_{k\leq k_2\wedge 0+O(1)} \|P_0 Q_{\geq 0}[\psi_1\Delta^{-1}\partial_j Q_{k\leq\cdot\leq C}P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_t L^2_x} \tag{5.18}$$

$$+ \sum_{k\leq k_2\wedge 0+O(1)}\sum_{m\geq C} 2^{-\varepsilon m}\|P_0 Q_m[Q_{\leq m-C}\psi_1\Delta^{-1}\partial_j\tilde{Q}_m P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_t L^2_x} \tag{5.19}$$

$$+ \sum_{k\leq k_2\wedge 0+O(1)}\sum_{m\geq C} \|P_0 Q_{\geq 0}[Q_{>m-C}\psi_1\Delta^{-1}\partial_j\tilde{Q}_m P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_t L^2_x} \tag{5.20}$$

First, by Lemma 4.19,

$$
\begin{aligned}
(5.18) &\lesssim \sum_{k\le k_2\wedge 0+O(1)} \|\psi_1\|_{L_t^\infty L_x^2}\, 2^{-k}\|Q_{k\le\cdot\le C}P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^\infty}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} \|\psi_1\|_{L_t^\infty L_x^2}\, \|Q_{k\le\cdot\le C}P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} \|\psi_1\|_{L_t^\infty L_x^2}\, 2^{\frac{k}{2}}2^{-\varepsilon k_2}\langle k_2-k\rangle^2\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{-\varepsilon|k_2|}\langle k_2\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Second, again by Lemma 4.19,

$$
\begin{aligned}
(5.19) &\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C} 2^{-\varepsilon m}\|P_0Q_m[Q_{\le m-C}\psi_1\Delta^{-1}\partial_j\tilde{Q}_mP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C} 2^{-\varepsilon m}\|\psi_1\|_{L_t^\infty L_x^2}\|\tilde{Q}_mP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} 2^{\frac{k}{2}}2^{-\varepsilon k_2}\langle k_2-k\rangle^2\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \lesssim 2^{-\varepsilon|k_2|}\langle k_2\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

and third,

$$
\begin{aligned}
(5.20) &\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C} \|Q_{>m-C}\psi_1\|_{L_t^2L_x^2}2^{-k}\|\tilde{Q}_mP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^\infty}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C} 2^{-(1-\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{\frac{m}{2}}\|\tilde{Q}_mP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C} 2^{-(\frac12-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{-\varepsilon m}\|\tilde{Q}_mP_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim 2^{-\varepsilon|k_2|}\langle k_2\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

where one argues as in the previous two cases to pass to the last line.
Thus, $A_0 = Q_{\le 0}$ for the remainder of Case 5. If $A_1 = I^c$, then necessarily $\tilde{A}_1 = I^c$ which implies $\alpha \neq 0$. Therefore,

$$
\begin{aligned}
&\|P_0Q_{\le 0}\partial^\beta[I^c\psi_1\Delta^{-1}\partial_jI^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C}\|Q_m\psi_1\Delta^{-1}\partial_jP_k\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^1L_x^2}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C}\|Q_m\psi_1\|_{L_t^2L_x^2}\,\|P_k\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\sum_{m\ge C}2^{-(1-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{-\varepsilon m}\|P_k\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^2} \qquad (5.21)\\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)}\|\psi_1\|_{S[k_1]}\,2^{\frac{k}{2}}2^{-\varepsilon k_2}\langle k_2-k\rangle^2\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{-\varepsilon|k_2|}\langle k_2\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is again admissible. So we may assume also that $A_1 = I$ which means that we can drop R_α from ψ_1. First, consider the case $\tilde{A}_1 = I^c$, whence we now face the expression

$$
P_0Q_{\le 0}\partial^\beta[\psi_1\Delta^{-1}\partial_jI^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]
$$

with the implicit frequency constraints of case 5. We write this as

$$
\begin{aligned}
&P_0Q_{\leq 0}\partial^\beta[\psi_1\Delta^{-1}\partial_j I^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\\
&=\sum_{k<O(1)}P_0Q_{\leq 0}\partial^\beta[\psi_1\Delta^{-1}\partial_j Q_{>k+C}P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\\
&=\sum_{k<O(1)}\sum_{O(1)>l>k+C}P_0Q_{\leq 0}\partial^\beta[Q_{<l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\\
&+\sum_{k<O(1)}\sum_{O(1)>l>k+C}P_0Q_{\leq 0}\partial^\beta[Q_{\geq l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\\
&+\sum_{k<O(1)}\sum_{l>O(1)}P_0Q_{\leq 0}\partial^\beta[Q_{\geq l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)
\end{aligned}
\tag{5.22}
$$

To estimate the first term of (5.22), we use

$$
\begin{aligned}
&\|\sum_{k<O(1)}\sum_{O(1)>l>k+C}P_0Q_{\leq 0}\partial^\beta[Q_{<l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{N[0]}\\
&\leq\sum_{k<O(1)}\sum_{O(1)>l>k+C}\|P_0Q_{l+O(1)\leq\cdot\leq 0}\partial^\beta[Q_{<l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{\dot{X}_0^{-1,-\frac12,1}}\\
&\lesssim\sum_{k<O(1)}\sum_{O(1)>l>k+C}2^{-(\frac12-\varepsilon)l}\|Q_{<l-C}\psi_1\|_{L_t^\infty L_x^2}2^{-\varepsilon l}\|\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^\infty}
\end{aligned}
\tag{5.23}
$$

Recalling the assumptions on the frequencies in case 5, and using Bernstein's inequality as well as Lemma 4.19, we can further bound the preceding by

$$
\lesssim\sum_{k<\min\{O(1),k_2\}}\sum_{O(1)>l>k+C}2^{-(\frac12-\varepsilon)l}2^{\frac{k}{2}}2^{-\varepsilon k_2}\|\psi_1\|_{S[k_1]}\prod_{j=2,3}\|P_{k_j}\psi_j\|_{S[k_j]}\lesssim 2^{-\varepsilon k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\tag{5.24}
$$

The second term of (5.22) is handled similarly, using

$$
\begin{aligned}
&\|\sum_{k<O(1)}\sum_{O(1)>l>k+C}P_0Q_{\leq 0}\partial^\beta[Q_{\geq l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{N[0]}\\
&\leq\sum_{k<O(1)}\sum_{O(1)>l>k+C}\|P_0Q_{\leq 0}\partial^\beta[Q_{\geq l-C}\psi_1\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^1\dot{H}^{-1}}\\
&\lesssim\sum_{k<O(1)}\sum_{O(1)>l>k+C}\|Q_{\geq l-C}\psi_1\|_{L^2_{t,.x}}\|\Delta^{-1}\partial_j Q_l P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^\infty}
\end{aligned}
\tag{5.25}
$$

and from here the estimate continues as for (5.23). The third term of (5.22) is handled identically, and we note here that one actually gains exponentially in l.

Finally, suppose at least one choice of $j = 2, 3$ satisfies $A_j = I^c$. Then necessarily, $A_2 = A_3 = I^c$ and

$$\begin{aligned}
&\|P_0 I \partial^\beta [I\psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{k \le k_2 \wedge 0 + O(1)} \|I\psi_1 \Delta^{-1}\partial_j I P_k \mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)]\|_{L^1_t L^2_x} \\
&\lesssim \sum_{k \le k_2 \wedge 0 + O(1)} \|\psi_1\|_{L^\infty_t L^2_x} 2^{-k} \|I P_k \mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)\|_{L^1_t L^\infty_x} \\
&\lesssim \sum_{k \le k_2 \wedge 0 + O(1)} \|\psi_1\|_{L^\infty_t L^2_x} 2^{k} \|I P_k \mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)\|_{L^1_t L^1_x} \\
&\lesssim 2^{k_2 \wedge 0} \|\psi_1\|_{L^\infty_t L^2_x} \sum_{m \ge k_2 + C} \|\mathcal{Q}_{\beta j}(Q_m\psi_2, \tilde{Q}_m\psi_3)\|_{L^1_t L^1_x} \\
&\lesssim 2^{k_2 \wedge 0} \|\psi_1\|_{L^\infty_t L^2_x} \sum_{m \ge k_2 + C} 2^{m-k_2} 2^{-2(1-\varepsilon)m} 2^{(1-2\varepsilon)k_2} \|\psi_2\|_{S[k_2]} \|\psi_3\|_{S[k_3]}
\end{aligned} \tag{5.26}$$

$$\lesssim 2^{-k_2 \vee 0} \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]} \tag{5.27}$$

as claimed.

Case 6: $O(1) = k_1 \ge k_2 + O(1) \ge k_3 + C$. This case proceeds similarly to Case 5. We begin with $A_0 = I^c$ and $\tilde{A}_1 = I$ (in fact, $A_0 = Q_{\ge 0}$ suffices here as usual). Moreover, we will drop R_α from ψ_1 which amounts to excluding the case $A_1 = I^c$ and $\alpha = 0$ but nothing else. Then, from Lemma 4.23,

$$\begin{aligned}
&\|P_0 I^c \partial^\beta [\psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \lesssim \|\psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^2_t L^2_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} 2^{-k_2} \|I \tilde{P}_{k_2} \mathcal{Q}_{\beta j}(\psi_2, \psi_3)\|_{L^2_t L^\infty_x} \lesssim 2^{(\frac{1}{2}-\varepsilon)k_3 + \varepsilon k_2} \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is better than needed. Now suppose $\alpha = 0$ and $A_0 = A_1 = I^c$, which implies that $\tilde{A}_1 = I$. Then

$$\begin{aligned}
&\|P_0 I^c \partial^\beta [I^c R_0 \psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m \ge 0} 2^{-\varepsilon m} \|P_0 Q_m [\tilde{Q}_m R_0 \psi_1 \, \Delta^{-1}\partial_j \tilde{P}_{k_2} I \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^2_t L^2_x} \\
&\lesssim \sum_{m \ge 0} 2^{(1-\varepsilon)m} \|\tilde{Q}_m \psi_1\|_{L^2_t L^2_x} 2^{-k_2} \|\tilde{P}_{k_2} I \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^\infty_t L^\infty_x} \\
&\lesssim 2^{\frac{k_2}{2}} 2^{(\frac{1}{2}-\varepsilon)k_3 + \varepsilon k_2} \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Next, consider the case $A_0 = I^c$, and $\tilde{A}_1 = I^c$. As before, we can drop $A_1 R_\alpha$ in this case. Then

$$\|P_0 I^c \partial^\beta [\psi_1 \Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]}$$

$$\lesssim \|P_0 Q_{\ge 0}[\psi_1 \Delta^{-1}\partial_j Q_{k \le \cdot \le C} \tilde{P}_{k_2} \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^2_t L^2_x} \tag{5.28}$$

$$+ \sum_{m \ge C} 2^{-\varepsilon m} \|P_0 Q_m [Q_{\le m - C} \psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \tilde{P}_{k_2} \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^2_t L^2_x} \tag{5.29}$$

$$+ \sum_{m \ge C} \|P_0 Q_{\ge 0} [Q_{> m - C} \psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \tilde{P}_{k_2} \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{L^2_t L^2_x} \tag{5.30}$$

First, by Lemma 4.24,

$$
\begin{aligned}
(5.28) &\lesssim \|\psi_1\|_{L_t^\infty L_x^2}\, 2^{-k_2}\|Q_{k_2-O(1)\le\cdot\le C}\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^\infty}\\
&\lesssim \|\psi_1\|_{L_t^\infty L_x^2}\, \|Q_{k_2-O(1)\le\cdot\le C}\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim 2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Second, again by Lemma 4.24,

$$
\begin{aligned}
(5.29) &\lesssim \sum_{m\ge C} 2^{-\varepsilon m}\|P_0Q_m[Q_{\le m-C}\psi_1\Delta^{-1}\partial_j\tilde{Q}_m\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{m\ge C} 2^{-\varepsilon m}\|\psi_1\|_{L_t^\infty L_x^2}\|\tilde{Q}_m\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^\infty}\\
&\lesssim 2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

and third,

$$
\begin{aligned}
(5.30) &\lesssim \sum_{m\ge C}\|Q_{>m-C}\psi_1\|_{L_t^2L_x^2}2^{-k_2}\|\tilde{Q}_m\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^\infty}\\
&\lesssim \sum_{m\ge C}2^{-(1-\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{\frac{m}{2}}\|\tilde{Q}_m\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim \sum_{m\ge C}2^{-(\frac12-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{-\varepsilon m}\|\tilde{Q}_m\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim 2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

where one argues as in the previous two cases to pass to the last line.
Thus, $A_0 = Q_{\le 0}$ for the remainder of Case 6. If $A_1 = I^c$, then necessarily $\tilde{A}_1 = I^c$ which implies $\alpha\neq 0$. Therefore,

$$
\begin{aligned}
&\|P_0Q_{\le0}\partial^\beta[I^c\psi_1\Delta^{-1}\partial_jI^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{m\ge C}\|Q_m\psi_1\Delta^{-1}\partial_j\tilde{P}_{k_2}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^1L_x^2}\\
&\lesssim \sum_{m\ge C}\|Q_m\psi_1\|_{L_t^2L_x^2}\,\|\tilde{P}_{k_2}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim \sum_{m\ge C}2^{-(1-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{-\varepsilon m}\|\tilde{P}_{k_2}\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
\tag{5.31}
$$

which is again admissible. So we may assume also that $A_1 = I$ which means that we can drop R_α from ψ_1.
Next, assume $\tilde{A}_1 = I^c$. Then write

$$
\begin{aligned}
&P_0Q_{\le0}\partial^\beta[\psi_1\Delta^{-1}\partial_jI^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\\
&=\sum_{l>k_2+C}P_0Q_{\le0}\partial^\beta[Q_{<l-C}\psi_1\Delta^{-1}\partial_jP_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\\
&+\sum_{l>k_2+C}P_0Q_{\le0}\partial^\beta[Q_{\ge l-C}\psi_1\Delta^{-1}\partial_jP_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]
\end{aligned}
\tag{5.32}
$$

The first term we estimate by using Lemma 4.24: we get

$$
\begin{aligned}
&\Big\| \sum_{l>k_2+C} P_0 Q_{\le 0}\partial^\beta[Q_{<l-C}\psi_1\Delta^{-1}\partial_j P_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\Big\|_{N[0]} \\
&\le \sum_{O(1)>l>k_2+C} \|P_0 Q_{l-C\le\cdot\le 0}\partial^\beta[Q_{<l-C}\psi_1\Delta^{-1}\partial_j P_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{O(1)>l>k_2+C} 2^{-(\frac12-\varepsilon)l}\|Q_{<l-C}\psi_1\|_{L^\infty_t L^2_x} 2^{-\varepsilon l}\|\Delta^{-1}\partial_j P_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^2_t L^\infty_x} \\
&\lesssim \sum_{O(1)>l>k_2+C} 2^{-(\frac12-\varepsilon)l}2^{(\frac12-\varepsilon)k_3}\prod_{j=1}^3\|P_{k_j}\psi_j\|_{S[k_j]} \lesssim 2^{(\frac12-\varepsilon)(k_2-k_3)}\prod_{j=1}^3\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}
\tag{5.33}
$$

The second term in (5.32) is more of the same, and estimated using

$$
\begin{aligned}
&\|P_0 Q_{\le 0}\partial^\beta[Q_{\ge l-C}\psi_1\Delta^{-1}\partial_j P_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{N[0]} \\
&\lesssim \|[Q_{\ge l-C}\psi_1\|_{L^2_{t,x}}\|\Delta^{-1}\partial_j P_{k_2+O(1)}Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_t L^\infty_x}
\end{aligned}
\tag{5.34}
$$

from which point the estimate is concluded as in the preceding case. This leaves the cases $A_2 = I^c$ or $A_3 = I^c$ to be considered. In the former case, necessarily $A_2 = A_3 = I^c$ and

$$
\begin{aligned}
&\|P_0 I\partial^\beta[I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)]\|_{N[0]} \lesssim \|I\psi_1\Delta^{-1}\partial_j I\tilde P_{k_2}\mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)]\|_{L^1_t L^2_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x}\, 2^{-k_2}\|I\tilde P_{k_2}\mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)\|_{L^1_t L^\infty_x} \lesssim \|\psi_1\|_{L^\infty_t L^2_x}\,\|I\tilde P_{k_2}\mathcal{Q}_{\beta j}(I^c\psi_2, I^c\psi_3)\|_{L^1_t L^2_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x}\sum_{m\ge k_2}\|\mathcal{Q}_{\beta j}(Q_m\psi_2,\tilde Q_m\psi_3)\|_{L^1_t L^2_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x}\sum_{m\ge k_2}\big(\|\nabla_{t,x}|\nabla|^{-1}Q_m\psi_2\|_{L^2_t L^2_x}\|\tilde Q_m\psi_3\|_{L^2_t L^\infty_x} + \|Q_m\psi_2\|_{L^2_t L^2_x}\|\nabla_{t,x}|\nabla|^{-1}\tilde Q_m\psi_3\|_{L^2_t L^\infty_x}\big) \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x}\sum_{m\ge k_2} 2^{-(1-2\varepsilon)m}2^{(\frac12-\varepsilon)k_2}2^{(\frac12-\varepsilon)k_3}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{(\frac12-\varepsilon)(k_3-k_2)}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is acceptable. The one remaining case is $A_0 = A_1 = \tilde A_1 = A_2 = I$ and $A_3 = I^c$. Of course one may also assume that $\psi_3 = Q_{\le k_2+C}\psi_3$. Then we write

$$
P_0 I\partial^\beta[I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)] = P_0 I[\partial^\beta I\psi_1\,\tilde P_{k_2}\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)] \tag{5.35}
$$

$$
+ P_0 I[I\psi_1\,\Delta^{-1}\partial_j\partial^\beta\tilde P_{k_2}I\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)] \tag{5.36}
$$

The term on the right-hand side of (5.35) is difficult. More specifically, the methods that we have employed up to this point do not seem to yield the necessary bound. However, Tao's trilinear estimate (5.1) implies that

$$
\|\partial^\beta P_0\psi_1\, R_\beta\psi_3\,\psi_2\|_{N[0]} \lesssim 2^{\sigma(k_3-k_2)}2^{k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \tag{5.37}
$$

for some constant $\sigma > 0$ as well as

$$
\|\partial^\beta P_0\psi_1\, R_\beta\psi_2\,\psi_3\|_{N[0]} \lesssim 2^{k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \tag{5.38}
$$

Since $2^{k_2}\tilde P_{k_2}\Delta^{-1}\partial_j I$ can be replaced by the convolution by a measure and all norms involved are translation invariant, these estimates imply (5.35).

The analysis of (5.36) is easier and similar to the considerations at the end of Case 2. More precisely, we first reduce the modulation of the entire output by means of Lemma 4.23:

$$\begin{aligned}
&\|P_0 Q_{(1-3\varepsilon)k_3\le\cdot\le C}[I\psi_1\Delta^{-1}\partial_j\partial^\beta\tilde{P}_{k_2}I\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\|_{N[0]}\\
&\lesssim 2^{-\frac12(1-3\varepsilon)k_3}\|\psi_1\|_{L_t^\infty L_x^2}\,\|I\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)\|_{L_t^2L_x^\infty}\\
&\lesssim 2^{-\frac12(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{(\frac12-\varepsilon)k_3}2^{(1+\varepsilon)k_2}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{(1+\varepsilon)k_2}2^{\frac{\varepsilon}{2}k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Next, we reduce the modulation of ψ_1:

$$\begin{aligned}
&\|P_0 Q_{\le(1-3\varepsilon)k_3}[Q_{\ge(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j\partial^\beta I\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\|_{N[0]}\\
&\lesssim \|P_0 Q_{\le(1-3\varepsilon)k_3}[Q_{\ge(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j\partial^\beta I\tilde{P}_{k_2}\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\|_{L_t^1L_x^2}\\
&\lesssim 2^{k_2}\|Q_{\ge(1-3\varepsilon)k_3}\psi_1\|_{L_t^2L_x^2}\,\|\tilde{P}_{k_2}I\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{-\frac12(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{(\frac12-\varepsilon)k_3}2^{(1+\varepsilon)k_2}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{\frac{\varepsilon}{2}k_3+(1+\varepsilon)k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Finally, we reduce the modulation of the interior null-form using Corollary 4.14:

$$\begin{aligned}
&\|P_0 Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j\partial^\beta\tilde{P}_{k_2}Q_{k_3\le\cdot\le k_2+C}\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\|_{N[0]}\\
&\lesssim \|\psi_1\|_{S[k_1]}\sum_{k_3\le\ell\le k_2+C}2^{\frac{\ell-k_2}{4}}2^{-\frac{\ell}{2}}\|\tilde{P}_{k_2}Q_\ell\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{(\frac14-\varepsilon)(k_3-k_2)}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is again admissible. After these preparations, we are faced with the following decomposition:

$$\begin{aligned}
&P_0 Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\\
&= P_0 Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial^\beta\partial_j Q_{\le k_3}\mathcal{Q}_{\beta j}(Q_{k_3\le\cdot\le k_2+C}\psi_2,Q_{k_3+C\le\cdot\le k_2+C}\psi_3)]\\
&= \sum_{\kappa,\kappa'\in\mathcal{C}_\ell}P_{0,\kappa}Q_{\le(1-3\varepsilon)k_3}[P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(Q_{k_3\le\cdot\le k_2+C}\psi_2,Q_{k_3+C\le\cdot\le k_2+C}\psi_3)]
\end{aligned}$$

where $\ell=\frac12[k_2+(1-3\varepsilon)k_3]$ and $\mathrm{dist}(\kappa,\kappa')\lesssim 2^\ell$. Placing the entire expression in $L_t^1L_x^2$ and using Bernstein's inequality results in the following estimate: with $J:=Q_{k_3\le\cdot\le k_2+C}$,

$$\begin{aligned}
&\|P_0 Q_{\le(1-3\varepsilon)k_3}[Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\|_{L_t^1L_x^2}\\
&\le \Big\|\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_\ell}\|P_{0,\kappa}[P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2,J\psi_3)]\|_{L_x^2}^2\Big)^{\frac12}\Big\|_{L_t^1}\\
&\le 2^{\frac{\ell}{2}}\Big\|\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_\ell}\|P_{0,\kappa}[P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2,J\psi_3)]\|_{L_x^1}^2\Big)^{\frac12}\Big\|_{L_t^1}\\
&\le 2^{\frac{\ell}{2}}\Big\|\Big(\sum_{\kappa'\in\mathcal{C}_\ell}\|P_{k_1,\kappa'}Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\|_{L_x^2}^2\|\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2,J\psi_3)]\|_{L_x^2}^2\Big)^{\frac12}\Big\|_{L_t^1}\\
&\le 2^{\frac{\ell}{2}}\|Q_{\le(1-3\varepsilon)k_3-k_1}\psi_1\|_{L_t^\infty L_x^2}\|\Delta^{-1}\partial_j\partial^\beta Q_{\le k_3}\mathcal{Q}_{\beta j}(J\psi_2,J\psi_3)]\|_{L_t^1L_x^2}\\
&\lesssim 2^{\frac{\ell}{2}}\|\psi_1\|_{S[k_1]}\,\|\nabla_{t,x}|\nabla|^{-1}J\psi_2\|_{L_t^2L_x^2}\|\nabla_{t,x}|\nabla|^{-1}J\psi_3\|_{L_t^2L_x^\infty}\\
&\lesssim 2^{\frac14(1-3\varepsilon)k_3+\frac{k_2}{4}}\|\psi_1\|_{S[k_1]}\,2^{-\frac{k_3}{2}}\|\psi_2\|_{S[k_2]}\,2^{(\frac12-\varepsilon)k_3}2^{\varepsilon k_2}\|\psi_3\|_{S[k_3]}
\end{aligned}$$

which is again admissible for small $\varepsilon>0$.

Case 7: $k_1 = O(1) \geq k_3 + O(1) \geq k_2 + C$. This case is symmetric to the previous one.

Case 8: $k_2 = O(1), \max(k_1, k_3) \leq -C$. We begin with $A_0 = Q_{\geq 0}$ and $\tilde{A}_1 = I$, and we drop R_α from ψ_1 excluding the case $A_1 = I^c$ and $\alpha = 0$ but nothing else. Then, from Lemma 4.23,

$$\begin{aligned}
&\|P_0 I^c \partial^\beta[\psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \lesssim \|\psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim \|\psi_1\|_{L_t^\infty L_x^\infty} \|I\tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2 L_x^2} \lesssim 2^{k_1} 2^{(\frac{1}{2}-\varepsilon)k_3} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is better than needed. Now suppose $\alpha = 0$ and $A_0 = A_1 = I^c$, which implies that $\tilde{A}_1 = I$. Then by Lemma 4.23

$$\begin{aligned}
&\|P_0 I^c \partial^\beta[I^c R_0 \psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\geq 0} 2^{-\varepsilon m} \|P_0 Q_m[\tilde{Q}_m R_0 \psi_1 \, \Delta^{-1}\partial_j \tilde{P}_0 I \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m\geq 0} 2^{-\varepsilon m} \|\tilde{Q}_m \nabla_{t,x}|\nabla|^{-1}\psi_1\|_{L_t^2 L_x^\infty} \|\Delta^{-1}\partial_j \tilde{P}_0 I \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^2} \\
&\lesssim \sum_{m\geq 0} 2^{(1-\varepsilon)m} \|\tilde{Q}_m \psi_1\|_{L_t^2 L_x^2} \|\tilde{P}_0 I \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{(\frac{1}{2}-\varepsilon)(k_1+k_3)} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Next, consider the case $A_0 = I^c$, and $\tilde{A}_1 = I^c$. As before, we can drop $A_1 R_\alpha$ in this case. Then

$$\|P_0 I^c \partial^\beta[\psi_1 \Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}$$

$$\lesssim \sum_{m\geq C} 2^{-\varepsilon m} \|P_0 Q_m[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \tag{5.39}$$

$$+ \sum_{m\geq C} \|P_0 Q_{\geq 0}[Q_{>m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \tag{5.40}$$

First, by Lemma 4.24,

$$\begin{aligned}
(5.39) &\lesssim \sum_{m\geq C} 2^{-\varepsilon m} \|P_0 Q_m[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{m\geq C} 2^{-\varepsilon m} 2^{k_1} \|\psi_1\|_{L_t^\infty L_x^2} \|\tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_1 + (\frac{1}{2}-\varepsilon)k_3} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

and second,

$$\begin{aligned}
(5.40) &\lesssim \sum_{m\geq C} 2^{k_1} \|Q_{>m-C}\psi_1\|_{L_t^2 L_x^2} \|\tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^2} \\
&\lesssim \sum_{m\geq C} 2^{(\frac{3}{2}-\varepsilon)k_1} 2^{-(1-\varepsilon)m} \|\psi_1\|_{S[k_1]} \, 2^{\frac{m}{2}} \|\tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_1} \sum_{m\geq C} 2^{-(\frac{1}{2}-2\varepsilon)m} \|\psi_1\|_{S[k_1]} \, 2^{-\varepsilon m} \|\tilde{Q}_m \tilde{P}_0 \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim 2^{k_1 + (\frac{1}{2}-\varepsilon)k_3} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

where one argues as in the previous two cases to pass to the last line.

Thus, $A_0 = Q_{\leq 0}$ for the remainder of Case 8. If $\tilde{A}_1 = I^c$, then necessarily $A_1 = I^c$ which implies $\alpha \neq 0$. Therefore,

$$\begin{aligned}
&\|P_0 Q_{\leq 0}\partial^\beta[I^c\psi_1\Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\geq C} \|\tilde{Q}_m\psi_1\Delta^{-1}\partial_j\tilde{P}_0 Q_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^1_tL^2_x} \\
&\lesssim \sum_{m\geq C} 2^{k_1}\|\tilde{Q}_m\psi_1\|_{L^2_tL^2_x}\,\|\tilde{P}_0 Q_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_tL^2_x} \\
&\lesssim 2^{k_1}\sum_{m\geq C} 2^{-(1-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{-\varepsilon m}\|\tilde{P}_0\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_tL^2_x} \\
&\lesssim 2^{k_1}2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is again admissible. So we may assume also that $\tilde{A}_1 = I$. Now suppose that $A_1 = I^c$. Then we can take $A_1 = Q_{k_1\leq\cdot\leq C}$ whence

$$\begin{aligned}
&\|P_0 Q_{\leq 0}\partial^\beta[Q_{k_1\leq\cdot\leq C}R_\alpha\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \|Q_{k_1\leq\cdot\leq C}R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^1_tL^2_x} \\
&\lesssim \|Q_{k_1\leq\cdot\leq C}\nabla_{t,x}|\nabla|^{-1}\psi_1\|_{L^2_tL^\infty_x}\,\|\tilde{P}_0 I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_tL^2_x} \\
&\lesssim 2^{(\frac12-\varepsilon)k_1}\|\psi_1\|_{S[k_1]}\,2^{(\frac12-\varepsilon)k_3}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \\
&\lesssim 2^{(\frac12-\varepsilon)(k_1+k_3)}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

So we may assume for the remainder of this case that $A_1 = I$ which means that we can drop R_α from ψ_1. This leaves the cases $A_2 = I^c$ or $A_3 = I^c$ to be considered. In the former case, necessarily $A_2 = A_3 = I^c$ and

$$\begin{aligned}
&\|P_0 I\partial^\beta[I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I^c\psi_2,I^c\psi_3)]\|_{N[0]} \lesssim \|I\psi_1\Delta^{-1}\partial_j I\tilde{P}_0\mathcal{Q}_{\beta j}(I^c\psi_2,I^c\psi_3)]\|_{L^1_tL^2_x} \\
&\lesssim 2^{k_1}\|\psi_1\|_{L^\infty_tL^2_x}\,\|I\tilde{P}_0\mathcal{Q}_{\beta j}(I^c\psi_2,I^c\psi_3)\|_{L^1_tL^2_x} \lesssim 2^{k_1}\|\psi_1\|_{L^\infty_tL^2_x}\,\|I\tilde{P}_0\mathcal{Q}_{\beta j}(I^c\psi_2,I^c\psi_3)\|_{L^1_tL^2_x} \\
&\lesssim 2^{k_1}\|\psi_1\|_{L^\infty_tL^2_x}\sum_{m\geq 0}\|\mathcal{Q}_{\beta j}(Q_m\psi_2,\tilde{Q}_m\psi_3)\|_{L^1_tL^2_x} \\
&\lesssim 2^{k_1}\|\psi_1\|_{L^\infty_tL^2_x}\sum_{m\geq 0}\big(\|\nabla_{t,x}|\nabla|^{-1}Q_m\psi_2\|_{L^2_tL^2_x}\|\tilde{Q}_m\psi_3\|_{L^2_tL^\infty_x} + \|Q_m\psi_2\|_{L^2_tL^2_x}\|\nabla_{t,x}|\nabla|^{-1}\tilde{Q}_m\psi_3\|_{L^2_tL^\infty_x}\big) \\
&\lesssim 2^{k_1}\|\psi_1\|_{L^\infty_tL^2_x}\sum_{m\geq 0}2^{-(1-2\varepsilon)m}2^{(\frac12-\varepsilon)k_3}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{k_1+(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is acceptable. The one remaining case is $A_0 = A_1 = \tilde{A}_1 = A_2 = I$ and $A_3 = I^c$. Of course one may also assume that $\psi_3 = Q_{\leq C}\psi_3$. The analysis in this case is similar to the considerations at the end of Case 2. More precisely, we first reduce the modulation of the entire output by means of Lemma 4.23:

$$\begin{aligned}
&\|P_0\partial^\beta Q_{(1-3\varepsilon)k_3\leq\cdot\leq C}[I\psi_1\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)]\|_{N[0]} \\
&\lesssim 2^{-\frac12(1-3\varepsilon)k_3}\|\psi_1\|_{L^\infty_tL^\infty_x}\,\|I\mathcal{Q}_{\beta j}(I\psi_2,I^c\psi_3)\|_{L^2_tL^2_x} \\
&\lesssim 2^{k_1}2^{-\frac12(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{(\frac12-\varepsilon)k_3}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \\
&\lesssim 2^{k_1}2^{\frac{\varepsilon}{2}k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Next, we reduce the modulation of ψ_1:

$$\begin{aligned}
&\|P_0 Q_{\leq (1-3\varepsilon)k_3}\partial^\beta[Q_{\geq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j I\tilde{P}_0\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{N[0]}\\
&\lesssim \|P_0\partial^\beta Q_{\leq (1-3\varepsilon)k_3}[Q_{\geq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j I\tilde{P}_0\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{L^1_tL^2_x}\\
&\lesssim 2^{k_1}\|Q_{\geq(1-3\varepsilon)k_3}\psi_1\|_{L^2_tL^2_x}\,\|\tilde{P}_0 I\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)\|_{L^2_tL^2_x}\\
&\lesssim 2^{k_1-\frac12(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{(\frac12-\varepsilon)k_3}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{\frac{\varepsilon}{2}k_3+k_1}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Finally, we reduce the modulation of the interior null-form using Corollary 4.14:

$$\begin{aligned}
&\|P_0 Q_{\leq (1-3\varepsilon)k_3}[Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j \tilde{P}_0 Q_{k_3\leq\cdot\leq C}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{N[0]}\\
&\lesssim 2^{k_1}\|\psi_1\|_{S[k_1]}\sum_{k_3\leq\ell\leq C}2^{\frac{\ell-k_1}{4}}2^{-\frac{\ell}{2}}\|\tilde{P}_0 Q_\ell\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)\|_{L^2_tL^2_x}\\
&\lesssim 2^{\frac{3k_1}{4}+(\frac14-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is again admissible. After these preparations, we are faced with the following decomposition:

$$\begin{aligned}
&P_0\partial^\beta Q_{\leq(1-3\varepsilon)k_3}[Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\\
&= P_0 Q_{\leq(1-3\varepsilon)k_3}\partial^\beta[Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(Q_{k_3\leq\cdot\leq k_2+C}\psi_2, Q_{k_3+C\leq\cdot\leq k_2+C}\psi_3)]\\
&= \sum_{\kappa,\kappa'\in\mathcal{C}_\ell}P_{0,\kappa}Q_{\leq(1-3\varepsilon)k_3}\partial^\beta[P_{k_1,\kappa'}Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(Q_{k_3\leq\cdot\leq k_2+C}\psi_2, Q_{k_3+C\leq\cdot\leq k_2+C}\psi_3)]
\end{aligned}$$

where $\ell = \frac12[(1-3\varepsilon)k_3 - k_1]\wedge 0$ and $\mathrm{dist}(\kappa,\kappa')\lesssim 2^\ell$. Placing the entire expression in $L^1_tL^2_x$ and using Bernstein's inequality results in the following estimate:

$$\begin{aligned}
&\|P_0 Q_{\leq(1-3\varepsilon)k_3}[Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(I\psi_2, I^c\psi_3)]\|_{L^1_tL^2_x}\\
&\leq \Big\|\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_\ell}\|P_{0,\kappa}[P_{k_1,\kappa'}Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(Q_{k_3\leq\cdot\leq C}\psi_2, Q_{k_3+C\leq\cdot\leq C}\psi_3)]\|^2_{L^2_x}\Big)^{\frac12}\Big\|_{L^1_t}\\
&\leq \Big\|\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_\ell}\|P_{k_1,\kappa'}Q_{\leq(1-3\varepsilon)k_3}\psi_1\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(Q_{k_3\leq\cdot\leq C}\psi_2, Q_{k_3+C\leq\cdot\leq C}\psi_3)\|^2_{L^2_x}\Big)^{\frac12}\Big\|_{L^1_t}\\
&\leq \Big\|\Big(\sum_{\kappa'\in\mathcal{C}_\ell}\|P_{k_1,\kappa'}Q_{\leq(1-3\varepsilon)k_3}\psi_1\|^2_{L^\infty_x}\|\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(Q_{k_3\leq\cdot\leq C}\psi_2, Q_{k_3+C\leq\cdot\leq C}\psi_3)]\|^2_{L^2_x}\Big)^{\frac12}\Big\|_{L^1_t}\\
&\leq 2^{\frac{\ell}{2}}2^{k_1}\|Q_{\leq(1-3\varepsilon)k_3}\psi_1\|_{L^\infty_tL^2_x}\|\Delta^{-1}\partial_j Q_{\leq k_3}\mathcal{Q}_{\beta j}(Q_{k_3\leq\cdot\leq C}\psi_2, Q_{k_3+C\leq\cdot\leq C}\psi_3)]\|_{L^1_tL^2_x}\\
&\lesssim 2^{k_1+\frac{\ell}{2}}\|\psi_1\|_{S[k_1]}\,\|\nabla_{t,x}|\nabla|^{-1}Q_{k_3\leq\cdot\leq C}\psi_2\|_{L^2_tL^2_x}\|\nabla_{t,x}|\nabla|^{-1}Q_{k_3+C\leq\cdot\leq C}\psi_3\|_{L^2_tL^\infty_x}\\
&\lesssim 2^{\frac{3k_1}{4}+\frac14(1-3\varepsilon)k_3}\|\psi_1\|_{S[k_1]}\,2^{-\frac{k_3}{2}}\|\psi_2\|_{S[k_2]}\,2^{(\frac12-\varepsilon)k_3}\|\psi_3\|_{S[k_3]}
\end{aligned}$$

which is again admissible for small $\varepsilon > 0$.

Case 9: $k_3 = O(1), \max(k_1, k_2) \leq -C$. Symmetric to Case 8. □

It is important to realize that Lemma 5.1 yields the following statement, which is really a corollary of its proof rather than its lemma.

Corollary 5.2. *Let ψ_i be Schwarz functions adapted to k_i for $i = 0, 1, 2$. Then for any $\alpha, \beta = 0, 1, 2$, and $j = 1, 2$,*

$$\|P_0\nabla_{t,x}A_0[A_1R_\alpha\psi_1\Delta^{-1}\partial_j\tilde{A}_1\mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{N[0]}\lesssim w(k_1,k_2,k_3)\prod_{i=1}^3\|\psi_i\|_{S[k_i]}$$

where A_i and $\tilde{A}_1$ are either I or I^c, with at least one being I^c. Moreover, we impose the following restrictions:

- *if $A_1 = \tilde{A}_1 = I^c$ then $\alpha = 0$ is excluded*
- *if $k_1 = O(1) > k_2 \geq k_3 + C$, then $A_0 = A_1 = \tilde{A}_1 = A_2 = I$, $A_3 = I^c$ is excluded*
- *if $k_1 = O(1) > k_3 \geq k_2 + C$, then $A_0 = A_1 = \tilde{A}_1 = A_3 = I$, $A_2 = I^c$ is excluded*

In particular,

$$\|P_0 \nabla_{t,x}[\psi_1 \Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(\psi_2, \psi_3)]\|_{N[0]} \lesssim w(k_1,k_2,k_3) \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]} \tag{5.41}$$

Proof. Note that the first exclusion in our list is precisely the exclusion in Lemma 5.1. The only real difference between this statement and that of Lemma 5.1 lies with the fact that we no longer require the outer most derivative to be ∂^β. But this mattered only in one case, namely when we applied Tao's bound (5.1) in Cases 6 and 7 above. Moreover, inspection of the argument in those cases reveals that the $\partial^\beta\phi\partial_\beta\psi$ null-form was needed only in those instances which are excluded as the second and third conditions of our above list (in fact, the modulations were narrowed down much more before any need for (5.1) arose). The final statement is an immediate consequence of the first one, since we removed R_α altogether (which eliminates the first exclusion) and since the other two exclusions do not arise due to $\tilde{A}_1 = I^c$. Therefore, one simply sums over all choices of A_0, A_1, A_2 and A_3. ☐

In fact, the proof of Lemma 5.1 makes no use of the fact that $\Delta^{-1}\partial_j$ contains the same index as the null-form $\mathcal{Q}_{\beta j}$. But the strengthening resulting from replacing $\Delta^{-1}\partial_j$ by $|\nabla|^{-1}$, say, is of no benefit to us so we do not carry it out. The following variant of Lemma 5.1 covers the other two types of trilinear nonlinearities arising in the Coulomb gauged wave-map system.

Lemma 5.3. *Let ψ_i be Schwarz functions adapted to k_i for $i = 0,1,2$. Then for any $\alpha = 0,1,2$, $j = 1,2$,*

$$\|P_0\partial^\beta A_0[A_1 R_\beta \psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\alpha j}(A_2\psi_2, A_3\psi_3)]\|_{N[0]} \lesssim w(k_1,k_2,k_3) \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]} \tag{5.42}$$

$$\|P_0\partial^\alpha A_0[A_1 R^\beta \psi_1 \Delta^{-1}\partial_j I \mathcal{Q}_{\beta j}(A_2\psi_2, A_3\psi_3)]\|_{N[0]} \lesssim w(k_1,k_2,k_3) \prod_{i=1}^{3} \|\psi_i\|_{S[k_i]} \tag{5.43}$$

where A_i are either I or I^c, with at least one being I^c.

Proof. Both these bounds follow from Corollary 5.2 provided we are not in those cases described as Items 2 and 3 in the list of exclusions (observe that the first exclusion does not arise due to our limitation to $\tilde{A}_1 = I$). So let us consider the second exclusion $k_1 = O(1) > k_2 \geq k_3 + C$ and $A_0 = A_1 = \tilde{A}_1 = A_2 = I$, $A_3 = I^c$ (the third one being symmetric to this case). Then (5.43) is an immediate consequence of (5.1), see (5.37) and (5.38) above. As for (5.42), observe that due to the analysis of (5.36) we may assume that the outer ∂^β derivative hits ψ_1. Hence, it suffices to bound

$$\|P_0 I[I\partial^\beta R_\beta \psi_1 \Delta^{-1}\partial_j I\mathcal{Q}_{\alpha j}(I\psi_2, I^c\psi_3)]\|_{N[0]}$$

However, due to the property that $\|\Box I P_0 \phi\|_{L^2_t L^2_x} \lesssim \|\phi\|_{S[0]}$ and $\partial^\beta\partial_\beta = \Box$, this is easy:

$$\begin{aligned}
&\|P_0 I[Q_{\leq C}\partial^\beta R_\beta \psi_1 \Delta^{-1}\partial_j I\mathcal{Q}_{\alpha j}(I\psi_2, I^c\psi_3)]\|_{N[0]} \\
&\lesssim \|P_0 I[Q_{\leq C}\partial^\beta R_\beta \psi_1 \Delta^{-1}\partial_j I\mathcal{Q}_{\alpha j}(I\psi_2, I^c\psi_3)]\|_{L^1_t L^2_x} \\
&\lesssim \|Q_{\leq C}\partial^\beta R_\beta \psi_1\|_{L^2_t L^2_x} \|\Delta^{-1}\partial_j \tilde{P}_{k_2} I\mathcal{Q}_{\alpha j}(I\psi_2, I^c\psi_3)]\|_{L^2_t L^\infty_x} \\
&\lesssim \|\psi_1\|_{S[k_1]}\, 2^{\frac12(1-3\varepsilon)k_3} 2^{\varepsilon k_2} \|\psi_2\|_{S[k_2]} \|\psi_3\|_{S[k_3]}
\end{aligned}$$

as desired. ☐

The following technical corollary will be important later.

Corollary 5.4. *For some absolute constant $\sigma_0 > 0$, and arbitrary Schwartz functions ψ_i,*

$$\sum_{j=1}^{2} \|P_0\nabla_{t,x}[\psi_1\Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \lesssim K^2 \sup_{k\in\mathbb{Z}} \max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]} \tag{5.44}$$

provided $\max_{i=1,2,3}\sum_{k\in\mathbb{Z}}\|P_k\psi_i\|^2_{S[k]} \le K^2$ *and with an absolute implicit constant. Moreover, given any* $\delta > 0$ *there exists a constant* $L = L(\delta) \gg 1$ *such that*

$$\sum\nolimits'_{k_1,k_2,k_3} \sum_{j=1}^{2} \|P_0\nabla_{t,x}[P_{k_1}\psi_1\Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(P_{k_2}\psi_2,P_{k_3}\psi_3)]\|_{N[0]} \le \delta\, K^2 \sup_{k\in\mathbb{Z}} \max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}$$

where the sum $\sum'_{k_1,k_2,k_3}$ *extends over all* k_1,k_2,k_3 outside of the range

$$|k_1| \le L, \quad k_2,k_3 \le L, \quad |k_2-k_3| \le L \tag{5.45}$$

Further, if $\sum''_{k_1,k_2,k_3}$ *denotes the sum* over *this range, then*

$$\begin{aligned}&\sum\nolimits''_{k_1,k_2,k_3} \sum_{k\le k_2-L'} \sum_{j=1}^{2} \|P_0\nabla_{t,x}[P_{k_1}\psi_1\Delta^{-1}\partial_j I^c P_k \mathcal{Q}_{\beta j}(P_{k_2}\psi_2,P_{k_3}\psi_3)]\|_{N[0]}\\ &\le \delta\, K^2 \sup_{k\in\mathbb{Z}} \max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}\end{aligned}$$

where $L' = L'(L,\delta)$ *is a large constant.*
Finally, given $\delta > 0$, *there exists* $C > 1$ *large enough such that we have*

$$\begin{aligned}&\sum_{k_1,2,3}\sum_{k<-C} \|P_0\nabla_{t,x}[P_{k_1}\psi_1\Delta^{-1}\partial_j I^c P_{<-k}Q_{>k+C}\mathcal{Q}_{\beta j}(P_{k_2}\psi_2,P_{k_3}\psi_3)]\|_{N[0]}\\ &\le \delta\, K^2 \sup_{k\in\mathbb{Z}} \max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}\end{aligned}$$

Proof. Write $\psi_i = \sum_{k_i\in\mathbb{Z}} P_{k_i}\psi_i$ for $1 \le i \le 3$. In view of the definition of the weights $w(k_1,k_2,k_3)$, summing (5.41) over all choices of k_1,k_2,k_3 yields (5.44). The second statement follows immediately from the fact that the weights $w(k_1,k_2,k_3)$ gain some smallness outside of the range (5.45) (namely $2^{-\delta L}$). For the third statement one needs to observe that in Case 5 — which is the one specified by (5.45) but of course with a range specified by the constant L — an extra gain can be obtained by restricting k to sufficiently small values compared to k_2,k_3. □

5.2. **Trilinear estimates for hyperbolic S-waves.** The following lemma finally proves the trilinear estimates in the "hyperbolic" case. The argument will rely on the following trilinear null-form expansion from [22]:

$$\begin{aligned}2\partial^\beta\psi_1\,\Delta^{-1}\partial_j\mathcal{Q}_{\beta j}(\psi_2,\psi_3) &= (\Box\psi_1)|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3 - \Box(\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3\\ &\quad + \psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3 + \Box(\psi_1\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3))\\ &\quad - (\Box\psi_1)\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3) - \psi_1\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\end{aligned} \tag{5.46}$$

as well as its "dual" form

$$\begin{aligned}2\partial^\beta[\psi_1\,\Delta^{-1}\partial_j\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] &= -\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3) + \Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2\\ &\quad - \psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3 - (\Box\psi_1)\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\\ &\quad + \Box(\psi_1\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)) + \psi_1\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\end{aligned} \tag{5.47}$$

Strictly speaking, we shall want to apply these identities to the trilinear expression

$$\partial^\beta[\psi_1\,\Delta^{-1}\partial_j I P_k\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]$$

for some P_k. In the case of (5.47) the operator IP_k can be inserted in front of any product involving ψ_2 and ψ_3 which is the case for all but the second term on the right-hand side of (5.47), i.e.,

$\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2$ (and similarly for (5.46)). Since IP_k is disposable, it takes the form of convolution with a measure ν_k with mass $\|\nu_k\|\lesssim 1$. Thus, the second term needs to be replaced by the convolution

$$\int \Box(\psi_1|\nabla|^{-1}\psi_3(\cdot-y))|\nabla|^{-1}\psi_2(\cdot-y)\nu_k(dy) \tag{5.48}$$

The logic will be that any estimate that we make on $\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2$ in the context of the $S[k]$ and $N[k]$ spaces will equally well apply to this convolution since all norms are translation invariant. We shall use this observation repeatedly in what follows without any further comment. Finally, the weights $w(k_1,k_2,k_3)$ are those specified at the beginning of this section.

Lemma 5.5. *Let ψ_j be adapted to k_j, for $j=1,2,3$. Then*

$$\Big\|\sum_{j=1}^2 P_0I\partial^\beta[IR_\alpha\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)]\Big\|_{N[0]}\lesssim w(k_1,k_2,k_3)\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \tag{5.49}$$

$$\Big\|\sum_{j=1}^2 P_0I\partial_\alpha[IR^\beta\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)]\Big\|_{N[0]}\lesssim w(k_1,k_2,k_3)\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \tag{5.50}$$

$$\Big\|\sum_{j=1}^2 P_0I\partial^\beta[IR_\beta\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\Big\|_{N[0]}\lesssim w(k_1,k_2,k_3)\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \tag{5.51}$$

for any $\alpha=0,1,2$.

Proof. We begin with (5.49). Due to the I in front of ψ_1 we shall drop the R_α operator. Also, it will be understood in this proof that $\psi_i=Q_{\le k_i+C}\psi_i$ for $1\le i\le 3$ and we will often drop the I-operator in front of the input functions.
Case 1: $0\le k_1\le k_2+O(1)=k_3+O(1)$. By Lemma 4.17,

$$\begin{aligned}\|P_0I\partial^\beta[Q_{\ge0}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)]\|_{N[0]}&\lesssim\|P_0I\partial^\beta[Q_{\ge0}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)]\|_{L_t^1L_x^2}\\&\lesssim\|Q_{\ge0}\psi_1\|_{L_t^2L_x^2}\,2^{-k_1}\|\tilde P_{k_1}I\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)]\|_{L_t^2L_x^2}\\&\lesssim 2^{-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned} \tag{5.52}$$

So it suffices to consider

$$P_0I\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)]=P_0Q_{\le C}\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_jQ_{\le C}\mathcal{Q}_{\beta j}(I\psi_2,I\psi_3)] \tag{5.53}$$

One can also limit the modulations of ψ_2,ψ_3 further. Indeed, by (4.42) of Lemma 4.13 and Corollary 4.14,

$$\begin{aligned}&\|P_0Q_{\le C}\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_jI\tilde P_{k_1}\mathcal{Q}_{\beta j}(Q_{\ge\varepsilon k_2}I\psi_2,I\psi_3)]\|_{N[0]}\\&\lesssim 2^{-k_1}\|\psi_1\nabla_{x,t}|\nabla|^{-1}I\psi_3\|_{\dot X^{0,\frac12,1}_{k_3}}\|Q_{\ge\varepsilon k_2}\nabla_{x,t}|\nabla|^{-1}I\psi_2\|_{\dot X^{0,-\frac12,1}_{k_2}}\\&\lesssim 2^{-\varepsilon k_2}\langle k_2-k_1\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned} \tag{5.54}$$

which is admissible. Note that we replaced $\Delta^{-1}\partial_j\tilde P_{k_1}$ by 2^{-k_1} as explained in the paragraph preceding this lemma. Thus, assume that $\psi_1=Q_{\le C}\psi_1$, $\psi_j=Q_{\le\varepsilon k_j}\psi_j$ for $j=2,3$, apply the identity (5.47), and estimate the six terms on the right-hand side of (5.47) in the order in which they appear. First, by the

Strichartz component (2.14),

$$\begin{aligned}\|P_0 I\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{N[0]} &\lesssim \|P_0 I\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_0^{0,-\frac12,1}}\\ &\lesssim \|\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3\|_{L_t^2L_x^2}\\ &\lesssim \|\psi\|_{L_t^\infty L_x^2}2^{-k_2}\|\psi_2\|_{L_t^4L_x^\infty}2^{-k_3}\|\psi_3\|_{L_t^4L_x^\infty}\\ &\lesssim 2^{-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Second, by (4.40) of Lemma 4.13 and Lemma 4.11,

$$\begin{aligned}\|P_0 I[\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2]\|_{N[0]} &\lesssim \langle k_3\rangle\|\tilde P_{k_3}Q_{\le\varepsilon k_3}\Box(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_3}^{0,-\frac12,1}}\||\nabla|^{-1}\psi_2\|_{S[k_2]}\\ &\lesssim \langle k_3\rangle\|\tilde P_{k_3}Q_{\le\varepsilon k_3}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_3}^{0,\frac12,1}}\|\psi_2\|_{S[k_2]}\\ &\lesssim 2^{k_1-k_3}2^{\frac14(\varepsilon k_3-k_1)}\langle k_3\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Here we use that the restriction on the modulation of the output and the modulation of ψ_2 allow us to restrict the modulation of $\Box(\psi_1|\nabla|^{-1}\psi_3)$ to size $< 2^{\varepsilon k_3}$. Third, by (4.42) and Lemma 4.11,

$$\begin{aligned}&\|P_0 I[\psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]}\\ &\lesssim \|\tilde P_{k_3}Q_{\le\varepsilon k_3}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_3}^{0,\frac12,1}}\sum_{j\le\varepsilon k_2}2^{\frac14 j\wedge 0}\|\Box Q_j(|\nabla|^{-1}\psi_2)\|_{\dot X_{k_2}^{0,-\frac12,\infty}}\\ &\lesssim 2^{k_1-k_3}2^{\frac14(3\varepsilon k_3-k_1)}\langle k_3\rangle^2\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Fourth, again by (4.42) and Lemma 4.11,

$$\begin{aligned}&\|P_0 I[(\Box\psi_1)\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}\\ &\lesssim 2^{-k_1}\sum_{\ell\le C}2^{\frac{\ell}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|\tilde P_{k_1}Q_{\le C}[R_j\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_{k_3}^{0,\frac12,1}}\\ &\lesssim 2^{\frac{k_1-k_2}{4}}2^{-\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Fifth, with $\ell = k_1 - k_2$,

$$\begin{aligned}&\|P_0 I\Box[\psi_1\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}\\ &\lesssim \|\psi_1\Delta^{-1}\partial_j\tilde P_{k_1}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{L_t^2L_x^2}\\ &\lesssim \|\psi_1\|_{L_t^\infty L_x^2}2^{-k_1}\sum_{c\in\mathcal D_{k_2,\ell}}\|P_cR_j\psi_2|\nabla|^{-1}P_{-c}\psi_3\|_{L_t^2L_x^\infty}\\ &\lesssim \|\psi_1\|_{L_t^\infty L_x^2}2^{-k_1-k_3}\Big(\sum_{c\in\mathcal D_{k_2,\ell}}\|P_cR_j\psi_2\|^2_{L_t^4L_x^\infty}\Big)^{\frac12}\Big(\sum_{c\in\mathcal D_{k_2,\ell}}\|P_{-c}\psi_3\|^2_{L_t^4L_x^\infty}\Big)^{\frac12}\\ &\lesssim \|\psi_1\|_{L_t^\infty L_x^2}2^{-k_1-k_2}2^{(1-2\varepsilon)\ell}2^{\frac{3k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\end{aligned}$$

which is admissible for small $\varepsilon > 0$. The sixth and final term is estimated by means of (4.40) and Lemma 4.11:

$$\begin{aligned}\|P_0 I[\psi_1 \Box \Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} &\lesssim \langle k_1\rangle \|\psi_1\|_{S[k_1]} \|\tilde{P}_{k_1} Q_{\le C}\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot{X}_{k_1}^{0,-\frac12,1}} \\ &\lesssim \langle k_1\rangle \|\psi_1\|_{S[k_1]} \|\tilde{P}_{k_1} Q_{\le C}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot{X}_{k_1}^{0,\frac12,1}} \\ &\lesssim \langle k_1\rangle \|\psi_1\|_{S[k_1]}\, 2^{\frac{5(k_1-k_2)}{4}} 2^{-\frac{k_2}{4}} \|\psi_2\|_{S[k_2]} \|\psi_3\|_{S[k_3]}\end{aligned}$$

which concludes Case 1.

Case 2: $0 \le k_1 = k_3 + O(1), k_2 \le k_3 - C$. By Lemma 4.23,

$$\begin{aligned}\|P_0 I\partial^\beta[Q_{\ge 0}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{N[0]} &\lesssim \|P_0 I\partial^\beta[Q_{\ge 0}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{L_t^1 L_x^2} \\ &\lesssim \|Q_{\ge 0}\psi_1\|_{L_t^2 L_x^2}\, 2^{-k_1} \|\tilde{P}_{k_1} I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{L_t^2 L_x^2} \\ &\lesssim 2^{(\frac12-\varepsilon)k_2} 2^{-(1-\varepsilon)k_1} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}\end{aligned}$$

So it suffices to consider

$$P_0 I\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)] = P_0 Q_{\le C}\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_j Q_{\le C}\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)] \tag{5.55}$$

One can also limit the modulation of ψ_3 further. Indeed, by (4.42) of Lemma 4.13 and Corollary 4.14,

$$\begin{aligned}&\|P_0 Q_{\le C}\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_j I\tilde{P}_{k_1}\mathcal{Q}_{\beta j}(I\psi_2, Q_{\ge \varepsilon k_3} I\psi_3)]\|_{N[0]} \\ &\lesssim 2^{-k_1}\|\tilde{P}_{k_3}\psi_1\nabla_{x,t}|\nabla|^{-1}I\psi_2\|_{\dot{X}_{k_3}^{0,\frac12,1}} \|Q_{\ge\varepsilon k_3}\nabla_{x,t}|\nabla|^{-1}I\psi_3\|_{\dot{X}_{k_3}^{0,-\frac12,1}} \\ &\lesssim 2^{k_2-k_1}\langle k_1-k_2\rangle 2^{-\varepsilon k_1}\prod_{i=1}^3 \|\psi_i\|_{S[k_i]}\end{aligned} \tag{5.56}$$

which is admissible. As explained in Case 1, we replaced $\Delta^{-1}\partial_j\tilde{P}_{k_1}$ by 2^{-k_1}. If $0 \le k_2$, then we can similarly reduce the modulation of the small frequency term, cf. (5.54):

$$\begin{aligned}&\|P_0 Q_{\le C}\partial^\beta[Q_{<0}\psi_1\Delta^{-1}\partial_j I\tilde{P}_{k_1}\mathcal{Q}_{\beta j}(Q_{\ge\varepsilon k_2} I\psi_2, I\psi_3)]\|_{N[0]} \\ &\lesssim 2^{-k_1}\|\tilde{P}_{k_2}[\psi_1\nabla_{x,t}|\nabla|^{-1}I\psi_3]\|_{\dot{X}_{k_2}^{0,\frac12,1}} \|Q_{\ge\varepsilon k_2}\nabla_{x,t}|\nabla|^{-1}I\psi_2\|_{\dot{X}_{k_2}^{0,-\frac12,1}} \\ &\lesssim 2^{\frac{k_2-k_1}{4}} 2^{-\varepsilon k_2}\prod_{i=1}^3 \|\psi_i\|_{S[k_i]}\end{aligned}$$

As a final preparation, we limit the modulation of the output in case $k_2 \le 0$. In fact, by Lemma 4.23,

$$\begin{aligned}&\|P_0 Q_{(1-3\varepsilon)k_2\le\cdot\le C}\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{N[0]} \\ &\lesssim \|P_0 Q_{(1-3\varepsilon)k_2\le\cdot\le C}\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{\dot{X}_{k_2}^{0,-\frac12,1}} \\ &\lesssim 2^{-\frac12(1-3\varepsilon)k_2}\|\psi_1\|_{L_t^\infty L_x^2}\, 2^{-k_1}\|\tilde{P}_{k_1} I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{L_t^2 L_x^2} \\ &\lesssim 2^{\frac12\varepsilon k_2} 2^{-(1-\varepsilon)k_1}\prod_{i=1}^3 \|\psi_i\|_{S[k_i]}\end{aligned}$$

Thus, for the remainder of this case we assume that $\psi_1 = Q_{\le C}\psi_1$, $\psi_2 = Q_{\le \varepsilon k_2\wedge k_2}\psi_2$, and $\psi_3 = Q_{\le\varepsilon k_3}\psi_3$. Moreover, the output is restricted by $Q_{\le(1-3\varepsilon)k_2\wedge C}$. We now estimate the six terms on the right-hand side of (5.47). First, by the Strichartz component (2.14),

$$\begin{aligned}\|P_0 Q_{\le(1-3\varepsilon)k_2\wedge C}\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{N[0]} &\lesssim \|P_0 Q_{\le(1-3\varepsilon)k_2\wedge C}\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_0^{0,-\frac12,1}} \\ &\lesssim 2^{\frac12(1-3\varepsilon)k_2\wedge 0}\|\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3\|_{L^2_tL^2_x} \\ &\lesssim 2^{\frac12(1-3\varepsilon)k_2\wedge 0}\|\psi_1\|_{L^\infty_tL^2_x}2^{-k_2}\|\psi_2\|_{L^4_tL^\infty_x}2^{-k_3}\|\psi_3\|_{L^4_tL^\infty_x} \\ &\lesssim 2^{\frac12(1-3\varepsilon)k_2\wedge 0}2^{-\frac{k_1}{4}}2^{-\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

which is admissible. Second, by (4.40) of Lemma 4.13 and Lemma 4.11,

$$\begin{aligned}&\|P_0 Q_{\le(1-3\varepsilon)k_2\wedge C}[\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2]\|_{N[0]} \\ &\lesssim 2^{k_2\wedge 0}\langle k_1\rangle\|\tilde P_{k_2\vee 0}Q_{\le(1-3\varepsilon)k_2\wedge\varepsilon k_2}\Box(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_2\vee 0}^{0,-\frac12,1}}\||\nabla|^{-1}\psi_2\|_{S[k_2]} \\ &\lesssim 2^{-k_2\vee 0}\langle k_1\rangle\|\tilde P_{k_2\vee 0}Q_{\le(1-3\varepsilon)k_2\wedge\varepsilon k_2}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_2\vee 0}^{0,\frac12,1}}\|\psi_2\|_{S[k_2]} \\ &\lesssim 2^{-k_2\vee 0}\langle k_1\rangle 2^{\frac{k_2\vee 0-k_1}{4}}2^{\frac14[(1-3\varepsilon)k_2\wedge\varepsilon k_2-k_1]}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

which is again admissible. Third, by (4.42) and Lemma 4.11,

$$\begin{aligned}&\|P_0 Q_{\le(1-3\varepsilon)k_2\wedge C}[\psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\ &\lesssim \|\tilde P_{k_2\vee 0}Q_{\le(1-3\varepsilon)k_2\wedge\varepsilon k_2}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_2\vee 0}^{0,\frac12,1}}\langle k_2\vee 0\rangle\|\Box(|\nabla|^{-1}Q_{\le\varepsilon k_2\wedge k_2}\psi_2)\|_{\dot X_{k_2}^{0,-\frac12,\infty}} \\ &\lesssim \langle k_2\vee 0\rangle 2^{\frac{k_2\vee 0-k_1}{4}}2^{\frac14[(1-3\varepsilon)k_2\wedge\varepsilon k_2-k_1]}2^{\frac{\varepsilon k_2\wedge k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Fourth, again by (4.42) and Lemma 4.11,

$$\begin{aligned}&\|P_0 I[(\Box\psi_1)\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\ &\lesssim 2^{-k_1}\sum_{\ell\le C}2^{\frac{\ell}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|\tilde P_{k_1}Q_{\le C}[R_j\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_{k_3}^{0,\frac12,1}} \\ &\lesssim 2^{k_2-k_1}2^{-\frac{k_1}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Fifth,

$$\begin{aligned}&\|P_0 I\Box[\psi_1\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \lesssim \|\psi_1\Delta^{-1}\partial_j\tilde P_{k_1}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{L^2_tL^2_x} \\ &\lesssim \|\psi_1\|_{L^\infty_tL^2_x}2^{-k_1}\|R_j\psi_2|\nabla|^{-1}\psi_3\|_{L^2_tL^\infty_x} \lesssim 2^{-2k_1}\|\psi_1\|_{L^\infty_tL^2_x}\|R_j\psi_2\|_{L^4_tL^\infty_x}\|\psi_3\|_{L^4_tL^\infty_x} \\ &\lesssim 2^{\frac{3k_2}{4}}2^{-\frac{5k_1}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

The sixth and final term is estimated by means of (4.40) and Lemma 4.11:

$$\begin{aligned}&\|P_0 I[\psi_1\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\ &\lesssim \langle k_1\rangle\|\psi_1\|_{S[k_1]}\|\tilde P_{k_1}Q_{\le\varepsilon k_3}\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_{k_1}^{0,-\frac12,1}} \\ &\lesssim \langle k_1\rangle\|\psi_1\|_{S[k_1]}\|\tilde P_{k_1}Q_{\le\varepsilon k_3}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_{k_1}^{0,\frac12,1}} \\ &\lesssim \langle k_1\rangle 2^{k_2-k_1}2^{-\frac14(1-\varepsilon)k_1}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

which concludes Case 2.

Case 3: $0 \leq k_1 = k_2 + O(1), k_3 \leq k_2 - C$. This is symmetric to the preceding.

Case 4: $O(1) \leq k_2 = k_3 + O(1), k_1 \leq -C$. This case proceeds similarly to Case 1. Following (5.54), we begin by limiting the modulations of ψ_2, ψ_3 to $2^{\varepsilon k_2}$. Indeed, by (4.42) of Lemma 4.13 and Corollary 4.14,

$$\begin{aligned}
&\|P_0 Q_{\leq C}\partial^\beta[I\psi_1\Delta^{-1}\partial_j I\tilde{P}_0\mathcal{Q}_{\beta j}(Q_{\geq \varepsilon k_2}I\psi_2, I\psi_3)]\|_{N[0]} \\
&\lesssim 2^{k_1-k_2}\|\psi_1\nabla_{x,t}|\nabla|^{-1}I\psi_3\|_{\dot{X}_{k_3}^{0,\frac12,1}}\|Q_{\geq\varepsilon k_2}\nabla_{x,t}|\nabla|^{-1}I\psi_2\|_{\dot{X}_{k_2}^{0,-\frac12,1}} \\
&\lesssim 2^{2k_1-k_2}2^{-\frac12\varepsilon k_2}\langle k_2-k_1\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible. Next, we limit the modulation of the output: by Lemma 4.17,

$$\begin{aligned}
&\|P_0 Q_{k_1\leq\cdot\leq C}\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{N[0]} \lesssim \|P_0 Q_{k_1\leq\cdot\leq C}\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{\dot{X}_{k_2}^{0,-\frac12,1}} \\
&\lesssim 2^{-\frac{k_1}{2}}\|\psi_1\|_{L_t^\infty L_x^\infty}\|\tilde{P}_0 I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{L_t^2L_x^2} \lesssim 2^{\frac{k_1-k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

We now again estimate the six terms on the right-hand side of (5.47). First, by the Strichartz component (2.14),

$$\begin{aligned}
\|P_0 Q_{\leq k_1}\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{N[0]} &\lesssim \|P_0 Q_{\leq k_1}\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{\dot{X}_0^{0,-\frac12,1}} \\
&\lesssim 2^{\frac{k_1}{2}}\|\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3\|_{L_t^2L_x^2} \\
&\lesssim 2^{\frac{k_1}{2}}\|\psi\|_{L_t^\infty L_x^2}2^{-k_2}\|\psi_2\|_{L_t^4L_x^\infty}2^{-k_3}\|\psi_3\|_{L_t^4L_x^\infty} \\
&\lesssim 2^{\frac{k_1-k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Second, by (4.40) of Lemma 4.13 and Lemma 4.11,

$$\begin{aligned}
&\|P_0 I[\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2]\|_{N[0]} \\
&\lesssim \langle k_3\rangle\|\tilde{P}_{k_3}Q_{\leq\varepsilon k_3}\Box(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot{X}_{k_3}^{0,-\frac12,1}}\||\nabla|^{-1}\psi_2\|_{S[k_2]} \\
&\lesssim \langle k_3\rangle\|\tilde{P}_{k_3}Q_{\leq\varepsilon k_3}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot{X}_{k_3}^{0,\frac12,1}}\|\psi_2\|_{S[k_2]} \\
&\lesssim 2^{k_1-k_3}2^{\frac14(\varepsilon k_3-k_1)}\langle k_3-k_1\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Third, by (4.42) and Lemma 4.11,

$$\begin{aligned}
&\|P_0 I[\psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \langle k_2\rangle\|\tilde{P}_{k_3}Q_{\leq\varepsilon k_3}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot{X}_{k_3}^{0,\frac12,1}}\|\Box(|\nabla|^{-1}\psi_2)\|_{\dot{X}_{k_2}^{0,-\frac12,\infty}} \\
&\lesssim 2^{k_1-k_3}2^{\frac14(3\varepsilon k_3-k_1)}\langle k_2\rangle^2\langle k_1\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Fourth, again by (4.42) and Lemma 4.11,

$$\begin{aligned}
&\|P_0 I[(\Box\psi_1)\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{\ell\leq k_1+C}2^{\frac{\ell-k_1}{4}}\|\Box Q_\ell\psi_1\|_{\dot{X}_{k_1}^{0,-\frac12,\infty}}\|\tilde{P}_0 Q_{\leq C}[R_j\psi_2|\nabla|^{-1}\psi_3]\|_{\dot{X}_{k_3}^{0,\frac12,1}} \\
&\lesssim 2^{k_1-\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Fifth, with $\ell = -k_2$,

$$\begin{aligned}
&\|P_0 Q_{\le k_1} \Box[\psi_1 \Delta^{-1}\partial_j(R_j\psi_2 |\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim 2^{k_1} \|\psi_1 \Delta^{-1}\partial_j \tilde{P}_{k_1}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{L^2_t L^2_x} \\
&\lesssim 2^{k_1} \|\psi_1\|_{L^\infty_t L^2_x} \sum_{c\in\mathcal{D}_{k_2,\ell}} \|P_c R_j \psi_2 |\nabla|^{-1} P_{-c}\psi_3\|_{L^2_t L^\infty_x} \\
&\lesssim 2^{k_1} \|\psi_1\|_{L^\infty_t L^2_x} 2^{-k_3} \Big(\sum_{c\in\mathcal{D}_{k_2,\ell}} \|P_c R_j\psi_2\|^2_{L^4_t L^\infty_x}\Big)^{\frac12} \Big(\sum_{c\in\mathcal{D}_{k_2,\ell}} \|P_{-c}\psi_3\|^2_{L^4_t L^\infty_x}\Big)^{\frac12} \\
&\lesssim 2^{k_1-k_2} 2^{(1-2\varepsilon)\ell} 2^{\frac{3k_2}{2}} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible for small $\varepsilon > 0$. The sixth and final term is estimated by means of (4.40) and Lemma 4.11:

$$\begin{aligned}
&\|P_0 I[\psi_1 \Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim 2^{k_1}\|\psi_1\|_{S[k_1]} \|\tilde{P}_0 Q_{\le k_1}\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot{X}^{0,-\frac12,1}_{k_1}} \\
&\lesssim 2^{2k_1}\|\psi_1\|_{S[k_1]} \|\tilde{P}_0(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot{X}^{0,\frac12,1}_{0}} \\
&\lesssim 2^{2k_1} 2^{-\frac{k_2}{4}} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

which concludes Case 4.

Case 5: $O(1) = k_1$, $k_2 = k_3 + O(1)$. We start with the decomposition

$$P_0\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] = \sum_{k\le k_2\wedge 0+O(1)} P_0\partial^\beta[\psi_1\Delta^{-1}\partial_j P_k I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] \tag{5.57}$$

We first limit the modulation of ψ_1:

$$\begin{aligned}
&\sum_{k\le k_2\wedge 0+O(1)} \|P_0\partial^\beta Q_{>k} I[Q_{>k+C}I\psi_1\ \Delta^{-1}\partial_j P_k I(R_\beta\psi_2 R_j\psi_3 - R_j\psi_2 R_\beta\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} \|P_0\partial^\beta Q_{>k} I[Q_{>k+C}I\psi_1\ [\Delta^{-1}\partial^2_{j\beta} P_k I(|\nabla|^{-1}\psi_2 R_j\psi_3) - P_k I(|\nabla|^{-1}\psi_2 R_\beta\psi_3)]]\|_{\dot{X}^{0,-\frac12,1}} \\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} 2^{-\frac{k}{2}} \|Q_{>k+C}\psi_1\|_{L^2_t L^2_x} \|\Delta^{-1}\partial^2_{j\beta} P_k I(|\nabla|^{-1}\psi_2 R_j\psi_3) - P_k I(|\nabla|^{-1}\psi_2 R_\beta\psi_3)]\|_{L^\infty_t L^\infty_x} \\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} 2^{k}\|\psi_1\|_{S[k_1]} \|\Delta^{-1}\partial^2_{j\beta} P_k I(|\nabla|^{-1}\psi_2 R_j\psi_3) - P_k I(|\nabla|^{-1}\psi_2 R_\beta\psi_3)]\|_{L^\infty_t L^1_x} \\
&\lesssim \sum_{k\le k_2\wedge 0+O(1)} 2^{k-k_2}\|\psi_1\|_{S[k_1]} \|\psi_2\|_{L^\infty_t L^2_x}\|\psi_3\|_{L^\infty_t L^2_x} \lesssim 2^{-k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned} \tag{5.58}$$

Hence, if the inner output has frequency $\sim 2^k$ then we may assume that ψ_1 has modulation $\lesssim 2^k$. As usual, we apply (5.47). First, by the Strichartz component (2.14),

$$\begin{aligned}
&\sum_{k\le k_2\wedge 0+C}\|P_0 I\Box(Q_{\le k}\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{N[0]}\\
&\lesssim \sum_{k\le k_2\wedge 0+C}\|P_0 Q_{\le k+C}\Box(Q_{\le k}\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{\dot X_0^{0,-\frac12,1}}\\
&\lesssim \sum_{k\le k_2\wedge 0+C} 2^{\frac{k}{2}}\|Q_{\le k}\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{L_t^2L_x^2}\lesssim \sum_{k\le k_2\wedge 0+C}2^{\frac{k-k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\\
&\lesssim 2^{-\frac12 k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

For the second term, we can assume that $\psi_1 = Q_{\le k_2\wedge 0+C}\psi_1$, see above. Then, by (4.40) of Lemma 4.13 and Lemma 4.11,

$$\begin{aligned}
&\|P_0 I[\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2]\|_{N[0]}\\
&\lesssim 2^{k_2\wedge 0}\sum_{j\le k_2\wedge 0+C}2^{\frac{j-k_2\wedge 0}{4}}\|\tilde P_{k_2\vee 0}\Box Q_j(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_2\vee 0}^{0,-\frac12,\infty}}\||\nabla|^{-1}\psi_2\|_{S[k_2]}\\
&\lesssim 2^{-k_2\vee 0}\|\tilde P_0 Q_{\le k_2\wedge 0+C}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_2\vee 0}^{0,\frac12,1}}\|\psi_2\|_{S[k_2]}\\
&\lesssim 2^{-2k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Third, by (4.42) and Lemma 4.11,

$$\begin{aligned}
&\|P_0 I[Q_{\le k_2\wedge 0+C}\psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]}\\
&\lesssim \sum_{j\le k_2\wedge 0+C}2^{\frac{j-k_2\wedge 0}{4}}\|\tilde P_{k_2\vee 0}Q_{\le k_2\wedge 0+C}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_{k_2\vee 0}^{0,\frac12,1}}\|\Box Q_j(|\nabla|^{-1}\psi_2)\|_{\dot X_{k_2}^{0,-\frac12,\infty}}\\
&\lesssim 2^{-k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Fourth, again by (4.42) and Lemma 4.11,

$$\begin{aligned}
&\sum_{k\le k_2\wedge 0+C}\|P_0 I[(\Box Q_{\le k+C}\psi_1)\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{k\le k_2\wedge 0+C}\sum_{\ell\le C}2^{\frac{\ell}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|\tilde P_k Q_{\le k+C}[R_j\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_k^{0,\frac12,1}}\\
&\lesssim \sum_{k\le k_2\wedge 0+C}\|\psi_1\|_{\dot X_{k_1}^{0,\frac12,\infty}}2^{\frac{k-k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}2^{-\frac12 k_2\vee 0}\lesssim\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Fifth, with $\ell = k - k_2$,

$$
\begin{aligned}
&\sum_{k\le k_2\wedge 0+C} \|P_0 I\Box[Q_{\le k+C}\psi_1 \Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{k\le k_2\wedge 0+C} 2^{\frac{k}{2}} \|Q_{\le k+C}\psi_1 \Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{L^2_t L^2_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{k\le k_2\wedge 0+C} 2^{-\frac{k}{2}} \sum_{c\in\mathcal{D}_{k_2,\ell}} \|P_c R_j\psi_2 |\nabla|^{-1} P_{-c}\psi_3\|_{L^2_t L^\infty_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{k\le k_2\wedge 0+C} 2^{-\frac{k}{2}-k_3} \Big(\sum_{c\in\mathcal{D}_{k_2,\ell}} \|P_c R_j\psi_2\|^2_{L^4_t L^\infty_x}\Big)^{\frac12} \Big(\sum_{c\in\mathcal{D}_{k_2,\ell}} \|P_{-c}\psi_3\|^2_{L^4_t L^\infty_x}\Big)^{\frac12} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{k\le k_2\wedge 0+C} 2^{(\frac12-2\varepsilon)(k-k_2)} \|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{-(\frac12-2\varepsilon)k_2\vee 0} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is admissible for small $\varepsilon > 0$. The sixth and final term is estimated by means of (4.40) and Lemma 4.11:

$$
\begin{aligned}
&\sum_{k\le k_2\wedge 0+C} \|P_0 I[Q_{\le k+C}\psi_1 \Box\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim \|\psi_1\|_{S[k_1]} \sum_{k\le k_2\wedge 0+C} 2^k \|P_k Q_{\le k+C}\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_k^{0,-\frac12,1}} \\
&\lesssim \|\psi_1\|_{S[k_1]} \sum_{k\le k_2\wedge 0+C} 2^k \|P_k Q_{\le k+C}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_k^{0,\frac12,1}} \\
&\lesssim \|\psi_1\|_{S[k_1]} \sum_{k\le k_2\wedge 0+C} 2^k\, 2^{\frac{k-k_2}{2}} \|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim 2^{-\frac12 k_2\vee 0} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which concludes Case 5.

Case 6: $O(1) = k_1 \ge k_2 + O(1) \ge k_3 + C$. Since Lemma 4.23 implies that

$$
\begin{aligned}
\|P_0\partial^\beta[Q_{>k_2}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L^1_t L^2_x} &\lesssim \|Q_{>k_2}\psi_1\|_{L^2_t L^2_x} 2^{-k_2} \|I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L^2_t L^\infty_x} \\
&\lesssim 2^{(\frac12-\varepsilon)(k_3-k_2)} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

we may assume that $\psi_1 = Q_{\le k_2}\psi_1$. Next, we reduce matters to (5.1). More precisely,

$$
\begin{aligned}
&(5.59) \qquad P_0 I\partial^\beta[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] = P_0 I[\partial^\beta\psi_1\, \tilde P_{k_2}\Delta^{-1}\partial_j I\tilde P_{k_2}\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] \\
&(5.60) \qquad\qquad + P_0 I[I\psi_1\,\Delta^{-1}\partial_j\partial^\beta \tilde P_{k_2} I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]
\end{aligned}
$$

The term in (5.59) satisfies the bounds (5.37) and (5.38), whereas (5.60) is expanded further:

$$
\begin{aligned}
&(5.61) \qquad P_0 I[I\psi_1\,\Delta^{-1}\partial_j\partial^\beta\tilde P_{k_2} I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] = P_0 I[I\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I(\Box|\nabla|^{-1}\psi_2 R_j\psi_3 - R_j\psi_2\Box|\nabla|^{-1}\psi_3 \\
&(5.62) \qquad\qquad + R_\beta\psi_2\partial^\beta R_j\psi_3 - \partial^\beta R_j\psi_2 R_\beta\psi_3)]
\end{aligned}
$$

The two terms in (5.62) are again controlled by (5.1). Consider the first term on the right-hand side of (5.61). Replacing $\Delta^{-1}\partial_j\tilde P_{k_2}$ by 2^{-k_2} as usual, one obtains from Lemmas 4.13 and 4.11,

$$
\begin{aligned}
\|\psi_1\,\Box|\nabla|^{-1}\psi_2\, R_j\psi_3\|_{N[0]} &\lesssim 2^{k_2} \sum_{j\le k_2+C} 2^{\frac{j-k_2}{4}} \|\Box Q_j|\nabla|^{-1}\psi_2\|_{\dot X_{k_2}^{0,-\frac12,\infty}} \|\psi_1\, R_j\psi_3\|_{\dot X_0^{0,\frac12,1}} \\
&\lesssim 2^{k_2}\|\psi_2\|_{\dot X_{k_2}^{0,\frac12,\infty}} 2^{k_3}\langle k_3\rangle \|\psi_1\|_{S[k_1]}\|\psi_3\|_{S[k_3]}
\end{aligned}
$$

which is more than enough. The second term in (5.61) is estimated similarly:

$$
\begin{aligned}
\|\psi_1\,\Box|\nabla|^{-1}\psi_3\,R_j\psi_2\|_{N[0]} &\lesssim 2^{k_3}\sum_{j\le k_3+C}2^{\frac{j-k_3}{4}}\|\Box Q_j|\nabla|^{-1}\psi_3\|_{\dot X_{k_3}^{0,-\frac12,\infty}}\|\psi_1\,R_j\psi_2\|_{\dot X_0^{0,\frac12,1}}\\
&\lesssim 2^{k_3}\|\psi_2\|_{\dot X_{k_2}^{0,\frac12,\infty}}2^{k_2}\langle k_2\rangle\|\psi_1\|_{S[k_1]}\|\psi_2\|_{S[k_2]}
\end{aligned}
$$

which concludes Case 6.

Case 7: $k_1 = O(1) \ge k_3 + O(1) \ge k_2 + C$. This case is symmetric to the previous one.

Case 8: $k_3 = O(1), \max(k_1, k_2) \le -C$. By Lemma 4.23,

$$
\begin{aligned}
&\|P_0 I\partial^\beta[Q_{\ge k_1+(1-3\varepsilon)k_2}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{N[0]}\\
&\lesssim \|P_0 I\partial^\beta[Q_{\ge k_1+(1-3\varepsilon)k_2}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{L_t^1L_x^2}\\
&\lesssim 2^{k_1}\|Q_{\ge k_1+(1-3\varepsilon)k_2}\psi_1\|_{L_t^2L_x^2}\,\|\tilde P_0 I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim 2^{\frac12\varepsilon k_2}2^{\frac{k_1}{2}}\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

A similar calculation shows that one can place $Q_{\le k_1+(1-3\varepsilon)k_2}$ in front of the entire output. So it suffices to consider

$$
\begin{aligned}
&P_0 Q_{\le k_1+(1-3\varepsilon)k_2}\partial^\beta[Q_{<k_1+(1-3\varepsilon)k_2}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]\\
&= P_0 Q_{\le k_1+(1-3\varepsilon)k_2}\partial^\beta[Q_{<k_1+(1-3\varepsilon)k_2}\psi_1\Delta^{-1}\partial_j Q_{\le k_1+C}\tilde P_0\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]
\end{aligned}
$$

We now stimate the six terms on the right-hand side of (5.47). First, by the Strichartz component (2.14),

$$
\begin{aligned}
&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{N[0]}\\
&\lesssim \|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}\Box(\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_0^{0,-\frac12,1}}\\
&\lesssim 2^{\frac12[(1-3\varepsilon)k_2+k_1]}\|\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3\|_{L_t^2L_x^2}\\
&\lesssim 2^{\frac12[(1-3\varepsilon)k_2+3k_1]}\|\psi_1\|_{L_t^\infty L_x^2}\|\psi_2\|_{L_t^4L_x^\infty}2^{-\frac{k_2}{4}}\|\psi_3\|_{L_t^4L_x^\infty}\\
&\lesssim 2^{\frac12[(1-3\varepsilon)k_2+3k_1]}2^{-\frac{k_2}{4}}\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is sufficient. Second, by (4.40) of Lemma 4.13 and Lemma 4.11,

$$
\begin{aligned}
&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[\Box(\psi_1|\nabla|^{-1}\psi_3)|\nabla|^{-1}\psi_2]\|_{N[0]}\\
&\lesssim 2^{k_2}\|\tilde P_0 Q_{\le(1-3\varepsilon)k_2}\Box(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_0^{0,-\frac12,1}}\||\nabla|^{-1}\psi_2\|_{S[k_2]}\\
&\lesssim \|\tilde P_0 Q_{\le(1-3\varepsilon)k_2}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_0^{0,\frac12,1}}\|\psi_2\|_{S[k_2]}\\
&\lesssim 2^{k_1}2^{\frac{(1-3\varepsilon)k_2-k_1}{4}\wedge 0}\langle k_1\rangle\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is admissible. Third, by (4.42) and Lemma 4.11,

$$
\begin{aligned}
&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[\psi_1\Box(|\nabla|^{-1}I\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]}\\
&\lesssim \|\tilde P_0 Q_{\le(1-3\varepsilon)k_2}(\psi_1|\nabla|^{-1}\psi_3)\|_{\dot X_0^{0,\frac12,1}}\;\|\Box(|\nabla|^{-1}I\psi_2)\|_{\dot X_{k_2}^{0,-\frac12,\infty}}\\
&\lesssim 2^{k_1}2^{\frac{(1-3\varepsilon)k_2-k_1}{4}\wedge 0}\langle k_1\rangle\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}
$$

Fourth, again by (4.42) and Lemma 4.11,

$$\begin{aligned}&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[(\Box Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1)\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}\\ &\lesssim 2^{k_1}\sum_{\ell\le k_1+(1-3\varepsilon)k_2} 2^{\frac{\ell-k_1}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|\tilde P_0 Q_{\le k_1+C}[R_j\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_{k_3}^{0,\frac12,1}}\\ &\lesssim 2^{2k_1}2^{\frac14(1-3\varepsilon)k_2}2^{k_2}\langle k_2\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

Fifth,

$$\begin{aligned}&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}\Box[\psi_1\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}\\ &\lesssim 2^{\frac12[(1-3\varepsilon)k_2+k_1]}\|\psi_1\Delta^{-1}\partial_j\tilde P_{k_1}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{L_t^2L_x^2}\\ &\lesssim 2^{\frac12[(1-3\varepsilon)k_2+k_1]}\|\psi_1\|_{L_t^\infty L_x^2}\|R_j\psi_2|\nabla|^{-1}\psi_3\|_{L_t^2L_x^\infty}\\ &\lesssim 2^{\frac12[(1-3\varepsilon)k_2+k_1]}\|\psi_1\|_{L_t^\infty L_x^2}\|R_j\psi_2\|_{L_t^4L_x^\infty}\|\psi_3\|_{L_t^4L_x^\infty}\\ &\lesssim 2^{\frac12[(1-3\varepsilon)k_2+k_1]}2^{\frac{3k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

The sixth and final term is estimated by means of (4.40) and Lemma 4.11:

$$\begin{aligned}&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}\\ &\lesssim 2^{k_1}\|\psi_1\|_{S[k_1]}\|\tilde P_0 Q_{\le k_1+C}\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_{k_1}^{0,-\frac12,1}}\\ &\lesssim 2^{k_1}\|\psi_1\|_{S[k_1]}\|\tilde P_{k_3}Q_{\le k_1}(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_{k_3}^{0,\frac12,1}}\\ &\lesssim 2^{k_1+k_2}\langle k_2\rangle\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

which concludes Case 8.

Case 9: $k_2=O(1),\max(k_1,k_3)\le -C$. Symmetric to Case 8.

Hence we are done with (5.49). Next, we turn to (5.50) which is similar; basically, one uses (5.46) instead of (5.47). First, one observes that any reductions in modulation which preceded application of (5.47) to (5.49) can equally well be carried out for (5.50) since these bounds only use Lemmas 4.17 and 4.23. Second, observe that the last four terms of (5.46) reappear as the last four terms of (5.47) up to the order and the choice of signs, both of which are irrelevant. Consequently, one only needs to verify that the first two terms of (5.46) satisfy the desired bounds.

Case 1: $0\le k_1\le k_2+O(1)=k_3+O(1)$. In this case the second terms in (5.46) and (5.47) satisfy the same bounds, whence it will suffice to bound the first term in (5.46). However, by (4.42) and Lemma 4.11,

$$\begin{aligned}&\|P_0 I[(\Box\psi_1)|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{N[0]}\\ &\lesssim\sum_{\ell\le C}2^{\frac{\ell}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|\tilde P_{k_1}Q_{\le C}[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_{k_3}^{0,\frac12,1}}\\ &\lesssim 2^{\frac{k_1-k_2}{4}}2^{-\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

which is admissible.

Case 2: $0\le k_1=k_3+O(1),k_2\le k_3-C$. Using the arguments from Case 2 above, we may assume that $\psi_1=Q_{\le(1-3\varepsilon)k_2\wedge 0-k_1}\psi_1$. In addition, it was shown there that it suffices to assume that $\psi_2=Q_{\le\varepsilon k_2\wedge k_2}\psi_2$,

$\psi_3 = Q_{\leq \varepsilon k_2}\psi_3$. First,

$$\begin{aligned}
&\|P_0 I(Q_{\leq(1-3\varepsilon)k_2\wedge 0-k_1}\Box\psi_1\; Q_{\leq C}[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{N[0]}\\
&\lesssim 2^{\frac{(1-3\varepsilon)k_2\wedge 0-k_1}{4}}\|\Box Q_{\leq(1-3\varepsilon)k_2\wedge 0-k_1}\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|Q_{\leq C}[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_{k_1}^{0,\frac12,1}}\\
&\lesssim 2^{\frac{(1-3\varepsilon)k_2\wedge 0-k_1}{4}}\langle k_1-k_2\rangle\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible. One may also restrict the modulation of the entire output by $Q_{\leq(1-3\varepsilon)k_2\wedge 0-k_1}$. Applying Lemma 4.13 and Lemma 4.11 to the second expression in (5.46) yields

$$\begin{aligned}
&\|P_0 Q_{\leq(1-3\varepsilon)k_2\wedge 0-k_1}[\Box Q_{\leq(1-3\varepsilon)k_2\wedge\varepsilon k_2}(\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]}\\
&\lesssim 2^{\frac{(1-3\varepsilon)k_2\wedge-(1-\varepsilon)k_2}{4}}\|\tilde P_{k_1}Q_{\leq(1-3\varepsilon)k_2\wedge\varepsilon k_2}\Box(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot X_{k_1}^{0,-\frac12,1}}\||\nabla|^{-1}\psi_3\|_{S[k_3]}\\
&\lesssim 2^{\frac{(1-3\varepsilon)k_2\wedge-(1-\varepsilon)k_2}{4}}\|Q_{\leq(1-3\varepsilon)k_2\wedge\varepsilon k_2}(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot X_{k_1}^{0,\frac12,1}}\|\psi_3\|_{S[k_3]}\\
&\lesssim 2^{\frac{(1-3\varepsilon)k_2\wedge-(1-\varepsilon)k_2}{4}}\langle k_1-k_2\rangle\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible provided $|k_2|>\gamma k_1$ for some $\gamma>0$. When this condition is violated, we have to work a little harder. First, since we may choose $\gamma>0$ arbitrarily small and the ensuing estimates won't be affected by our choice of γ, we may from now as well assume $k_2=O(1)$. With the modulation and frequency restrictions from above in place, and going back to the original (un-expanded) version of the term under consideration, write schematically

$$\nabla_{x,t}[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] = \psi_1(\nabla^{-1}\psi_3)\psi_2$$

where we suppress the action of convolution operators of bounded L^1-mass, as they don't affect our estimates. Then we get (for some $\delta_1>0$ small)

$$\begin{aligned}
&\|P_0 I[\psi_1 Q_{\geq-(1-\delta_1)k_1}(\nabla^{-1}\psi_3)\psi_2]\|_{N[0]}\\
&\lesssim\|\psi_1\psi_2\|_{\dot X_{k_1}^{0,\frac12,1}}\|(\nabla^{-1}Q_{\geq-(1-\delta_1)k_1}\psi_3\|_{\dot X_{k_3}^{0,-\frac12,1}}\lesssim 2^{-\delta_1 k_1}\prod_{j=1}^{3}\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}$$

In light of the modulation restrictions from earlier, we now reduce to estimating

$$P_0 I[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)\psi_2]$$

Decompose this expression into

$$P_0 I[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)\psi_2]$$

$$= P_0 I\big[P_{<-\delta_2 k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big] \tag{5.63}$$

$$+ P_0 I\big[P_{\geq-\delta_2 k_1}Q_{<-\delta_3 k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big] \tag{5.64}$$

$$+ P_0 I\big[P_{\geq-\delta_2 k_1}Q_{\geq-\delta_3 k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big] \tag{5.65}$$

where we pick $0<\delta_2<\delta_3<<1$. For the first term on the right, suppressing the action of the convolution operator $P_{<-\delta_2 k_1}$ and reverting to the original form, we have reduced to estimating the term

$$P_0\nabla_{x,t}[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2,\psi_3)] = \sum_{\kappa,\kappa'\in\mathcal{C}_{-\delta_2 k_1},\,\kappa\sim\kappa'} P_{0,\kappa'}\nabla_{x,t}[\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(P_{0,\kappa}\psi_2,\psi_3)],$$

where the point is of course that we can localize the output as well as the small-frequency input ψ_2 to approximately the same small angular sector. If we then make the null-form expansion as at the beginning of Case 2 above, we reduce to estimating

$$\begin{aligned}&\|\sum_{\kappa,\kappa'\in\mathcal{C}_{-\delta_2 k_1},\,\kappa\sim\kappa'}P_{0,\kappa'}Q_{\leq(1-3\varepsilon)k_2\wedge 0-k_1}(Q_{\leq(1-3\varepsilon)k_2\wedge\varepsilon k_2}\Box(\psi_1|\nabla|^{-1}P_{0,\kappa}\psi_2)|\nabla|^{-1}\psi_3])\|_{N[0]}\\&\lesssim\Big(\sum_{\kappa\in\mathcal{C}_{-\delta_2 k_1}}\|P_0Q_{\leq(1-3\varepsilon)k_2\wedge 0-k_1}(Q_{\leq(1-3\varepsilon)k_2\wedge\varepsilon k_2}\Box(\psi_1|\nabla|^{-1}P_{0,\kappa}\psi_2)|\nabla|^{-1}\psi_3])\|_{N[0]}^2\Big)^{\frac12}\\&\lesssim 2^{-\delta_4 k_1}\prod_{j=1}^3\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

where we have used (the proof of) Corollary 4.10, which concludes estimating the contribution of (5.63). As for that of (5.64), here we can write

$$\begin{aligned}&P_0I\big[P_{\geq-\delta_2k_1}Q_{<-\delta_3k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big]\\&=\sum_{\kappa,\kappa'\in\mathcal{C}_{-(1+\delta_2+\frac{\delta_3-\delta_2}{2})k_1}}P_0I\big[P_{\geq-\delta_2k_1}Q_{<-\delta_3k_1}[P_{k_1,\kappa}Q_{<-k_1}\psi_1(\nabla^{-1}P_{k_3,\kappa'}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big]\\&=\sum_{O(1)>l>-\delta_2k_1}\sum_{\kappa,\kappa'\in\mathcal{C}_{-k_1+l+\frac{l-k_1}{2}}}P_0I\big[P_lQ_{<-\delta_3k_1}[P_{k_1,\kappa}Q_{<-k_1}\psi_1(\nabla^{-1}P_{k_3,\kappa'}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big]\\&=\sum_{O(1)>l>-\delta_2k_1}\sum_{\kappa,\kappa'\in\mathcal{C}_{-k_1+l+\frac{-l-\delta_3k_1}{2}},\kappa\sim\mp\kappa'}P_0I\big[P_{l,\kappa''}Q_{<-\delta_3k_1}[P_{k_1,\kappa}Q_{<-k_1}\psi_1(\nabla^{-1}P_{k_3,\kappa'}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big]\end{aligned}$$

where in the last sum κ'' ranges over the $O(1)$ many caps in $\mathcal{C}_{\frac{-l-\delta_3k_1}{2}}$ such that either one of $\pm\kappa''$ is at distance $\lesssim 2^{\frac{-l-\delta_3k_1}{2}}$ from κ. As the operator $P_{l,\kappa''}Q_{<-\delta_3k_1}$ is given by convolution with a kernel of bounded L^1-mass, we can then again suppress it and revert to estimating the expression

$$\sum_{\kappa,\kappa'\in\mathcal{C}_{-k_1+l+\frac{-l-\delta_3k_1}{2}},\kappa\sim\mp\kappa'}P_0\nabla_{x,t}[P_{k_1,\kappa}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(\psi_2,P_{k_3,\kappa'}\psi_3)]$$

where we have suppressed the implicit dependence on l (coming from the suppressed action of $P_{l,\kappa''}Q_{<-\delta_3k_1}$). Due to the preceding identity, it is easy to see that we may write

$$\begin{aligned}&\sum_{\kappa,\kappa'\in\mathcal{C}_{-k_1+l+\frac{-l-\delta_3k_1}{2}},\kappa\sim\mp\kappa'}P_0\nabla_{x,t}[P_{k_1,\kappa}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(\psi_2,P_{k_3,\kappa'}\psi_3)]\\&=\sum_{\kappa,\kappa'\in\mathcal{C}_{-k_1+l+\frac{-l-\delta_3k_1}{2}},\kappa\sim\mp\kappa'}\sum_{\tilde\kappa_{1,2}\in\mathcal{C}_{\frac{-l-\delta_3k_1}{2}},\tilde\kappa_1\sim\pm\kappa+\tilde\kappa_2}P_{0,\tilde\kappa_2}\nabla_{x,t}[P_{k_1,\kappa}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(P_{0,\tilde\kappa_1}\psi_2,P_{k_3,\kappa'}\psi_3)]\end{aligned}$$

For fixed κ,κ', one can now again expand the null-form as at the beginning of Case 2, and as for (5.63), one then gets for fixed κ,κ' that

$$\begin{aligned}&\|\sum_{\tilde\kappa_{1,2}\in\mathcal{C}_{\frac{-l-\delta_3k_1}{2}},\tilde\kappa_1\sim\pm\kappa+\tilde\kappa_2}P_{0,\tilde\kappa_2}\nabla_{x,t}[P_{k_1,\kappa}\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\beta j}(P_{0,\tilde\kappa_1}\psi_2,P_{k_3,\kappa'}\psi_3)]\|_{N[0]}\\&\qquad\lesssim 2^{-\delta_5k_1}\|P_{k_1,\kappa}\psi_1\|_{S[k_1]}\|P_0\psi_2\|_{S[k_2]}\|P_{k_3,\kappa'}\psi_3\|_{S[k_3]}\end{aligned}$$

for some $\delta_5 > 0$ depending on $\delta_3 >> \delta_2$. One can then perform the summation over κ, κ' with the aid of Cauchy-Schwarz, and the summation over l only costs logarithmically. In conclusion, we get

$$\begin{aligned}
&\| \sum_{O(1)>l>-\delta_2 k_1} \sum_{\kappa,\kappa'\in\mathcal{C}_{-k_1+l+\frac{-l-\delta_3 k_1}{2}},\kappa\sim\mp\kappa'} P_0\nabla_{x,t}[P_{k_1,\kappa}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(\psi_2, P_{k_3,\kappa'}\psi_3)]\|_{N[0]} \\
&\lesssim 2^{-\delta_6 k_1}\prod_{j=1}^{3}\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}$$

for some small $\delta_6 > 0$, which concludes the contribution of (5.64). Finally, we consider the contribution of (5.65). Here we take advantage of Lemma 4.11, which gives

$$\|\big[P_l Q_{\geq -\delta_3 k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big]\|_{\dot{X}_l^{0,\frac{1}{2},\infty}} \lesssim 2^{\frac{l-k_1}{4}}\|Q_{<-k_1}\psi_1\|_{S[k_1]}\|Q_{<-(1-\delta_1)k_1}\psi_3\|_{S[k_3]}$$

Then we use Lemma 4.13, which gives

$$\begin{aligned}
&\|P_0 I\big[P_{O(1)\geq\cdot\geq-\delta_2 k_1}Q_{\geq-\delta_3 k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\psi_2\big]\|_{N[0]} \\
&\lesssim \sum_{O(1)>l>-\delta_2 k_1}\|P_l Q_{\geq-\delta_3 k_1}[Q_{<-k_1}\psi_1(\nabla^{-1}Q_{<-(1-\delta_1)k_1}\psi_3)]\|_{\dot{X}_l^{0,-\frac{1}{2},1}}\|\psi_2\|_{S[k_2]} \\
&\lesssim 2^{(\delta_3-\frac{1}{4})k_1}\prod_{j=1}^{3}\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}$$

Case 3: $0 \leq k_1 = k_2 + O(1), k_3 \leq k_2 - C$. This is symmetric to Case 2.

Case 4: $O(1) \leq k_2 = k_3 + O(1), k_1 \leq -C$. This is similar to Case 1. Indeed, the second terms in (5.46) and (5.47) satisfy the same bounds, whence it will suffice to bound the first term in (5.46). However, by (4.42) and Lemma 4.11,

$$\begin{aligned}
&\|P_0 I[(\Box I\psi_1)|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \sum_{\ell\leq k_1+C}2^{\frac{\ell-k_1}{4}}\|\Box Q_\ell\psi_1\|_{\dot{X}_{k_1}^{0,-\frac{1}{2},\infty}}\|\tilde{P}_0 Q_{\leq C}[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{\dot{X}_{k_2}^{0,\frac{1}{2},1}} \\
&\lesssim 2^{k_1-k_2}\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible.

Case 5: $O(1) = k_1,\ k_2 = k_3 + O(1)$. Here again it suffices to only consider the first term in (5.46). Moreover, (5.57) and (5.58) apply whence that first term is bounded by the Strichartz component (2.14):

$$\begin{aligned}
&\sum_{k\leq k_2\wedge 0+C}\|P_0 I(Q_{\leq k}\Box\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{N[0]} \\
&\lesssim \sum_{k\leq k_2\wedge 0+C}\|P_0 Q_{\leq k+C}(Q_{\leq k}\Box\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{L_t^1 L_x^2} \\
&\lesssim \sum_{k\leq k_2\wedge 0+C}\|Q_{\leq k}\Box\psi_1\|_{L_t^2 L_x^2}\|P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{L_t^2 L_x^\infty} \\
&\lesssim \sum_{k\leq k_2\wedge 0+C}2^{k-\frac{k_2}{2}}\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]} \lesssim 2^{-\frac{k_2\vee 0}{2}}\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Case 6: $O(1) = k_1 \geq k_2 + O(1) \geq k_3 + C$. Here one basically starts from (5.59), which can be handled via (5.1).

Case 7: $k_1 = O(1) \geq k_3 + O(1) \geq k_2 + C$. This case is symmetric to the previous one.

Case 8: $k_3 = O(1), \max(k_1, k_2) \le -C$. As in Case 8 above, one first shows that one can place $Q_{\le k_1+(1-3\varepsilon)k_2}$ in front of the entire output, as well as in front of ψ_1. So it suffices to consider

$$\begin{aligned}
&P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[Q_{<k_1+(1-3\varepsilon)k_2}\partial^\beta \psi_1 \Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)] \\
&= P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[Q_{<k_1+(1-3\varepsilon)k_2}\partial^\beta \psi_1 \Delta^{-1}\partial_j Q_{\le k_1+C}\tilde{P}_0\mathcal{Q}_{\beta j}(I\psi_2, I\psi_3)]
\end{aligned}$$

We now stimate the first two terms on the right-hand side of (5.46). First, by the Strichartz component (2.14),

$$\begin{aligned}
&\|P_0 I(Q_{\le k_1+(1-3\varepsilon)k_2}\Box\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{N[0]} \\
&\lesssim \|P_0 I(Q_{\le k_1+(1-3\varepsilon)k_2}\Box\psi_1|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3)\|_{L_t^1 L_x^2} \\
&\lesssim 2^{(1-3\varepsilon)k_2+k_1}\|\psi_1\|_{L_t^\infty L_x^2}\||\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3\|_{L_t^2 L_x^\infty} \\
&\lesssim 2^{(1-3\varepsilon)k_2+k_1}\|\psi_1\|_{L_t^\infty L_x^2}\|\psi_2\|_{L_t^4 L_x^\infty}2^{-\frac{k_2}{4}}\|\psi_3\|_{L_t^4 L_x^\infty} \\
&\lesssim 2^{(1-3\varepsilon)k_2+k_1}2^{-\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is sufficient. Second, by (4.40) of Lemma 4.13 and Lemma 4.11, and assuming first that $k_1 = k_2 + O(1)$,

$$\begin{aligned}
&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[\Box(Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \sum_{k\le k_1+C}\|P_0[\Box Q_{\le k_1+C}P_k(Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \sum_{k\le k_1+C}\|P_k Q_{\le k_1+C}\Box(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot{X}_k^{0,-\frac{1}{2},1}}\|\,|\nabla|^{-1}\psi_3\|_{S[k_3]} \\
&\lesssim \sum_{k\le k_1+C}2^k\|P_k Q_{\le k_1+C}(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot{X}_k^{0,\frac{1}{2},1}}\|\psi_3\|_{S[k_3]} \\
&\lesssim \sum_{k\le k_1+C}2^k 2^{\frac{k-k_1}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible. If $k_2 < k_1 - C$, then by the same lemmas,

$$\begin{aligned}
&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[\Box(Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \|P_0[\Box Q_{\le(1-3\varepsilon)k_2+C}\tilde{P}_{k_1}(Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \|\tilde{P}_{k_1}Q_{\le(1-3\varepsilon)k_2+C}\Box(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot{X}_{k_1}^{0,-\frac{1}{2},1}}\|\,|\nabla|^{-1}\psi_3\|_{S[k_3]} \\
&\lesssim 2^{k_1}\|\tilde{P}_{k_1}Q_{\le(1-3\varepsilon)k_2+C}(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot{X}_{k_1}^{0,\frac{1}{2},1}}\|\psi_3\|_{S[k_3]} \\
&\lesssim 2^{k_1}2^{\frac{(1-3\varepsilon)k_2-k_1}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is again admissible. Finally, if $k_1 < k_2 - C$, then arguing analogously yields

$$\begin{aligned}
&\|P_0 Q_{\le k_1+(1-3\varepsilon)k_2}[\Box(Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \|P_0[\Box Q_{\le k_2}\tilde{P}_{k_2}(Q_{\le k_1+(1-3\varepsilon)k_2}\psi_1|\nabla|^{-1}\psi_2)|\nabla|^{-1}\psi_3]\|_{N[0]} \\
&\lesssim \|\tilde{P}_{k_2}Q_{\le k_2}\Box(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot{X}_{k_2}^{0,-\frac{1}{2},1}}\|\,|\nabla|^{-1}\psi_3\|_{S[k_3]} \\
&\lesssim 2^{k_2}\|\tilde{P}_{k_2}Q_{\le k_2}(\psi_1|\nabla|^{-1}\psi_2)\|_{\dot{X}_{k_2}^{0,\frac{1}{2},1}}\|\psi_3\|_{S[k_3]} \\
&\lesssim 2^{k_1}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which concludes this case.

Case 9: $k_2 = O(1), \max(k_1, k_3) \leq -C$. Symmetric to Case 8. This concludes the analysis of (5.50).

Neither of the identities (5.46) or (5.47) applies to (5.51). Hence, (5.51) requires somewhat different arguments.
Case 1: $0 \leq k_1 \leq k_2 + O(1) = k_3 + O(1)$. As in (5.52) one sees that it suffices to consider $\psi_1 = Q_{\leq 0}\psi_1$. Then $\mathcal{Q}_{\alpha j} = Q_{\leq C}\mathcal{Q}_{\alpha j}$ and we split

$$\begin{aligned} & P_0 I\partial^\beta[IR_\beta\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)] \\ & = \sum_{\ell\leq C} P_0 Q_{\leq \ell - C}\partial^\beta[R_\beta Q_{\leq \ell-C}\psi_1\Delta^{-1}\partial_j Q_\ell \mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)] \qquad (5.66)\\ & + \sum_{\ell - C\leq \ell_1\leq C} P_0 Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq \ell_1}\psi_1\Delta^{-1}\partial_j Q_\ell \mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)] \qquad (5.67)\\ & + \sum_{\ell - C\leq \ell_2\leq C} P_0 Q_{<\ell_2}\partial^\beta[R_\beta Q_{\ell_2}\psi_1\Delta^{-1}\partial_j Q_\ell \mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)] \qquad (5.68) \end{aligned}$$

Decomposing (5.66) via Lemma 4.1 into caps of size $2^{\frac{\ell}{2}}$ yields

$$\begin{aligned} & P_0 Q_{\leq \ell - C}\partial^\beta[R_\beta Q_{\leq \ell-C}\psi_1\Delta^{-1}\partial_j Q_\ell \mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)] \\ & = \sum_{\kappa\sim\kappa'\in\mathcal{C}_{\frac{\ell}{2}}} P_{0,\kappa} Q_{\leq \ell - C}\partial^\beta[R_\beta Q_{\leq \ell-C}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_j Q_\ell \mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)] \end{aligned}$$

where $\kappa\sim\kappa'$ denotes that these caps have distance about $2^{\frac{\ell}{2}}$. Hence we gain a factor of 2^ℓ from the nullform involving ∂^β and R_β. From (2.29) one now obtains

$$\begin{aligned} \|(5.66)\|_{N[0]} & \lesssim \sum_{\ell\leq C}\Big(\sum_{\kappa\sim\kappa'\in\mathcal{C}_{\frac{\ell}{2}}} \|P_{0,\kappa} Q_{\leq \ell - C}\partial^\beta[R_\beta Q_{\leq \ell-C}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_j Q_\ell \mathcal{Q}_{\alpha j}(I\psi_2, I\psi_3)]\|_{\mathrm{NF}[\kappa]}^2\Big)^{\frac{1}{2}} \\ & \lesssim \sum_{\ell\leq C} 2^\ell 2^{-\frac{\ell}{4}}2^{\frac{k_1}{2}}\Big(\sum_{\kappa\in\mathcal{C}_{\frac{\ell}{2}}}\|P_{k_1,\kappa}Q_{\leq\ell-C}\psi_1\|_{S[\kappa]}^2 2^{-2k_1}\|Q_\ell\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{L_t^2L_x^2}^2\Big)^{\frac{1}{2}} \\ & \lesssim \sum_{\ell\leq C} 2^{\frac{3\ell}{4}}2^{\frac{k_1}{2}}\|\psi_1\|_{S[k_1]}2^{-\frac{k_1}{4+}}2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \\ & \lesssim 2^{\frac{k_1}{4-}-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \end{aligned}$$

Here we also used Lemma 2.7 as well as Lemma 4.17. The expressions in (5.67) are decomposed into caps of size $2^{\frac{\ell_1}{2}}$ but without separation. Therefore, with a gain of 2^{ℓ_1} from the outer null-form,

$$\begin{aligned} \|(5.67)\|_{N[0]} & \lesssim \sum_{\ell-C\leq\ell_1\leq C}\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{\ell_1}{2}}}\|P_{0,\kappa}Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_j Q_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|_{\dot X_0^{0,-\frac{1}{2},1}}^2\Big)^{\frac{1}{2}} \\ & \lesssim \sum_{\ell-C\leq\ell_1\leq C} 2^{-\frac{\ell_1}{2}}\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{\ell_1}{2}}}\|P_{0,\kappa}Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_j Q_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|_{L_t^2L_x^2}^2\Big)^{\frac{1}{2}} \\ & \lesssim \sum_{\ell-C\leq\ell_1\leq C}\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{\ell_1}{2}}}\|P_{0,\kappa}Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_j Q_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|_{L_t^2L_x^1}^2\Big)^{\frac{1}{2}} \qquad (5.69) \end{aligned}$$

To pass to (5.69) one invokes the improved Bernstein estimate of Lemma 2.1. Hence, this can be further bounded by

$$\lesssim \sum_{\ell-C\leq\ell_1\leq C} 2^{\ell_1}\Big(\sum_{\kappa'\in\mathcal{C}_{\frac{\ell_1}{2}}} \|Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\|^2_{L^\infty_t L^2_x}\, 2^{-2k_1}\|Q_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|^2_{L^2_tL^2_x}\Big)^{\frac12}$$
$$\lesssim \sum_{\ell-C\leq\ell_1\leq C} 2^{\ell_1}\Big(\sum_{\kappa'\in\mathcal{C}_{\frac{\ell_1}{2}}} \|Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\|^2_{S[k_1,\kappa']}\Big)^{\frac12} 2^{-k_1}2^{\frac{\ell-k_1}{4+}}2^{k_1-\frac{k_2}{2}}\|Q_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|_{L^2_tL^2_x}$$
$$\lesssim 2^{-\frac{k_1}{4-}-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}$$

For (5.68) one proceeds similarly, performing a cap decomposition and placing the entire expression in $L^1_tL^2_x$. We skip the details.

Case 2: $0\leq k_1=k_3+O(1), k_2\leq k_3-C$. This is essentially the same as the preceding with Lemma 4.23 replacing Lemma 4.17.

Case 3: $0\leq k_1=k_2+O(1), k_3\leq k_2-C$. This is symmetric to the preceding.

Case 4: $O(1)\leq k_2=k_3+O(1), k_1\leq -C$. This is very similar to Case 1. First, one checks that the entire output can be restricted by $Q_{\leq k_1}$. This implies that due to the I-operator in front of ψ_1, the decomposition (5.66)–(5.68) continues to hold but with $\ell\leq k_1+C$:

$$P_0I\partial^\beta[IR_\beta\psi_1\Delta^{-1}\partial_jI\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]$$
$$=\sum_{\ell\leq k_1+C}P_0Q_{\leq\ell-C}\partial^\beta[R_\beta Q_{\leq\ell-C}\psi_1\Delta^{-1}\partial_jQ_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)] \tag{5.70}$$
$$+\sum_{\ell-C\leq\ell_1\leq k_1+C}P_0Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq\ell_1}\psi_1\Delta^{-1}\partial_jQ_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)] \tag{5.71}$$
$$+\sum_{\ell-C\leq\ell_2\leq k_1+C}P_0Q_{<\ell_2}\partial^\beta[R_\beta Q_{\ell_2}\psi_1\Delta^{-1}\partial_jQ_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)] \tag{5.72}$$

One can again decompose (5.70) into caps, but of size $2^{\frac{\ell-k_1}{2}}$. Therefore,

$$\|(5.70)\|_{N[0]}\lesssim\sum_{\ell\leq k_1+C}\Big(\sum_{\kappa\sim\kappa'\in\mathcal{C}_{\frac{\ell-k_1}{2}}}\|P_{0,\kappa}Q_{\leq\ell-C}\partial^\beta[R_\beta Q_{\leq\ell-C}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_jQ_\ell\tilde P_0\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|^2_{\mathrm{NF}[\kappa]}\Big)^{\frac12}$$
$$\lesssim\sum_{\ell\leq k_1+C}2^{\frac{3(\ell-k_1)}{4}}2^{\frac{k_1}{2}}\Big(\sum_{\kappa\in\mathcal{C}_{\frac{\ell}{2}}}\|P_{k_1,\kappa}Q_{\leq\ell-C}\psi_1\|^2_{S[\kappa]}\|Q_\ell\tilde P_0\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|^2_{L^2_tL^2_x}\Big)^{\frac12}$$
$$\lesssim\sum_{\ell\leq k_1+C}2^{\frac{3(\ell-k_1)}{4}}2^{\frac{k_1}{2}}\|\psi_1\|_{S[k_1]}2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}$$
$$\lesssim 2^{\frac{k_1}{2}-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}$$

which is admissible. Furthermore, $\|(5.71)\|_{N[0]}$ is bounded by

$$\lesssim \sum_{\ell-C\leq\ell_1\leq k_1+C}\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{\ell_1-k_1}{2}}}\|P_{0,\kappa}Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_jQ_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|^2_{\dot X_0^{0,-\frac12,1}}\Big)^{\frac12}$$

$$\lesssim \sum_{\ell-C\leq\ell_1\leq k_1+C}2^{-\frac{\ell_1}{2}}\Big(\sum_{\kappa,\kappa'\in\mathcal{C}_{\frac{\ell_1-k_1}{2}}}\|P_{0,\kappa}Q_{\ell_1}\partial^\beta[R_\beta Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\Delta^{-1}\partial_jQ_\ell\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|^2_{L^2_tL^2_x}\Big)^{\frac12}$$

$$\lesssim \sum_{\ell-C\leq\ell_1\leq k_1+C}2^{\frac{\ell_1}{2}-k_1}\Big(\sum_{\kappa'\in\mathcal{C}_{\frac{\ell_1-k_1}{2}}}\|Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\|^2_{L^\infty_tL^\infty_x}\,\|Q_\ell\tilde P_0\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|^2_{L^2_tL^2_x}\Big)^{\frac12}$$

$$\lesssim \sum_{\ell-C\leq\ell_1\leq k_1+C}2^{\frac{\ell_1}{2}}2^{\frac{\ell_1-k_1}{4}}\Big(\sum_{\kappa'\in\mathcal{C}_{\frac{\ell_1-k_1}{2}}}\|Q_{\leq\ell_1}P_{k_1,\kappa'}\psi_1\|^2_{S[k_1,\kappa']}\Big)^{\frac12}2^{\frac{\ell}{4+}-\frac{k_2}{2}}\|Q_\ell\tilde P_0\mathcal{Q}_{\alpha j}(I\psi_2,I\psi_3)]\|_{L^2_tL^2_x}$$

$$\lesssim 2^{\frac{3k_1}{4+}-\frac{k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}$$

Finally, (5.72) is similar to the previous estimate and we skip it.

Case 5: $O(1)=k_1,\ k_2=k_3+O(1)$. We apply (5.57) and reduce the modulation of ψ_1 via (5.58) to $\psi_1=Q_{\leq k}\psi_1$. Furthermore,

$$P_0\partial^\beta I[R_\beta Q_{\leq k}\psi_1\,\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] = P_0I[\Box|\nabla|^{-1}Q_{\leq k}\psi_1\,\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \tag{5.73}$$
$$+\,P_0I[R_\beta Q_{\leq k}\psi_1\,\Delta^{-1}\partial_j\partial^\beta P_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \tag{5.74}$$

Lemmas 4.13 and 4.17 imply the following bound on (5.73):

$$\sum_{k\leq k_2\wedge 0+C}\|P_0I[\Box|\nabla|^{-1}Q_{\leq k}\psi_1\,\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]}$$
$$\lesssim \sum_{k\leq k_2\wedge 0+C}\sum_{m\leq k}2^{\frac{m-k}{4}}\|\Box|\nabla|^{-1}Q_m\psi_1\|_{\dot X_0^{0,-\frac12,1}}\|\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{\dot X_k^{0,\frac12,1}}$$
$$\lesssim \sum_{k\leq k_2\wedge 0+C}2^{\frac{k-k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\lesssim 2^{-\frac12 k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}$$

which is admissible. The second term (5.74) needs to be expanded as follows:

$$2R_\beta Q_{\leq k}\psi_1\,\Delta^{-1}\partial_j\partial^\beta P_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)=\Box[Q_{\leq k}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \tag{5.75}$$
$$-\,\Box Q_{\leq k}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) \tag{5.76}$$
$$-\,Q_{\leq k}|\nabla|^{-1}\psi_1\,\Box\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) \tag{5.77}$$

We just dealt with the term (5.76). Since the modulation of the entire output is $\lesssim 2^k$, one concludes that

$$(5.75)\lesssim \sum_{k\leq k_2\wedge 0+C}\|\Box[Q_{\leq k}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{\dot X_0^{0,-\frac12,1}}$$
$$\lesssim \sum_{k\leq k_2\wedge 0+C}2^{\frac{k}{2}}\|\psi_1\|_{L^\infty_tL^2_x}\|\Delta^{-1}\partial_jP_kI\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{L^2_tL^\infty_x}$$
$$\lesssim \sum_{k\leq k_2\wedge 0+C}2^{\frac{k-k_2}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\lesssim 2^{-\frac12 k_2\vee 0}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}$$

as well as, from Lemma 4.13,

$$
\begin{aligned}
(5.77) &\lesssim \sum_{k\le k_2\wedge 0+C} \|Q_{\le k}|\nabla|^{-1}\psi_1 \,\Box\Delta^{-1}\partial_j P_k I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{k\le k_2\wedge 0+C} \|\psi_1\|_{L_t^\infty L_x^2} \|\Box\Delta^{-1}\partial_j P_k I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{\dot X_0^{0,-\frac12,1}} \\
&\lesssim \sum_{k\le k_2\wedge 0+C} 2^{\frac{3k-k_2}{2}} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]} \lesssim 2^{-\frac12 k_2\vee 0} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is sufficient.

Case 6: $O(1) = k_1 \ge k_2 + O(1) \ge k_3 + C$. As before, one reduces the modulation of ψ_1 to $\psi_1 = Q_{\le k_2}\psi_1$. Furthermore,

$$
\begin{aligned}
&(5.78) \qquad P_0\partial^\beta I[R_\beta Q_{\le k_2}\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] = P_0 I[\Box|\nabla|^{-1}Q_{\le k_2}\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \\
&(5.79) \qquad\qquad + P_0 I[R_\beta Q_{\le k_2}\psi_1\,\Delta^{-1}\partial_j\partial^\beta\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]
\end{aligned}
$$

Lemmas 4.13 and 4.23 imply the following bound on (5.78):

$$
\begin{aligned}
&\|P_0 I[\Box|\nabla|^{-1}Q_{\le k_2}\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{m\le k_2} 2^{\frac{m-k_2}{4}} \|\Box|\nabla|^{-1}Q_m\psi_1\|_{\dot X_0^{0,-\frac12,1}} \|\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{\dot X_{k_2}^{0,\frac12,1}} \\
&\lesssim 2^{(\frac12-\varepsilon)(k_3-k_2)} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which is admissible. The second term (5.79) needs to be expanded as follows:

$$
\begin{aligned}
&(5.80) \qquad 2R_\beta Q_{\le k_2}\psi_1\,\Delta^{-1}\partial_j\partial^\beta\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) = \Box[Q_{\le k_2}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \\
&(5.81) \qquad\qquad - \Box Q_{\le k_2}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) \\
&(5.82) \qquad\qquad - Q_{\le k_2}|\nabla|^{-1}\psi_1\,\Box\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)
\end{aligned}
$$

We just dealt with the term (5.81). Since the modulation of the entire output is $\lesssim 2^{k_2}$, one concludes that

$$
\begin{aligned}
(5.80) &\lesssim \|\Box[Q_{\le k_2}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{\dot X_0^{0,-\frac12,1}} \\
&\lesssim 2^{\frac{k_2}{2}} \|\psi_1\|_{L_t^\infty L_x^2} \|\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{L_t^2 L_x^\infty} \\
&\lesssim 2^{(\frac12-\varepsilon)(k_3-k_2)} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]} \lesssim \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

as well as, from Lemma 4.13,

$$
\begin{aligned}
(5.82) &\lesssim \|Q_{\le k}|\nabla|^{-1}\psi_1\,\Box\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]} \\
&\lesssim \|\psi_1\|_{S[k_1]} \|\Box\Delta^{-1}\partial_j\tilde P_{k_2} I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{\dot X_0^{0,-\frac12,1}} \\
&\lesssim 2^{(\frac12-\varepsilon)(k_3-k_2)} 2^{k_2} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}
$$

which concludes Case 6.

Case 7: $k_1 = O(1) \ge k_3 + O(1) \ge k_2 + C$. This case is symmetric to the previous one.

Case 8: $k_3 = O(1), \max(k_1, k_2) \le -C$. The modulation of the output can be reduced to $Q_{\le k_1}$:

$$\begin{aligned}
&\|P_0\partial^\beta I Q_{\ge k_1}[R_\beta Q_{\le k_1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim 2^{-\frac{k_1}{2}}\|\psi_1\|_{L_t^\infty L_x^\infty}\|\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{L_t^2L_x^2}\\
&\lesssim 2^{\frac{k_1}{2}}2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Similarly, the input ψ_1 can be reduced to $Q_{\le k_1}\psi_1$. As in Case 6,

$$P_0\partial^\beta Q_{\le k_1}[R_\beta Q_{\le k_1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 Q_{\le k_1}\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]$$

$$= P_0Q_{\le k_1}[\Box|\nabla|^{-1}Q_{\le k_1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 Q_{\le k_1}\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \tag{5.83}$$

$$+ P_0Q_{\le k_1}[R_\beta Q_{\le k_1}\psi_1\,\Delta^{-1}\partial_j\partial^\beta\tilde{P}_0 Q_{\le k_1}\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \tag{5.84}$$

Lemmas 4.13 and 4.23 imply the following bound on (5.83):

$$\begin{aligned}
&\|P_0Q_{\le k_1}[\Box|\nabla|^{-1}Q_{\le k_1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{m\le k_1}2^{\frac{m-k_1}{4}}\|\Box|\nabla|^{-1}Q_m\psi_1\|_{\dot X_{k_1}^{0,-\frac12,1}}\|\tilde{P}_0 Q_{\le k_1}\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{\dot X_0^{0,\frac12,1}}\\
&\lesssim 2^{\frac{k_1}{2}}2^{(\frac12-\varepsilon)k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which is admissible. The second term (5.84) needs to be expanded as follows:

$$2R_\beta Q_{\le k_1}\psi_1\,\Delta^{-1}\partial_j\partial^\beta\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) = \Box[Q_{\le k_1}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)] \tag{5.85}$$

$$-\Box Q_{\le k_1}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) \tag{5.86}$$

$$-Q_{\le k_1}|\nabla|^{-1}\psi_1\,\Box\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3) \tag{5.87}$$

We just dealt with the term (5.86). Next,

$$\begin{aligned}
(5.85) &\lesssim \|\Box Q_{\le k_1}[Q_{\le k_1}|\nabla|^{-1}\psi_1\,\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{\dot X_0^{0,-\frac12,1}}\\
&\lesssim 2^{\frac{k_1}{2}}\||\nabla|^{-1}\psi_1\|_{L_t^\infty L_x^\infty}\|\Delta^{-1}\partial_j\tilde{P}_0 I\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{L_t^2L_x^2}\\
&\lesssim 2^{\frac{k_1}{2}}2^{(\frac12-\varepsilon)k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

as well as, from Lemma 4.13,

$$\begin{aligned}
(5.87) &\lesssim \|Q_{\le k_1}|\nabla|^{-1}\psi_1\,\Box\Delta^{-1}\partial_j\tilde{P}_0 Q_{\le k_1}\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)]\|_{N[0]}\\
&\lesssim \|\psi_1\|_{S[k_1]}\|\Box\Delta^{-1}\partial_j\tilde{P}_0 Q_{\le k_1}\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\|_{\dot X_0^{0,-\frac12,1}}\\
&\lesssim 2^{(\frac12-\varepsilon)k_3}2^{\frac{k_1}{2}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

which concludes Case 8.

Case 9: $k_2 = O(1), \max(k_1, k_3) \le -C$. Symmetric to Case 8. □

Remark 5.6. It follows from the high-low-low interaction case of the proof of Lemma 5.5 that for some $\sigma > 0$,

$$\big\|P_0 I[P_{k_1}I\psi_1\,\partial^\beta\partial_j\Delta^{-1}P_k I\mathcal{Q}_{\beta j}(P_{k_2}\psi_2, P_{k_3}\psi_3)\big\|_{N[0]} \lesssim 2^{\sigma k}\,w(k_1,k_2,k_3)\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \tag{5.88}$$

provided $k_1 = O(1)$, $k \le k_2 = k_3 + O(1) \le O(1)$.

In effect, for later use, we also mention the following lemma, which is proved using identical reasoning:

Lemma 5.7. *Assume $k_1 = O(1)$. Then we have the bounds*

$$\|P_0\big[IR^\beta\psi_1\partial_\beta P_k\Delta^{-1}\partial_j\mathcal{Q}_{\alpha j}(\psi_2,\psi_3)\big]\|_{N[0]} \lesssim 2^{\sigma k}w(k_1,k_2,k_3)\prod_{i=1}^{3}\|P_{k_i}\psi_i\|_{S[k_i]}$$

$$\|P_0\big[IR^\beta\psi_1\partial_\alpha P_k\Delta^{-1}\partial_j\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\big]\|_{N[0]} \lesssim 2^{\sigma k}w(k_1,k_2,k_3)\prod_{i=1}^{3}\|P_{k_i}\psi_i\|_{S[k_i]}$$

for suitable $\sigma > 0$.

5.3. **Improved trilinear estimates with angular alignment.** We conclude this section on trilinear bounds with a technical result which we shall require in several instances, such as the blow-up criterion of the following section. By Corollary 5.4, one gains extra smallness outside of the parameter range (5.45); note that the latter describes precisely Case 5 in the proof of Lemmas 5.1 and 5.5 which is the high-low-low case of interactions. In fact, the exact same gain as in that corollary can also be obtained for the trilinear expressions of Lemma 5.5.

Corollary 5.8. *The nonlinearities of Lemma 5.5 satisfy the estimates of Corollary 5.4. I.e., given $\delta > 0$ there exist L, L' large so that the δ–gains in the sum over $\sum'_{k_1,k_2,k_3}$ as well as $\sum''_{k_1,k_2,k_3}$ with $k \le k_2 - L'$, are obtained for the three types of trilinear null-forms in Lemma 5.5.*

Proof. As in the case of Corollary 5.4, this follows from the form of the weights $w(k_1, k_2, k_3)$ as well as from the fact that an extra gain in Case 5 of Lemma 5.5 was obtained when $k < k_2 - L'$. □

However, one cannot gain smallness in the high-low-low case without further assumptions. In this section we shall prove that *angular alignment* between the Fourier support of at least two of the inputs implies smallness in this case.

We start with the contributions by $I^c\mathcal{Q}_{\beta j}$. In Corollary 5.4 we isolated one case where smallness cannot be obtained without any further assumptions. It was given by the sum $\sum''_{k_1,k_2,k_3\in\mathbb{Z}}$ over the range (5.45) together with $k_2 - L' \le k \le k_2 + O(1)$. Recall that L and L' are very large depending on δ. Throughout this section, ψ_i will be Schwartz functions satisfying

$$\max_{i=1,2,3}\sum_{k\in\mathbb{Z}}\|P_k\psi_i\|^2_{S[k]} \le K^2$$

for some constant K. We shall use $\sum''_{k_1,k_2,k_3\in\mathbb{Z}}$ repeatedly in the sense that it was defined earlier.

Lemma 5.9. *Given any $\delta > 0$ there exists $m_0(\delta)$ large and negative such that*

$$\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}}\ {\sum}''_{k_1,k_2,k_3\in\mathbb{Z}}\ \sum_{k=k_2-L'}^{k_2+O(1)}\sum_{j=1}^{2}\|P_0\nabla_{t,x}[P_{k_1}\psi_1\Delta^{-1}\partial_j I^cP_k\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[0]}$$
$$\le \delta\, K^2\sup_{k\in\mathbb{Z}}\max_{i=1,2,3}2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}$$

as well as

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\le 2^{m_0}}}\ {\sum}''_{k_1,k_2,k_3\in\mathbb{Z}}\ \sum_{k=k_2-L'}^{k_2+O(1)}\sum_{j=1}^{2}\|P_0\nabla_{t,x}[P_{k_1,\kappa_1}\psi_1\Delta^{-1}\partial_j I^cP_k\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2, P_{k_3}\psi_3)]\|_{N[0]}$$
$$+\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}}\ {\sum}''_{k_1,k_2,k_3\in\mathbb{Z}}\ \sum_{k=k_2-L'}^{k_2+O(1)}\sum_{j=1}^{2}\|P_0\nabla_{t,x}[P_{k_1,\kappa_1}\psi_1\Delta^{-1}\partial_j I^cP_k\mathcal{Q}_{\beta j}(P_{k_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[0]}$$
$$\le \delta\, K^2\sup_{k\in\mathbb{Z}}\max_{i=1,2,3}2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}$$

Proof. The proof simply consists in verifying that the argument in Case 5 of Lemma 5.1 allows for this extra gain. We first consider angular alignment between ψ_1 and ψ_2. In this case, we will need to repeat the argument of Case 5, obtaining the gain from Bernstein's inequality. First, restrict the output by $Q_{\geq 0}$ and assume that $\psi_1 = P_{k_1,\kappa_1}\psi_1$ and $\psi_2 = P_{k_2,\kappa_2}\psi_2$ with fixed caps κ_1, κ_2. In the end, one verifies that it is possible to sum over these caps. Then

$$\|P_0 Q_{\geq 0}\partial^\beta[\psi_1 \Delta^{-1}\partial_j I^c \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}$$

$$\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} \|P_0 Q_{\geq 0}[\psi_1 \Delta^{-1}\partial_j Q_{k\leq\cdot\leq C} P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \tag{5.89}$$

$$+ \sum_{k=k_2-L'}^{k_2+O(1)} \sum_{m\geq C} 2^{-\varepsilon m}\|P_0 Q_m[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \tag{5.90}$$

$$+ \sum_{k=k_2-L'}^{k_2+O(1)} \sum_{m\geq C} \|P_0 Q_{\geq 0}[Q_{>m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \tag{5.91}$$

First, by Lemma 4.19, and with M large but finite and $\frac{1}{p}+\frac{1}{M}=\frac{1}{2}$,

$$\begin{aligned}
(5.89) &\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} \|\psi_1\|_{L_t^\infty L_x^p}\,\|Q_{k\leq\cdot\leq C} P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^M} \\
&\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} 2^{m_0(\frac{1}{2}-\frac{1}{p})}\|\psi_1\|_{L_t^\infty L_x^2}\, 2^{\frac{k}{2}} 2^{-\varepsilon k_2} 2^{|k_2|\frac{2}{M}}\,\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \leq \delta \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Since $p>2$ one can take m_0 large and negative to obtain the final estimate here. Second, again by Lemma 4.19,

$$\begin{aligned}
(5.90) &\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} \sum_{m\geq C} 2^{-\varepsilon m}\|P_0 Q_m[Q_{\leq m-C}\psi_1 \Delta^{-1}\partial_j \tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} \sum_{m\geq C} 2^{-\varepsilon m}\|\psi_1\|_{L_t^\infty L_x^p}\; 2^{-k}\|\tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^M} \\
&\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} 2^{m_0(\frac{1}{2}-\frac{1}{p})} 2^{\frac{k}{2}} 2^{-\varepsilon k_2} 2^{|k_2|\frac{2}{M}} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]} \leq \delta \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

and third,

$$\begin{aligned}
(5.91) &\lesssim \sum_{k=k_2-L'}^{k_2+O(1)} \sum_{m\geq C} \|Q_{>m-C}\psi_1\|_{L_t^2 L_x^p} 2^{-k}\|\tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^\infty L_x^M} \\
&\lesssim 2^{|k_2|\frac{2}{M}} \sum_{k=k_2-L'}^{k_2+O(1)} 2^{m_0(\frac{1}{2}-\frac{1}{p})} \sum_{m\geq C} 2^{-(1-\varepsilon)m}\|\psi_1\|_{S[k_1]}\, 2^{\frac{m}{2}}\,\|\tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^M} \\
&\lesssim 2^{m_0(\frac{1}{2}-\frac{1}{p})} 2^{|k_2|\frac{2}{M}} \sum_{k=k_2-L'}^{k_2+O(1)} \sum_{m\geq C} 2^{-(\frac{1}{2}-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\, 2^{-\varepsilon m}\|\tilde{Q}_m P_k \mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2 L_x^2} \\
&\leq \delta \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

where one argues as in the previous two cases to pass to the last line. Next, suppose the output is limited by $Q_{\leq 0}$. Then

$$
\begin{aligned}
&\|P_0 Q_{\leq 0}\partial^\beta[I^c\psi_1\Delta^{-1}\partial_j I^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \lesssim \sum_{k=k_2-L'}^{k_2+O(1)}\sum_{m\geq C}\|Q_m\psi_1\Delta^{-1}\partial_j P_k\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^1L_x^2} \\
&\lesssim \sum_{k=k_2-L'}^{k_2+O(1)}\sum_{m\geq C}\|Q_m\psi_1\|_{L_t^2L_x^p}\,2^{-k}\|P_k\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^M} \\
&\lesssim 2^{m_0(\frac{1}{2}-\frac{1}{p})}\sum_{k=k_2-L'}^{k_2+O(1)}\sum_{m\geq C}2^{-(1-2\varepsilon)m}\|\psi_1\|_{S[k_1]}\,2^{-\varepsilon m}2^{|k_2|\frac{2}{M}}\|P_k\tilde{Q}_m\mathcal{Q}_{\beta j}(\psi_2,\psi_3)\|_{L_t^2L_x^2} \\
&\lesssim 2^{m_0(\frac{1}{2}-\frac{1}{p})}\sum_{k=k_2-L'}^{k_2+O(1)}\|\psi_1\|_{S[k_1]}\,2^{\frac{k}{2}}2^{-\varepsilon k_2}2^{|k_2|\frac{2}{M}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\leq\delta\prod_{i=1}^{3}\|\psi_i\|_{S[k_i]}
\end{aligned}
\tag{5.92}
$$

which is again admissible. On the other hand, assume now that $\psi_1 = I\psi_1$. Then, as we may suppose that $k = k_2 + O(1) = k_3 + O(1)$, we get

$$
\begin{aligned}
&\|P_0 Q_{\leq 0}\partial^\beta[I\psi_1\Delta^{-1}\partial_j I^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]} \leq \|P_0 Q_{\leq 0}\partial^\beta[I\psi_1\Delta^{-1}\partial_j I^c\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{\dot{X}_0^{-1,-\frac{1}{2},1}+L_t^1\dot{H}^{-1}} \\
&\lesssim \sum_{O(1)>l\geq k_2+C_1} 2^{-\frac{l}{2}}\|Q_{<l-10}\psi_1\|_{L_t^\infty L_x^2}\|\Delta^{-1}\partial_j Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^\infty} \\
&+ \sum_{O(1)>l\geq k_2+C_1} 2^{-\frac{l}{2}}\|Q_{\geq l-10}\psi_1\|_{L_t^\infty L_x^2}\|\Delta^{-1}\partial_j Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{L_t^2L_x^\infty} \\
&+ \sum_{k_2+C<l<k_2+C_1}\|P_0 Q_{\leq 0}\partial^\beta[I\psi_1\Delta^{-1}\partial_j Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]}
\end{aligned}
\tag{5.93}
$$

Here we have chosen C_1 large enough depending on δ, while C is as in the definition of I^c. Then using Lemma 4.19, we infer that the first two amongst the last three preceding terms are bounded by

$$\lesssim \delta\prod_{j=1}^{3}\|P_{k_j}\psi_j\|_{S[k_j]}$$

and summation over the angular sectors/frequencies is straightforward to give the bound of the lemma. On the other hand, for the last expression

$$\sum_{k_2+C<l<k_2+C_1}\|P_0 Q_{\leq 0}\partial^\beta[I\psi_1\Delta^{-1}\partial_j Q_l\mathcal{Q}_{\beta j}(\psi_2,\psi_3)]\|_{N[0]},$$

we use Lemma 4.17 to give the same bound. To conclude the case of angular alignment between ψ_1,ψ_2, we sum over κ_1,κ_2 using Cauchy-Schwarz, Lemma 2.18, and Corollary 4.21.
Finally, consider the case where ψ_2 and ψ_3 are aligned on the Fourier side. Using Lemmas 4.18 and 4.22 instead of Lemmas 4.17 and 4.19, respectively, one immediately verifies that the desired gain can indeed be obtained. The only exception here is the estimate (5.26). But this case is excluded here as it involves $I\mathcal{Q}_{\beta j}$ and not $I^c\mathcal{Q}_{\beta j}$. □

Next, we need to obtain an analogous statement in the hyperbolic regime of the inner nullform. As in Corollary 5.4, Lemma 5.5 implies the following result.

Corollary 5.10. *Let $\delta > 0$ be small. Then*

$$\begin{aligned}
&\sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k\le k_2-L'} \Big\| \sum_{j=1}^{2} P_0\partial^\beta[R_\alpha P_{k_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(P_{k_2}\psi_2,P_{k_3}\psi_3)]\Big\|_{N[0]}\\
&+\sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k\le k_2-L'} \Big\| \sum_{j=1}^{2} P_0\partial_\alpha[R^\beta P_{k_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(P_{k_2}\psi_2,P_{k_3}\psi_3)]\Big\|_{N[0]}\\
&+\sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k\le k_2-L'} \Big\| \sum_{j=1}^{2} P_0\partial_\beta[R^\beta P_{k_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\alpha j}(P_{k_2}\psi_2,P_{k_3}\psi_3)]\Big\|_{N[0]}\\
&\le \delta\, K^2 \sup_{k\in\mathbb{Z}}\max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}
\end{aligned}$$

where $L' = L'(L,\delta)$ is a large constant.

Next, we need to obtain an improvement in the range (5.45) under the additional assumption of angular alignment.

Lemma 5.11. *For any $\delta > 0$ there exists $m_0(\delta)$, a large negative constant, such that*

$$\begin{aligned}
&\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k=k_2-L'}^{k_2+O(1)} \Big\| \sum_{j=1}^{2} P_0\partial^\beta[R_\alpha P_{k_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2,P_{k_3,\kappa_3}\psi_3)]\Big\|_{N[0]}\\
&\le \delta\, K^2 \sup_{k\in\mathbb{Z}}\max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}
\end{aligned}$$

as well as

$$\begin{aligned}
&\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k=k_2-L'}^{k_2+O(1)} \Big\| \sum_{j=1}^{2} P_0\partial_\alpha[R^\beta P_{k_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2,P_{k_3,\kappa_3}\psi_3)]\Big\|_{N[0]}\\
&+\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k=k_2-L'}^{k_2+O(1)} \Big\| \sum_{j=1}^{2} P_0\partial_\beta[R^\beta P_{k_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2,P_{k_3,\kappa_3}\psi_3)]\Big\|_{N[0]}\\
&\le \delta\, K^2 \sup_{k\in\mathbb{Z}}\max_{i=1,2,3} 2^{-\sigma_0|k|}\|P_k\psi_i\|_{S[k]}
\end{aligned}$$

for any $\alpha = 0,1,2$. An analogous statement holds in case ψ_1,ψ_2 or ψ_1,ψ_3 are similarly aligned.

Proof. We begin with the first trilinear form, and also assume alignment between ψ_2 and ψ_3. We first reduce ourselves to the purely hyperbolic case, i.e., when all inputs are restricted by the operator I, as well as the entire output. Without further mention, implicit constants are allowed to depend on L, L'. In particular, we assume that k, k_1, k_2, k_3 are fixed in the range we are summing over. In the notation of

Lemma 5.1, if $A_0 = I^c$, then $A_1 = I^c$ and by Lemma 4.18,

$$\begin{aligned}
&\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \|P_0 I^c \partial^\beta[I^c\nabla_{t,x}\psi_1 \Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \sum_{m\ge 0} 2^{-\varepsilon m}\|P_0 Q_m[\tilde{Q}_m\nabla_{t,x}\psi_1\, \Delta^{-1}\partial_j P_k I\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{L^2_t L^2_x} \\
&\lesssim \sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \sum_{m\ge 0} 2^{(1-\varepsilon)m}\|\tilde{Q}_m\psi_1\|_{L^2_t L^2_x} 2^{-k}\|P_k I\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{L^\infty_t L^\infty_x} \\
&\lesssim \sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\le 2^{m_0}}} \|\psi_1\|_{S[k_1]}\, 2^{\frac{k}{2}}\|IP_k\mathcal{Q}_{\beta j}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)\|_{L^2_t L^2_x} \\
&\lesssim \delta 2^{\frac{k}{2}}\|\psi_1\|_{S[k_1]} 2^{\frac{k_2}{2}}\|P_{k_2}\psi_2\|_{S[k_2]}\|P_{k_3}\psi_3\|_{S[k_3]} \le \delta\|\psi_1\|_{S[k_1]}\|P_{k_2}\psi_2\|_{S[k_2]}\|P_{k_3}\psi_3\|_{S[k_3]}
\end{aligned}$$

Summing over $k_1 = O(1)$, $k_2 = k_3 + O(1)$ yields the desired gain. Hence, we can assume that $A_0 = I$ as well as $A_1 = I$. If $A_2 = I^c$, then also $A_3 = I^c$ and

$$\begin{aligned}
&\|P_0 I\partial^\beta[I\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I^c P_{k_2,\kappa_2}\psi_2, I^c P_{k_3,\kappa_3}\psi_3)]\|_{N[0]} \\
&\lesssim \|I\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(I^c P_{k_2,\kappa_2}\psi_2, I^c P_{k_3,\kappa_3}\psi_3)]\|_{L^1_t L^2_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} 2^{-k_2}\|IP_k\mathcal{Q}_{\beta j}(I^c P_{k_2,\kappa_2}\psi_2, I^c P_{k_3,\kappa_3}\psi_3)\|_{L^1_t L^\infty_x} \\
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{m\ge k_2+C} 2^{-k_2}\|\mathcal{Q}_{\beta j}(Q_m P_{k_2,\kappa_2}\psi_2, \tilde{Q}_m P_{k_3,\kappa_3}\psi_3)\|_{L^1_t L^\infty_x}
\end{aligned}$$

In the last inequality we use that we may assume $k = k_2 + O(1)$. Splitting the modulations of the last two inputs dyadically yields

$$\begin{aligned}
&\lesssim \|\psi_1\|_{L^\infty_t L^2_x} \sum_{m\ge k_2+C} 2^{m-2k_2}\|P_{k_2,\kappa_2}Q_m\psi_2\|_{L^2_t L^\infty_x}\|P_{k_3,\kappa_3}\tilde{Q}_m\psi_3\|_{L^2_t L^\infty_x} \\
&\lesssim 2^{m_0+k_2}\|\psi_1\|_{L^\infty_t L^2_x} \sum_{m\ge k_2+C} 2^{m-k_2}2^{-2(1-\varepsilon)m}2^{(1-2\varepsilon)k_2}\|P_{k_2,\kappa_2}Q_m\psi_2\|_{\dot{X}^{-\frac{1}{2}+\varepsilon,1-\varepsilon,\infty}_{k_2}}\|P_{k_3,\kappa_3}\tilde{Q}_m\psi_3\|_{\dot{X}^{-\frac{1}{2}+\varepsilon,1-\varepsilon,\infty}_{k_3}} \\
&\lesssim 2^{m_0}\|\psi_1\|_{S[k_1]}\|P_{k_2,\kappa_2}\psi_2\|_{\dot{X}^{-\frac{1}{2}+\varepsilon,1-\varepsilon,2}_{k_2}}\|P_{k_3,\kappa_3}\psi_3\|_{\dot{X}^{-\frac{1}{2}+\varepsilon,1-\varepsilon,2}_{k_3}}
\end{aligned}$$

Summing over the caps κ_2, κ_3 and $k_1 = O(1)$, $k_2 = k_3 + O(1)$ yields the desired gain.
We may therefore assume that $A_0 = A_1 = A_2 = A_3 = I$, which reduces us to the trilinear nullform expansion (5.47) restricted to Case 5 of Lemma 5.5. Beginning with the first of the trilinear nonlinearities and for the case of aligned ψ_2, ψ_3, we now modify the analysis of Case 5 from that lemma. For ease of notation we will fix caps κ_2, κ_3 and drop the projections P_{k_i,κ_i}. In the end, an application of the Cauchy-Schwarz inequality will allow for summation over the caps. We first limit the modulation of ψ_1:

$$\begin{aligned}
&\|P_0\partial^\beta Q_{>k}I[Q_{>k+C}I\psi_1\ \Delta^{-1}\partial_j P_k I(R_\beta\psi_2 R_j\psi_3 - R_j\psi_2 R_\beta\psi_3)]\|_{N[0]} \\
&\lesssim \|P_0\partial^\beta Q_{>k}I[Q_{>k+C}I\psi_1\ [\Delta^{-1}\partial^2_{j\beta}P_k I(|\nabla|^{-1}\psi_2 R_j\psi_3) - P_k I(|\nabla|^{-1}\psi_2 R_\beta\psi_3)]]\|_{\dot{X}^{0,-\frac{1}{2},1}} \\
&\lesssim 2^{-\frac{k}{2}}\|Q_{>k+C}\psi_1\|_{L^2_t L^2_x}\|\Delta^{-1}\partial^2_{j\beta}P_k I(|\nabla|^{-1}\psi_2 R_j\psi_3) - P_k I(|\nabla|^{-1}\psi_2 R_\beta\psi_3)\|_{L^\infty_t L^\infty_x} \\
&\lesssim 2^{-k}\|\psi_1\|_{S[k_1]}2^{k_2}2^{m_0}\|\psi_1\|_{S[k_1]}\|\psi_2\|_{L^\infty_t L^2_x}\|\psi_3\|_{L^\infty_t L^2_x} \le \delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

where the gain is a result of Bernstein's inequality. Summation over κ_2, κ_3 is admissible here in view of Lemma 2.18. Hence, if the inner output has frequency $\sim 2^k$ then we may assume that ψ_1 has modulation $\lesssim 2^k$. Next, we apply (5.47) and bound the six terms on the right-hand side of that identity one by one.

Previously, we estimated the first term by means of the Strichartz component (2.14). However, this does not seem to yield the angular improvement so we use a different argument:

$$\begin{aligned}
&\|P_0 I\Box(Q_{\leq k}\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{N[0]}\\
&\lesssim \sum_{a\leq k+C}\|P_0 Q_a(Q_{\leq k}\psi_1 P_k I[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3])\|_{\dot{X}_0^{0,\frac12,1}}\\
&\lesssim \sum_{a\leq j\leq k+C} 2^{\frac{a}{2}}\|Q_{\leq k}\psi_1\|_{L_t^\infty L_x^2}\,2^k\|P_k Q_j[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{L_t^2L_x^2}\\
&+\sum_{j\leq a\leq k+C} 2^k 2^{\frac{a-k}{4}}\|Q_{\leq k}\psi_1\|_{S[k_1]}\|P_kQ_j[|\nabla|^{-1}\psi_2|\nabla|^{-1}\psi_3]\|_{\dot{X}_k^{0,\frac12,1}}
\end{aligned} \tag{5.94}$$

Lemma 4.11 was used to pass to the last line. By Corollary 4.10 one can continue as follows:

$$\begin{aligned}
&\lesssim \sum_{j\leq k+C} 2^{\frac{j}{2}}\|\psi_1\|_{L_t^\infty L_x^2}\,2^k\delta 2^{-\frac{j-k_2}{3}}2^{-\frac{3k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&+\sum_{j\leq k+C}2^k\|\psi_1\|_{S[k_1]}\,\delta 2^{-\frac{j-k_2}{3}}2^{-\frac{3k_2}{2}}2^{\frac{j}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\leq \delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned} \tag{5.95}$$

Moreover, Corollary 4.10 shows that this bound allows for summation over the caps.

For the second term, we can assume that $\psi_1 = Q_{\leq k_2+C}\psi_1$, see above. Then, by Corollary 4.15 as well as Corollary 4.10, and some large constant M,

$$\begin{aligned}
&\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\leq 2^{m_0}}}\|P_0 I[\Box(\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)|\nabla|^{-1}P_{k_2,\kappa_2}\psi_2]\|_{N[0]}\\
&\lesssim 2^{k_2}|m_0|\sum_{j\leq k_2+C}2^{\frac{j-k_2}{4}}\Big(\sum_{\kappa_3\in\mathcal{C}_{m_0}}\|\tilde{P}_0\Box Q_j(\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|^2_{\dot{X}_0^{0,-\frac12,\infty}}\Big)^{\frac12}\||\nabla|^{-1}\psi_2\|_{S[k_2]}\\
&\lesssim 2^{\frac{m_0}{M}}|m_0|\sum_{j\leq k_2+C}2^{\frac{j-k_2}{4}}2^{\frac{k_2-j}{3}}2^{\frac{k_2}{2}}2^{\frac{j}{2}}2^{-k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\lesssim\delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Third, by Lemma 4.13 and (4.33) of Corollary 4.10,

$$\begin{aligned}
&\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\leq 2^{m_0}}}\|P_0 I[Q_{\leq k_2+C}\psi_1\Box(|\nabla|^{-1}P_{k_2,\kappa_2}\psi_2)|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3]\|_{N[0]}\\
&\lesssim\sum_{j\leq k_2+C}2^{\frac{j-k_2}{4}}\sum_{\substack{\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_2,\kappa_3)\leq 2^{m_0}}}\|\tilde{P}_0Q_{\leq k_2+C}(\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|_{\dot{X}_0^{0,\frac12,1}}\|\Box Q_j(|\nabla|^{-1}P_{k_2,\kappa_2}\psi_2)\|_{\dot{X}_{k_2}^{0,-\frac12,\infty}}\\
&\lesssim\Big(\sum_{\kappa_3\in\mathcal{C}_{m_0}}\|\tilde{P}_0Q_{\leq k_2+C}(\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|^2_{\dot{X}_0^{0,\frac12,1}}\Big)^{\frac12}\|\psi_2\|_{S[k_2]}\\
&\lesssim\sum_{\ell\leq k_2+C}\Big(\sum_{\kappa_3\in\mathcal{C}_{m_0}}\|\tilde{P}_0Q_\ell(\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|^2_{\dot{X}_0^{0,\frac12,1}}\Big)^{\frac12}\|\psi_2\|_{S[k_2]}\\
&\lesssim\sum_{\ell\leq k_2+C}\delta 2^{\frac{\ell}{2}}2^{\frac{k_2-\ell}{3}}2^{\frac{k_3}{2}}2^{-k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\lesssim\delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

The summation over the caps was carried out explicitly for the second and third terms since it requires some care. Fourth, by (4.42) and Corollary 4.10,

$$\begin{aligned}
&\|P_0 I[(\Box Q_{\le k+C}\psi_1)\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{\ell\le k+C} 2^{\frac{\ell}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\|\tilde P_k Q_{\le k+C}[R_j\psi_2|\nabla|^{-1}\psi_3]\|_{\dot X_k^{0,\frac12,1}} \\
&\lesssim \delta\sum_{\ell\le k+C}\sum_{m\le k+C} 2^{\frac{\ell}{4}}\|\psi_1\|_{\dot X_{k_1}^{0,\frac12,\infty}}\, 2^{\frac{k-j}{3}}2^{\frac{k_2}{2}}2^{\frac{m}{2}}\|\psi_2\|_{S[k_2]}\||\nabla|^{-1}\psi_3\|_{S[k_3]} \lesssim \delta 2^{\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Since $k = k_1 + O(1) = k_2 + O(1)$, the fifth term

$$\|P_0 I\Box[Q_{\le k+C}\psi_1\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}$$

is bounded exactly like the first, see (5.94), (5.95). The sixth and final term is estimated by means of (4.40) and Corollary 4.10:

$$\begin{aligned}
&\|P_0 I[Q_{\le k+C}\psi_1\Box\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]} \\
&\lesssim \|\psi_1\|_{S[k_1]}\, 2^k\|P_k Q_{\le k+C}\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_k^{0,-\frac12,1}} \\
&\lesssim \|\psi_1\|_{S[k_1]}\sum_{m\le k+C} 2^k\|P_k Q_m(R_j\psi_2|\nabla|^{-1}\psi_3)\|_{\dot X_k^{0,\frac12,1}} \\
&\lesssim \delta\|\psi_1\|_{S[k_1]}\sum_{m\le k+C} 2^k\, 2^{\frac{k-m}{3}}2^{\frac{m}{2}}2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim \delta 2^{k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

as claimed.

We now repeat this analysis for the case of alignment between ψ_1 and ψ_3 (the remaining case being symmetric). We again begin with the reduction of various modulations. Using the notation of Lemma 5.1, if $A_0 = I^c$, then $A_1 = I^c$. By (4.52) of Lemma 4.17 and with $\frac12 = \frac1p + \frac1q$ where $q < \infty$ is very large,

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_3\in\mathcal C_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|P_0 I^c\partial^\beta[I^c\nabla_{t,x}P_{k_1,\kappa_1}\psi_1\Delta^{-1}\partial_j I\mathcal Q_{\beta j}(\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[0]} \\
&\lesssim \sum_{\substack{\kappa_1,\kappa_3\in\mathcal C_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}}\sum_{m\ge 0} 2^{-\varepsilon m}\|P_0 Q_m[\tilde Q_m\nabla_{t,x}P_{k_1,\kappa_1}\psi_1\,\Delta^{-1}\partial_j P_k I\mathcal Q_{\beta j}(\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{L_t^2L_x^2} \\
&\lesssim \sum_{m\ge 0} 2^{(1-\varepsilon)m}\sum_{\substack{\kappa_1,\kappa_3\in\mathcal C_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|\tilde Q_m P_{k_1,\kappa_1}\psi_1\|_{L_t^2L_x^p}\, 2^{-k}\|P_k I\mathcal Q_{\beta j}(\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{L_t^\infty L_x^q} \\
&\lesssim 2^{m_0(\frac12-\frac1p)}\sum_{m\ge 0} 2^{(1-\varepsilon)m}\Big(\sum_{\kappa_1\in\mathcal C_{m_0}}\|P_{k_1,\kappa_1}\tilde Q_m\psi_1\|_{L_t^2L_x^2}^2\Big)^{\frac12} 2^{(\frac12-\frac2q)k}\Big(\sum_{\kappa_3\in\mathcal C_{m_0}}\|IP_k\mathcal Q_{\beta j}(\psi_2, P_{k_3,\kappa_3}\psi_3)\|_{L_t^2L_x^2}^2\Big)^{\frac12} \\
&\lesssim |m_0|\, 2^{m_0(\frac12-\frac1p)}2^{(1-\frac2q)k}\|\psi_1\|_{\dot X_0^{0,1-\varepsilon,2}}\|P_{k_2}\psi_2\|_{S[k_2]}\|P_{k_3}\psi_3\|_{S[k_3]} \lesssim \delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Hence, we can assume that $A_0 = I$ as well as $A_1 = I$. If $A_2 = I^c$, then also $A_3 = I^c$ and

$$\begin{aligned}
&\|P_0 I\partial^\beta[IP_{k_1,\kappa_1}\psi_1\Delta^{-1}\partial_j I\mathcal{Q}_{\beta j}(I^c\psi_2, I^cP_{k_3,\kappa_3}\psi_3)]\|_{N[0]}\\
&\lesssim \sum_{m\geq k_2+C}\|IP_{k_1,\kappa_1}\psi_1\Delta^{-1}\partial_j IP_k\mathcal{Q}_{\beta j}(I^c\psi_2, I^cP_{k_3,\kappa_3}\psi_3)]\|_{L^1_tL^2_x}\\
&\lesssim \|P_{k_1,\kappa_1}\psi_1\|_{L^\infty_tL^2_x}\,2^{-k}\|IP_k\mathcal{Q}_{\beta j}(I^c\psi_2, I^cP_{k_3,\kappa_3}\psi_3)\|_{L^1_tL^\infty_x}\\
&\lesssim \|P_{k_1,\kappa_1}\psi_1\|_{L^\infty_tL^2_x}\sum_{m\geq k_2+C}2^{-k_2}\|\mathcal{Q}_{\beta j}(Q_m\psi_2,\tilde{Q}_mP_{k_3,\kappa_3}\psi_3)\|_{L^1_tL^\infty_x}\\
&\lesssim \|P_{k_1,\kappa_1}\psi_1\|_{L^\infty_tL^2_x}\sum_{m\geq k_2+C}2^{m-2k_2}\|Q_m\psi_2\|_{L^2_tL^\infty_x}\|P_{k_3,\kappa_3}\tilde{Q}_m\psi_3\|_{L^2_tL^\infty_x}\\
&\lesssim 2^{m_0+k_2}\|P_{k_1,\kappa_1}\psi_1\|_{L^\infty_tL^2_x}\sum_{m\geq k_2+C}2^{m-k_2}2^{-2(1-\varepsilon)m}2^{(1-2\varepsilon)k_2}\|Q_m\psi_2\|_{\dot{X}_{k_2}^{-\frac12+\varepsilon,1-\varepsilon,\infty}}\|P_{k_3,\kappa_3}\tilde{Q}_m\psi_3\|_{\dot{X}_{k_3}^{-\frac12+\varepsilon,1-\varepsilon,\infty}}\\
&\lesssim 2^{m_0}\|P_{k_1,\kappa_1}\psi_1\|_{L^\infty_tL^2_x}\|\psi_2\|_{\dot{X}_{k_2}^{-\frac12+\varepsilon,1-\varepsilon,2}}\|P_{k_3,\kappa_3}\psi_3\|_{\dot{X}_{k_3}^{-\frac12+\varepsilon,1-\varepsilon,2}}
\end{aligned}$$

Summing over the caps κ_1, κ_3 and $k_1 = O(1)$, $k_2 = k_3 + O(1)$ yields the desired gain. For ψ_1 one uses Lemma 2.18. As before, this reduces us to the trilinear nullform expansion (5.47). By the estimate (5.94), it suffices to consider $P_{\leq k+C}\psi_1$ if the inner output has frequency $\sim 2^k$. Beginning with the first term on the right-hand side of (5.47), one has

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\leq 2^{m_0}}}\|P_0I\Box(Q_{\leq k}P_{k_1,\kappa_1}\psi_1\;P_kI[|\nabla|^{-1}\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3])\|_{N[0]}\\
&\lesssim \sum_{a\leq k+C}\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\leq 2^{m_0}}}\|P_0Q_a(Q_{\leq k}P_{k_1,\kappa_1}\psi_1P_kI[|\nabla|^{-1}\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3])\|_{\dot{X}_0^{0,\frac12,1}}\\
&\lesssim \sum_{a\leq j\leq k+C}2^{\frac{a}{2}}2^k\Big(\sum_{\kappa_1}\|Q_{\leq k}P_{k_1,\kappa_1}\psi_1\|^2_{L^\infty_tL^2_x}\Big)^{\frac12}\Big(\sum_{\kappa_1}\|P_kQ_j[|\nabla|^{-1}\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3]\|^2_{L^2_tL^2_x}\Big)^{\frac12}\\
&\quad+\sum_{j\leq a\leq k+C}2^{\frac{3k}{4}}2^{\frac{a}{4}}\|\psi_1\|_{S[k_1]}\Big(\sum_{\kappa_3}\|P_kQ_j[|\nabla|^{-1}\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3]\|^2_{\dot{X}_k^{0,\frac12,\infty}}\Big)^{\frac12}
\end{aligned}\tag{5.96}$$

Corollary 4.8 was used to pass to the last line. By Lemma 2.18 and Corollary 4.10 one can continue as follows:

$$\begin{aligned}
&\lesssim \delta\sum_{j\leq k+C}2^{\frac{j}{2}}\|\psi_1\|_{L^\infty_tL^2_x}2^k2^{-\frac{j-k_2}{3}}2^{-\frac{3k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\\
&\quad+\delta\sum_{j\leq k+C}2^k\|\psi_1\|_{S[k_1]}2^{-\frac{j-k_2}{3}}2^{-\frac{3k_2}{2}}2^{\frac{j}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]}\leq\delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}\tag{5.97}$$

Moreover, Corollary 4.10 shows that this bound allows for summation over the caps. For the second term, we can assume that $\psi_1 = Q_{\leq k_2+C}\psi_1$, see above. Then, by Lemma 4.13 as well as Corollary 4.10,

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\leq 2^{m_0}}}\|P_0I[\Box(P_{k_1,\kappa_1}\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)|\nabla|^{-1}\psi_2]\|_{N[0]}\\
&\lesssim 2^{k_2}\sum_{j\leq k_2+C}2^{\frac{j-k_2}{4}}\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\leq 2^{m_0}}}\|\tilde{P}_0\Box Q_j(P_{k_1,\kappa_1}\psi_1\,|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|_{\dot{X}_0^{0,-\frac12,\infty}}\||\nabla|^{-1}\psi_2\|_{S[k_2]}\\
&\lesssim \delta\sum_{j\leq k_2+C}2^{\frac{j-k_2}{4}}2^{\frac{k_2-j}{3}}2^{\frac{k_2}{2}}2^{\frac{j}{2}}2^{-k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\lesssim\delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Third, by Lemma 4.13 and (4.33) of Corollary 4.10,

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|P_0 I[Q_{\le k_2+C}P_{k_1,\kappa_1}\psi_1\Box(|\nabla|^{-1}\psi_2)|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3]\|_{N[0]}\\
&\lesssim \sum_{j\le k_2+C} 2^{\frac{j-k_2}{4}} \sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|\tilde P_0 Q_{\le k_2+C}(P_{k_1,\kappa_1}\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|_{\dot X_0^{0,\frac12,1}}\|\Box Q_j(|\nabla|^{-1}\psi_2)\|_{\dot X_{k_2}^{0,-\frac12,\infty}}\\
&\lesssim \sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|\tilde P_0 Q_{\le k_2+C}(P_{k_1,\kappa_1}\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|_{\dot X_0^{0,\frac12,1}}\|\psi_2\|_{S[k_2]}\\
&\lesssim \sum_{\ell\le k_2+C}\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|\tilde P_0 Q_\ell(P_{k_1,\kappa_1}\psi_1|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|_{\dot X_0^{0,\frac12,1}}\|\psi_2\|_{S[k_2]}\\
&\lesssim \sum_{\ell\le k_2+C} \delta 2^{\frac{\ell}{2}}2^{\frac{k_2-\ell}{3}}2^{\frac{k_3}{2}}2^{-k_3}\prod_{i=1}^3\|\psi_i\|_{S[k_i]} \lesssim \delta\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Fourth, by Lemma 4.13, Cauchy-Schwarz applied to the cap-sum, and Corollary 4.10,

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|P_0 I[(\Box Q_{\le k+C}P_{k_1,\kappa_1}\psi_1)\Delta^{-1}\partial_j P_k I(R_j\psi_2\,|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)]\|_{N[0]}\\
&\lesssim |m_0|\sum_{\ell\le k+C} 2^{\frac{\ell}{4}}\|\Box Q_\ell\psi_1\|_{\dot X_{k_1}^{0,-\frac12,\infty}}\Big(\sum_{\kappa_3\in\mathcal{C}_{m_0}}\|\tilde P_k Q_{\le k+C}[R_j\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3]\|^2_{\dot X_k^{0,\frac12,1}}\Big)^{\frac12}\\
&\lesssim \delta\sum_{\ell\le k+C}\sum_{m\le k+C} 2^{\frac{\ell}{4}}\|\psi_1\|_{\dot X_{k_1}^{0,\frac12,\infty}} 2^{\frac{k-m}{3}}2^{\frac{k_2}{2}}2^{\frac{m}{2}}\|\psi_2\|_{S[k_2]}\||\nabla|^{-1}\psi_3\|_{S[k_3]} \lesssim \delta 2^{\frac{k_2}{4}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

Since $k = k_1 + O(1) = k_2 + O(1)$, the fifth term

$$\|P_0 I\Box[Q_{\le k+C}\psi_1\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}\psi_3)]\|_{N[0]}$$

is bounded exactly like the first, see (5.94), (5.95). The sixth and final term is estimated by means of Corollary 4.15 and Corollary 4.10:

$$\begin{aligned}
&\sum_{\substack{\kappa_1,\kappa_3\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_3)\le 2^{m_0}}} \|P_0 I[Q_{\le k+C}P_{k_1,\kappa_1}\psi_1\Box\Delta^{-1}\partial_j P_k I(R_j\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)]\|_{N[0]}\\
&\lesssim |m_0|\|\psi_1\|_{S[k_1]}\,2^k\sum_{m\le k+C}\Big(\sum_{\kappa_3\in\mathcal{C}_{m_0}}\|P_k Q_m\Box\Delta^{-1}\partial_j(R_j\psi_2|\nabla|^{-1}P_{k_3,\kappa_3}\psi_3)\|^2_{\dot X_k^{0,-\frac12,1}}\Big)^{\frac12}\\
&\lesssim \delta\|\psi_1\|_{S[k_1]}\sum_{m\le k+C} 2^k\,2^{\frac{k-m}{3}}2^{\frac{m}{2}}2^{-\frac{k_2}{2}}\|\psi_2\|_{S[k_2]}\|\psi_3\|_{S[k_3]} \lesssim \delta 2^{k_2}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}
\end{aligned}$$

as claimed. The other two types of trilinear null-forms are similar and left to the reader. □

Remark 5.12. The proof of the preceding estimates actually leads to a slightly better result: letting $P_0 F(P_{k_1}\psi_1, P_{k_2}\psi_2, P_{k_3}\psi_3)$ be a frequency localized trilinear null-form as above, then given any $\delta > 0$, there exists some $l_0 \le -100$ such that we can write

$$P_0 F(P_{k_1}\psi_1, P_{k_2}\psi_2, P_{k_3}\psi_3) = F_1 + F_2$$

where F_1 is a sum of energy, $\dot{X}^{s,b,q}$, as well as wave-packet atoms of scale $l \geq l_0$ (where scale refers to the size 2^l of the caps κ used), with the bound

$$\|F_1\|_{N[0]} \lesssim w(k_1, k_2, k_3) \prod_{j=1}^{3} \|P_{k_j}\psi_j\|_{S[k_j]}$$

and universal implied constant (independent of δ), while we also have

$$\|F_1\|_{N[0]} \lesssim \delta w(k_1, k_2, k_3) \prod_{j=1}^{3} \|P_{k_j}\psi_j\|_{S[k_j]}$$

The reason for this is that whenever a wave-packet atom of extremely fine scale is being used to estimate some constituent of P_0F, one gains a small exponential power in that scale.

6. Quintilinear and higher nonlinearities

Here we detail the estimates needed in order to control the higher order error terms generated by the process described in Section 3. This section is quite technical but the main point here is that the higher order terms, while still somewhat complicated, are much easier to estimate than the trilinear null-forms, and only require a very mild null-structure. We start with the lowest order errors, of quintilinear type. These are either of first or second type, see the discussion in Section 3. We commence with those of the first type, which can be schematically written as

$$\nabla_{x,t}[\psi\nabla^{-1}(R_\nu\psi\nabla^{-1}(\psi\nabla^{-1}Q_{\mu j}(\psi, \psi)))],$$

where not both ν, μ are simultaneously zero. Assume that $\nu = 0, \mu \neq 0$, the remaining cases being treated analogously. The following lemma is then representative for the higher order errors, for a universal $\delta > 0$.

Lemma 6.1. *We have the estimates*

$$\|\nabla_{x,t}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{N[0]}$$
$$\lesssim 2^{\delta[\min_{j\neq 0}\{r_j,k_j\} - \max_{j\neq 0}\{r_j,k_j\}]} \prod_{i=0}^{4} \|P_{k_i}\psi_i\|_{S[k_i]}, \quad r_1 < -10$$
$$\|\nabla_{x,t}[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{N[0]}$$
$$\lesssim 2^{\delta k_0} 2^{\delta[\min\{r_j,k_j\} - \max\{r_j,k_j\}]} \prod_{i=0}^{4} \|P_{k_i}\psi_i\|_{S[k_i]}, \quad r_1 \in [-10, 10]$$
$$\|\nabla_{x,t}P_0[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{N[0]}$$
$$\lesssim 2^{-\delta k_0} 2^{\delta[\min\{r_j,k_j\} - \max\{r_j,k_j\}]} \prod_{i=0}^{4} \|P_{k_i}\psi_i\|_{S[k_i]}, \quad r_1 > 10$$

All implied constants are universal.

Proof. All three inequalities are proved similarly, and we treat here the high-low case in detail, i.e., the first of them. We first deal with the elliptic cases:

(i): Output in elliptic regime. This is the expression (we have included the gratuitous cutoff $P_{[-5,5]}$ in light of $r_1 < -10$)

$$\nabla_{x,t}P_{[-5,5]}Q_{>10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]$$
$$= \sum_{l>10} \nabla_{x,t}P_{[-5,5]}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]$$

Now distinguish between further cases:

(i1): $\max\{k_1,\dots,k_4\} \ll l$, $R_0P_{k_1}\psi_1 = R_0P_{k_1}Q_{<l-100}\psi_1$. In this case at least one other factor $P_{k_j}\psi_j$ has modulation at least 2^{l-10}. For argument' s sake, let this be $P_{k_2}\psi_2 = P_{k_2}Q_{>l-10}\psi_2$ (the other cases being similar), so we now reduce to estimating

$$\sum_{l>10}\nabla_{x,t}P_{[-5,5]}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{<l-100}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}Q_{>l-10}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))],$$

where we also make the further assumptions of case (i1). Freezing l for now, we estimate this expression as follows: first, note that we get

$$\begin{aligned}&\|\nabla^{-1}P_{r_2}(P_{k_2}Q_{>l-10}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4))\|_{L_t^2\dot H_x^{\frac12}}\\&\lesssim 2^{(1-\epsilon)(k_2-l)}2^{[\min\{r_{2,3},k_{2,3,4}\}-\max\{r_{2,3},k_{2,3,4}\}]}\prod_{j=2}^4\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

This follows by straightforward usage of Bernstein's inequality and the definition of $S[k]$, as well as exploiting the null-structure of Q_{jk}. Furthermore, we have

$$\|R_0P_{k_1}Q_{<l-100}\psi_1\|_{L^2_{t,x}}\lesssim 2^{\epsilon(l-k_1)}2^{-\frac{k_1}{2}}\|P_{k_1}\psi\|_{S[k_1]},$$

where $\epsilon>0$ is as in the definition of $S[k]$, which implies that

$$\|R_0P_{k_1}Q_{<l-100}\psi_1\|_{L_t^\infty L_x^2}\lesssim 2^{\epsilon(l-k_1)}2^{\frac{l-k_1}{2}}\|P_{k_1}\psi\|_{S[k_1]}$$

¿From here we get

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{<l-100}\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}Q_{>l-10}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{\dot X_0^{-\frac12+\epsilon,-1-\epsilon,2}}\\&\lesssim 2^{-\epsilon l}\|P_0\psi_0\|_{L_t^\infty L_x^2}\\&\times\|\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{<l-100}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}Q_{>l-10}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty}\\&\lesssim 2^{\frac{r_1}{2}}2^{\min\{k_1-\min\{r_{1,2}\},0\}}2^{\frac{\min\{\min\{r_{1,2}\}-k_1,0\}}{2}}\|R_0P_{k_1}Q_{<l-100}\psi_1\|_{L_t^\infty L_x^2}\\&\qquad\times\|\nabla^{-1}P_{r_2}(P_{k_2}Q_{>l-10}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4))\|_{L_t^2\dot H_x^{\frac12}}\end{aligned}$$

Substituting the bounds from before, this is bounded by

$$\begin{aligned}&\lesssim 2^{-\varepsilon l}2^{\frac{r_1}{2}}2^{\min\{k_1-\min\{r_{1,2}\},0\}}2^{\frac{\min\{\min\{r_{1,2}\}-k_1,0\}}{2}}\\&\times 2^{\epsilon(l-k_1)}2^{\frac{l-k_1}{2}}\|P_{k_1}\psi\|_{S[k_1]}2^{(1-\epsilon)(k_2-l)}2^{[\min\{r_{2,3},k_{2,3,4}\}-\max\{r_{2,3},k_{2,3,4}\}]}\prod_{j=2}^4\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

This is equivalent to an estimate of the form claimed in the lemma, with an extra gain $2^{-\epsilon l}$ which allows us to sum over $l>10$.
(i2): $\max\{k_1,\dots,k_4\}\ll l$, $R_0P_{k_1}\psi_1 = R_0P_{k_1}Q_{[l-100.l+100]}\psi_1$. The estimate here is similar except that square summation over l is made possible since we have

$$P_{k_1}Q_{>k_1}\psi_1\in\dot X_{k_1}^{-\frac12+\epsilon,1-\epsilon,2}$$

(i3): $\max\{k_1,\dots,k_4\}\ll l$, $R_0P_{k_1}\psi_1 = R_0P_{k_1}Q_{\gg l+100]}\psi_1$. This is again similar. Fixing the modulation of $R_0P_{k_1}Q_{\gg l+100]}\psi_1$ to size 2^{l_1}, $l_1>l+100$, there is at least one other input which has modulation at least comparable to 2^{l_1}. Then one proceeds as in case (i1).

(i4): $\max\{k_1,\dots,k_4\}>l+O(1)$, $R_0P_{k_1}\psi_1 = R_0P_{k_1}Q_{<\max\{k_{1,2,3,4}\}}\psi_1$.
Here we obtain a gain in $\min\{r_{1,2,3},k_{1,2,3,4}\}-\max\{r_{1,2,3},k_{1,2,3,4}\}$, which suffices to offset the loss due to

the possibly large modulation of $R_0 P_{k_1} Q_{<\max\{k_{1,2,3,4}\}}\psi_1$. Specifically, write

$$\begin{aligned}
&\nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{<\max\{k_{1,2,3,4}\}}P_{k_1}\psi_1\\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\\
&= \nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{<k_1}P_{k_1}\psi_1\\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\\
&+ \nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{[k_1,\max\{k_{1,2,3,4}\}]}P_{k_1}\psi_1\\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]
\end{aligned}$$

Here we use the inequalities

$$\begin{aligned}
&\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4))\|_{L_t^2\dot{H}_x^{\frac{1}{2}}}\\
&\lesssim 2^{\frac{1}{2}[\min\{r_{2,3},k_{1,2,3,4}\}-\max\{r_{2,3},k_{1,2,3,4}\}]}\prod_{i=1}^4\|P_{k_i}\psi_i\|_{S[k_i]},
\end{aligned}$$

$$\begin{aligned}
&\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4))\|_{L_t^\infty L_x^2}\\
&\lesssim 2^{\min\{r_{2,3},k_{1,2,3,4}\}-\max\{r_{2,3},k_{1,2,3,4}\}}\prod_{i=1}^4\|P_{k_i}\psi_i\|_{S[k_i]}.
\end{aligned}$$

Then we can estimate

$$\begin{aligned}
&\|\nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{[k_1,\max\{k_{1,2,3,4}\}]}P_{k_1}\psi_1\\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,-1-\epsilon,2}}\\
&\lesssim 2^{-\epsilon l}\|P_0\psi_0\|_{L_t^\infty L_x^2}\\
&\times\|\nabla^{-1}P_{r_1}(R_0 Q_{[k_1,\max\{k_{1,2,3,4}\}]}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty}
\end{aligned}$$

To conclude the contribution of this term, one then checks, using standard Littlewood-Paley trichotomy, that

$$\begin{aligned}
&\|\nabla^{-1}P_{r_1}(R_0 Q_{[k_1,\max\{k_{1,2,3,4}\}]}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty}\\
&\lesssim 2^{\frac{\min\{r_{1,2},k_1\}}{2}}2^{\frac{\min\{r_{1,2},k_1\}-\max\{r_{1,2},k_1\}}{2}}\|R_0 Q_{[k_1,\max\{k_{1,2,3,4}\}]}P_{k_1}\psi_1\|_{L_t^2\dot{H}_x^{\frac{1}{2}}}\\
&\qquad\qquad \times\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4))\|_{L_t^\infty L_x^2}
\end{aligned}$$

Combining with the bound from above, and furthermore assuming the ϵ in the definition of $\|.\|_S$ to be small enough, we conclude that for suitable $\delta > 0$ we have

$$\begin{aligned}
&\|\nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{[k_1,\max\{k_{1,2,3,4}\}]}P_{k_1}\psi_1\\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,-1-\epsilon,2}}\\
&\lesssim 2^{-\epsilon l}2^{\delta[\min_{j\neq 0}\{r_j,k_j\}-\max_{j\neq 0}\{r_j,k_j\}]}\prod_{i=0}^4\|P_{k_i}\psi_i\|_{S[k_i]}
\end{aligned}$$

and square summing over $10 < l < \max\{k_{1,2,3,4}\}$ yields the desired bound.
For the term

$$\begin{aligned}
&\nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{<k_1}P_{k_1}\psi_1\\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]
\end{aligned}$$

from further above, estimate

$$\begin{aligned}&\|\nabla_{x,t}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,-1-\epsilon,2}}\\&\lesssim 2^{-\epsilon l}\|P_0\psi_0\|_{L_t^\infty L_x^2}\\&\times\|\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty},\end{aligned}$$

and we have

$$\begin{aligned}&\|\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty}\\&\lesssim 2^{\frac{\min\{r_{1,2},k_1\}}{2}}2^{\frac{\min\{r_{1,2},k_1\}-\max\{r_{1,2},k_1\}}{2}}\|P_{k_1}\psi_1\|_{L_t^\infty L_x^2}\\&\qquad\times\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4))\|_{L_t^2\dot{H}^{\frac{1}{2}}},\end{aligned}$$

which in conjunction with the bound from above

$$\begin{aligned}&\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4))\|_{L_t^2\dot{H}_x^{\frac{1}{2}}}\\&\lesssim 2^{\frac{1}{2}[\min\{r_{2,3},k_{1,2,3,4}\}-\max\{r_{2,3},k_{1,2,3,4}\}]}\prod_{i=1}^4\|P_{k_i}\psi_i\|_{S[k_i]},\end{aligned}$$

implies that

$$\begin{aligned}&\|\nabla_{x,t}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\\&\qquad\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,-1-\epsilon,2}}\\&\lesssim 2^{-\epsilon l}2^{\frac{\min\{r_{1,2},k_1\}}{2}}2^{\frac{\min\{r_{1,2},k_1\}-\max\{r_{1,2},k_1\}}{2}}\\&\qquad\times 2^{\frac{1}{2}[\min\{r_{2,3},k_{1,2,3,4}\}-\max\{r_{2,3},k_{1,2,3,4}\}]}\prod_{i=0}^4\|P_{k_i}\psi_i\|_{S[k_i]}\end{aligned}$$

Square summing over $l>10$ yields a bound as claimed in the lemma with $\delta=\frac{1}{2}$.

(i5): $\max\{k_1,\dots,k_4\}>l+O(1)$, $R_0P_{k_1}\psi_1=R_0P_{k_1}Q_{\gg\max\{k_{1,2,3,4}\}}\psi_1$.

Freeze the modulation of $R_0P_{k_1}\psi_1$ to dyadic value $2^{l_1}\gg 2^{\max\{k_{1,2,3,4}\}}$. Here there must be at least one other input with modulation at least comparable to 2^{l_1}. Let this input be $P_{k_2}\psi_2$ for definitiveness' sake, the other cases being treated similarly. Thus consider the term

$$\nabla_{x,t}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0Q_{l_1}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}Q_{\geq l_1+O(1)}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]$$

Assuming a high-low frequency cascade $r_1 \ll k_1 \ll k_2$, we can estimate this by (using Bernstein's inequality)

$$\begin{aligned}
&\|\nabla_{x,t} Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 Q_{l_1}P_{k_1}\psi_1 \\
&\qquad\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}Q_{>l_1+O(1)}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,-1-\epsilon,2}} \\
&\lesssim 2^{l(\frac{1}{2}-\epsilon)}\|P_0\psi_0\|_{L_t^\infty L_x^2} \\
&\times \|\nabla^{-1}P_{r_1}(R_0 Q_{l_1}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}Q_{>l_1+O(1)}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))\|_{L_t^1 L_x^\infty} \\
&\lesssim 2^{l(\frac{1}{2}-\epsilon)+r_1}2^{\min\{r_3,k_{3,4}\}-\max\{r_3,k_{3,4}\}}2^{k_1-k_2} \\
&\qquad \times \|P_0\psi_0\|_{L_t^\infty L_x^2}\|R_0 Q_{l_1}P_{k_1}\psi_1\|_{L_{t,x}^2}\|P_{k_2}Q_{>l_1+O(1)}\psi_2\|_{L_{t,x}^2}\prod_{j=3,4}\|P_{k_j}\psi_j\|_{S[k_j]} \\
&\lesssim 2^{l(\frac{1}{2}-\epsilon)}2^{\min\{r_3,k_{3,4}\}-\max\{r_3,k_{3,4}\}}2^{\epsilon(l_1-k_1)}2^{r_1-\frac{k_1+k_2}{2}}2^{k_1-k_2}2^{(1-\epsilon)(k_2-l_1)}\prod_{j=0}^{4}\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}$$

Summing over $l_1 \gg \max\{k_{1,2,3,4}\} > l + O(1)$, one obtains a bound of the form claimed in the lemma with $\delta = \frac{1}{2} - \varepsilon$ in the particular case at hand. The remaining frequency interactions, while keeping our assumptions on the modulations, are treated similarly. This concludes the *elliptic case (i).*

(ii): Output in hyperbolic regime. Now we consider the expression

$$\begin{aligned}
&\nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 P_{k_1}\psi_1 \\
&\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]
\end{aligned}$$

We decompose this into

$$\begin{aligned}
&\nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 P_{k_1}\psi_1 \\
&\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))] \\
&= \nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 P_{k_1}Q_{<k_1}\psi_1 \qquad (6.1)\\
&\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))] \\
&+ \nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 P_{k_1}Q_{[k_1,\max\{k_{1,2,3,4}\}+O(1)]}\psi_1 \qquad (6.2)\\
&\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))] \\
&+ \nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0 P_{k_1}Q_{\gg\max\{k_{1,2,3,4}\}}\psi_1 \qquad (6.3)\\
&\qquad \times \nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]
\end{aligned}$$

To estimate the first expression (6.1) on the right, we exploit the fact that we control sharp Strichartz norms, in addition to the basic null-form bilinear estimate controlling $Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)$. The key is the fact that we have the almost sharp Klainerman-Tataru norm built into S. To see this, consider the most difficult case, a high-low frequency cascade corresponding to $r_1 \ll k_1 \ll k_2$. We estimate the expression by starting from the inside:

$$\begin{aligned}
&\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4))\|_{L_t^{\frac{4}{3}}L_x^2} \\
&= 2^{-r_2}\sum_{c_{1,2}\in\mathcal{D}_{k_2,r_2},\mathrm{dist}(c_1,-c_2)\lesssim r_2}\|P_{r_2}(P_{k_2,c_1}\psi_2\nabla^{-1}P_{r_3,c_2}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4))\|_{L_t^{\frac{4}{3}}L_x^2} \\
&\lesssim 2^{-r_2}(\sum_{c_1\in\mathcal{D}_{k_2,r_2}}\|P_{k_2,c_1}\psi_2\|_{L_t^4 L_x^\infty}^2)^{\frac{1}{2}}\|\nabla^{-1}P_{r_3,c_2}Q_{jk}(P_{k_3}\psi_3, P_{k_4}\psi_4)\|_{L_{t,x}^2}
\end{aligned}$$

Here we have used Cauchy-Schwarz and Plancherel's theorem. Then using the definition of $\|.\|_S$, we can bound this by

$$2^{-r_2}(\sum_{c_1\in\mathcal{D}_{k_2,r_2}}\|P_{k_2,c_1}\psi_2\|_{L_t^4L_x^\infty}^2)^{\frac12}\|\nabla^{-1}P_{r_3,c_2}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)\|_{L^2_{t,x}}$$

$$\lesssim 2^{-r_2}2^{\frac{k_2}{4}}2^{\frac{r_2-k_2}{2+}}2^{\frac{\min\{r_3,k_{3,4}\}-\max\{r_3,k_{3,4}\}}{2}}\prod_{j=2}^4\|P_{k_j}\psi_2\|_{S[k_j]}$$

Turning to the full expression further above, we then get for the contribution of this term to the hyperbolic part of the output

$$\begin{aligned}
&\|\nabla_{x,t}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\\
&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{L_t^1\dot H_x^{-1}}\\
&\lesssim\|P_0\psi_0\|_{L_t^\infty L_x^2}\|\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^1L_x^\infty}\\
&\lesssim\|P_0\psi_0\|_{L_t^\infty}\\
&\times\|\nabla^{-1}P_{r_1}(\sum_{c_{1,2}\in R_{k_1,r_1},\,\mathrm{dist}(c_1,-c_2)\lesssim 2^{r_1}}R_0Q_{<k_1}P_{k_1,c_1}\psi_1\nabla^{-1}P_{r_2,c_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^1L_x^\infty}
\end{aligned}$$

We intend to substitute the intermediate bound from above for

$$\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4))\|_{L_t^{\frac43}L_x^2},$$

where we can exploit that, by Minkowski's and Plancherel's inequality, we have

$$(\sum_{c_1\in R_{r_2,c_2}}\|P_{r_2,c_2}F\|^2_{L_t^{\frac43}L_x^2})^{\frac12}\lesssim\|P_{r_2}F\|_{L_t^{\frac43}L_x^2}.$$

Thus we can estimate, using Cauchy-Schwarz, Bernstein's inequality and the preceding observation

$$\begin{aligned}
&\|\nabla^{-1}P_{r_1}(\sum_{c_{1,2}\in R_{k_1,r_1},\,\mathrm{dist}(c_1,-c_2)\lesssim 2^{r_1}}R_0Q_{<k_1}P_{k_1,c_1}\psi_1\\
&\qquad\times\nabla^{-1}P_{r_2,c_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^1L_x^\infty}\\
&\lesssim(\sum_{c_1\in R_{k_1,r_1}}\|P_{k_1,c_1}\psi_1\|^2_{L_t^4L_x^\infty})^{\frac12}\|\nabla^{-1}P_{r_2,c_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^{\frac43}L_x^2}\\
&\lesssim 2^{\frac{r_1-k_1}{2+}}2^{\frac34 k_1}2^{-r_2}2^{\frac{k_2}{4}}2^{\frac{r_2-k_2}{2+}}2^{\frac{\min\{r_3,k_{3,4}\}-\max\{r_3,k_{3,4}\}}{2}}\prod_{j=1}^4\|P_{k_j}\psi_2\|_{S[k_j]}
\end{aligned}$$

But by our assumption $r_1\ll k_1\ll k_2$ we have $r_2=k_1+O(1)$, whence we can replace the above bound by

$$\begin{aligned}
&\|\nabla_{x,t}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0Q_{<k_1}P_{k_1}\psi_1\\
&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{L_t^1\dot H_x^{-1}}\\
&\lesssim 2^{\frac{r_1-k_1}{2+}}2^{\frac{r_2-k_2}{4+}}2^{\frac{\min\{r_3,k_{3,4}\}-\max\{r_3,k_{3,4}\}}{2}}\prod_{j=0}^4\|P_{k_j}\psi_2\|_{S[k_j]},
\end{aligned}$$

and this is again enough to yield the statement of the lemma (here with $\delta=\frac{1}{4+}$). The remaining frequency interactions can be handled similarly.

Next, consider the second term (6.2) above, i.e.,

$$\begin{aligned}
&\nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{[k_1,\max\{k_{1,2,3,4}\}+O(1)]}\psi_1\\
&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]
\end{aligned}$$

This is much simpler: we get

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{[k_1,\max\{k_{1,2,3,4}\}+O(1)]}\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{L_t^1\dot{H}^{-1}}\\&\lesssim\|P_0\psi_0\|_{L_t^\infty L_x^2}\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{[k_1,\max\{k_{1,2,3,4}\}+O(1)]}\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^1L_x^\infty}\end{aligned}$$

For definitiveness' sake, we again assume that $r_1 \ll k_1 \ll k_2$, the remaining cases being similar. Then we get

$$\begin{aligned}&\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{[k_1,\max\{k_{1,2,3,4}\}+O(1)]}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^1L_x^\infty}\\&\lesssim 2^{r_1}\|R_0P_{k_1}Q_{[k_1,\max\{k_{1,2,3,4}\}+O(1)]}\psi_1\|_{L_{t,x}^2}\|\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)\|_{L_t^2L_x^2}\\&\lesssim 2^{r_1-k_1}2^{\epsilon(\max\{k_{1,2,3,4}\}-k_1)}2^{\frac{r_2-\max\{k_{3,4}\}}{2}}\prod_{j=0}^4\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

This corresponds to a bound as in the lemma with $\delta=\frac12-\epsilon$, where we recall ϵ is as in the definition of $\|\cdot\|_{S[k]}$. The remaining frequency interactions for this term are treated similarly.

Finally, consider the last term above

$$\begin{aligned}&\nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_{\gg\max\{k_{1,2,3,4}\}}\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\end{aligned}$$

Here we again need to compensate for the losses coming from estimating $R_0P_{k_1}Q_{\gg\max\{k_{1,2,3,4}\}}\psi_1$. Freeze its modulation to dyadic size 2^l. Then either at least one other input has at least comparable modulation, or else the output has modulation $\sim 2^l$ (in which case necessarily $l<O(1)$. In the latter case, one then estimates (where $l\gg\max\{k_{1,2,3,4}\}$ and we assume all other inputs to be at much lower modulation)

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_{<10}[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_l\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{N[0]}\\&=\|\nabla_{x,t}P_{[-5,5]}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_l\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{N[0]}\\&\leq\|\nabla_{x,t}P_{[-5,5]}Q_l[P_0\psi_0\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_l\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))]\|_{\dot{X}_0^{-1,-\frac12,1}}\\&\lesssim 2^{-\frac{l}{2}}\|P_0\psi_0\|_{L_t^\infty L_x^2}\|\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_l\psi_1\\&\qquad\times\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty}\end{aligned}$$

Here the second factor above is estimated by

$$\begin{aligned}&\|\nabla^{-1}P_{r_1}(R_0P_{k_1}Q_l\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{jk}(P_{k_3}\psi_3,P_{k_4}\psi_4)))\|_{L_t^2L_x^\infty}\\&\lesssim 2^{\frac{\min\{r_{1,2},k_1\}}{2}}2^{\frac{\min\{r_1,r_2,k_1\}-\max\{r_1,r_2,k_1\}}{2}}2^{\frac{\min\{r_{2,3},k_{2,3,4}\}-\max\{r_{2,3},k_{2,3,4}\}}{2}}2^{\epsilon(l-k_1)}\prod_{j=1}^4\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

Inserting this bound into the last inequality but one and summing over $l\gg\max\{k_{1,2,3,4}\}$ results in a bound as in the lemma with $\delta=\frac12-\epsilon$.

The case when at least one further input has at least modulation at least comparable to 2^l is similar, one places the output into $L_t^1\dot{H}^{-1}$.

This completes the proof of the first inequality of the lemma. The remaining ones are treated by an identical procedure. □

In a similar vein, one has estimates controlling the second kind of quintilinear term. We state the

Lemma 6.2. *For the second type of quintilinear null-form, we have the following estimates for suitable $\delta > 0$:*

$$\|\nabla_{x,t}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{N[0]}$$
$$\lesssim 2^{\delta r_1}2^{\delta[\min\{0,k_{1,2,3},s_1\}-\max\{0,k_{1,2,3},s_1\}]}2^{\delta[\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}]}\prod_{j=1}^{5}\|P_{k_j}\psi_j\|_{S[k_j]}, \quad r_1 < -10$$

$$\|\nabla_{x,t}P_0[\big(P_{s_1}[\nabla^{-1}(P_{k_1}\psi_1 P_{s_2}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{N[0]}$$
$$\lesssim 2^{\delta s_1}2^{\delta[\min\{s_{1,2},k_{1,2,3}\}-\max\{s1,2,k_{1,2,3}\}]}2^{\delta[\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}]}\prod_{j=1}^{5}\|P_{k_j}\psi_j\|_{S[k_j]}, \quad r_1 \in [-10,10]$$

$$\|\nabla_{x,t}P_0[\big(P_{s_1}[\nabla^{-1}(P_{k_1}\psi_1 P_{s_2}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{N[0]}$$
$$\lesssim 2^{-\delta s_1}2^{\delta[\min\{s_{1,2},k_{1,2,3}\}-\max\{s_{1,2},k_{1,2,3}\}]}2^{\delta[\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}]}\prod_{j=1}^{5}\|P_{k_j}\psi_j\|_{S[k_j]}, \quad r_1 > 10$$

Proof. We verify this again for the first inequality above, the other ones following a similar pattern. As usual, we distinguish between elliptic and hyperbolic output components:

(i): Output in elliptic regime. This is the expression

$$\nabla_{x,t}P_{[-5,5]}Q_{>10}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]$$
$$\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)$$

As usual the only slight complication arises due to the fact that w may have $\nu = 0$. Freeze the modulation of the output to dyadic size 2^l, $l > 10$. Then one re-iterates the same steps as in the preceding proof:

(i1): $\max\{k_{1,2,3}\} \ll l$*, time derivative falls on term with modulation* $< 2^{l-100}$. In this case at least one additional input (which is not hit by a time derivative) has modulation $> 2^{l-10}$. For example, assume this is $P_{k_1}\psi_1 = P_{k_1}Q_{>l-10}\psi_1$, the other cases being treated similarly. Then assuming a high-low scenario, say, i.e., $k_1 \gg 1$, we have (using Bernstein's inequality)

$$\|P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]\|_{L_t^1L_x^2}$$
$$\lesssim \|P_{k_1}Q_{>l-10}\psi_1\|_{L_{t,x}^2}\|P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3)\|_{L_{t,x}^2}$$
$$\lesssim 2^{-k_1}2^{(1-\epsilon)(k_1-l)}2^{\epsilon(l-k_2)}\prod_{j=1}^{3}\|P_{k_j}\psi_j\|_{S[k_j]}$$

Substituting this into the full expression, we obtain for the output the bound

$$\|\nabla_{x,t}P_{[-5,5]}Q_l[\big(P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]$$
$$\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{\dot{X}_0^{-\frac{1}{2}+\epsilon,-1-\epsilon,2}}$$
$$\lesssim 2^{-\epsilon l}\|P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]\|_{L_{t,x}^2}$$
$$\times \|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\|_{L_{t,x}^\infty}$$
$$\lesssim 2^{(\frac{1}{2}-\epsilon)l}\|P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]\|_{L_t^1L_x^2}$$
$$\times \|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\|_{L_{t,x}^\infty}$$
$$\lesssim 2^{(\frac{1}{2}-\epsilon)l}2^{-k_1}2^{(1-\epsilon)(k_1-l)}2^{\epsilon(l-k_2)}2^{r_1}2^{\frac{\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}}{2}}\prod_{j=1}^{5}\|P_{k_j}\psi_j\|_{S[k_j]}$$

Summing over $l \gg \max\{k_{1,2,3}\}$, the desired inequality of the first type of the lemma follows in this case. The remaining frequency interactions within

$$P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))]$$

are handled similarly.

(i2): $\max\{k_{1,2,3}\} \ll l$, time derivative falls on term with modulation $\sim 2^l$. In this case, we place the time derivative term into $L^2_{t,x}$, and are guaranteed gains in the maximal occurring frequency: for example, consider the term (arising upon unraveling the inner $Q_{\nu j}$ null-structure with $\nu = 0$)

$$P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}(P_{k_2}Q_{l+O(1)}R_0\psi_2 P_{k_3}\psi_3))]$$

In the high-high case $k_2 \gg s_1$, one can then estimate

$$\begin{aligned}
&\|P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}(P_{k_2}Q_{l+O(1)}R_0\psi_2 P_{k_3}\psi_3))]\|_{L^2_{t,x}} \\
&\lesssim 2^{\frac{\min\{0,k_1,s_1\}-\max\{0,k_1,s_1\}}{2}} 2^{\frac{\min\{s_1,k_2,k_3\}-\max\{s_1,k_2,k_3\}}{2}} \\
&\qquad \times \|P_{k_1}\psi_1\|_{L^\infty_t L^2_x} \|P_{k_2}Q_{l+O(1)}R_0\psi_2\|_{L^2_t \dot{H}^{\frac12}} \|P_{k_3}\psi_3\|_{L^\infty_t L^2_x}
\end{aligned}$$

From here one estimates the full expression by

$$\begin{aligned}
&\|\nabla_{x,t}P_{[-5,5]}Q_l\big[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}(P_{k_2}Q_{l+O(1)}R_0\psi_2 P_{k_3}\psi_3))] \\
&\qquad\qquad \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{\dot{X}_0^{-\frac12+\epsilon,-1-\epsilon,2}} \\
&\lesssim 2^{-\epsilon l}\|P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}(P_{k_2}Q_{l+O(1)}R_0\psi_2 P_{k_3}\psi_3))]\|_{L^2_{t,x}} \\
&\qquad\qquad \times \|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\|_{L^\infty_{t,x}} \\
&\lesssim 2^{-\epsilon l}2^{\epsilon(l-k_2)}2^{\frac{\min\{0,k_1,s_1\}-\max\{0,k_1,s_1\}}{2}} 2^{\frac{\min\{s_1,k_2,k_3\}-\max\{s_1,k_2,k_3\}}{2}} \\
&\qquad \times \|P_{k_1}\psi_1\|_{L^\infty_t L^2_x}\|P_{k_2}Q_{l+O(1)}\psi_2\|_{\dot{X}_{k_2}^{-\frac12+\epsilon,1-\epsilon,1}}\|P_{k_3}\psi_3\|_{L^\infty_t L^2_x} \\
&\qquad\qquad \times 2^{r_1}2^{\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}}\prod_{j=4,5}\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}$$

One may sum here to obtain a bound of the type as in the first inequality of the lemma, with $\delta = \frac12 - \varepsilon$. The remaining frequency interactions within

$$P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}(P_{k_2}Q_{l+O(1)}R_0\psi_2 P_{k_3}\psi_3))]$$

are again handled similarly.

(i3): $\max\{k_{1,2,3}\} \ll l$, time derivative falls on term with modulation $\gg 2^l$. In this case at least one additional term has at least comparable modulation, and one argues as in case (i1).

(i4): $\max\{k_{1,2,3}\} > l + O(1)$, time derivative falls on term with modulation $< \max\{k_{1,2,3}\} + O(1)$. Here the losses coming from the time derivative are easily counteracted by the gains in the large frequencies: first, one reduces the inputs $P_{k_{2,3}}\psi_{2,3}$ to the elliptic regimes. To do so, note that we have

$$\begin{aligned}
&\|P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}(P_{k_2}Q_{[k_2,\max\{k_{1,2,3}\}}R_0\psi_2 P_{k_3}\psi_3))]\|_{L^2_{t,x}} \\
&\lesssim 2^{\frac{\min\{0,s_1,k_{1,2,3}\}-\max\{0,s_1,k_{1,2,3}\}}{2}} 2^{\epsilon(\max\{k_{1,2,3}\}-k_2)}\prod_{j=1}^{3}\|P_{k_j}\psi_j\|_{S[k_j]},
\end{aligned}$$

and inserting this into the full expression is easily seen to yield the desired inequality. Hence we have reduced this case to the expression

$$\nabla_{x,t}P_{[-5,5]}Q_l[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3))] \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)$$

Of course in the present case at least one of $k_{2,3} > l + O(1)$. Assume that we have a high-high-low type situation in

$$P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3),$$

i.e., $s_1 \ll k_2$, this being the most delicate case. We distinguish between two cases:

(a): modulation of $P_{s_1}\dots$ is less than $2^{l+O(1)}$. In this case, we may "pull out" a (time)- derivative from the $Q_{\nu j}$-null-form, using the simple identity

$$R_\nu\psi^1 R_j\psi^2 - R_j\psi^1 R_\nu\psi^2 = \partial_\nu[\nabla^{-1}\psi^1 R_j\psi_2] - \partial_j[\nabla^{-1}\psi^1 R_\nu\psi_2]$$

Hence in this case we can estimate

$$\begin{aligned}
&\|\nabla_{x,t}P_{[-5,5]}Q_l[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}Q_{<l+O(1)}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3))] \\
&\qquad\qquad \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{\dot X_0^{-\frac12+\epsilon,-1-\epsilon,2}} \\
&\lesssim 2^{-\epsilon l}\|P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}Q_{<l+O(1)}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3))]\|_{L_t^\infty L_x^2} \\
&\qquad\qquad \times \|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\|_{L_t^2 L_x^\infty} \\
&\lesssim 2^{-\epsilon l}2^{l-k_2}2^{\frac{r_1}{2}}2^{\frac{\min\{r_1,k_{3,4}\}-\max\{r_1,k_{3,4}\}}{2}}2^{\min\{s_1,k_1,0\}}\prod_{j=1}^{5}\|P_{k_j}\psi_j\|_{S[k_j]}
\end{aligned}$$

One can sum over $l < \max\{k_{1,2,3}\}$ to get the desired first inequality of the lemma in the case at hand.

(b): modulation of $P_{s_1}\dots$ is $\gg 2^l$. In this case the modulation of the first input $P_{k_1}\psi_1$ needs to be comparable to that of

$$P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3))]$$

Hence we can write this contribution as

$$\sum_{l_1\gg l}\nabla_{x,t}P_{[-5,5]}Q_l[\big(P_0[\nabla^{-1}(P_{k_1}Q_{l_1+O(1)}\psi_1 P_{s_1}Q_{l_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3))] \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)$$

To estimate it, we use

$$\|P_{s_1}Q_{l_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3)\|_{L_{t,x}^2} \lesssim 2^{l_1-k_2}2^{-\frac{s_1}{2}}\prod_{i=2,3}\|P_{k_i}Q_{<k_i}\psi_2\|_{S[k_i]}$$

We the insert this bound into the full expression. In case that $k_1 > O(1)$, we can estimate

$$\begin{aligned}
&\|\nabla_{x,t}P_{[-5,5]}Q_l[\big(P_0[\nabla^{-1}(P_{k_1}Q_{l_1+O(1)}\psi_1 P_{s_1}Q_{l_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3))] \\
&\qquad\qquad \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)]]\|_{\dot X_0^{-\frac12+\epsilon,-1-\epsilon,2}} \\
&\lesssim 2^{-\epsilon l}2^{\frac{l}{2}}\|P_{k_1}Q_{l_1+O(1)}\psi_1\|_{L_{t,x}^2}\|P_{s_1}Q_{l_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2 P_{k_3}Q_{<k_3}\psi_3)\|_{L_{t,x}^2} \\
&\qquad\qquad \times \|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\|_{L_{t,x}^\infty}
\end{aligned}$$

In case that $l_1 < k_1 + O(1)$, we can bound this by

$$\begin{aligned}
&2^{-\epsilon l}2^{\frac{l}{2}}\|P_{k_1}Q_{l_1+O(1)}\psi_1\|_{L^2_{t,x}}\|P_{s_1}Q_{l_1}\nabla^{-1}Q_{\nu j}(P_{k_2}Q_{<k_2}\psi_2P_{k_3}Q_{<k_3}\psi_3)\|_{L^2_{t,x}}\\
&\qquad\qquad\times\|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\|_{L^\infty_{t,x}}\\
&\lesssim 2^{-\epsilon l}2^{\frac{l}{2}}2^{-\frac{l_1}{2}}\|P_{k_1}\psi_1\|_{S[k_1]}2^{l_1-k_2}2^{-\frac{s_1}{2}}\prod_{i=2,3}\|P_{k_i}Q_{<k_i}\psi_2\|_{S[k_i]}\\
&\qquad\qquad\times\|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\|_{L^\infty_{t,x}}\\
&\lesssim 2^{-\epsilon l}2^{\frac{l}{2}}2^{-\frac{l_1}{2}}2^{l_1-k_2}2^{-\frac{s_1}{2}}2^{r_1}2^{\min\{k_3,k_4\}-\max\{k_3,k_4\}}\prod_{i=1}^{5}\|P_{k_i}\psi_i\|_{S[k_i]}
\end{aligned}$$

Summing over $k_1+O(1) > l_1 \gg l$ and then over $l > O(1)$ results in a bound as in the first inequality of the lemma with $\delta = \frac{1}{2}+\epsilon$.
Next, still in the case $k_1 > O(1)$, if $l_1 > k_1 + O(1)$, one proceeds as before but uses

$$\|P_{k_1}Q_{l_1+O(1)}\psi_1\|_{L^2_{t,x}} \lesssim 2^{-\frac{k_1}{2}}2^{(1-\epsilon)(k_1-l_1)}\|P_{k_1}\psi\|_{S[k_1]}$$

One obtains a final bound with the same $\delta = \frac{1}{2}+\epsilon$ as in the preceding case.

In the case $k_1 < O(1)$, one simply places $P_{k_1}Q_{l_1+O(1)}\psi_1$ into $L^2_tL^\infty_x$, thereby gaining an additional factor 2^{k_1}. We omit the details.

(i5): $\max\{k_{1,2,3}\} > l + O(1)$, *time derivative falls on a term with modulation* $>> \max\{k_{1,2,3}\}$. This is similar to the preceding an omitted.

This concludes case (i), when the output is in the elliptic regime.

(ii): Output in hyperbolic regime. This is the expression

$$\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)$$

To treat it, we decompose

$$\begin{aligned}
P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))] =& P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\
&+P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}\nabla^{-1}IQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]
\end{aligned}$$

(iia): contribution of the elliptic type term. This is the expression

$$\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)$$

We shall treat the case $s_1 \ll -10$, i.e., the case of a high-low interaction within

$$P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}\nabla^{-1}Q_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]$$

The remaining cases are again more of the same. Now freeze the modulation of the expression

$$P_{s_1}\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))$$

to size 2^l, $l \gg s_1$. Then decompose the corresponding full expression into the following:

$$\begin{aligned}
&\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)\\
&=\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\
&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big) \qquad (6.4)\\
&+\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{<l-10}\psi_1P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\
&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big) \qquad (6.5)
\end{aligned}$$

The first term (6.4) on the right can then be estimated by

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{>l-10}\psi_1 P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)\|_{L^1_t\dot H^{-1}}\\&\lesssim \|P_{k_1}Q_{>l-10}\psi_1\|_{L^2_{t,x}}\|P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))\|_{L^2_tL^\infty_x}\\&\qquad\qquad\times\|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\|_{L^\infty_{t,x}}\end{aligned}$$

Then from Lemma 4.18 and Bernstein's inequality we infer that provided $k_2\gg s_1$, we have

$$2^{-\frac{s_1}{2}}\|P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))\|_{L^2_tL^\infty_x}\lesssim 2^{\epsilon l}2^{-\epsilon\max\{s_1,k_{2,3}\}}\max\{k_2-s_1,1\}^2\prod_{j=2,3}\|P_{k_j}\psi_j\|_{S[k_j]}$$

Inserting this into the preceding bound we infer that

$$\begin{aligned}&\|P_{k_1}Q_{>l-10}\psi_1\|_{L^2_{t,x}}\|P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))\|_{L^2_tL^\infty_x}\\&\qquad\qquad\times\|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\|_{L^\infty_{t,x}}\\&\lesssim 2^{\frac{s_1-l}{2}}2^{\epsilon l}2^{-\epsilon\max\{s_1,k_{2,3}\}}\max\{k_2-s_1,1\}^2 2^{r_1}2^{\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}}\prod_{j=1}^5\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

Summing over $l>s_1$ yields the bound of the first inequality of the lemma with $\delta=\epsilon-$. On the other hand, when $k_2=s_1+O(1)$, say, one can use Lemma 4.23 instead, which then gives the desired inequality with $\delta=\frac12-\epsilon$.

Next consider (6.5). Here we distinguish between the cases $l<r_1+O(1)$ and $l\gg r_1$. In the former case, as before assuming $s_1<-10$, we get

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{<l-10}\psi_1 P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)\|_{L^1_t\dot H^{1}}\\&\lesssim \|P_{k_1}Q_{<l-10}\psi_1\|_{L^\infty_tL^2_x}\|P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3)\|_{L^2_tL^\infty_x}\\&\qquad\qquad\times\|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\|_{L^2_tL^\infty_x}\end{aligned}$$

Using Lemma 4.18-4.23 again, we obtain the bound

$$\lesssim 2^{\frac{s_1+r_1}{2}}2^{\epsilon(l-\max\{k_2,k_3\})}|s_1-k_2|^2 2^{\frac{r_1}{2}}2^{\frac{\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}}{2}}\prod_{j=1}^5\|P_{k_j}\psi_j\|_{S[k_j]}$$

One may sum here over $s_1<l<r_1+O(1)$ to get the desired first inequality of the lemma with $\delta=\epsilon-$. Next, consider the case $l\gg r_1$. But in this case we can write

$$\begin{aligned}&\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{<l-10}\psi_1 P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)\\&=\nabla_{x,t}P_{[-5,5]}Q_{[l-10,10]}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{<l-10}\psi_1 P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)\end{aligned}$$

But this we can then estimate via the $\|.\|_{\dot X_0^{-1,-\frac12,1}}$-norm of the output, i.e., it suffices to bound

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_{[l-10,10]}[\big(P_0[\nabla^{-1}(P_{k_1}Q_{<l-10}\psi_1 P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))]\\&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\big)\|_{\dot X_0^{-1,-\frac12,1}}\\&\lesssim 2^{-\frac{l}{2}}\|P_{s_1}Q_l\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3)\|_{L^2_tL^\infty_x}\|P_{k_1}Q_{<l-10}\psi_1\|_{L^\infty_tL^2_x}\|\\&\qquad\qquad\times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4,P_{k_5}\psi_5)\|_{L^\infty_{t,x}}\end{aligned}$$

From here the estimates are continued in a fashion identical to the ones used to control (6.4). This completes estimating the contribution of $P_{s_1}\nabla^{-1}I^cQ_{\nu j}(P_{k_2}\psi_2,P_{k_3}\psi_3))$.

(iib): contribution of the hyperbolic type term. Next we consider the contribution of

$$P_{s_1}\nabla^{-1}IQ_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3)),$$

which is the expression

$$\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}IQ_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))] \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)$$

We shall again make the reduction $s_1 < -10$, the remaining frequency interactions being treated analogously. This is accomplished using Lemma 4.16. We obtain

$$\begin{aligned}&\|\nabla_{x,t}P_{[-5,5]}Q_{<10}[\big(P_0[\nabla^{-1}(P_{k_1}\psi_1 P_{s_1}\nabla^{-1}IQ_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3))] \\ &\qquad\qquad \times P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\big)\|_{L_t^1\dot{H}^{-1}} \\ &\lesssim \|P_{k_1}\psi_1\|_{L_t^\infty L_x^2}\|P_{s_1}\nabla^{-1}IQ_{\nu j}(P_{k_2}\psi_2, P_{k_3}\psi_3)\|_{L_t^2L_x^\infty}\|P_{r_1}\nabla^{-1}IQ_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)\|_{L_t^2L_x^\infty} \\ &\lesssim 2^{\frac{s_1+r_1}{2}}2^{\frac{\min\{s_1,k_{2,3}\}-\max\{s_1,k_{2,3}\}}{2}}2^{\frac{\min\{r_1,k_{4,5}\}-\max\{r_1,k_{4,5}\}}{2}}\prod_{j=1}^5\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

This is as desired with $\delta = \frac{1}{2}$. □

6.0.1. *Error terms of order higher than five.* Here we consider the errors generated by repeated application of Hodge decompositions, which are of higher than quintic degree. We recall that they arise when we apply repeated Hodge decompositions to the second and third input in

$$\nabla_{x,t}[\psi\nabla^{-1}(\psi^2)]$$

or else to the second and third input in

$$\nabla_{x,t}[\nabla^{-1}[\psi\nabla^{-1}(\psi^2)]\nabla^{-1}IQ_{\nu j}(\psi,\psi)]$$

To simplify the discussion, we shall call terms that arise in the first situation 'of the first type', while those in that arise in the second situation will be called of 'second type'. In either case, we associate a binary graph with each such expression as in the discussion above, see section 3. We call expressions whose associated graph has only directed subgraphs of length at most three 'short', and those with directed graphs of length at least four 'long'. For technical reasons, it will be most convenient to organize the 'short' and 'long' higher order terms into suitable sums, which are easier to estimate. Specifically, note that each of these higher order terms consists of nested terms of the form

$$\ldots\nabla^{-1}P_{s_1}[P_{k_1}R_\nu\psi P_{r_1}\nabla^{-1}(P_{k_2}\psi P_{s_2}\nabla^{-1}P_{s_2}[\ldots]], \tag{6.6}$$

here the case of a node with one outgoing edge, or alternatively

$$\ldots\nabla^{-1}P_{s_1}[P_{s_2}\nabla^{-1}[P_{k_1}\psi\nabla^{-1}P_{r_1}[\ldots]]\nabla^{-1}(P_{k_2}\psi P_{s_3}\nabla^{-1}[\ldots]] \tag{6.7}$$

in case of two outgoing edges.

It is the first type of expression which may cause some mild difficulties due to the presence of the R_ν-operator, which for $\nu = 0$ may be formally unbounded. However, re-combining a term of type (6.6) with a suitable term of the form (6.7) and using the relation

$$R_\nu\psi + \chi_\nu = \psi_\nu, \quad \psi = -\sum_{k=1,2}R_k\psi_k,$$

we replace each such 'intermediate' gradient term (i.e., not contributing to one of the innermost $Q_{\nu j}$ null-forms in case of 'short' expressions) $R_\nu\psi$ by its non-gradient counterpart ψ_ν. We shall call the resulting expressions 'reduced'. Thus for example the (short) quintilinear expression

$$\nabla_{x,t}[P_{k_1}\psi_1\nabla^{-1}P_{r_1}[P_{k_2}R_\nu\psi_2 P_{r_2}[P_{k_3}\psi\nabla^{-1}P_{r_3}Q_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)]]]$$

has reduced version

$$\nabla_{x,t}[P_{k_1}\psi_1\nabla^{-1}P_{r_1}[P_{k_2}\psi_{2\nu}P_{r_2}[P_{k_3}\psi\nabla^{-1}P_{r_3}Q_{\mu j}(P_{k_4}\psi_4, P_{k_5}\psi_5)]]]$$

Now we can formulate

Proposition 6.3. *Let*

$$P_0F_{2l+1}(\psi), \quad l=2,3,4$$

be short reduced higher order expression of first type at frequency ~ 1. *We can write it in nested form*

$$\begin{aligned}&P_0F_{2l+1}(\psi)=\\&\nabla_{x,t}P_0\big[P_{k_1}\psi_1\nabla^{-1}P_{r_1}\big[\dots\nabla^{-1}P_{r_j}\big[P_{k_{j+1}}\psi_{j+1}\nabla^{-1}P_{r_{j+1}}\big[P_{k_{j+2}}\psi_{j+2}\\&\qquad\times\nabla^{-1}\big[\dots\nabla^{-1}P_{r_{2l-1}}Q_{\mu j}(P_{k_{2l}}\psi_{2l},P_{k_{2l+1}}\psi_{2l+1})\big]\big]\big]\big]\big]\end{aligned}$$

Then we have the following bounds:

(1) If $r_1\ll -10$, *we have*

$$\|P_0F_{2l+1}(\psi)\|_{N[0]}\lesssim 2^{\delta[\min\{k_2,\dots,k_{2l+1},r_1,\dots,r_{2l-1}\}-\max\{k_2,\dots,k_{2l+1},r_1,\dots,r_{2l-1}\}}\prod_{j=1}^{2l+1}\|P_{k_j}\psi_j\|_{S[k_j]}$$

for a suitable constant $\delta>0$.

(2) If $r_1\in[-10,10]$, *we get the bound*

$$\|P_0F_{2l+1}(\psi)\|_{N[0]}\lesssim 2^{\delta k_1}2^{\delta[\min\{k_2,\dots,k_{2l+1},r_1,\dots,r_{2l-1}\}-\max\{k_2,\dots,k_{2l+1},r_1,\dots,r_{2l-1}\}}\prod_{j=1}^{2l+1}\|P_{k_j}\psi_j\|_{S[k_j]}$$

(3) If $r_1>10$, *we have*

$$\|P_0F_{2l+1}(\psi)\|_{N[0]}\lesssim 2^{-\delta k_1}2^{\delta[\min\{k_2,\dots,k_{2l+1},r_1,\dots,r_{2l-1}\}-\max\{k_2,\dots,k_{2l+1},r_1,\dots,r_{2l-1}\}}\prod_{j=1}^{2l+1}\|P_{k_j}\psi_j\|_{S[k_j]}$$

The proof of this follows the exact same pattern as the one for Proposition 6.1, and is omitted. In fact, for $l>2$, one no longer needs to use the sharp improved Strichartz endpoint as in the case $l=2$.

In a similar vein, we have the analogue of Proposition 6.2. A short reduced expression of the second type can be written as

(6.8)

$$\begin{aligned}P_0F_{2l+3}(\psi)=&\nabla_{x,t}P_0\big[\nabla^{-1}P_{r_1}\big[\dots\nabla^{-1}P_{r_j}\big[P_{k_{j+1}}\psi_{j+1}\nabla^{-1}P_{r_{j+1}}\big[P_{k_{j+2}}\psi_{j+2}\\&\times\nabla^{-1}\big[\dots\nabla^{-1}P_{r_{2l-1}}Q_{\mu j}(P_{k_{2l}}\psi_{2l},P_{k_{2l+1}}\psi_{2l+1})\big]\big]\big]\big]P_{s_1}Q_{\nu j}\nabla^{-1}(P_{k_{2l+2}}\psi_{2l+2},P_{k_{2l+3}}\psi_{2l+3}\big],\end{aligned}$$

where $l=1,2,3$. Then we have

Proposition 6.4. *Using the representation* (6.8), *let* $P_0F_{2l+3}(\psi)$ *be a short reduced term at frequency* ~ 1 *of the the second type. Then the following hold:*

(1) If $s_1<-10$, *we have*

$$\begin{aligned}\|P_0F_{2l+3}(\psi)\|_{N[0]}\lesssim&2^{\delta s_1}2^{\delta[\min\{s_1,k_{2l+2},k_{2l+3}\}-\max\{s_1,k_{2l+2},k_{2l+3}\}]}\\&\times 2^{\delta[\min\{k_2,\dots,k_{2l},r_1,\dots,r_{2l-1}\}-\max\{k_2,\dots,k_{2l},r_1,\dots,r_{2l-1}\}}\prod_{j=1}^{2l+3}\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

(2) If $s_1\in[-10,10]$, *we have*

$$\begin{aligned}\|P_0F_{2l+3}(\psi)\|_{N[0]}\lesssim&2^{\delta r_1}2^{\delta[\min\{s_1,k_{2l+2},k_{2l+3}\}-\max\{s_1,k_{2l+2},k_{2l+3}\}]}\\&\times 2^{\delta[\min\{k_2,\dots,k_{2l},r_1,\dots,r_{2l-1}\}-\max\{k_2,\dots,k_{2l},r_1,\dots,r_{2l-1}\}}\prod_{j=1}^{2l+3}\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

(3) If $s_1 > 10$, we have

$$\|P_0F_{2l+3}(\psi)\|_{N[0]} \lesssim 2^{-\delta r_1} 2^{\delta[\min\{s_1,k_{2l+2},k_{2l+3}\}-\max\{s_1,k_{2l+2},k_{2l+3}\}]} \times 2^{\delta[\min\{k_2,\ldots,k_{2l},r_1,\ldots,r_{2l-1}\}-\max\{k_2,\ldots,k_{2l},r_1,\ldots,r_{2l-1}\}} \prod_{j=1}^{2l+3} \|P_{k_j}\psi_j\|_{S[k_j]}$$

Again the proof is similar to the one of Proposition 6.2.
Note that in order to estimate the expressions of short type, we still need to to use a little bit of null-structure to make them amenable to estimation by the S-spaces. This is no longer the case for 'long' expressions $P_0F_11(\psi)$ of reduced type: write such an expression as

$$P_0F_{11}(\psi) = \nabla_{x,t}P_0\big[P_{k_1}\psi_1\nabla^{-1}P_{r_1}\big[\ldots\nabla^{-1}P_{r_j}\big[P_{k_{j+1}}\psi_{j+1}\nabla^{-1}P_{r_{j+1}}\big[P_{k_{j+2}}\psi_{j+2} \times \nabla^{-1}\big[\ldots\nabla^{-1}P_{r_9}(P_{k_{10}}\psi_{10}P_{k_{11}}\psi_{11})\big]\big]\big]\big]\big]$$

if it is of first type or

$$P_0F_{11}(\psi) = \nabla_{x,t}P_0\big[\nabla^{-1}P_{r_1}\big[\ldots\nabla^{-1}P_{r_j}\big[P_{k_{j+1}}\psi_{j+1}\nabla^{-1}P_{r_{j+1}}\big[P_{k_{j+2}}\psi_{j+2} \times \nabla^{-1}\big[\ldots\nabla^{-1}P_{r_9}(P_{k_{10}}\psi_{10}P_{k_{11}}\psi_{11})\big]\big]\big]\big]P_{s_1}\nabla^{-1}(P_{k_{12}}\psi_{12}P_{k_{13}}\psi_{13})\big],$$

if it is of the second type. Note that the innermost bilinear expressions

$$\nabla^{-1}P_{r_9}(P_{k_{10}}\psi_{10}P_{k_{11}}\psi_{11})$$

are no longer null-forms.

Proposition 6.5. *Let $P_0F_{11}(\psi)$ be a long expression of either first or second type, written as in immediately preceding. The if $P_0F_{11}(\psi)$ is of the first type, we have if $r_1 < -10$*

$$\|P_0F_{11}(\psi)\|_{N[0]} \lesssim 2^{\delta[\min\{k_2,\ldots,k_9,r_1,\ldots,r_9\}-\max\{k_2,\ldots,k_9,r_1,\ldots,r_9\}]} 2^{\delta[\min\{k_{10},k_{11}\}-\max\{k_{10},k_{11}\}]} \prod_{j=1}^{11} \|P_{k_j}\psi_j\|_{S[k_j]}$$

Thus by contrast to Proposition 6.3 case (1), we have an extra factor

$$2^{\delta[\min\{k_{10},k_{11}\}-\max\{k_{10},k_{11}\}]}$$

whence we cannot gain in case of high-high interactions in the innermost expression

$$\nabla^{-1}P_{r_9}(P_{k_{10}}\psi_{10}P_{k_{11}}\psi_{11})$$

Similarly, if P_0F_{11} is of the second type, we get

$$\|P_0F_{11}(\psi)\|_{N[0]} \lesssim 2^{\delta[\min\{k_2,\ldots,k_9,r_1,\ldots,r_9\}-\max\{k_2,\ldots,k_9,r_1,\ldots,r_9\}} 2^{\delta[\min\{k_{10},k_{11}\}-\max\{k_{10},k_{11}\}]} \times 2^{-\delta|s_1|} 2^{\delta[\min\{s_1,k_{12},k_{13}\}-\max\{s_1,k_{12},k_{13}\}]} \prod_{j=1}^{13} \|P_{k_j}\psi_j\|_{S[k_j]}$$

Proof. This is purely an application of our available Strichartz norms: indeed, for expressions of the first type, we have for suitable $\delta > 0$

$$\|P_{r_8}[P_{k_9}\psi_9\nabla^{-1}P_{r_9}(P_{k_{10}}\psi_{10}P_{k_{11}}\psi_{11})]\|_{L_t^8 L_x^{\frac{4}{3}+}} \lesssim 2^{(\frac{3}{8}+\nu)r_8} 2^{\delta[\min\{r_{8,9},k_9\}-\max\{r_{8,9},k_9\}]} 2^{\delta[\min\{k_{10},k_{11}\}-\max\{k_{10},k_{11}\}]} \prod_{j=9}^{11} \|P_{k_j}\psi_j\|_{S[k_j]}$$

where we define

$$\nu = \frac{3}{4} - (\frac{4}{3}+)^{-1}$$

Further, we have for $p = 1, 2, \ldots, 7$ and suitable $\delta_p > 0$

$$2^{-\nu_{p+1}r_1}\|P_{r_1}[P_k\psi\nabla^{-1}P_{r_2}F]\|_{L_t^{\frac{8}{p+1}} L_x^{\frac{4}{3}+}} \lesssim 2^{\delta_p[\min\{r_{1,2},k\}-\max\{r_{1,2},k\}]}[2^{-\nu_p r_2}\|P_{r_2}F\|_{L_t^{\frac{8}{p+1}} L_x^{\frac{4}{3}+}}]$$

where scaling dictates

$$\nu_p = 2 - \frac{p}{8} - 2(\frac{4}{3}+)^{-1}$$

The proposition follows by applying these two inequalities sufficiently often. The argument for expressions of the second type is similar. □

Remark 6.6. We note that in the estimates above, we have not used wave-packet atoms.

Remark 6.7. Using the arguments of the present subsection, it is straightforward to obtain the following refinement: given $\delta > 0$, there exists $C > 1$ large enough, such that if $P_0 F_{2i+1} = P_0[P_{k_1}\psi_1 P_{r_1}\nabla^{-1}[\dots]]$ is of the first type, $i \geq 2$, and we specialize to $k_1 = O(1)$ and

$$P_0\tilde{F}_{2i+1} := \sum_{r<-C} P_0[P_{k_1}\psi_1 P_{r_1} Q_{>r_1+2C}\nabla^{-1}[\dots]],$$

then we can improve the bounds of the preceding propositions by a factor δ.

7. Some basic perturbative results

This section develops some of the basic perturbative theory required for our work. More precisely, we introduce a norm locally on some time interval $(-T_0, T_1)$ which we denoted by $\|\psi\|_{S(-T_0,T_1)}$ with the property that its finiteness insures that the gauged wave map ψ can be continued outside of that time interval. The second topic we discuss is the issue of defining wave maps with data which are merely of energy class. This is accomplished by means of passing to the limit in energy of smooth wave maps.

7.1. **A blow-up criterion.** Assume we are given a wave map $\mathbf{u} : (-T_0, T_1) \times \mathbb{R}^2 \to \mathbb{H}^2$ with Schwartz data at time $t = 0$, by which we mean that the derivative components ϕ^i_α, $i = 1, 2$, $\alpha = 0, 1, 2$, and thus also the Coulomb components ψ^i_α, are Schwartz functions at time $t = 0$. These functions will then also be Schwartz on fixed time slices on the maximal interval of existence $(-T_0, T_1) \times \mathbb{R}^2$. The following norm will provide us with sufficient control for long time existence and scattering.

Definition 7.1. *For any Schwartz function on* $(-T_0, T_1) \times \mathbb{R}^2$ *set*

$$\|\psi\|_S := \Big(\sum_{k\in\mathbb{Z}} \|P_k\psi\|^2_{S[k]((-T_0,T_1)\times\mathbb{R}^2)}\Big)^{\frac{1}{2}}$$

Here

$$\|P_k\psi\|_{S[k]((-T_0,T_1)\times\mathbb{R}^2)} := \sup_{T<T_1, T'<T_0} \|P_k\psi_i\|_{S[k]([-T',T]\times\mathbb{R}^2)}$$

where the local norms are those from (2.67) using the $\|\!|\cdot|\!\|$*-norm.*

The goal of this section is to prove the following result.

Proposition 7.2. *Let* $(-T_0, T_1)$ *be the maximal interval of existence for the wave map* $\mathbf{u}$ *in the smooth sense. If* $\|\psi\|_S < \infty$ *then necessarily* $T_0 = T_1 = \infty$*. Moreover, the wave map scatters at infinity, i.e., the components* ψ, ϕ *approach free waves in the energy topology as* $t \to \pm\infty$*.*

The strategy for proving the theorem will be to demonstrate an a priori bound

$$\sup_{t\in(-T_0,T_1)} \|\phi(t,\cdot)\|_{H^s} < \infty$$

for some $s > 0$, using the assumption $\|\psi\|_S < \infty$. By the Klainerman-Machedon local well-posedness theory, this implies that u may be extended smoothly to some interval $(-T_0, T_1 + \varepsilon)$ for $\varepsilon > 0$ provided $T_1 < \infty$, which contradicts minimality, and similarly for T_0. Once we know that u exists for $t \in (-\infty, \infty)$, scattering will follow by using a similar argument. To obtain a priori control over sub-critical norms, we use Tao's device of *frequency envelope*: for some $\delta_1 > 0$ depending only on certain a priori parameters specified later, define

$$c_k := (\sum_{\ell\in\mathbb{Z}} 2^{-\delta_1|k-\ell|} \|P_\ell\phi(0,\cdot)\|^2_{L^2_x})^{\frac{1}{2}} \tag{7.1}$$

Here as always we let $\phi = \{\phi^i_\alpha\}$ the vector of derivative components. Proposition 7.2 now follows from the following result.

Proposition 7.3. *Let $(-T_0, T_1)$ be the maximal interval of existence for the wave map $\mathbf{u}$ in the smooth sense. If $\|\psi\|_S < \infty$, then there exists a number $C_1 = C_1(u) < \infty$ (which may depend in a complicated fashion on the wave map $\mathbf{u}$, and not just its energy)*[15] *such that*

$$\|P_k\phi\|_{L^\infty_t((-T_0,T_1);L^2_x)} \leq C_1 c_k$$

In fact,

$$\|P_k\psi\|_{S[k]((-T_0,T_1)\times\mathbb{R}^2)} \leq C_1 c_k$$

To establish the existence of C_1, we shall cover the time interval $(-T_0, T_1)$ by a finite number of shorter open intervals I_j (which can still be very large): let $\|\psi\|_S < C_0$. Then

$$(-T_0, T_1) = \cup_{j=1}^{M_1} I_j, \quad M_1 = M_1(C_0), \tag{7.2}$$

where $\psi|_{I_j}$ will satisfy a suitable smallness property. The idea then is to bootstrap certain bounds on each I_j, beginning with the interval containing the time slice $t = 0$. More precisely, the intervals I_j will be chosen so that the wave map restricted to each I_j is *well approximated by a free wave.* While the error can be treated perturbatively, the free wave has better dispersive properties which we can exploit. All functions will be smooth in space and time and Schwartz functions on fixed time slices.

7.1.1. *Splitting the wave map on shorter time intervals.* We first derive a simple estimate on the nonlinearities appearing in (1.12) and (1.13). It will be based entirely on the Strichartz estimates, see Lemma 2.17. We will keep the time interval $(-T_0, T_1)$ from above fixed throughout.

Lemma 7.4. *Let $\max_{i=1,2,3}\|\psi_i\|_S < C_0$. Then*

$$\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L^M_t((-T_0,T_1);L^2_x)} \lesssim C_0^2 \sup_{i=1,2,3}\sup_{k\in\mathbb{Z}} 2^{-\frac{|k|}{M}}\|P_k\psi_i\|_{S[k]((-T_0,T_1)\times\mathbb{R}^2)}$$

provided M is large and with an absolute implicit constant. Alternatively, one has the bound

$$\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L^M_t((-T_0,T_1);L^2_x)} \lesssim \|\psi_2\|_{L^\infty_t L^2_x}\|\psi_3\|_{L^\infty_t L^2_x}\sup_{k\in\mathbb{Z}} 2^{-\frac{|k|}{M}}\|P_k\psi_1\|_{S[k]((-T_0,T_1)\times\mathbb{R}^2)}$$

with an absolute implicit constant.

Proof. Assume to begin with that ψ_i is adapted to $k_i \in \mathbb{Z}$. As in Section 5, we now consider all possible cases of interactions. Also, we shall drop the time interval $(-T_0, T_1)$ from our notation with the understanding that integration in time is to be restricted to this interval. Moreover, replacing each ψ_i by a globally defined Schwartz function $\tilde{\psi}_i$ with the property that

$$\|\tilde{\psi}_i\|_{S[k_i]} \leq 2\|\psi_i\|_{S[k_i]([-T',T]\times\mathbb{R}^2)}, \quad \tilde{\psi}_i|_{[-T',T]} = \psi_i|_{[-T',T]}$$

for some T', T as above, allows us to assume that the ψ_i are globally defined initially. Finally, fix any $M \geq 100$.

Case 1: $0 \leq k_1 \leq k_2 + O(1) = k_3 + O(1)$. Then

$$\begin{aligned}&\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L^M_t L^2_x} \lesssim \|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L^M_t L^1_x}\\ &\lesssim \|\psi_1\|_{L^M_t L^\infty_x} 2^{-k_1}\|\psi_2\|_{L^\infty_t L^2_x}\|\psi_3\|_{L^\infty_t L^2_x} \lesssim 2^{-\frac{k_1}{M}}\prod_{i=1}^3\|\psi_i\|_{S[k_i]}\end{aligned}$$

where the final estimate is from (2.37).

[15]This is an artifact of the proof to follow and will be improved in the following sections

Case 2: $0 \le k_1 = k_2 + O(1), k_3 \le k_2 - C$. If $k_3 \ge 0$, one proceeds as in Case 1. Otherwise,

$$\begin{aligned}
&\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \lesssim \|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^1} \\
&\lesssim \|\psi_1\|_{L_t^\infty L_x^2} 2^{-k_1} \|\psi_2\|_{L_t^\infty L_x^2} \|\psi_3\|_{L_t^M L_x^\infty} \\
&\lesssim 2^{-k_1} 2^{k_3(1-\frac{1}{M})} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

by (2.37) and Bernstein's inequality.

Case 3: $0 \le k_1 = k_3 + O(1), k_2 \le k_3 - C$. This case is symmetric to the previous one.

Case 4: $O(1) \le k_2 = k_3 + O(1), k_1 \le -C$. Here

$$\begin{aligned}
&\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \lesssim \|P_0(\psi_1|\nabla|^{-1}\tilde{P}_0(\psi_2\psi_3))\|_{L_t^M L_x^1} \\
&\lesssim \|\psi_1\|_{L_t^M L_x^\infty} \|\psi_2\|_{L_t^\infty L_x^2} \|\psi_3\|_{L_t^\infty L_x^2} \lesssim 2^{(1-\frac{1}{M})k_1} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Case 5: $k_1 = O(1),\ k_2 = k_3 + O(1)$. In this case we estimate

$$\begin{aligned}
&\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \lesssim \sum_{k \le k_2 \wedge 0 + C} \|P_0(\psi_1|\nabla|^{-1}P_k(\psi_2\psi_3))\|_{L_t^M L_x^2} \\
&\lesssim \sum_{k \le k_2 \wedge 0 + C} \|\psi_1\|_{L_t^M L_x^{2+}} \||\nabla|^{-1}P_k(\psi_2\psi_3)\|_{L_t^\infty L_x^N} \\
&\lesssim \sum_{k \le k_2 \wedge 0 + C} \|\psi_1\|_{S[k_1]} 2^{(1-\frac{2}{N})k} \|\psi_2\psi_3\|_{L_t^\infty L_x^1} \\
&\lesssim 2^{(1-\frac{2}{N})k_2 \wedge 0} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Case 6: $O(1) = k_1 \ge k_2 + O(1) \ge k_3 + C$. Here one has

$$\begin{aligned}
&\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \lesssim \|P_0(\psi_1|\nabla|^{-1}\tilde{P}_{k_1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \\
&\lesssim \|\psi_1\|_{L_t^M L_x^{2+}} \||\nabla|^{-1}\tilde{P}_{k_1}(\psi_2\psi_3)\|_{L_t^\infty L_x^N} \\
&\lesssim \|\psi_1\|_{S[k_1]} \|\psi_2\psi_3\|_{L_t^\infty L_x^2} \lesssim \|\psi_1\|_{S[k_1]} \|\psi_2\|_{L_t^\infty L_x^2} \|\psi_3\|_{L_t^\infty L_x^N} \\
&\lesssim 2^{(1-\frac{2}{N})k_3} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Case 7: $k_1 = O(1) \ge k_3 + O(1) \ge k_2 + C$. This case is symmetric to the previous one.

Case 8: $k_2 = O(1), \max(k_1, k_3) \le -C$. Finally, in this case the estimate reads

$$\begin{aligned}
&\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \lesssim \|P_0(\psi_1|\nabla|^{-1}\tilde{P}_0(\psi_2\psi_3))\|_{L_t^M L_x^2} \\
&\lesssim \|\psi_1\|_{L_t^M L_x^\infty} \|\psi_2\psi_3\|_{L_t^\infty L_x^2} \lesssim 2^{k_1(1-\frac{1}{M})} 2^{k_3} \prod_{i=1}^3 \|\psi_i\|_{S[k_i]}
\end{aligned}$$

Case 9: $k_3 = O(1), \max(k_1, k_2) \le -C$. This is symmetric to the preceding case.

We now drop the assumption on the frequency support of the inputs. Summing over all these cases yields the bound

$$\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M L_x^2} \lesssim \sup_{i=1,2,3} \sup_{k \in \mathbb{Z}} \left[2^{-\frac{|k|}{M}} \|P_k\psi_i\|_{S[k]}\right] \max_{j=1,2,3} \sum_{k \in \mathbb{Z}} \|P_k\psi_j\|_{S[k]}^2$$

which proves the first bound. The proof of the second estimate is implicit in the preceding and the lemma is proved. □

Remark 7.5. If ψ_2, ψ_3 are gauged wave maps with energy bounded by E, then the second bound of the preceding lemma becomes

$$\|P_0(\psi_1|\nabla|^{-1}(\psi_2\psi_3))\|_{L_t^M((-T_0,T_1);L_x^2)} \lesssim E^2 \sup_{k\in\mathbb{Z}} 2^{-\frac{|k|}{M}} \|P_k\psi_1\|_{S[k]((-T_0,T_1)\times\mathbb{R}^2)} \tag{7.3}$$

with an absolute implicit constant.

Our main goal here is to prove the following decomposition of the gauged wave map.

Lemma 7.6. *Let $\|\psi\|_S < C_0$. Given $\varepsilon_0 > 0$, there exist $M_1 = M_1(C_0, \varepsilon_0)$ many intervals I_j as in (7.2) with the following property: for each $I_j = (t_j, t_{j+1})$, there is a decomposition*

$$\psi|_{I_j} = \psi_L^{(j)} + \psi_{NL}^{(j)}, \quad \Box\psi_L^{(j)} = 0$$

which satisfies

$$\sum_{k\in\mathbb{Z}} \|P_k\psi_{NL}^{(j)}\|^2_{S[k](I_j\times\mathbb{R}^2)} < \varepsilon_0 \tag{7.4}$$

$$\|\nabla_{x,t}\psi_L^{(j)}\|_{L_t^\infty \dot{H}^{-1}} \le M_2(C_0,\varepsilon_0) \tag{7.5}$$

where the constant $M_2 = M_2(C_0,\varepsilon_0)$ satisfies $M_2 \lesssim C_0^3 \varepsilon_0^{-\frac{1}{M}}$ with $M \ge 100$ as in the preceding lemma. Moreover, $P_k\psi_{NL}^{(j)}$ and $P_k\psi_L^{(j)}$ are Schwartz functions for each $k \in \mathbb{Z}$. We also have the bounds

$$\|\nabla_{x,t}P_k\psi_L^{(j)}\|_{\dot{H}_x^{-1}} + \|P_k\psi_{NL}^{(j)}\|_{S[k](I_j\times\mathbb{R}^2)} \lesssim c_k \tag{7.6}$$

with implied constant depending on C_0, provided c_k is a sufficiently flat frequency envelope with $\|P_k\psi\|_{S[k]} \le c_k$.

Proof. The ψ_α satisfy the system (1.12)–(1.14). Consider the frequency component $P_0\psi_\alpha$.
Case 1: The underlying time interval $I = (-T_0, T_1)$ is very small, say $|I| < \varepsilon_1$ with an ε_1 that is to be determined. As explained in Section 2.5 one uses the div-curl system (1.12), (1.13) in this case. Schematically, this system takes the form

$$\partial_t P_0\psi = \nabla_x P_0\psi + P_0[\psi\nabla^{-1}(\psi^2)]$$

where we suppress the subscripts and also ignore the null-structure in the nonlinearity. Therefore,

$$\|P_0\psi(t) - P_0\psi(0)\|_{L_x^2} \le \Big\|\int_0^t \nabla_x P_0\psi(s,\cdot)\,ds\Big\|_{L_x^2} + \Big\|\int_0^t P_0[\psi\nabla^{-1}(\psi^2)](s,\cdot)\,ds\Big\|_{L_x^2} \tag{7.7}$$

For all $j \in \mathbb{Z}$ define

$$a_j := \sup_{k\in\mathbb{Z}} 2^{-\frac{|k-j|}{M}} \|P_k\psi\|_{S[k](I\times\mathbb{R}^2)} \lesssim C_0 \tag{7.8}$$

Clearly,

$$\Big\|\int_0^t \nabla_x P_0\psi(s,\cdot)\,ds\Big\|_{L_x^2} \le \varepsilon_1 \|P_0\psi\|_{S[0](I\times\mathbb{R}^2)} \le a_0\,\varepsilon_1 \tag{7.9}$$

Lemma 7.4 implies

$$\Big\|\int_0^t P_0[\psi\nabla^{-1}(\psi^2)](s,\cdot)\,ds\Big\|_{L_t^\infty(I;L_x^2)} \lesssim C_0^2\,a_0\,\varepsilon_1^{1-\frac{1}{M}}$$

$$\Big\|\int_0^t P_0[\psi\nabla^{-1}(\psi^2)](s,\cdot)\,ds\Big\|_{L_t^2(I;L_x^2)} \lesssim C_0^2\,a_0\,\varepsilon_1^{\frac{3}{2}-\frac{1}{M}}$$

¿From the div-curl system (1.12) and (1.13),

$$\|\partial_t P_0\psi\|_{L_t^2(I;L_x^2)} \le \|\nabla_x P_0\psi\|_{L_t^2(I;L_x^2)} + \|P_0[\psi\nabla^{-1}(\psi^2)]\|_{L_t^2(I;L_x^2)} \lesssim C_0^2\,a_0\,\varepsilon_1^{\frac{1}{2}-\frac{1}{M}}$$

where we assumed without loss of generality that $C_0 \ge 1$. We claim that these bounds imply that

$$\Big\|\int_0^t P_0[\psi\nabla^{-1}(\psi^2)]\Big\|_{S[0](I\times\mathbb{R}^2)} \ll \varepsilon_0 a_0 \tag{7.10}$$

provided ε_1 was chosen sufficiently small depending on ε_0. To see this, let $I' := [-T', T] \subset I = (-T_0, T_1)$ and pick any smooth bump function χ supported in I so that $\chi = 1$ on I' and with $0 \le \chi \le 1$. Moreover, let $\tilde{\chi}$ be any smooth compactly supported function with $\tilde{\chi} = 1$ on I (the choice of this function does not depend on I'). Then define

$$\tilde{\psi}(t) := \tilde{\chi}(t)\big[P_0\psi(0) + \int_0^t \chi(s)\partial_s P_0\psi(s)\,ds\big]$$

By construction, $\tilde{\psi}$ is a global Schwartz function so that $\tilde{\psi} = \psi$ on I'. Moreover, by the preceding bounds,

$$\|\tilde{\psi}\|_{L_t^2 L_x^2} + \|\partial_t \tilde{\psi}\|_{L_t^2 L_x^2} \ll \varepsilon_0\, a_0$$

provided ε_1 was chosen small enough (this smallness does not depend on the choice of I'). This now implies that

$$\|\tilde{\psi}\|_{\dot{X}_0^{0,\frac{1}{2},1}} \ll \varepsilon_0\, a_0$$

whence (7.10). In view of (7.9), (7.10) and (7.7),

$$\|P_0\psi(t) - P_0\psi(0)\|_{S[0](I\times\mathbb{R}^2)} \ll \varepsilon_0\, a_0 \tag{7.11}$$

We now define $P_0\psi_L$ to be the free wave with initial data $(P_0\psi(0), 0)$ at time $t = 0$. Clearly

$$\|P_0\nabla_{x,t}\psi_L\|_{L_t^\infty \dot{H}^{-1}} \lesssim \|P_0\psi(0)\|_{L_x^2}$$
$$\|P_0\psi_L - P_0\psi_L(0)\|_{S[0](I\times\mathbb{R}^2)} \ll \varepsilon_0\, a_0$$

The second inequality here implies that

$$\|P_0\psi - P_0\psi_L\|_{S[0](I_j\times\mathbb{R}^2)} \ll \varepsilon_0\, a_0$$

Thus in the present situation, we approximate $P_0\psi$ by the free wave $P_0\psi_L$ just described and the bounds which we just obtained should be viewed as versions of (7.4) and (7.5) on a fixed dyadic frequency block. Several remarks are in order: First, we shall of course need to construct ψ_L and ψ_{NL} for each such dyadic block P_k, and then obtain the global bounds required by (7.4) and (7.5). In this regard, any bound depending on a_j can easily be square-summed since

$$\sum_j a_j^2 \le C(M) \sum_{k\in\mathbb{Z}} \|P_k\psi\|^2_{S[k](I\times\mathbb{R}^2)} \le C(M)\, C_0^2$$

Second, the construction we just carried out applies to $P_k\psi$ equally well provided $|I| \le 2^{-k}\varepsilon_1$. Moreover, I can be any time interval on which ψ is defined — with any $t_0 \in I$ playing the role of $t = 0$ — and we shall indeed apply this exact same procedure to those intervals I_j which we are about to construct provided they satisfy this length restriction.

Case 2: The underlying time interval $I = (-T_0, T_1)$ satisfies $|I| > \varepsilon_1$ with ε_1 as in Case 1. To construct the I_j, we shall use the wave equation (1.14) for ψ_α. By means of Schwartz extensions and successive Hodge type decompositions of the ψ_α-components as explained above, the nonlinearity can be written as

$$\Box\psi_\alpha = F_\alpha(\psi) = F_\alpha^3(\psi) + F_\alpha^5(\psi) + F_\alpha^7(\psi) + F_\alpha^9(\psi) + F_\alpha^{11}(\psi), \tag{7.12}$$

where the superscripts denote the degree of multi-linearity, see Section 3. The contribution of the trilinear null-form $F_\alpha^3(\psi)$ here is in a sense the principal contribution, and causes the main technical difficulties. We now make the following claim: *There exists a cover $I = \bigcup_{j=1}^{M_1} I_j$ by open intervals I_j, $1 \le j \le M_1$, $M_1 = M_1(\varepsilon_0)$, such that*

$$\max_{1\le j\le M_1} \sum_{\ell\in\mathbb{Z}} \|P_\ell F_\alpha(\psi)\|^2_{N[\ell](I_j\times\mathbb{R}^2)} < \varepsilon_0 C_0^6 \tag{7.13}$$

This will be enough to ensure the conditions of the lemma, if we replace ε_0 by $C_0^{-6}\varepsilon_0$, whence the number of intervals will then also depend on C_0, the bound on $\|\psi\|_S$. We verify this for each of the different types of nonlinearities appearing on the right-hand side of (7.12) starting with the trilinear ones. Let us schematically write anyone of these trilinear expressions in the form $\nabla_{t,x}[\psi_1|\nabla|^{-1}I^c\mathcal{Q}(\psi_2,\psi_3)]$ or $\nabla_{t,x}[R\,\psi_1|\nabla|^{-1}I\mathcal{Q}(\psi_2,\psi_3)]$, where $\mathcal{Q}$ stands for the usual bilinear nullforms and R for a Riesz transform (each of the $\psi_i = \psi$ but it will be convenient to view these inputs as independent). Break up the inputs into dyadic frequency pieces: $\psi_i = \sum_{k_i} P_{k_i}\psi_i$ for $i = 1, 2, 3$. In view of our discussion in Section 5.3, it suffices to consider the high-low-low case $|k_2 - k_3| < L$, $k_2 < k_1 + L$ for some large $L = L(\varepsilon_0)$. In addition, it suffices to restrict attention to frequencies $k > k_2 - L'$ where P_k localizes the frequency of $\mathcal{Q}$ and $L' = L'(\delta)$ is large. Finally, one can assume angular separation between the inputs: there exists $m_0 = m_0(\varepsilon_0) \ll -1$ so that (7.13) reduces to the estimates

$$\max_{1\le j\le M_1} \sum_{\ell} \Big\| \sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \nabla_{t,x} P_\ell [P_{k_1,\kappa_1}\psi_1 |\nabla|^{-1} P_k I^c \mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\Big\|^2_{N[\ell](I_j\times\mathbb{R}^2)} < \varepsilon_0 C_0^6 \tag{7.14}$$

$$\max_{1\le j\le M_1} \sum_{\ell} \Big\| \sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \nabla_{t,x} P_\ell [P_{k_1,\kappa_1} R\,\psi_1 |\nabla|^{-1} P_k I \mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\Big\|^2_{N[\ell](I_j\times\mathbb{R}^2)} < \varepsilon_0 C_0^6 \tag{7.15}$$

where the sums extend over integers k, ℓ, k_1, k_2, k_3 as specified above and further $|k_1 - \ell| < L$, as well as over caps $\kappa_1, \kappa_2, \kappa_3 \in \mathcal{C}_{m_0}$ with $\mathrm{dist}(\kappa_i, \kappa_j) > 2^{m_0}$ for $i \ne j$. Let us first consider the case where the entire output is restricted by $Q_{<2m_0+\ell}$ in modulation, and the inputs are in the hyperbolic regime, i.e., $P_{k_i}\psi_i = Q_{\le k_i+C} P_{k_i}\psi_i$ where C is large depending on L. Then we bound (7.14) (and (7.15)) as follows, first on the whole time axis $\mathbb{R}$ (assuming as we may that the inputs have been suitably extended):

$$\begin{aligned}
&\sum_{\ell} \Big\| \sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \nabla_{t,x} Q_{<2m_0+\ell-C} P_\ell [P_{k_1,\kappa_1}\psi_1 |\nabla|^{-1} P_k I^c \mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\Big\|^2_{N[\ell]} \\
&\lesssim \sum_{\ell} \sum_{\kappa\in\mathcal{C}_{m_0}} \Big\| \sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} P_{\ell,\kappa} Q_{<2m_0+\ell-C} [P_{k_1,\kappa_1}\psi_1 |\nabla|^{-1} P_k I^c \mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\Big\|^2_{\mathrm{NF}[\kappa]}
\end{aligned} \tag{7.16}$$

Note that the $2^{2\ell}$-factor produced by the output is canceled against the scaling factor which is part of the $N[\ell]$-norm, see Definition 2.9. By the usual arguments involving disposable multipliers, we may replace $|\nabla|^{-1} P_k I^c$ by 2^{-k_2} (implicit constants are allowed to depend on L). Since the inputs are hyperbolic, we may also ignore the null-form $\mathcal{Q}$. For any κ,

$$\max_{i=2,3} \mathrm{dist}(\kappa, \kappa_i) \gtrsim 2^{m_0}$$

Let us assume that this happens for $i = 3$. Then by (2.29), followed by (2.30) and Cauchy-Schwarz(recall that $k_2 = k_3 + O(1)$),

$$\begin{aligned}
(7.16) &\le C(L, m_0) \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} 2^{-k_2} \|P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2\|^2_{L^2_t L^2_x} \|P_{k_3,\kappa_3}\psi_3\|^2_{S[k_3]} \\
&\le C(L, m_0) \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \|P_{k_1,\kappa_1}\psi_1\|^2_{S[k_1,\kappa_1]} \|P_{k_2,\kappa_2}\psi_2\|^2_{S[k_2,\kappa_2]} \|P_{k_3,\kappa_3}\psi_3\|^2_{S[k_3,\kappa_3]} \\
&\le C(L, m_0) \Big(\sum_{k\in\mathbb{Z}} \|P_k\psi\|^2_{S[k]} \Big)^3 \le C(L, m_0) C_0^6
\end{aligned}$$

Note that we are not assuming that $P_{k_i,\kappa_i}\psi_i$ are wave-packets, i.e., localized in modulation to $< 2^{2m_0+k_i}$ but only to modulations $< 2^{k_i+C}$. Therefore, to pass to the last line one needs to use Lemma 2.7 for the modulations between these two cut-offs. However, this only costs a factor of $\lesssim |m_0|$ which is admissible. We now rewrite the first line in this estimate in the form

$$(7.16) \le C(L, m_0) \int_{\mathbb{R}^3} \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} 2^{-k_2} |P_{k_1,\kappa_1}\psi_1(t,x) P_{k_2,\kappa_2}\psi_2(t,x)|^2 \|P_{k_3,\kappa_3}\psi_3\|^2_{S[k_3]}\, dtdx \le C(L, m_0) C_0^6$$

By the dominated convergence theorem, we can cover the line (and especially $(-T_0, T_1)$) into finitely many intervals I_j such that

$$C(L, m_0) \int_{I_j\times\mathbb{R}^2} \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} 2^{-k_2} |P_{k_1,\kappa_1}\psi_1(t,x) P_{k_2,\kappa_2}\psi_2(t,x)|^2 \|P_{k_3,\kappa_3}\psi_3\|^2_{S[k_3]}\, dtdx < \varepsilon_0 C_0^6 \tag{7.17}$$

for each I_j. Moreover, the number of these intervals is $\leq M_1(\varepsilon_0)$. Unfortunately, in the preceding calculations we happily suppressed the action of the nonlocal operator $P_k I^c$, which is given by convolution with a kernel (in both t, x) of bounded L^1-mass. Although this is only a technical nuisance, we quickly explain here how to deal with it for completeness' sake: consider the schematically written expression

$$\sum_{\ell} \| \sum_{k_1=\ell+O(1)} \sum_{\substack{k_2=k_3+O(1)\\ \leq k_1+O(1)}} \int_{\mathbb{R}^{2+1}} m(a)\, P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2(\cdot - a) P_{k_3,\kappa_3}\psi_3(\cdot - a)\|_{N[\ell]}^2$$

where we have $\|m(\cdot)\|_{L^1(\mathbb{R}^{2+1})} = O(1)$. Under the same assumptions on the frequency localizations as above, we estimate this by

$$\lesssim \sum_{\ell} \sum_{k_1=\ell+O(1)} \Big(\sum_{\substack{k_2=k_3+O(1)\\ \leq k_1+O(1)}} \int_{\mathbb{R}^{2+1}} m(a) \|P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2(\cdot - a)\|_{L^2_{t,x}}\, da\, \|P_{k_3,\kappa_3}\psi_3\|_{S[k_3]}\Big)^2 \tag{7.18}$$

where we have used Minkowski's inequality as well as the translation invariance of the norms used. We proceed by using Cauchy-Schwarz twice in a row, to bound the preceding by

$$\begin{aligned}
&\lesssim \sum_{k_1=\ell+O(1)} \Big(\sum_{\substack{k_2=k_3+O(1)\\ \leq k_1+O(1)}} \Big[\int_{\mathbb{R}^{2+1}} |m(a)| \|P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2(\cdot - a)\|_{L^2_{t,x}}^2\, da\Big]^{\frac{1}{2}} \|P_{k_3,\kappa_3}\psi_3\|_{S[k_3]}\Big)^2 \\
&\lesssim \sum_{k_1=\ell+O(1)} \Big(\int_{\mathbb{R}^{2+1}} |m(a)| \|P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2(\cdot - a)\|_{L^2_{t,x}}^2\, da\Big) \Big(\sum_{k_3} \|P_{k_3,\kappa_3}\psi_3\|_{S[k_3]}^2\Big) \\
&\lesssim \Big(\sum_{k_1 \geq k_2+O(1)} \int_{\mathbb{R}^{2+1}} |m(a)| \int_{\mathbb{R}^{2+1}} |P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2(\cdot - a)|^2\, dtdx\, da\Big) \Big(\sum_{k_3} \|P_{k_3,\kappa_3}\psi_3\|_{S[k_3]}^2\Big)
\end{aligned}$$

Thus, properly speaking, instead of smallness of

$$\int_{I_j \times \mathbb{R}^2} \sum_{k_1 \geq k_2+O(1)} 2^{-k_2} |P_{k_1,\kappa_1}\psi_1(t,x) P_{k_2,\kappa_2}\psi_2(t,x)|^2,$$

we need to achieve smallness of

$$\sum_{k_1 \geq k_2+O(1)} \int_{I_j \times \mathbb{R}^2} |m(a)| \int_{\mathbb{R}^{2+1}} |P_{k_1,\kappa_1}\psi_1 P_{k_2,\kappa_2}\psi_2(\cdot - a)|^2\, dtdx\, da,$$

which of course can be achieved identically, via divisibility of the inner integral and Fubini's theorem. We shall henceforth suppress this technicality and stick with the notationally simpler condition (7.17) as well as similar ones in the sequel, it being understood that we sometimes suppress harmless convolution operators.

Retracing our steps shows that the intervals I_j have the desired properties (7.14) and (7.15) under the modulation assumptions $P_{k_i}\psi_i = Q_{\leq k_i+C} P_{k_i}\psi_i$, and the additional assumption that the output is limited to size $\lesssim 2^{2m_0+\ell}$ (the Schwartz extensions implicit in (7.14) and (7.15) are simply obtained by multiplying the L^2_{tx} functions by smooth bump functions). The remaining cases where these modulation assumptions are violated are handled similarly. For example, consider (7.15) for outputs of modulations $\gtrsim 2^{\ell+C}$ but again on the whole time axis; here, it is easy to see that we may assume $k = k_j + O(1)$, $j = 1, 2, 3$, whence we reduce to estimating (with the preceding frequency restraint implicit in the summation)

$$
\begin{aligned}
&\sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \|\nabla_{t,x} P_\ell Q_{\geq \ell+C}[P_{k_1,\kappa_1} R\,\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[\ell]}^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \|\nabla_{t,x} P_\ell Q_{\geq \ell+C}[P_{k_1,\kappa_1} R\,\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{\dot{X}_\ell^{-\frac12+\varepsilon,-1-\varepsilon,2}}^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \Big(\sum_{m\geq \ell+C} 2^{-(\frac12-\varepsilon)\ell} 2^{(1-\varepsilon)m} 2^{-\ell} \|P_{k_1,\kappa_1} Q_m \psi_1\|_{L_t^2 L_x^2}\Big)^2\, 2^{-2k} \|P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)\|_{L_t^\infty L_x^\infty}^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} 2^{-2\ell} \|P_{k_1,\kappa_1}\psi_1\|_{\dot{X}_\ell^{-\frac12+\varepsilon,1-\varepsilon,2}}^2\, 2^k \|P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)\|_{L_t^2 L_x^2}^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} 2^{-2\ell} \|P_{k_1,\kappa_1}\psi_1\|_{\dot{X}_\ell^{-\frac12+\varepsilon,1-\varepsilon,2}}^2\, 2^k \|\mathcal{L}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)\|_{L_t^2 L_x^2}^2\\
&\leq C(L,m_0) C_0^6
\end{aligned}
$$

where the final bound again follows from (2.29) ($\mathcal{L}$ stands for the usual averaged space-time translation operator which arises via removal of disposable multipliers). Writing out the $L_t^2 L_x^2$-norm explicitly in the previous estimate allows us again to choose intervals I_j with the desired properties. The remaining case of output modulations Q_m with $2m_0+\ell\leq m\leq \ell+C$ is similar:

$$
\begin{aligned}
&\sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \|\nabla_{t,x} P_\ell Q_m[P_{k_1,\kappa_1} R\,\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[\ell]}^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \|\nabla_{t,x} P_\ell Q_m[P_{k_1,\kappa_1} R\,\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{\dot{X}_\ell^{-\frac12+\varepsilon,-1-\varepsilon,2}}^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} 2^{-\ell} \|P_{k_1,\kappa_1} Q_m\psi_1\|_{L_t^\infty L_x^2}^2\, \|P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)\|_{L_t^2 L_x^2}^2\\
&\leq C(L,m_0) C_0^6
\end{aligned}
$$

Due to the $L_t^2 L_x^2$-norm one can now proceed as before. Finally, suppose that the output as well as ψ_1 are hyperbolic, but that ψ_2 and ψ_3 are elliptic. Then, restricting the inner sum to $k=\ell+O(1)\geq k_2=k_3+O(1)$ as we may, we have

$$
\begin{aligned}
&\sum_\ell \|\sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \nabla_{t,x} P_\ell Q_{\leq \ell+C}[P_{k_1,\kappa_1} Q_{\leq k_1+C} R\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]\|_{N[\ell]}^2\\
&\lesssim \sum_\ell 2^{-2\ell} \|\sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \sum_{m\geq k_2+C} \nabla_{t,x} P_\ell Q_{\leq \ell+C}[P_{k_1,\kappa_1} Q_{\leq k_1+C} R\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2} Q_m\psi_2, P_{k_3,\kappa_3}\tilde{Q}_m\psi_3)]\|_{L_t^1 L_x^2}^2
\end{aligned}
$$

which can be further estimated as (using Cauchy-Schwarz for the third inequality)

$$\begin{aligned}
&\lesssim \sum_{\ell} \| \sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \sum_{m\geq k_2+C} P_\ell Q_{\leq \ell+C}[P_{k_1,\kappa_1} Q_{\leq k_1+C} R\psi_1 |\nabla|^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2} Q_m\psi_2, P_{k_3,\kappa_3}\tilde{Q}_m\psi_3)]\|^2_{L^1_t L^2_x}\\
&\lesssim \sum_{\ell} \Big(\sum_{\substack{k,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \sum_{m\geq k_2+C} \|P_{k_1,\kappa_1}\psi_1\|_{L^\infty_t L^2_x} \|\nabla^{-1} P_k I\mathcal{Q}(P_{k_2,\kappa_2} Q_m\psi_2, P_{k_3,\kappa_3}\tilde{Q}_m\psi_3)]\|_{L^1_t L^2_x}\Big)^2\\
&\lesssim \sum_{\substack{k,\ell,k_1,k_2,k_3\\ \kappa_1,\kappa_2,\kappa_3}} \sum_{m_{1,2}\geq k_2+C} \|P_{k_1,\kappa_1}\psi_1\|^2_{L^\infty_t L^2_x}\, 2^{2(m_1-k_2)} \|P_{k_2,\kappa_2} Q_{m_1}\psi_2\|^2_{L^2_t L^2_x} \|P_{k_3,\kappa_3}\tilde{Q}_{m_2}\psi_3\|^2_{L^2_t L^\infty_x}\\
&\leq C(L,m_0) C_0^6
\end{aligned}$$

by Bernstein's inequality and the definition of $S[k]$; to pass to the second line use that $P_\ell Q_{\leq \ell+C}$ is disposable. Partitioning $\mathbb{R}$ into finitely many intervals on which $\sum_{m\geq k} 2^{2(1-\varepsilon)m} 2^{-(1-2\varepsilon)k} \|P_k Q_m\psi_2\|^2_{L^2_{t,x}}$ is small allows us to obtain the desired conclusion as before. Alternatively, one can gain smallness here by taking C in $m \geq k_2 + C$ large; this will be important later (see Remark 7.8). We leave the analogous analysis of (7.14) to the reader.

The proof of the claim (7.13) for the higher degree nonlinearities is outlined in the Appendix.

A crucial feature of the construction of the intervals $\{I_j\}_{1\leq j\leq M_1}$ above is that is *universal*, i.e., it does not depend on the choice of the underlying frequency scale. We now conclude the proof of Lemma 7.6. Fix some I_j and localize ψ to frequency 2^k. If $|I_j| < \varepsilon_1 2^{-k}$, then $P_k\psi^{(j)}_{NL} := P_k\psi - P_k\psi_L$ satisfies the bound (7.4) by the analysis in Case 1. Otherwise, one represents the solution via (2.72). The bounds in Case 1 above then imply the estimate

$$\|(P_k\psi)\big|_{[t_j-\varepsilon_1 2^{-k}, t_j+\varepsilon_1 2^{-k}]}\|_{S[k]} \lesssim \|P_k\psi\|_{S[k]}$$

The free wave $P_k\psi_L$ at dyadic frequency 2^k is now defined as the free evolution in (2.72) with data at some $t_0 \in [t_j - \varepsilon_1 2^{-k}, t_j + \varepsilon_1 2^{-k}]$ where

$$\|(P_k\psi_\alpha,\, \partial_t P_k\psi_\alpha)(t_0,\cdot)\|_{L^2\times \dot{H}^{-1}} \lesssim \varepsilon_0^{-\frac{1}{M}} a_k,$$

whereas $P_k\psi^{(j)}_{NL}$ is the sum of the other two terms in that formula. That the preceding choice of t_0 is possible follows from Lemma 7.4. Summing over k now yields the claimed local splitting of ψ in Lemma 7.6. Finally, the proof of (7.6) is implicit in the preceding and we skip the details. ☐

Remark 7.7. Later we will apply (7.6) in the following context. If

$$\Big(\sum_{k>k_0} \|P_k\psi\|^2_{S[k]}\Big)^{\frac{1}{2}} < \delta_3$$

for some (very small) $\delta_3 > 0$, we have

$$\|\nabla_{x,t} P_{>k_0}\psi^{(j)}_L\|_{\dot{H}^{-1}_x} + \Big(\sum_{k>k_0} \|P_k\psi^{(j)}_{NL}\|^2_{S[k](I_j\times\mathbb{R}^2)}\Big)^{\frac{1}{2}} \lesssim \delta_3$$

where the implied constant depends on $\|\psi\|_S$.

Remark 7.8. The preceding proof can be easily modified to give the following result that will be important later: Let ψ be the gauged derivative components of an admissible wave map. Assume that we have an a priori bound of the form

$$\sum_{k_1>k_2} \sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2} \|P_{k_1,\kappa_1}\psi P_{k_2,\kappa_2}\psi\|^2_{L^2_{t,x}} + \sum_{k<l} [2^{(1-\varepsilon)l-(\frac{1}{2}-\varepsilon)k} \|P_k Q_l\psi\|_{L^2_{t,x}}]^2 < \Lambda \tag{7.19}$$

where m_0 is sufficiently large depending on the energy E of ψ. Then we can infer a bound of the form

$$\|\psi\|_S \lesssim C_2(E,\, m_0,\, \Lambda)$$

This is done by a bootstrap, with the desired smallness coming either from the intervals I_j or the gains from the angular alignment. Moreover, assume that for each $k_1 > k_2$ one has

$$\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} P_{k_1,\kappa_1}\psi P_{k_2,\kappa_2}\psi = f_{k_1,k_2} + g_{k_1,k_2}$$

$$P_k Q_{>k}\psi = h_k + i_k$$

with for some positive integer ν

$$\begin{aligned}\sum_{k_1>k_2} 2^{-k_2}\|f_{k_1,k_2}\|_{L^2_{t,x}}^2 + \sum_{k<l}[2^{(1-\varepsilon)l-(\frac12-\varepsilon)k}\|Q_l h_k\|_{L^2_{t,x}}]^2 < \Lambda\\ \sum_{k_1>k_2} 2^{-k_2}\|g_{k_1,k_2}\|_{L^2_{t,x}}^2 + \sum_{k<l}[2^{(1-\varepsilon)l-(\frac12-\varepsilon)k}\|Q_l i_k\|_{L^2_{t,x}}]^2 < \delta\|\psi\|_S^\nu\end{aligned} \tag{7.20}$$

where $\delta > 0$ is small depending on the energy and the integer ν, but independent of $\|\psi\|_S$. Then one can again conclude

$$\|\psi\|_S \lesssim C_2(E,\, m_0,\, \Lambda)$$

Note the the time intervals I_j are determined only by means of the f_{k_1,k_2} and not the g_{k_1,k_2}.

7.1.2. *Proof of Proposition 7.3.* Recall that we are making the assumption $\|\psi\|_S < C_0$. We first show that the wave map cannot break down in finite time, i.e., $T = T' = \infty$. Assume for example that $T < \infty$. For $\varepsilon_0 > 0$ a sufficiently small but absolute constant (which will be specified later), pick the $M_1(C_0, \varepsilon_0)$-many intervals I_j as in Lemma 7.6. It will suffice to consider that interval I_{j_0} which has T as its endpoints. Alternatively, starting with that interval I_j containing the initial time slice $t = 0$, one can inductively obtain control over the frequency-localized constituents of ψ, the $P_k\psi$.

Lemma 7.9. *Let $I_j = (t_j, t_{j+1})$ be an interval as in Lemma 7.6. Introduce the frequency envelope*

$$c_k := \Big(\sum_{\ell\in\mathbb{Z}} 2^{-\sigma_0|k-\ell|}\|P_\ell\psi(t_j,\cdot)\|_{L^2_x}^2\Big)^{\frac12}$$

where $\sigma_0 > 0$ is some small constant. Also, write $\psi|_{I_j} = \psi_L + \psi_{NL}$. Then there is a number $C_1 = C_1(\psi_L) < \infty$ with the property that

$$\|P_k\psi\|_{S[k](I_j\times\mathbb{R}^2)} \le C_1 c_k, \quad \forall\, k\in\mathbb{Z}$$

Proof. We prove this by splitting the interval I_j into a finite number of smaller intervals depending on ψ_L. Thus we shall write

$$I_j = \cup_i J_{ji}$$

for a finite number of smaller intervals depending on ψ_L. The exact definition of these intervals will be given later in the proof. On each J_{ji}, we now run a bootstrap argument, commencing with the *bootstrap assumption:*

$$\|P_k\psi\|_{S[k](J_j\times\mathbb{R}^2)} \le A(C_0)c_k$$

Here $A(C_0)$ is a number that depends purely on the a priori bound we are making on the wave map. We shall show that provided $A(C_0)$ is chosen large enough, the bootstrap assumption implies the better bound

$$\|P_k\psi\|_{S[k](J_j\times\mathbb{R}^2)} \le \frac{A(C_0)}{2}c_k$$

We prove this for each frequency mode. By scaling invariance, we may assume $k = 0$. As before, one needs to distinguish between $|J_j| < \varepsilon_1$ and the opposite case, where ε_1 is chosen sufficiently small. In the former case, one directly uses the div-curl system

$$\partial_t\psi = \nabla_x\psi + \psi\nabla^{-1}(\psi^2)$$

as in the previous section to obtain the desired conclusion for $P_0\psi$. Thus we can assume that the interval satisfies $|J_j| \ge \varepsilon_1$, which means we can control $\big(P_0\psi(t_0,\cdot), P_0\partial_t\psi(t_0,\cdot)\big)$ for some $t_0 \in J_j$ via

$$\|\big(P_0\psi(t_0,\cdot), P_0\partial_t\psi(t_0,\cdot)\big)\|_{L^2_x\times\dot H^{-1}} \lesssim A_1(C_0)c_0$$

for some constant $A_1(C_0)$, which is explicitly computable, independently of $A(C_0)$. Passing to the wave equation

$$\Box P_0\psi = P_0 F_\alpha(\psi) = \sum_{i=1}^{5} P_0 F_\alpha^{2i+1}(\psi)$$

via Schwartz extensions and Hodge decompositions as before, we first consider the principal terms $P_0F_\alpha^3(\psi)$. These terms can be schematically written as

$$P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi^2)]$$

More accurately, they are of the form

$$\nabla_{t,x}P_\ell[P_{k_1,\kappa_1}\psi_1|\nabla|^{-1}P_kI^c\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]$$
$$\nabla_{t,x}P_\ell[P_{k_1,\kappa_1}R\,\psi_1|\nabla|^{-1}P_kI\mathcal{Q}(P_{k_2,\kappa_2}\psi_2, P_{k_3,\kappa_3}\psi_3)]$$

with a Riesz projection R and a nullform $\mathcal{Q}$. Substituting the decomposition $\psi = \psi_L + \psi_{NL}$ into the inner null-form yields

$$P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi^2)] = P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L^2)] + P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L\psi_{NL})] + P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_{NL}^2)]$$

Note that the last term automatically has the desired smallness property if we choose ε_0 smaller than some absolute constant. Indeed, by (7.4), and the trilinear estimates of Section 5,

$$\|P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_{NL}^2)]\|_{N[0]} \lesssim \varepsilon_0^2 \sup_{k\in\mathbb{Z}} 2^{-\sigma_0\,|k|}\|P_k\psi\|_{S[k]} \lesssim \varepsilon_0^2 A(C_0)c_0 \ll A(C_0)c_0$$

for small ε_0. Next, for the mixed term $P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L\psi_{NL})]$, choosing ε_0 sufficiently small (depending on C_0), we can arrange in light of Lemma 7.6 and the trilinear estimates

$$\|P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L\psi_{NL})]\|_{N[0]} \lesssim A(C_0)C_0^3\varepsilon_0^{1-\frac{1}{M}}\ c_0 \ll A(C_0)\,c_0$$

The first term

$$P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L^2)]$$

requires a separate argument. In fact, we treat this term by decomposing the interval I_i into smaller ones. In order to select these intervals, first note that upon localizing the frequencies of the inputs according to

$$P_0\nabla_{x,t}[P_{k_1}\psi\nabla^{-1}P_k(P_{k_2}\psi_LP_{k_3}\psi_L)]$$

one obtains from the trilinear bounds of Section 5

$$\|P_0\nabla_{x,t}[P_{k_1}\psi\nabla^{-1}P_k(P_{k_2}\psi_LP_{k_3}\psi_L)]\|_{N[0]} \le \varepsilon_0\,2^{-\sigma|k_1|}\|P_{k_1}\psi_1\|_{S[k_1]} \ll A(C_0)\,c_0$$

in the following two cases: k_1, k_2, k_3 fall outside the range (5.45) (the high-low-low case), or, if they do fall in the range (5.45), then $k \le k_2 - L'$. Here L and L' are large constants depending on C_0, ε_0, due to the bounds on ψ_L from Lemma 7.6. Thus, denoting by

$$P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L^2)]'$$

the sum over all frequency interactions described by these conditions, one then obtains the estimate

$$\|P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi_L^2)]'\|_{N[0]} \ll A(C_0)\,c_0$$

Employing the notations of Section 5.3, it thus suffices to consider the sum of expressions

$$\sum\nolimits''_{k_1,k_2,k_3\in\mathbb{Z}}\sum_{k=k_2-L'}^{k_2+O(1)} P_0\nabla_{x,t}[P_{k_1}\psi\nabla^{-1}P_k(P_{k_2}\psi_LP_{k_3}\psi_L)],$$

where, of course, the implied constants may be quite large depending on C_0, ε_0. Furthermore, by the results of that section, we may assume that the inputs have pairwise angular separation on the Fourier side, and in particular we make this assumption for the free wave inputs $P_{k_2}\psi_L$ and $P_{k_3}\psi_L$. Thus we have now reduced ourselves to estimating

$$\sum_{\substack{\kappa_1,\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \max_{i\neq j}\mathrm{dist}(\kappa_i,\kappa_j)>2^{m_0}}} {\sum}''_{k_1,k_2,k_3\in\mathbb{Z}} \sum_{k=k_2-L'}^{k_2+O(1)} P_0\nabla_{x,t}[P_{k_1,\kappa_1}\psi\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)],$$

The next step is to exploit the dispersive properties of the expression

$$\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)$$

First, due to the energy bound for ψ_L, there exists some finite set $A\subset\mathbb{Z}$ so that

$$\sum_{\substack{\kappa_1,\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \max_{i\neq j}\mathrm{dist}(\kappa_i,\kappa_j)>2^{m_0}}} {\sum}''_{\substack{k_1,k_2,k_3\in\mathbb{Z}\\ k_2\notin A}} \sum_{k=k_2-L'}^{k_2+O(1)} \|P_0\nabla_{x,t}[P_{k_1,\kappa_1}\psi\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)]\|_{N[0]}$$
$$\leq \varepsilon_0 2^{-\sigma_0\,|k_1|}\|P_{k_1}\psi\|_{S[k_1]}$$

On the other hand, assume now that $k_2\in A$ and consider

$$\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)$$

where k,k_3 are chosen as in (5.45). Note that the set A depends on the dyadic frequency of the output, in this case frequency 2^0. Changing the frequency localizations of the output amounts to a rescaling of A. Nonetheless, one has the following estimates which are independent under rescaling:

$$\|\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L P_{k_3,\kappa_3}\psi_L)\|_{L_t^1L_x^\infty} < C_3(\psi_L,k_2)$$

In particular,

$$\sum_{k_2\in A}\|\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L P_{k_3,\kappa_3}\psi_L)\|_{L_t^1L_x^\infty} < C_4(\psi_L) < \infty$$

To prove these bounds, set $k_2=0$ by scaling invariance. But then $P_{k_2}\psi_L(t_j,\cdot)$ is a Schwartz function in the x-variable. Using the angular separation of the inputs it is now straightforward to see that

$$\|\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)\|_{L_t^1L_x^\infty} < C_3(\psi_L,k_2)$$

Indeed, this follows from stationary phase and the angular separation of the inputs. We now define the intervals J_j by requiring that

$$\sum_{\substack{\kappa_1,\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \max_{i\neq j}\mathrm{dist}(\kappa_i,\kappa_j)>2^{m_0}}} \sum_{\substack{k_2\in A\\ |k_2-k_3|<L}} \sum_{k=k_2-L'}^{k_2+O(1)} \|\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)]\|_{L_t^1L_x^\infty(J_j\times\mathbb{R}^2)} < \varepsilon_0$$

It is furthermore clear that we also obtain

$$\sum_{\substack{\kappa_1,\kappa_2,\kappa_3\in\mathcal{C}_{m_0}\\ \max_{i\neq j}\mathrm{dist}(\kappa_i,\kappa_j)>2^{m_0}}} {\sum}''_{\substack{k_1,k_2,k_3\in\mathbb{Z}\\ k_2\in A}} \sum_{k=k_2-L'}^{k_2+O(1)} \|P_0\nabla_{x,t}[P_{k_1,\kappa_1}\psi\nabla^{-1}P_k(P_{k_2,\kappa_2}\psi_L\ P_{k_3,\kappa_3}\psi_L)]\|_{N[0](J_j\times\mathbb{R}^2)}$$
$$\leq \varepsilon_0 2^{-\sigma_0\,|k_1|}\|P_{k_1}\psi\|_{S[k_1]} \ll A(C_0)c_0$$

which completes the bootstrap for the trilinear source terms.

The contribution of the higher order terms is dealt with in the appendix.

By applying the above bootstrap argument on each of the finitely many intervals J_{ji} comprising each I_j, the proof of Lemma 7.9 now follows. □

The proof of Proposition 7.3 can easily be concluded. Indeed, one infers from the Klainerman-Machedon criterion that $T = T' = \infty$. Moreover, we obtain a global a priori bound

$$\|P_k\psi\|_{S[k]} \le C_1 c_k$$

were the constant C_1 depends implicitly on ψ_L. As the above argument is fairly crude, we have no a priori way here of controlling this number. In section 9 we will refine this type of estimate Finally, Lemma 7.6 implies the scattering for large times.

7.2. **Control of wave-maps via a fixed L^2-profile.** A fundamental issue that we need to address is the *very definition* of wave maps with data that are in some sense only of energy class. To propagate such data under the wave map evolution, we shall use approximations by smooth wave maps each of which can be continued canonically. The following lemma justifies this procedure.

Lemma 7.10. *Let ϕ_α^n be the derivative components of a sequence of Schwartz class*[16] *wave maps $\mathbf{u}^n$: $(-T_0^n, T_1^n) \times \mathbb{R}^2 \to \mathbb{H}^2$ on their maximal time interval of existence and assume that the Coulomb components $\psi_\alpha^n(0,\cdot)$ satisfy*

$$\lim_{n\to\infty} \|\psi_\alpha^n(0,\cdot) - V_\alpha\|_{L_x^2} = 0$$

for some $V_\alpha \in L^2(\mathbb{R}^2)$. Denoting the collection of components V_α by V, there is a time $T_0 = T_0(V) > 0$ such that $\min(T_0^n, T_1^n) > T_0$ for all sufficiently large n and

$$\limsup_{n\to\infty} \|\psi_\alpha^n\|_{S((-T_0,T_0)\times\mathbb{R}^2)} \le C(V) < \infty$$

Furthermore, there is a constant $C_1(V)$ with the following property: defining the frequency envelope

$$c_k^{(n)} := \max_{\alpha=0,1,2} (\sum_{\ell\in\mathbb{Z}} 2^{-\sigma|k-\ell|} \|P_\ell\psi_\alpha^n\|_{L_x^2}^2)^{\frac{1}{2}}$$

for sufficiently small fixed $\sigma > 0$, one has for all $k \in \mathbb{Z}$ and all large n

$$\max_{\alpha=0,1,2} \|P_k\psi_\alpha^n\|_{S[k]((-T_0,T_0)\times\mathbb{R}^2)} \le C_1(V) c_k^{(n)}$$

Finally, the wave map propagations of the ψ_α^n converge on fixed time slices $t = t_0 \in [-T_0, T_0]$ in the L^2-topology, uniformly in time.

The proof of this lemma will occupy the remainder of this section. Before we begin with the proof, we discuss some related results and implications of Lemma 7.10. Most fundamental is the following stability result:

Proposition 7.11. *Let $\mathbf{u} : [-T_0, T_1] \times \mathbb{R}^2 \to \mathbb{H}^2$ be an admissible wave-map with gauged derivative components denoted by ψ. Assume that $\|\psi\|_{S([-T_0,T_1]\times\mathbb{R}^2)} = A < \infty$. Then there exists $\varepsilon_1 = \varepsilon_1(A) > 0$ with the following property: any other admissible wave-map $\mathbf{v}$ defined locally around $t = 0$ and with gauged derivative components $\tilde\psi$ satisfying $\|\psi(0)-\tilde\psi(0)\|_2 < \varepsilon < \varepsilon_1$ extends as an admissible wave-map to $[-T_0, T_1]$ and satisfies $\|\tilde\psi\|_{S([-T_0,T_1]\times\mathbb{R}^2)} < A + c(\varepsilon_1)$ where $c(\varepsilon) \to 0$ as $\varepsilon \to 0$. We also have the local Lipschitz type bound*

$$\|\tilde\psi - \psi\|_{S([-T_0,T_0]\times\mathbb{R}^2)} \lesssim \|\psi(0) - \tilde\psi(0)\|_{L^2}$$

for $\tilde\psi$ satisfying the above proximity condition, with implied constant depending on $\|\psi\|_{S([-T_0,T_0]\times\mathbb{R}^2)}$.

Proof. The proof will be given in Section 9.5, as it follows directly from the proof of Proposition 9.12. □

As a consequence, one has the following important continuation result.

Corollary 7.12. *Let $\{\psi^n\}_{n=1}^\infty$ be a sequence of Coulomb components of admissible wave maps $\mathbf{u}^n : I \to \mathbb{H}^2$ where I some fixed nonempty closed interval such that for some $t_0 \in I$ one has*

$$\lim_{n\to\infty} \|\psi_\alpha^n(t_0,\cdot) - V_\alpha\|_{L_x^2} = 0$$

[16]In the usual sense that $\phi_\alpha^n|_{t=const}$ is Schwartz on $\mathbb{R}^2$.

with $V_\alpha \in L^2(\mathbb{R}^2)$ as well as

$$\sup_n \|\psi^n\|_{S(I\times\mathbb{R}^2)} < \infty.$$

Then there exists a true extension $\tilde{I}$ of I (meaning that it extends by some positive distance beyond the endpoints of I in sofar as they are finite) to which each $\mathbf{u}^n$ can be continued as an admissible wave map provided n is large.

Proof. By Proposition 7.11 we can define $\lim_{n\to\infty} \psi^n_\alpha(t,.)$ in the L^2-sense for t an endpoint of I. By Lemma 7.10 the ψ^n_α extend beyond the (finite) endpoints for n large enough. □

We can use the preceding results to define wave maps with L^2 data at the level of the Coulomb gauge.

Definition 7.13. *Assume we are given a family $\{V_\alpha\}$, $\alpha = 0,1,2$, of $L^2(\mathbb{R}^2)$-functions, to be interpreted as data at time $t=0$. Also, assume we have*

$$V_\alpha = \lim_{n\to\infty} \psi^n_\alpha$$

where $\{\psi^n_\alpha\}$ are Coulomb components of admissible wave maps at time $t=0$. Determine $I = (-T_0, T_1) = \cup I_1$ to be the union of all open time intervals I_1 with the property that

$$\sup_{\tilde{I}\subset I_1,\, \tilde{I}\ closed} \liminf_{n\to\infty} \|\psi^n_\alpha\|_{S(\tilde{I}\times\mathbb{R}^2)} < \infty$$

Then we define the Coulomb wave maps propagation of $\{V_\alpha\}$ to be

$$\Psi^\infty_\alpha(t,x) := \lim_{n\to\infty} \psi^n_\alpha(t,x), \quad t\in I$$

We call $I\times\mathbb{R}^2$ the lifespan of the (Coulomb) wave maps evolution of $\{V_\alpha\}$.

It is of course important that the life span does not depend on the choice of sequence and, moreover, that the "solutions" V_α are unique. These statements follow from Proposition 7.11.

The aforementioned uniqueness properties are now immediate – indeed, simply mix any two sequences which converge to V_α. Moreover, we can characterize the life-span as follows.

Corollary 7.14. *Let V_α, $\{\psi^n_\alpha\}$, and I be as in Definition 7.13. Assume in addition that $I \neq (-\infty,\infty)$. Then*

$$\sup_{\substack{J\subset I\\ J\ closed}} \liminf_{n\to\infty} \|\psi^n_\alpha\|_{S(J\times\mathbb{R}^2)} = \infty \tag{7.21}$$

Proof. Suppose not. Let $I = (-T_0, T_1)$ where w. l. o. g. we assume $T_1 < \infty$. Then there exists a number $M < \infty$ with the property that for every closed $J \subset I$ with $0 \in J$ one has

$$\liminf_{n\to\infty} \|\psi^n_\alpha\|_{S(J\times\mathbb{R}^2)} < M$$

Now observe that

$$\limsup_{n,\, J\subset I} \|\psi^n_\alpha\|_{S(J\times\mathbb{R}^2)} = \infty,$$

where J ranges over the closed subsets of I. Indeed, if not, we have

$$\limsup_{n\to\infty} \|\psi^n_\alpha\|_{S([0,T_1]\times\mathbb{R}^2)} < \infty$$

But then by Corollary 7.12 one can extend ψ^n_α beyond the endpoint T_1 of I to some interval $\tilde{I}$ for n large enough while maintaining the finiteness of $\|\psi^n_\alpha\|_{S(\tilde{I}\times\mathbb{R}^2)}$, contradicting the definition of I.
Now pick ϵ_1 as in Proposition 7.11, with M replacing A, and pick $J\subset I$, n_0 large enough such that

$$\|\psi^{n_0}_\alpha\|_{S(J\times\mathbb{R}^2)} \gg M, \quad \sup_{n,m\geq n_0} \|(\psi^n_\alpha - \psi^m_\alpha)(0,\cdot)\|_{L^2} < \epsilon_1$$

But by our definition of M there exists $k_0 > n_0$ with the property that

$$\|\psi^{k_0}_\alpha\|_{S(J\times\mathbb{R}^2)} < M$$

and then applying Proposition 7.11 to $\psi^{k_0}_\alpha(0,\cdot)$ we obtain a contradiction. This proves the corollary. □

Another important property is to be able to ensure the a priori existence of wave maps flows "at infinity", i.e., the solution of the scattering problem. In this regard, we have the following result.

Proposition 7.15. *Assume we are given admissible data at time $t = 0$ of the form*

$$\psi_\alpha = \partial_\alpha\big(S(0-t_0)(\partial_t V, V)\big) + o_{L^2}(1), \quad \alpha = 0,\, 1,\, 2$$

Here $(\partial_t V, V) \in L^2 \times \dot{H}^1$ is a fixed profile. Then for $t_0 = t_0(\partial_t V, V) > 0$ large enough and $o_{L^2}(1)$ small enough, the wave map associated with ψ_α exists on $(-\infty, 0]$, is admissible there, and we have

$$\|\psi_\alpha\|_{S((-\infty,0]\times\mathbb{R}^2)} < \infty$$

Moreover, letting ψ^n_α be a sequence of admissible Coulomb components (i.e., associated with admissible maps) at time $t = 0$ satisfying

$$\psi^n_\alpha \to \partial_\alpha\big(S(0-t_0)(\partial_t V, V)\big)$$

for $(\partial_t V, V)$ as before and t_0 large enough also as before, the limit

$$\lim_{n\to\infty} \psi^n_\alpha(t,x) = \Psi^\infty_\alpha(t,x), \quad t \in (-\infty, 0]$$

exists independently of the particular sequence chosen. We call this the Coulomb wave maps evolution of the data

$$\partial_\alpha\big(S(0-t_0)(\partial_t V, V)\big)$$

at time $t = -\infty$. A similar construction applies at time $t = \infty$.

Corollary 7.16. *Assume that for a sequence of admissible Coulomb components ψ^n_α at time $t = 0$ we have*

$$\psi^n_\alpha = \partial_\alpha\big(S(t-t^n)(\partial_t V,\, V)\big) + o_{L^2}(1)$$

Then if $t_n \to \infty$, the limits

$$\lim_{n\to\infty} \psi^n_\alpha(t+t^n,\, x) = \Psi^\infty_\alpha(t,\, x)$$

exist in the L^2-sense on some interval $(-\infty, -C)$, uniformly on closed subintervals, for C large enough. We have

$$\limsup_{n\to\infty} \|\psi^n_\alpha(t+t^n,\, x)\|_{S((-\infty,-C_0]\times\mathbb{R}^2)} < \infty$$

for $C_0 > C$. We call the maximal interval $I = (-\infty, -C)$ for which these statements hold the lifespan of the limiting object Ψ^∞_α; here C may be negative or $-\infty$. A similar construction applies when $t_n \to -\infty$.

Both Proposition 7.15 as well as Corollary 7.16 will be proved in Section 9.8. Having defined limiting objects Ψ^n_α as in Lemma 7.13 (temporally bounded case) as well as Corollary 7.16, we can now define in obvious fashion the norms

$$\|\Psi^\infty_\alpha\|_{S(J\times\mathbb{R}^2)} = \lim_{n\to\infty} \|\psi^n_\alpha\|_{S(J\times\mathbb{R}^2)}$$

for $J \subset I$ closed, with I the lifespan of the limiting object. This is well-defined due to Proposition 7.11. We can then also state the following

Lemma 7.17. *Let Ψ^∞_α be as before, with lifespan I. Assume in addition that $I \neq (-\infty, \infty)$. Then*

$$\sup_{\substack{J\subset I\\ J \text{ closed}}} \|\Psi^\infty_\alpha\|_{S(J\times\mathbb{R}^2)} = \infty \tag{7.22}$$

The same conclusion holds for arbitrary I provided the sequence ψ^n_α is essentially singular[17].

We now turn to the proof of Lemma 7.10. We begin with the the lower bound on the life span of the ψ^n_α. In essence, this is a consequence of the fact that $\psi^n_\alpha \to V_\alpha$ in L^2 implies a uniform non-concentration property of the energy of the ψ^n_α. This then allows one to approximate the corresponding wave maps with derivative components ϕ^n_α on small discs — with radii depending only on the limiting "profile" V — by small energy smooth wave maps; the small energy theory and finite propagation speed then imply a uniform lower bound on the life span. Technically speaking, restricting to small scales requires some care

[17] Recall the definition in section 1.4

since localizing the wave map by applying a smooth cutoff does not necessarily decrease the energy. To see this, let χ be a cutoff to a small ball B of size r. Then the first term on the right-hand side of

$$\int_{\mathbb{R}^2} |\nabla(\chi\phi)(x)|^2\,dx \lesssim r^{-2}\int_B |\phi(x)|^2\,dx + \int_B |\nabla\phi(x)|^2\,dx \tag{7.23}$$

does in general not become small as $r \to 0$.
Let $\varepsilon_0 > 0$ be the cutoff such that smooth data with energy less than ε_0 result in global wave maps. More precisely, we will rely on the following result by the first author, see [22].

Theorem 7.18. *Given smooth initial data $(\mathbf{x}, \mathbf{y})[0] : 0 \times \mathbb{R}^2 \to \mathbb{H}^2$ which are sufficiently small in the sense that*

$$\int_{0\times\mathbb{R}^2} \sum_{\alpha=0}^{2} \big[\big(\frac{\partial_\alpha \mathbf{x}}{\mathbf{y}}\big)^2 + \big(\frac{\partial_\alpha \mathbf{y}}{\mathbf{y}}\big)^2\big]\,dx_1dx_2 < \varepsilon_0^2$$

where $\varepsilon_0 > 0$ is a small absolute constant, there exists a unique classical wave map from $\mathbb{R}^{2+1}$ to $\mathbb{H}^2$ extending these data globally in time. Moreover, one has the bound $\sum_{\alpha=0}^{2} \|\psi_\alpha\|_{S(\mathbb{R}^{1+2})} \le C\varepsilon_0$ where C is an absolute constant.

Denoting the actual map at time $t = 0$ giving rise (together with the time derivatives) to $\phi_\alpha^n, \psi_\alpha^n$, by $(\mathbf{x}, \mathbf{y})(0, \cdot) : \mathbb{R}^2 \to \mathbb{H}^2$, where we have omitted the superscript n for simplicity, we now consider a "re-normalized" map, subject to a choice of $x_0 \in \mathbb{R}^2$ and $r_0 > 0$,

$$(\mathbf{x}_1, \mathbf{y}_1) := \Big(\chi_{[|x-x_0|<r_0]}\frac{\mathbf{x}-\mathbf{x}_0}{\mathbf{y}_0},\ e^{\chi_{[|x-x_0|<2r_0]}\log[\frac{\mathbf{y}}{\mathbf{y}_0}(0,\cdot)]}\Big) \tag{7.24}$$

Here $\chi_{[|x-x_0|<r_0]}$ is a smooth cutoff to the disk $D_{x_0,r_0} := \{|x-x_0| < r_0\}$ which equals one on $|x-x_0| < \frac{r_0}{2}$, say, and

$$\mathbf{x}_0 := \fint_{[|x-x_0|<r_0]} \mathbf{x}(x)\,dx_1dx_2, \qquad \mathbf{y}_0 := \exp\big(\fint_{[|x-x_0|<2r_0]} \log\mathbf{y}(x)\,dx_1dx_2\big)$$

with $\fint_B := |B|^{-1}\int_B$. Note that we have chosen the cutoffs on the two components differently − the one on the second component is slightly larger than the first. This is merely a technical convenience which amounts to $\mathbf{y}_1 = \frac{\mathbf{y}}{\mathbf{y}_0}$ when $\nabla\chi_{[|x-x_0|<r_0]} \neq 0$. Lemma 7.21 below verifies the desired smallness of energy property for these data. We begin with a basic imbedding lemma which we shall need in the proof of that lemma. Even though we only require the case $d = 2$, we formulate this lemma in any dimension.

Lemma 7.19. $\dot{B}^{\frac{d}{2}}_{2,\infty}(\mathbb{R}^d) \hookrightarrow \mathrm{BMO}(\mathbb{R}^d)$.

Proof. By duality, it suffices to prove that $\mathcal{H}^1 \hookrightarrow \dot{B}^{-\frac{d}{2}}_{2,1}(\mathbb{R}^d)$ for the Hardy space $\mathcal{H}^1$. Thus, we need to show that

$$\sum_{j\in\mathbb{Z}} 2^{-j\frac{d}{2}}\|P_j\phi\|_2 \le C(d)$$

for any ϕ which is an $\mathcal{H}^1(\mathbb{R}^d)$ atom. Here P_j are the usual Littlewood-Paley projections to frequencies of size 2^j. By scaling and translation invariance we may assume that $\mathrm{supp}(\phi) \subset B(0,1)$, $|\phi| \le 1$ and $\int \phi(x)\,dx = 0$. If $j \ge 0$, then we use that

$$\|P_j\phi\|_2 \le \|\phi\|_2 \le C(d)$$

If $j \le 0$, then writing $P_j\phi = 2^{jd}\psi(2^j\cdot) * \phi$ we conclude that

$$|P_j\phi(x)| \le C\int_{B(0,1)} 2^{j(d+1)}\int_0^1 |\nabla\psi|(2^j(x-ty))\,dt|\phi(y)|\,dy$$

which implies that

$$\|P_j\phi\|_2 \le C\int_{B(0,1)} 2^{j(d+1)}\|\nabla\psi\|_2 2^{-j\frac{d}{2}}|\phi(y)|\,dy \le C(d)2^{j(\frac{d}{2}+1)}$$

and we are done. □

The importance of BMO in this context lies with the fact that exponentiation maps small balls in BMO into the A_p-class. Recall that w is an A_p-weight in the sense of Muckenhaupt (see Chapter 7 in [7] or Stein [44] for all this standard material) provided

$$|Q|^{-1}\int_Q w(x)\,dx\,\Big(|Q|^{-1}\int_Q w^{1-p'}(x)\,dx\Big)^{p-1}\le A_p(w) \tag{7.25}$$

uniformly for all cubes $Q\subset\mathbb{R}^d$ for some constant $A_p(w)$. Here $1<p<\infty$ and $p'=\frac{p}{p-1}$ as usual. Note that $A_p\subset A_q$ if $p\le q$. The A_1 class is defined as all w with $Mw\le Cw$ a.e., where M is the Hardy-Littlewood maximal operator. At the other end one has $A_\infty:=\bigcup_{1\le p<\infty}A_p$, which is characterized by the estimate

$$\frac{w(S)}{w(Q)}\le C\Big(\frac{|S|}{|Q|}\Big)^\delta \tag{7.26}$$

for all $S\subset Q$ (this is deep and requires the "reverse Hölder inequality"). Here C and $\delta>0$ only depend on w. ¿From the John-Nirenberg inequality, $w=e^\phi$ is an A_p weight for some $1<p<\infty$ provided $\|\phi\|_{\mathrm{BMO}}<r_0$ is small enough and the A_p-constant $A_p(w)$ in (7.25) only depends on r_0.

Lemma 7.20. *Let $\|\varphi\|_{L^2}\le A$ and set $w:=e^{(-\Delta)^{-\frac12}\varphi}$. Then for any $1<p<\infty$ one has $A_p(w)\le C(p,A)$ where the latter constant only depends on p and A.*

Proof. For any $\delta>0$,

$$\#\{j\in\mathbb{Z}\mid \|P_j\varphi\|_2\ge\delta\}\le\delta^{-2}A^2$$

In particular, for any $\delta>0$ there is a decomposition $\varphi=\widetilde{\varphi}+(\varphi-\widetilde{\varphi})$ so that $\|(-\Delta)^{-\frac12}\widetilde{\varphi}\|_\infty\le C\delta^{-2}A^3$ and, by Lemma 7.19,

$$\|(-\Delta)^{-\frac12}(\varphi-\widetilde{\varphi})\|_{\mathrm{BMO}}\le C\delta$$

By the John-Nirenberg inequality one may choose δ small depending on $p\in(1,\infty)$ such that $\exp\big((-\Delta)^{-\frac12}(\varphi-\widetilde{\varphi})\big)\in A_p$ with some absolute A_p-constant. Since

$$\big\|e^{(-\Delta)^{-\frac12}\widetilde{\varphi}}\big\|_\infty\le e^{C\delta^{-2}A^3}$$

we are done. □

The importance of A_p weights lies with the fact that the Hardy-Littlewood maximal operator M as well as Calderon-Zygmund operators T are bounded on $L^p(w\,dx)$ with constants that only depend on the dimension and the A_p constant from (7.25) (and T in case of a singular integral) provided $1<p<\infty$. In the present context, we will require a version of Poincaré's inequality with A_2 weights. Now for the small energy lemma.

Lemma 7.21. *Let $(\mathbf{x}_1^n,\mathbf{y}_1^n)$ be as in (7.24) applied to $(\mathbf{x}^n,\mathbf{y}^n)$. Then given $\varepsilon_0>0$, we can pick $r_0>0$ small enough such that*

$$\Big\|\frac{\nabla\mathbf{x}_1^n}{\mathbf{y}_1^n}\Big\|_{L^2_x}+\Big\|\frac{\nabla\mathbf{y}_1^n}{\mathbf{y}_1^n}\Big\|_{L^2_x}\ll\varepsilon_0$$

Here ∇ is the spatial gradient and r_0 does not depend on n. Since one can clearly also arrange

$$\|\chi_{[|x-x_0|<r_0]}\phi_0^n\|_{L^2_x}\ll\varepsilon_0,$$

we have now achieved smallness of the energy of these data. Moreover, $r_0>0$ can be chosen uniformly *in $x_0\in\mathbb{R}^2$.*

Proof. We assume as we may (by rescaling) that $\|\phi_\alpha^1\|_2+\|\phi_\alpha^2\|_2\le1$ for $\alpha=0,1,2$. We shall also suppress the time dependence and drop the superscript n. In view of (7.23) it suffices to estimate the contributions of those terms in which the derivatives falls on the cutoff χ in (7.24). Starting with the component $\mathbf{y}_1 = \chi_{[|x-x_0|<r_0]}\frac{\mathbf{y}(x)}{\mathbf{y}(x_0)}$, note that Poincaré's inequality implies that, with $B := \{|x-x_0| < r_0\}$,

$$r_0^{-2}\int_B \Big|\log\big[\frac{\mathbf{y}(x)}{\mathbf{y}_0}\big]\Big|^2\,dx_1dx_2 \lesssim \int_B \Big|\frac{\nabla\mathbf{y}(x)}{\mathbf{y}(x)}\Big|^2\,dx_1dx_2$$
$$\lesssim \sum_{\alpha=1,2}\int_B |\phi_\alpha(x)|^2\,dx_1dx_2 = \sum_{\alpha=1,2}\int_B |\psi_\alpha(x)|^2\,dx_1dx_2 \ll \varepsilon_0^2$$

uniformly in n provided r_0 is small enough. Here we used the relation (1.4) and that the gauge change is given by multiplication by a unimodular factor. For the $\mathbf{x}_1$-component, we make the preliminary observation that $\mathbf{y} \in A_p$ for any $1 < p < \infty$. Indeed, by Lemma 7.19, for any $1 \le M < \infty$ we can find $C = C(M)$ so large that

$$\|P_{\mathbb{R}\setminus[-C,C]}\log\mathbf{y}\|_{\mathrm{BMO}} = \|P_{\mathbb{R}\setminus[-C,C]}\sum_{j=1,2}\Delta^{-1}\partial_j\phi_j^2\|_{\mathrm{BMO}} \ll M^{-1}$$

which implies that $\mathbf{y}_2 := \exp\big(P_{\mathbb{R}\setminus[-C(p),C(p)]}\log\mathbf{y}\big) \in A_p$ for any $1 < p < \infty$ with a suitable $C(p)$. Since Lemma 7.19 implies that $\|\mathbf{y}\mathbf{y}_2^{-1}\|_\infty \le C$, the claim follows. We now use the following weighted Poincaré inequality, see Theorem 1.5 in [8]: for any $w \in A_2$, and ball B of radius $r > 0$,

$$\int_B |f(x)-(f)_B|^2\,w(x)dx \le C(w)r^2\int_B |\nabla f(x)|^2\,w(x)dx, \qquad (f)_B := \fint_B f(x)\,dx$$

Consequently, with $w = \mathbf{y}^{-2} \in A_2$, and in view of our definition of $\mathbf{x}_0$,

$$r_0^{-2}\int_B \Big|\frac{\mathbf{x}(x)-\mathbf{x}_0}{\mathbf{y}(x)}\Big|^2\,dx_1dx_2 \lesssim \int_B \Big|\frac{\nabla\mathbf{x}(x)}{\mathbf{y}(x)}\Big|^2\,dx_1dx_2 \lesssim \sum_{j=1,2}\int_B |\phi_j^1(x)|^2\,dx_1dx_2$$

By our choice of cutoffs in (7.24) we are done. To obtain the final statement of the proof, simply note that we can always find $r_0 > 0$ such that

$$\sup_{x_0\in\mathbb{R}^2}\sum_{\alpha=0}^2\int_{D_{x_0,r_0}} |V_\alpha(x)|^2\,dx_1dx_2 \ll \varepsilon_0^2$$

Consequently, for all sufficiently large n,

$$\sup_{x_0\in\mathbb{R}^2}\sum_{\alpha=0}^2\int_{D_{x_0,r_0}} |\psi_\alpha^n(x)|^2\,dx_1dx_2 \ll \varepsilon_0^2$$

and therefore also

$$\sup_{x_0\in\mathbb{R}^2}\sum_{\alpha=0}^2\int_{D_{x_0,r_0}} |\phi_\alpha^n(x)|^2\,dx_1dx_2 \ll \varepsilon_0^2$$

for all large n, which is all that is needed for the proof. □

We will also require an analogous result on small energy outside of a big ball. Thus, let $R_0 \gg 1$ be large and define

$$(\mathbf{x}_2,\mathbf{y}_2) := \Big(\chi_{[|x|>R_0]}\frac{\mathbf{x}-\mathbf{x}_0}{\mathbf{y}_0},\ e^{\chi_{[|x|>\frac{R_0}{2}]}\log[\frac{\mathbf{y}}{\mathbf{y}_0}(0,\cdot)]}\Big) \tag{7.27}$$

Here $\chi_{[|x|>R_0]}$ is a smooth cutoff to the set $\{|x| > R_0\}$ which equals one on $|x| > 2R_0$, say, and

$$\mathbf{x}_0 := \fint_{[R_0<|x|<2R_0]}\mathbf{x}(x)\,dx_1dx_2, \qquad \mathbf{y}_0 := \exp\Big(\fint_{[\frac{R_0}{2}<|x|<R_0]}\log\mathbf{y}(x)\,dx_1dx_2\Big)$$

In analogy to (7.24) the construction here is such that $\mathbf{y}_2 = \frac{\mathbf{y}}{\mathbf{y}_0}$ on the set $\{\nabla\chi_{[|x|>R_0]} \ne 0\}$.

Lemma 7.22. *Let $(\mathbf{x}_2^n, \mathbf{y}_2^n)$ be as in (7.27) applied to $(\mathbf{x}^n, \mathbf{y}^n)$. Then given $\varepsilon_0 > 0$, we can pick $R_0 > 0$ large enough such that*

$$\|\frac{\nabla \mathbf{x}_2^n}{\mathbf{y}_2^n}\|_{L_x^2} + \|\frac{\nabla \mathbf{y}_2^n}{\mathbf{y}_2^n}\|_{L_x^2} \ll \varepsilon_0$$

Here ∇ is the spatial gradient and R_0 does not depend on n. Since one can clearly also arrange

$$\|\chi_{[|x|>R_0]}\phi_0^n\|_{L_x^2} \ll \varepsilon_0,$$

we have now achieved smallness of the energy of these data.

Proof. The argument is completely analogous to the one for Lemma 7.21. The only difference is that one uses the following Poincaré inequalities on annuli instead of disks: for any $R_0 > 0$,

$$\int_{R_0<|x|<2R_0} |f(x) - (f)_{R_0}|^2 \, w(x)dx \le CR_0^2 \int_{R_0<|x|<2R_0} |\nabla f(x)|^2 \, w(x)dx$$

for any A_2-weight w and a constant C which only depends on the A_2 constant of w. As usual $(f)_{R_0}$ denotes the average of f over the annulus. For $w = 1$ this is of course standard, and for general w it follows from [8]. □

Next, we wish to establish control over the ψ_α^n in the S-norm on a nonempty time interval $(-T_0, T_0)$ uniformly in n. The idea is to apply Theorem 7.18 to the finitely many small energy maps given by Lemma 7.21 and then to reconstruct and also bound the original sequence ψ_α^n in terms of these constituents. The latter of course relies on finite propagation speed and involves smooth partitions of unity. In order to handle partitions of unity, we need to derive estimates of the form

$$\|\chi\psi\|_S \le C(\chi)\|\psi\|_S$$

for Schwartz functions χ and some constant $C(\chi)$. Due to issues having to do with the slow decay as well as limited regularity of the logarithmic potential $\Delta^{-1}\partial\phi$ which appears in the phase of the gauge change, we will need to allow for a larger class of functions χ. The following lemma is tailored to such purposes.

Lemma 7.23. *Let $\chi \in C^\infty(\mathbb{R}^{2+1})$ satisfy the following properties[18]: for some constant A*

- $\max_{k=0,1} \max_{|\alpha|\le 100} \|\partial_t^k \nabla_x^\alpha \chi\|_{L_t^q L_x^p} \le A$ *for all* $2 < p \le \infty$, $1 \le q \le \infty$
- $\|\langle\tau\rangle \max(|\xi|, |\xi|^{100})\widehat{\chi}(\tau, \xi)\|_{L_\tau^q L_\xi^\infty} \le A$ *for all* $2 \le q \le \infty$

Then there exists an absolute constant C_0 such that $\|\chi\psi\|_S \le C_0 A\|\psi\|_S$ for any Schwartz function ψ. The S-norm here is defined in terms of the $\|\!|\cdot|\!\|$-norm, and it can be either localized to some interval in time or be defined globally in time.

Proof. It suffices to consider global in time estimates. We need to prove

$$\|\chi\psi\|_S := \Big(\sum_{k\in\mathbb{Z}} \|\!|P_k(\chi\psi)|\!\|_{S[k]}^2\Big)^{\frac{1}{2}} \lesssim A\Big(\sum_{\ell\in\mathbb{Z}} \|\!|P_\ell\psi|\!\|_{S[\ell]}^2\Big)^{\frac{1}{2}} \tag{7.28}$$

Written out, the left-hand side here means

$$\Big(\sum_{k\in\mathbb{Z}} \|P_k(\chi\psi)\|_{L_t^\infty L_x^2}^2 + \|P_k\Box(\chi\psi)\|_{N[k]}^2\Big)^{\frac{1}{2}} \tag{7.29}$$

and we shall write $\|\cdot\|$ instead of $\|\!|\cdot|\!\|$. We begin with the energy component of the norm. If $k \le C$, then by Bernstein's inequality

$$\|P_k(\chi\psi)\|_{L_t^\infty L_x^2} \lesssim 2^{\frac{2k}{3}} \|\chi\|_{L_t^\infty L_x^3}\|\psi\|_{L_t^\infty L_x^2} \lesssim 2^{\frac{2k}{3}} \|\chi\|_{L_t^\infty L_x^3}\|\psi\|_S \lesssim 2^{\frac{2k}{3}} A\|\psi\|_S \tag{7.30}$$

[18]The logarithmic potential in (1.11) decays like $|z|^{-1}$ (but in general no faster) which explains why we need $p > 2$ in the first condition. Since one in fact has asymptotic equality with $|z|^{-1}$ up to a multiplicative constant, it follows that the Fourier transform of this potential around zero exhibits a $|\xi|^{-1}$-singularity, which explains the second condition. Finally, we cannot control more than one time derivative of (1.11), and showing that one time derivative can be controlled in terms of the energy alone is nontrivial and requires the div-curl system for ϕ, see Corollary 7.25.

Here we used that

$$\|\psi\|_{L_t^\infty L_x^2} \le \Big(\sum_{k\in\mathbb{Z}} \|P_k\psi\|^2_{L_t^\infty L_x^2}\Big)^{\frac12} \le \Big(\sum_{k\in\mathbb{Z}} \|P_k\psi\|^2_{S[k]}\Big)^{\frac12} \le \|\psi\|_S$$

On the other hand, if $k \ge C$, then

$$\begin{aligned}\|P_k(\chi\psi)\|_{L_t^\infty L_x^2} &\le \|P_k(P_{\le k-10}\chi\,\tilde P_k\psi)\|_{L_t^\infty L_x^2} + \|P_k(P_{>k-10}\chi\,\psi)\|_{L_t^\infty L_x^2}\\ &\lesssim \|\chi\|_{L_t^\infty L_x^\infty}\|\tilde P_k\psi\|_{L_t^\infty L_x^2} + \|P_{>k-10}\chi\|_{L^\infty_{t,x}}\|\psi\|_{L_t^\infty L_x^2}\\ &\lesssim (\|\chi\|_{L_t^\infty L_x^\infty} + \|\nabla\chi\|_{L_t^\infty L_x^\infty})(\|\tilde P_k\psi\|_{S[k+O(1)]} + 2^{-k}\|\psi\|_S)\\ &\lesssim A(\|\tilde P_k\psi\|_{S[k+O(1)]} + 2^{-k}\|\psi\|_S)\end{aligned}$$

where we used the reverse Bernstein inequality

$$\|P_{>k-10}\chi\|_{L_t^\infty L_x^\infty} \lesssim 2^{-k}\|\nabla\chi\|_{L_t^\infty L_x^\infty}$$

Square-summing now implies the desired bound. It remains to bound the second term in (7.29). First,

$$\begin{aligned}\sum_{k\le C}\|P_kQ_{\le k+C}\Box(\chi\psi)\|^2_{N[k]} &\lesssim \sum_{k\le C}\|P_kQ_{\le k+C}\Box(\chi\psi)\|^2_{\dot X_k^{-1,-\frac12,1}}\\ &\lesssim \sum_{k\le C} 2^k\|P_k(\chi\psi)\|^2_{L_t^2L_x^2} \lesssim \|\chi\|^2_{L_t^2L_x^\infty}\|\psi\|_S^2 \lesssim A^2\|\psi\|_S^2\end{aligned}$$

Second,

$$\begin{aligned}\sum_{k\le C}\|P_kQ_{>k+C}\Box(\chi\psi)\|^2_{N[k]} &\lesssim \sum_{k\le C}\|P_kQ_{>k+C}\Box(\chi\psi)\|^2_{\dot X_k^{-\frac12+\varepsilon,-1-\varepsilon,2}}\\ &\lesssim \sum_{k\le C}\sum_{j>k+C} 2^{-(1-2\varepsilon)k}2^{2j(1-\varepsilon)}\|P_kQ_j(\chi\psi)\|^2_{L_t^2L_x^2}\\ &\lesssim \sum_{k\le C}\sum_{j>k+C} 2^{2\varepsilon k}2^{2j(1-\varepsilon)}\|P_kQ_j(\chi\psi)\|^2_{L_t^2L_x^{\frac43}}\end{aligned} \tag{7.31}$$

The sum $\sum_{k+C<j<C}$ does not pose a problem since in these cases P_kQ_j is disposable whence

$$\|P_kQ_j(\chi\psi)\|_{L_t^2L_x^{\frac43}} \lesssim \|\chi\|_{L_t^2L_x^4}\|\psi\|_{L_t^\infty L_x^2} \lesssim A\|\psi\|_S$$

which can be summed in this range. So we restrict our attention to $j \ge C$. We can assume furthermore that $\chi = P_{\le j-C}\chi$ as otherwise

$$\|P_kQ_j(P_{>j-C}\chi\,\psi)\|_{L_t^2L_x^{\frac43}} \lesssim 2^{-j}\|\nabla P_{>j-C}\chi\|_{L_t^2L_x^4}\|\psi\|_{L_t^\infty L_x^2} \lesssim 2^{-j}A\|\psi\|_S$$

makes a summable contribution to (7.31). We now split $\chi = Q_{\ge j-C}\chi + Q_{<j-C}\chi$. On the one hand,

$$\begin{aligned}\|P_kQ_j(Q_{\ge j-C}\chi\,\psi)\|_{L_t^2L_x^{\frac43}} &\lesssim \|Q_{\ge j-C}\chi\|_{L_t^2L_x^4}\,\|\psi\|_{L_t^\infty L_x^2} \lesssim 2^{-j}\|\nabla_{t,x}\chi\|_{L_t^2L_x^4}\,\|\psi\|_{L_t^\infty L_x^2}\\ &\lesssim 2^{-j}\,A\|\psi\|_S\end{aligned}$$

which makes a summable contribution to (7.31), and on the other hand,

$$\begin{aligned}\|P_kQ_j(Q_{<j-C}\chi\,\psi)\|^2_{L_t^2L_x^{\frac43}} &\lesssim \|P_kQ_j(Q_{<j-C}\chi\,P_{\le j-C}\tilde Q_j\psi)\|^2_{L_t^2L_x^{\frac43}}\\ &\lesssim A^2\sum_{\ell\le j-C}(1+2^\ell)^{-2}\|P_\ell\tilde Q_j\psi\|^2_{L_t^2L_x^2}\end{aligned}$$

Substituting this bound into (7.31) yields

$$\begin{aligned}
&\sum_{k\le C}\sum_{j>k+C} 2^{2\varepsilon k}2^{2j(1-\varepsilon)}A^2\sum_{\ell\le j-C}(1+2^\ell)^{-2}\|P_\ell\tilde Q_j\psi\|^2_{L^2_tL^2_x}\\
&\lesssim A^2\sum_\ell\sum_{j\ge \ell+C}2^{2j(1-\varepsilon)}(1+2^\ell)^{-2}\|P_\ell\tilde Q_j\psi\|^2_{L^2_tL^2_x}\\
&\lesssim A^2\sum_\ell\|P_\ell\psi\|^2_{S[\ell]}\lesssim A^2\|\psi\|_S^2
\end{aligned}$$

which is admissible. We may therefore assume that $k\ge C$. To proceed we need to control $\|P_\ell\chi\|_{S[\ell]}$. If $\ell\le 0$, then

$$\begin{aligned}
\|P_\ell\chi\|_{S[\ell]}&\lesssim \|P_\ell Q_{\le\ell}\chi\|_{\dot X_\ell^{0,\frac12,1}}+\|P_\ell Q_{>\ell}\chi\|_{\dot X_\ell^{-\frac12+\varepsilon,1-\varepsilon,2}}\\
&\lesssim 2^{\frac{\ell}{2}}\|P_\ell Q_{\le\ell}\chi\|_{L^2_tL^2_x}+2^{(-\frac12+\varepsilon)\ell}\|P_\ell Q_{\ell<\cdot<0}\chi\|_{L^2_tL^2_x}+2^{(-\frac12+\varepsilon)\ell}\|P_\ell Q_{>0}\,\partial_t\chi\|_{L^2_tL^2_x}\\
&\lesssim 2^{\frac{\ell}{2}}\Big(\int\sup_{|\eta|\sim 2^\ell}(|\eta|^2|\widehat\chi(\tau,\eta)|^2)\,d\tau\Big)^{\frac12}+2^{(-\frac12+\varepsilon)\ell}\Big(\int\langle\tau\rangle^2\sup_{|\eta|\sim2^\ell}(|\eta|^2|\widehat\chi(\tau,\eta)|^2)\,d\tau\Big)^{\frac12}\\
&\lesssim A\,2^{(-\frac12+\varepsilon)\ell} \qquad (7.32)
\end{aligned}$$

whereas if $\ell\ge 0$, then

$$\begin{aligned}
\|P_\ell\chi\|_{S[\ell]}&\lesssim \|P_\ell Q_{\le\ell}\chi\|_{\dot X_\ell^{0,\frac12,1}}+\|P_\ell Q_{>\ell}\chi\|_{\dot X_\ell^{-\frac12+\varepsilon,1-\varepsilon,2}}\\
&\lesssim 2^{\frac{\ell}{2}}\|P_\ell Q_{\le\ell}\chi\|_{L^2_tL^2_x}+2^{(-\frac12+\varepsilon)\ell}\|P_\ell Q_{>\ell}\,\partial_t\chi\|_{L^2_tL^2_x}\\
&\lesssim 2^{-10\ell}\Big(\int\langle\tau\rangle^2\sup_{|\eta|\sim2^\ell}(|\eta|^{22}|\widehat\chi(\tau,\eta)|^2)\,d\tau\Big)^{\frac12}\lesssim A\,2^{-10\ell} \qquad (7.33)
\end{aligned}$$

Using these bounds and applying Lemma 4.11 one obtains

$$\begin{aligned}
\sum_{k>C}\|P_kQ_{\le C}\Box(\chi\psi)\|^2_{N[k]}&\lesssim\sum_{k>C}\|P_kQ_{\le C}\Box(\chi\psi)\|^2_{\dot X_k^{-1,-\frac12,1}}\lesssim\sum_{k>C}\Big(\sum_{j\le C}\|P_kQ_j(\chi\psi)\|_{\dot X_k^{0,\frac12,\infty}}\Big)^2\\
&\lesssim\sum_{k>C}\Big(\sum_{k_1,k_2}2^{\frac34 k_1\wedge k_2}2^{\frac{k-k_1\vee k_2}{4}}\|P_{k_1}\chi\|_{S[k_1]}\|P_{k_2}\psi\|_{S[k_2]}\Big)^2\lesssim A^2\|\psi\|_S^2 \qquad (7.34)
\end{aligned}$$

The sum over k,k_1,k_2 here respects the usual trichotomy. We remark that one needs to limit the output here to modulations $Q_{\le C}$ as one would otherwise encounter logarithmic divergences. Next, one estimates

$$\begin{aligned}
\sum_{k>C}\|P_kQ_{\ge k}\Box(\chi\psi)\|^2_{N[k]}&\lesssim\sum_{k>C}\sum_{j\ge k}\|P_kQ_j\Box(\chi\psi)\|^2_{\dot X_k^{-\frac12+\varepsilon,-1-\varepsilon,2}}\\
&\lesssim\sum_{k>C}\sum_{j\ge k}2^{-(1-2\varepsilon)k}2^{2(1-\varepsilon)j}\|P_kQ_j(\chi\psi)\|^2_{L^2_tL^2_x} \qquad (7.35)
\end{aligned}$$

The contribution of $P_{\ge j-C}\chi$ is bounded by

$$\|P_kQ_j(P_{\ge j-C}\chi\,\psi)\|_{L^2_tL^2_x}\lesssim\|P_{\ge j-C}\chi\|_{L^2_tL^\infty_x}\|\psi\|_{L^\infty_tL^2_x}\lesssim A\,2^{-10j}\|\psi\|_S$$

which can be summed in (7.35). Furthermore,

$$\|P_kQ_j(P_{\le j-C}Q_{\ge j-C}\chi\,\psi)\|_{L^2_tL^2_x}\lesssim 2^{-j}\|P_{\le j-C}Q_{\ge j-C}\partial_t\chi\|_{L^2_tL^\infty_x}\|\psi\|_{L^\infty_tL^2_x}\lesssim A\,2^{-j}\|\psi\|_S$$

which is again sufficient for (7.35). Finally,

$$\begin{aligned}
\|P_kQ_j(P_{\le j-C}Q_{\le j-C}\chi\,\psi)\|_{L^2_tL^2_x}&\lesssim\|P_kQ_j(P_{\le k-C}Q_{\le j-C}\chi\,\tilde P_k\tilde Q_j\psi)\|_{L^2_tL^2_x} \qquad (7.36)\\
&\quad+\|P_kQ_j(P_{k-C<\cdot\le j-C}Q_{\le j-C}\chi\,P_{\le j}\tilde Q_j\psi)\|_{L^2_tL^2_x} \qquad (7.37)
\end{aligned}$$

Substituting (7.36) into (7.35) yields the estimate

$$\sum_{k>C}\sum_{j\ge k}2^{-(1-2\varepsilon)k}2^{2(1-\varepsilon)j}\|\chi\|^2_{L^\infty_{t,x}}\|\tilde P_k\tilde Q_j\psi)\|^2_{L^2_tL^2_x}\lesssim A^2\|\psi\|^2_S$$

Similarly, after a further frequency decomposition of ψ, substituting (7.37) into (7.35) leads to the same estimate. In summary, we are left to consider the output under the modulation constraint $Q_{C<\cdot<k}$. As a first reduction, we limit the modulation of χ (since we cannot control $\Box\chi$):

$$\begin{aligned}\|P_kQ_{\le k}\Box(Q_{>\frac{3k}{4}}\chi\,\psi)\|_{N[k]} &\lesssim \|P_kQ_{\le k}\Box(Q_{>\frac{3k}{4}}\chi\,\psi)\|_{\dot{X}_k^{-1,-\frac12,1}} \lesssim 2^{\frac{k}{2}}\|P_k(Q_{>\frac{3k}{4}}\chi\,\psi)\|_{L^2_tL^2_x}\\ &\lesssim 2^{\frac{k}{2}}\|P_k(P_{>k-10}Q_{>\frac{3k}{4}}\chi\,\psi)\|_{L^2_tL^2_x} + 2^{\frac{k}{2}}\|P_k(P_{\le k-10}Q_{>\frac{3k}{4}}\chi\,\tilde P_k\psi)\|_{L^2_tL^2_x}\\ &\lesssim 2^{\frac{k}{2}}\|\chi_{[|\xi|\gtrsim 2^k]}\widehat{\chi}(\tau,\xi)\|_{L^1_\xi L^2_\tau}\|\psi\|_S + 2^{\frac{k}{2}}\|\chi_{[||\tau|-|\xi||\gtrsim 2^{\frac{3k}{4}}]}\widehat{\chi}(\tau,\xi)\|_{L^1_\xi L^2_\tau}\|\tilde P_k\psi\|_{S[k]}\\ &\lesssim 2^{-k}\||\xi|^{100}\widehat{\chi}(\tau,\xi)\|_{L^2_\tau L^\infty_\xi}\|\psi\|_S + 2^{-\frac{k}{4}}\|\langle\tau\rangle|\xi|\vee|\xi|^{100}\widehat{\chi}(\tau,\xi)\|_{L^2_\tau L^\infty_\xi}\|\tilde P_k\psi\|_{S[k]}\\ &\lesssim A\,2^{-\frac{k}{4}}\|\psi\|_S\end{aligned}$$

which is admissible. We now estimate each of the three terms on the right-hand side of

$$\begin{aligned}\|P_kQ_{>C}\Box(Q_{\le\frac{3k}{4}}\chi\,\psi)\|_{N[k]} &\lesssim \|P_kQ_{>C}(\Box Q_{\le\frac{3k}{4}}\chi\,\psi)\|_{N[k]} + \|P_kQ_{>C}(\partial_\alpha Q_{\le\frac{3k}{4}}\chi\,\partial^\alpha\psi)\|_{N[k]}\\ &\quad+\|P_kQ_{>C}(Q_{\le\frac{3k}{4}}\chi\,\Box\psi)\|_{N[k]}\end{aligned}\tag{7.38}$$

First,

$$\begin{aligned}\|P_kQ_{>C}(\Box Q_{\le\frac{3k}{4}}\chi\,\psi)\|_{N[k]} &\lesssim 2^{-k}\|\Box Q_{\le\frac{3k}{4}}\chi\,\psi\|_{L^1_tL^2_x} \lesssim 2^{-k}\|\Box Q_{\le\frac{3k}{4}}\chi\|_{L^1_tL^\infty_x}\|\psi\|_S\\ &\lesssim 2^{-\frac{k}{4}}\max_{k=0,1}\max_{|\alpha|\le 1}\|\partial_t^k\nabla^\alpha\chi\|_{L^1_tL^\infty_x}\|\psi\|_S \lesssim A\,2^{-\frac{k}{4}}\|\psi\|_S\end{aligned}$$

which is admissible. Second, by estimate (29) in [57] as well as (7.32) and (7.33), and with k,k_1,k_2 respecting the usual trichotomy,

$$\begin{aligned}\|P_kQ_{>C}(\partial_\alpha Q_{\le\frac{3k}{4}}\chi\,\partial^\alpha\psi)\|_{N[k]} &\lesssim 2^{-k}\sum_{k_1,k_2}2^{k_1+k_2}\|P_{k_1}\chi\|_{S[k_1]}\|P_{k_2}\psi\|_{S[k_2]}\\ &\lesssim A2^{-k}\sum_{k_1,k_2}2^{k_1+k_2}\min(2^{-(\frac12-\varepsilon)k_1},2^{-10k_1})\|P_{k_2}\psi\|_{S[k_2]}\\ &\lesssim A(2^{-k}\|\psi\|_S+\|P_k\psi\|_{S[k]})\end{aligned}$$

which is again square-summable in k. As for the third term in (7.38) we are reduced to showing the bound

$$\sum_{k\ge C}\|P_kQ_{>C}(Q_{\le\frac{3k}{4}}\chi\,F)\|^2_{N[k]} \lesssim A^2\sum_{\ell\in\mathbb{Z}}\|P_\ell F\|^2_{N[\ell]}\tag{7.39}$$

This bound in turn follows via Schur's lemma from the following claim:

$$\|P_kQ_{>C}(Q_{\le\frac{3k}{4}}\chi\,P_\ell F)\|_{N[k]} \lesssim A2^{-\frac14|k-\ell|}\|P_\ell F\|_{N[\ell]}\tag{7.40}$$

If $j\le\ell$ and $j\le C$, then by (2.32) one always has the bound

$$\begin{aligned}\|P_kQ_{>C}(Q_{\le\frac{3k}{4}}\chi\,P_\ell Q_{\le j}F)\|_{N[k]} &\le 2^{-k}\|P_k(Q_{\le\frac{3k}{4}}\chi\,P_\ell Q_{\le j}F)\|_{L^1_tL^2_x} \lesssim 2^{-k}\|Q_{\le\frac{3k}{4}}\chi\|_{L^2_tL^\infty_x}\|P_\ell Q_{\le j}F\|_{L^2_tL^2_x}\\ &\lesssim 2^{\ell-k}2^{\frac{j}{2}}\|\widehat{\chi}(\tau,\xi)\|_{L^2_\tau L^1_\xi}\|P_\ell Q_{\le j}F\|_{N[\ell]} \lesssim A\,2^{\ell-k}\|F\|_{N[\ell]}\end{aligned}$$

which agrees with (7.40) provided $\ell\le k+C$. On the other hand, if $\ell\ge k+C$ but still $j\le C$ the same estimate holds with an additional high-high gain of $2^{-100\ell}$ coming from χ which is of course more than sufficient for (7.40). Finally, if $C\ge j\ge\ell$, then an additional Bernstein gain yields

$$\begin{aligned}\|P_kQ_{>C}(Q_{\le\frac{3k}{4}}\chi\,P_\ell Q_{\le j}F)\|_{N[k]} &\le 2^{-k}\|P_k(Q_{\le\frac{3k}{4}}\chi\,P_\ell Q_{\le j}F)\|_{L^1_tL^2_x} \lesssim 2^{-k}\|Q_{\le\frac{3k}{4}}\chi\|_{L^2_tL^4_x}\|P_\ell Q_{\le j}F\|_{L^2_tL^4_x}\\ &\lesssim 2^{\frac{\ell}{2}-k}\|\widehat{\chi}(\tau,\xi)\|_{L^{\frac43}_\xi L^2_\tau}\|P_\ell Q_{\le j}F\|_{L^2_tL^2_x} \lesssim A\,2^{\ell-k}\|F\|_{N[\ell]}\end{aligned}$$

as desired. Therefore, the claim (7.40) holds provided $F=Q_{\le C}F$. Let us now verify (7.40) for each of the four types of $N[\ell]$-atoms with the additional assumption that $F\ne Q_{\le C}F$. If F is an energy atom, then

$$\|P_kQ_{>C}(Q_{\le\frac{3k}{4}}\chi P_\ell F)\|_{N[k]} \lesssim 2^{-k}\|Q_{\le\frac{3k}{4}}\chi P_\ell F\|_{L^1_tL^2_x} \lesssim 2^{\ell-k}\|Q_{\le\frac{3k}{4}}\chi\|_{L^\infty_tL^\infty_x}\|P_\ell F\|_{N[\ell]}$$

which is sufficient if $\ell \le k+C$ and if $\ell > k+C$ then

$$\|P_k Q_{>C}(Q_{\le \frac{3k}{4}}\chi P_\ell F)\|_{N[k]} \lesssim 2^{-k}\|Q_{\le \frac{3k}{4}}\chi P_\ell F\|_{L^1_t L^2_x} \lesssim 2^{-\ell-k}\|Q_{\le \frac{3k}{4}}\chi\|_{L^\infty_t L^\infty_x}\|P_\ell F\|_{N[\ell]}$$

which is more than sufficient. Here we used the estimate

$$\|Q_{\le \frac{3k}{4}}\chi\|_{L^\infty_t L^\infty_x} \lesssim \|\widehat{\chi}(\tau,\xi)\|_{L^1_\tau L^1_\xi} \lesssim \|\langle\tau\rangle|\xi| \vee |\xi|^{100}\widehat{\chi}(\tau,\xi)\|_{L^2_\tau L^\infty_\xi} \lesssim A \tag{7.41}$$

For the remaining atoms we first make the simplifying assumption that $\widehat{\chi}(\tau,\xi)$ is supported on $|\tau|+|\xi| \lesssim 1$. Now suppose that $P_\ell Q_j F = F$ with $j > C$. If $\|F\|_{L^2_t L^2_x} \le 2^\ell 2^{\frac{j}{2}}$ and $j \le \ell$, then χ essentially does not change the Fourier support of F. Thus, $\ell = k + O(1)$ and

$$\|P_k Q_{>C}(\chi F)\|_{N[k]} \lesssim 2^{-\ell}2^{-\frac{j}{2}}\|\chi F\|_{L^2_t L^2_x} \lesssim A\|F\|_{\dot{X}_\ell^{-1,-\frac12,1}} \tag{7.42}$$

as desired. On the other hand, if $j > \ell$ and $\|F\|_{L^2_t L^2_x} \le 2^{\ell(\frac12-\varepsilon)}2^{j(1+\varepsilon)}$, then we need to distinguish the case $\ell \le C$ from $\ell > C$. In the latter case, one argues as in (7.42). In the former case, the modulation of the output is essentially 2^j and $k \le C$ which is excluded. It remains to consider the null-frame atoms. Thus, $F = \sum_{\kappa\in\mathcal{C}_m} F_\kappa$ where $F_\kappa = P_{\ell,\kappa}Q_{\le \ell+2m}F_\kappa$ and $m \le -100$. Due to $F \ne Q_{\le C}F$, one has $\ell+m \ge \ell+2m \ge C$ which implies that the Fourier support of χF_κ is essentially that of F_κ. Therefore, $\chi F = \sum_\kappa \chi F_\kappa$ can be treated as a wave-packet atom satisfying the bounds

$$\sum_\kappa \|\chi F_\kappa\|^2_{\mathrm{NF}[\kappa]} \lesssim A^2 \sum_\kappa \|F_\kappa\|^2_{\mathrm{NF}[\kappa]}$$

Since $k = \ell + O(1)$ we are done. Next, suppose that $\widehat{\chi}(\tau,\xi)$ is supported on $|\tau| \sim 2^n$ with $n \ge 10$ and $|\xi| \lesssim 1$. Then $\|\chi\|_{L^\infty_t L^\infty_x} \lesssim 2^{-n}A$. Start with a wave-packet atom F of the type we just considered. If $n \le k+2m+10$, then χF_κ has essentially the same Fourier support as F_κ whence

$$\|\chi F\|_{N[k]} \lesssim 2^{-k}\Big(\sum_\kappa \|\chi F_\kappa\|^2_{\mathrm{NF}[\kappa]}\Big)^{\frac12} \lesssim A2^{-n}2^{-\ell}\Big(\sum_\kappa \|F_\kappa\|^2_{\mathrm{NF}[\kappa]}\Big)^{\frac12} \lesssim 2^{-n}A\|F\|_{N[\ell]}$$

which is summable in $n \ge 10$. If $n > k+2m+10$, then χF has modulation of size 2^n. If $k \ge n$, then

$$\begin{aligned}
\|\chi F\|_{N[k]} &\lesssim \|\chi F\|_{\dot{X}_k^{-1,-\frac12,1}} \lesssim 2^{-\frac{n}{2}-k}\|\chi F\|_{L^2_t L^2_x}\\
&\lesssim A\,2^{-\frac{3n}{2}-k}\|F\|_{L^2_t L^2_x} \lesssim A\,2^{-\frac{3n}{2}-k}2^{\frac{3\ell}{2}}\|F\|_{\dot{X}_\ell^{-1,-\frac12,\infty}}\\
&\lesssim A\,2^{-n}\|F\|_{N[\ell]}
\end{aligned}$$

where we used (2.32) and $\ell = k + O(1)$. If $k < n$, then

$$\begin{aligned}
\|\chi F\|_{N[k]} &\lesssim \|\chi F\|_{\dot{X}_k^{-\frac12+\varepsilon,-1-\varepsilon,2}} \lesssim 2^{-n(1+\varepsilon)}2^{-k(\frac12-\varepsilon)}\|\chi F\|_{L^2_t L^2_x}\\
&\lesssim A\,2^{-n(2+\varepsilon)}2^{-k(\frac12-\varepsilon)}\|F\|_{L^2_t L^2_x} \lesssim A\,2^{-n(2+\varepsilon)}2^{k(1+\varepsilon)}\|F\|_{\dot{X}_\ell^{-1,-\frac12,\infty}}\\
&\lesssim A\,2^{-n}\|F\|_{N[\ell]}
\end{aligned}$$

Now suppose that F is a $\dot{X}_\ell^{-1,-\frac12,1}$-atom with $F = P_\ell Q_j F$. If $j > n+10$, then χF is the same kind of atom and one argues as before gaining a factor of 2^{-n}. If $j = n + O(1)$, then

$$\begin{aligned}
\|\chi F\|_{N[k]} &\lesssim 2^{-k}\|\chi F\|_{L^1_t L^2_x} \lesssim 2^{-\ell}\|\chi\|_{L^2_t L^\infty_x}\|F\|_{L^2_t L^2_x}\\
&\lesssim 2^{-n}A2^{\frac{j}{2}}\|F\|_{\dot{X}_\ell^{-1,-\frac12,1}} \lesssim 2^{-\frac{n}{2}}A
\end{aligned}$$

Finally, if $n > j+10$, then χF has modulation of size 2^n. If $n \le \ell$, then

$$\|\chi F\|_{N[k]} \lesssim \|\chi F\|_{\dot{X}_k^{-1,-\frac12,1}} \lesssim 2^{-\ell-\frac{n}{2}}\|\chi F\|_{L^2_t L^2_x} \lesssim 2^{-n}A$$

whereas in case $n \ge \ell$, one checks similarly that

$$\|\chi F\|_{N[k]} \lesssim \|\chi F\|_{\dot{X}_k^{-\frac12+\varepsilon,-1-\varepsilon,2}} \lesssim 2^{-n}A$$

as desired. If F is a $\dot{X}_\ell^{-\frac{1}{2}+\varepsilon,-1-\varepsilon,2}$-atom with $F = P_\ell Q_j F$, then analogous arguments lead to a bound of $\|\chi F\|_{N[\ell]} \lesssim 2^{-\varepsilon n} A$ which is again summable in $n \geq 0$.

Finally, one needs to consider the case where $\widehat{\chi}(\tau,\xi)$ is supported on $|\xi| \sim 2^n$ with $n \geq 10$, say. However, this is easier due to the rapid decay of $\widehat{\chi}$ in ξ. We leave those details to the reader. □

Remark 7.24. Lemma 7.23 of course applies to any space-time Schwartz function χ. Moreover, one can check that the exact same conclusions of Lemma 7.23 hold for any Schwartz function χ which only depends on t and x alone; the only difference is that C_0A needs to be replaced by $C(\chi)$.

It is now a simple matter to prove that the ψ_α^n have uniformly controlled S norms on some time interval $(-T_0, T_0)$ where $T_0 = T_0(V)$.

Corollary 7.25. *Under the assumptions of Lemma 7.10 there exists a time $T_0 = T_0(V) > 0$ such that*

$$\max_{\alpha=0,1,2} \|\psi_\alpha^n\|_{S((-T_0,T_0)\times\mathbb{R}^2)} \leq C(V) < \infty \tag{7.43}$$

uniformly in large n.

Proof. Pick $r_0 > 0$ small enough and R_0 large enough according to Lemmas 7.21 and 7.22, respectively. In view of (7.24), Theorem 7.18, and finite propagation speed, patching up the local evolutions of $(\mathbf{x}_1^n, \mathbf{y}_1^n)$ shows that the evolution of $(\mathbf{x}^n, \mathbf{y}^n)$ exists on some time interval $(-T_0, T_0)$ uniformly in large n; in fact, one can take $T_0 = r_0$. Note that this part of the argument does not require $(\mathbf{x}_2^n, \mathbf{y}_2^n)$. These functions are needed to obtain uniform control over $\|\psi_\alpha^n\|_{S((-T_0,T_0)\times\mathbb{R}^2)}$, to which we now turn. The ϕ_α^n of the original sequence agree with the $\tilde{\phi}_\alpha^n$ obtained from (7.24) on the cone $K_{x_0,r_0} := \{(t,x) \mid |x - x_0| < r_0 - t,\ 0 \leq t < r_0\}$. This follows from the construction of $(\mathbf{x}_1^n, \mathbf{y}_1^n)$ and finite propagation speed. Note that the $\tilde{\phi}_\alpha^n$ exist globally in $\mathbb{R}^{1+2}$ but agree with ϕ_α^n only on K_{x_0,r_0}. A similar observation applies to $(\mathbf{x}_2^n, \mathbf{y}_2^n)$ on the set $K_{R_0,T_0} := \{|x| > R_0 + t, 0 \leq t < T_0\}$. Cover $\mathbb{R}^2$ by finitely many $D_j := D(x_j, r_0)$ as well as the complement of $D_0 := D(0, R_0)$. This can be done is such a fashion that there exists a smooth and finite partition of unity $1 = \sum_{j=1}^J \chi_j$ on $[0, T_0] \times \mathbb{R}^2$ such that each χ_j is entirely supported in either a cone K_{x_j,r_0} or within K_{R_0,T_0}. Thus

$$\psi_\alpha^n = \sum_j \chi_j \psi_\alpha^n = \sum_j \chi_j \widetilde{\psi}_\alpha^{n,j} e^{i\Delta^{-1}\partial \mathrm{Re}(\tilde{\phi}^{n,j} - \phi^n)}$$

Here $\tilde{\phi}_\alpha^{n,j}$ are the derivative components of the *small energy* wave maps which were constructed by means of Lemmas 7.21 and 7.22, and $\widetilde{\psi}_\alpha^{n,j}$ are their gauged counterparts. If χ_j has compact support, we now claim that

$$\widetilde{\chi}_j^n := \chi_j\, e^{i\Delta^{-1}\partial \mathrm{Re}(\tilde{\phi}^{n,j} - \phi^n)}$$

satisfies the hypotheses of Lemma 7.23 with a constant A that can be chosen uniformly in n. The compact support assumption in time can of course be fulfilled. Since for each j and all n

$$\chi_j(\tilde{\phi}_\alpha^{n,j} - \phi_\alpha^n) = 0$$

it follows from the uniform L^2 bound on $\tilde{\phi}_\alpha^{n,j}$ and ϕ_α^n that

$$\Delta^{-1}\partial \mathrm{Re}(\tilde{\phi}_\alpha^{n,j} - \phi_\alpha^n)(t,x) = \frac{1}{2\pi} \int \frac{x-y}{|x-y|^2} \mathrm{Re}(\tilde{\phi}_\alpha^{n,j} - \phi_\alpha^n)(t,y)\, dy \tag{7.44}$$

is a smooth function relative to x on the support of χ_j with uniform L^∞ bounds on the derivatives (uniform here means relative to large n). Indeed,

$$\|\chi_j \nabla_x^\alpha \Delta^{-1}\partial \mathrm{Re}(\tilde{\phi}_\beta^{n,j} - \phi_\beta^n)(t,x)\|_{L_x^\infty} \leq C_\alpha \|(\tilde{\phi}_\beta^{n,j} - \phi_\beta^n)(t,x)\|_{L_x^2} \leq C_\alpha\, E \tag{7.45}$$

where E governs the energy uniformly in t. It turns out that we can also incorporate one time derivative into these bounds (but not necessarily any higher regularity in time). This follows from the div-curl system for ϕ_α, see (1.6)–(1.9). Indeed, if $\alpha \neq 0$ then plugging (1.6) into

$$\partial_t \Delta^{-1}\partial \mathrm{Re}(\tilde{\phi}_\alpha^{n,j} - \phi_\alpha^n)(x) = \frac{1}{2\pi} \int \frac{x-y}{|x-y|^2}\, \partial_0 \mathrm{Re}(\tilde{\phi}_\alpha^{n,j} - \phi_\alpha^n)(t,y)\, dy$$

leads to an expression which is of the schematic form

$$\int \frac{x-y}{|x-y|^2}\,\partial_\alpha \mathrm{Re}(\tilde{\phi}_0^{n,j}-\phi_0^n)(t,y)\,dy + \int \frac{x-y}{|x-y|^2}\big[(\tilde{\phi}^{n,j})^2-(\phi^n)^2\big](t,y)\,dy$$

Integrating by parts in the first integral moves the derivative from the ϕ's onto the kernel which allows for the same estimate as in (7.45). As for the second integral on the right-hand side, one has

$$\Big\|\nabla_x^\alpha\big[\chi_j \int \frac{x-y}{|x-y|^2}\big[(\tilde{\phi}^{n,j})^2-(\phi^n)^2\big](t,y)\,dy\big]\Big\|_{L_x^\infty} \le C_\alpha \|(\tilde{\phi}^{n,j})^2-(\phi^n)^2(t)\|_{L_x^1} \le C_\alpha\, E$$

as desired. If $\alpha = 0$, then one uses (1.9) to arrive at the same conclusion. This establishes our claim concerning the hypotheses of Lemma 7.23; in fact, we obtained stronger conclusions as far as the conditions for large x or small ξ are concerned. Now let us consider the cut-off function χ_j with unbounded support, which we may assume is χ_0. We can arrange the partition of unity so that $\chi_0(t,x) = \chi_{00}(x)\chi_{01}(t)$ with χ_{01} smooth and supported in $(-1,1)$ and with $1-\chi_{00}$ smooth and compactly supported in $\mathbb{R}^2$. With $\widetilde{\chi}_0^n$ defined as above, we now claim that

$$\chi_{01}(t) - \widetilde{\chi}_0^n(t,x) = \chi_{01}(t)\big(1-\chi_{00}(x)\,e^{i\Delta^{-1}\partial \mathrm{Re}(\tilde{\phi}^{n,j}-\phi^n)(t,x)}\big)$$

satisfies the requirements of Lemma 7.23 with a constant A that is controlled uniformly in n. First,

$$\chi_{01}(t)\chi_{00}(x)\,\mathrm{Re}(\tilde{\phi}_\alpha^{n,j}-\phi_\alpha^n)(t,x) = 0$$

which shows as before that $\widetilde{\chi}_0^n(t,x)$ is smooth in x with derivatives that are uniformly bounded in L_x^∞ relative to n. In addition, the same arguments involving the div-curl system allow us to place one ∂_t on $\widetilde{\chi}_0^n(t,x)$ without destroying these conclusions. As for the asymptotic behavior in $x\to\infty$ and $\xi \to 0$, one simply expands

$$\frac{x-y}{|x-y|^2} = \frac{x}{|x|^2} + O(\frac{y}{|x|^2})$$

inside the integral in (7.44) which is sufficient due to $|y| \lesssim R_0$. In conclusion, by Lemma 7.23 and Remark 7.24

$$\|\psi_\alpha^n\|_{S(-T_0,T_0)} \le \sum_j \|\widetilde{\chi}_j^n \widetilde{\psi}_\alpha^{n,j}\|_{S(-T_0,T_0)} \le \sum_j C(\widetilde{\chi}_j^n)\|\widetilde{\psi}_\alpha^{n,j}\|_S \le \sum_j C(\widetilde{\chi}_j^n) C\varepsilon_0$$

is finite uniformly in n. □

The preceding corollary concludes the proof of Lemma 7.10 up to the assertion about the frequency envelope at the end. This will be proved in Section 9.5.

We close this section with an important strengthening of the bound on ψ_L from Lemma 7.6. More specifically, we prove that the intervals I_j can be chosen in such a way that the estimate (7.5) only depends on the energy of ψ. This will play an important role later on. In order to achieve this property, we require an improvement over Lemma 7.4. We begin with the following technical statements which allow us to make a better choice of the intervals I_j in the proof of Lemma 7.6.

Lemma 7.26. *Let $\|\psi\|_S < C_0$ and $\varepsilon_0 > 0$ be arbitrary, with ψ defined on $\mathbb{R}^{2+1}$. Then there exists a partition of $\mathbb{R}$ into intervals $\{I_j\}_{j=1}^M$ which depend on ψ but with $M = M(\varepsilon_0, C_0)$ and which satisfy*

$$\max_{1\le j\le M} \sum_{k\in\mathbb{Z}} \|P_k(\psi\ |\nabla|^{-1}\psi^2)\|^2_{L_t^2(I_j;\dot{H}^{-\frac12})} \le \varepsilon_0$$

where $\nabla = \nabla_x$ and $\psi\ |\nabla|^{-1}\psi^2$ is schematic notation which stands for any one of the nonlinearities appearing on the right-hand side of the div-curl system (1.12)*,* (1.13)*.*

Proof. It of course suffices to show that

$$\sum_{k\in\mathbb{Z}} \|P_k[\psi_1\ |\nabla|^{-1}(\psi_2\psi_3)]\|^2_{L_t^2(\mathbb{R};\dot{H}^{-\frac12})} \lesssim \prod_{i=1}^3 \|\psi_i\|_S^2 \tag{7.46}$$

We begin with the case where $\psi_2\psi_3$ is replaced by $I^c\psi_2\cdot\psi_3$. It is easy to see that

$$\|P_k(I^c\psi_2\cdot\psi_3)\|_{L^2_{t,x}} \lesssim 2^{\frac{k}{2}}\|\psi_2\|_S\|\psi_3\|_S$$

Then by the usual trichotomy,

$$\begin{aligned}\sum_{k\in\mathbb{Z}} 2^{-k}\|P_k(\psi_1\,|\nabla|^{-1}(I^c\psi_2\cdot\psi_3))\|^2_{L^2_{t,x}} &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\|P_k\psi_1\,|\nabla|^{-1}P_{<k-5}(I^c\psi_2\cdot\psi_3)\|^2_{L^2_{t,x}}\\ &+\sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k}\|P_k[\tilde P_\ell\psi_1\,|\nabla|^{-1}P_\ell(I^c\psi_2\cdot\psi_3)]\|_{L^2_{t,x}}\Big)^2\\ &+\sum_{k\in\mathbb{Z}} 2^{-k}\|P_{<k-5}\psi_1\,|\nabla|^{-1}P_k(I^c\psi_2\cdot\psi_3)\|^2_{L^2_{t,x}} \lesssim \prod_{i=1}^3\|\psi_i\|_S^2\end{aligned}$$

Hence, we may assume that the two inner inputs are both hyperbolic, i.e., $\psi_i = Q_{\le k_i}\psi_i$ for $i = 2, 3$. Now implement the Hodge decomposition for the inputs of $|\nabla|^{-1}(\psi^2)$, i.e., write

$$\psi_\alpha = R_\alpha\psi + \chi_\alpha$$

We begin by considering the resulting trilinear expressions, more specifically the one where the inner null-form is hyperbolic: Suppressing the indices on ψ for simplicity,

$$\sum_{k\in\mathbb{Z}}\|P_k(\psi|\nabla|^{-1}I\mathcal{Q}_{\alpha\beta}(\psi,\psi))\|^2_{L^2_t(\mathbb{R};\dot H^{-\frac12})} \le C\|\psi\|_S^6 \tag{7.47}$$

where $\mathcal{Q}_{\alpha j}$ is the null-form from Definition 4.16. As usual, this splits into the high-low, high-high, and low-high cases:

$$\begin{aligned}\sum_{k\in\mathbb{Z}} 2^{-k}\|P_k(\psi|\nabla|^{-1}I\mathcal{Q}_{\alpha\beta}(\psi,\psi))\|^2_{L^2_{t,x}} &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\|P_k(\tilde P_k\psi|\nabla|^{-1}P_{<k-5}I\mathcal{Q}_{\alpha\beta}(\psi,\psi))\|^2_{L^2_{t,x}}\\ &+\sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k}\|P_k(P_\ell\psi\tilde P_\ell|\nabla|^{-1}I\mathcal{Q}_{\alpha\beta}(\psi,\psi))\|_{L^2_{t,x}}\Big)^2\\ &+\sum_{k\in\mathbb{Z}} 2^{-k}\|P_k(P_{<k-5}\psi\tilde P_k|\nabla|^{-1}I\mathcal{Q}_{\alpha\beta}(\psi,\psi))\|^2_{L^2_{t,x}}\\ &=: A+B+C\end{aligned}$$

Next, one writes $A \le A_1 + A_2 + A_3$ reflecting the high-high, high-low, and low-high decomposition of the $\mathcal{Q}_{\alpha\beta}$-nullform. Thus, by Lemma 4.17,

$$\begin{aligned}A_1 &\le \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_0<k-5}\;\sum_{k_2=k_3+O(1)>k_0-5}\|P_k(\tilde P_k\psi|\nabla|^{-1}P_{k_0}I\mathcal{Q}_{\alpha\beta}(P_{k_2}\psi,P_{k_3}\psi))\|_{L^2_{t,x}}\Big)^2\\ &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_0<k-5}\;\sum_{k_2=k_3+O(1)>k_0-5}\|\tilde P_k\psi\|_{L^\infty_tL^2_x}\|P_{k_0}I\mathcal{Q}_{\alpha\beta}(P_{k_2}\psi,P_{k_3}\psi))\|_{L^2_{t,x}}\Big)^2\\ &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\|\tilde P_k\psi\|^2_{L^\infty_tL^2_x}\Big(\sum_{k_0<k-5}\;\sum_{k_2=k_3+O(1)>k_0-5}2^{k_0-\frac{k_2}{2}}\|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6\end{aligned}$$

Similarly, by Lemma 4.23,

$$\begin{aligned}A_2 &\le \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_2+O(1)=k_0<k-5}\;\sum_{k_3<k_0-5}\|P_k(\tilde P_k\psi|\nabla|^{-1}P_{k_0}I\mathcal{Q}_{\alpha\beta}(P_{k_2}\psi,P_{k_3}\psi))\|_{L^2_{t,x}}\Big)^2\\ &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_2+O(1)=k_0<k-5}\;\sum_{k_3<k_0-5}\|\tilde P_k\psi\|_{L^\infty_tL^2_x}\|P_{k_0}I\mathcal{Q}_{\alpha\beta}(P_{k_2}\psi,P_{k_3}\psi))\|_{L^2_{t,x}}\Big)^2\\ &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\|\tilde P_k\psi\|^2_{L^\infty_tL^2_x}\Big(\sum_{k_2+O(1)=k_0<k-5}\;\sum_{k_3<k_0-5}2^{(\frac12-\varepsilon)k_3}2^{\varepsilon k_0}\|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6\end{aligned}$$

This concludes the bound on A since A_3 is of course symmetric to A_2. Next, with $B \le B_1 + B_2 + B_3$ via the same trichotomy,

$$B_1 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k} 2^k \|P_\ell\psi\|_{L_t^\infty L_x^2}\, 2^{-\ell} \sum_{k_2=k_3+O(1)>\ell-5} \|\tilde{P}_\ell I \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi))\|_{L_t^2 L_x^2}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k} 2^k \|P_\ell\psi\|_{L_t^\infty L_x^2}\, 2^{-\ell} \sum_{k_2=k_3+O(1)>\ell-5} 2^{\ell-\frac{k_2}{2}} \|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6$$

by Lemma 4.17, whereas

$$B_2 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k} 2^k \|P_\ell\psi\|_{L_t^\infty L_x^2}\, 2^{-\ell} \sum_{\ell=k_2+O(1)>k_3-5} \|\tilde{P}_\ell I \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi))\|_{L_t^2 L_x^2}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k} 2^k \|P_\ell\psi\|_{L_t^\infty L_x^2}\, 2^{-\ell} \sum_{\ell=k_2+O(1)>k_3-5} 2^{(\frac{1}{2}-\varepsilon)k_3} 2^{\varepsilon k_2} \|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6$$

by Lemma 4.23. The low-high case of (7.47) is treated in an analogous fashion and we skip it.

Next, we treat the case where the inner null-form is elliptic. Then the desired bound reads

$$\sum_{k\in\mathbb{Z}} \|P_k(\psi\,|\nabla|^{-1} I^c \mathcal{Q}_{\alpha\beta}(\psi,\psi))\|^2_{L_t^2(\mathbb{R};\dot{H}^{-\frac{1}{2}})} \le C\|\psi\|_S^6 \tag{7.48}$$

As before, $A \le A_1 + A_2 + A_3$ reflecting the high-high, high-low, and low-high decomposition of the $\mathcal{Q}_{\alpha\beta}$-nullform. *We will first exclude the contributions by opposing high-high interactions in the null-form*, cf. Remark 4.20. Hence, by (4.55) *without* the $\langle k_1 - k\rangle^2$ loss,

$$A_1 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_0<k-5}\ \sum_{k_2=k_3+O(1)>k_0-5} \|P_k(\tilde{P}_k\psi\,|\nabla|^{-1} P_{k_0} Q_{>k_0} \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi))\|_{L^2_{t,x}}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_0<k-5}\ \sum_{k_2=k_3+O(1)>k_0-5} \|P_k\psi\|_{L_t^\infty L_x^2} \|P_{k_0} Q_{>k_0} \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi)\|_{L^2_{t,x}}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_0<k-5}\ \sum_{k_2=k_3+O(1)>k_0-5} \|P_k\psi\|_{L_t^\infty L_x^2} 2^{\frac{k_0}{2}} \|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6$$

For A_2 one proceeds similarly, using Lemma 4.24 instead. In fact, due to the hyperbolic nature of ψ_1, ψ_3,

$$A_2 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_2+O(1)=k_0<k-5}\ \sum_{k_3<k_0-5} \|P_k(\tilde{P}_k\psi\,|\nabla|^{-1} P_{k_0} I^c \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi))\|_{L^2_{t,x}}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_2+O(1)=k_0<k-5}\ \sum_{k_3<k_0-5} \|P_k\psi\|_{L_t^\infty L_x^2} \|P_{k_0} \tilde{Q}_{k_0} \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi)\|_{L^2_{t,x}}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_2+O(1)=k_0<k-5}\ \sum_{k_3<k_0-5} \|P_k\psi\|_{L_t^\infty L_x^2} 2^{\frac{k_3}{2}} \|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6$$

This concludes the high-low case A. In the high-high case we write $B \le B_1 + B_2 + B_3$ as before. Therefore,

$$B_1 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k} 2^k \|P_\ell\psi\|_{L_t^\infty L_x^2}\, 2^{-\ell} \sum_{k_2=k_3+O(1)>\ell-5} \|\tilde{P}_\ell Q_{>\ell} \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi)\|_{L^2_{t,x}}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} \Big(\sum_{\ell>k} 2^{\frac{k}{2}} \|P_\ell\psi\|_{S[\ell]}\, 2^{-\ell} \sum_{k_2=k_3+O(1)>\ell-5} 2^{\frac{\ell}{2}} \|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6$$

by Lemma 4.19, whereas

$$B_2 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{\ell>k} 2^k \|P_\ell Q_j \psi\|_{L^2_{t,x}}\, 2^{-\ell} \sum_{\ell=k_2+O(1)>k_3-5} \|\tilde{P}_\ell \tilde{Q}_\ell \mathcal{Q}_{\alpha\beta}(P_{k_2}\psi, P_{k_3}\psi)\|_{L^2_t L^2_x}\Big)^2$$
$$\lesssim \sum_{k\in\mathbb{Z}} \Big(\sum_{\ell>k} 2^{\frac{k}{2}} \|P_\ell \psi\|_{S[\ell]}\, 2^{-\ell} \sum_{\ell=k_2+O(1)>k_3-5} 2^{\frac{k_3}{2}} \|P_{k_2}\psi\|_{S[k_2]} \|P_{k_3}\psi\|_{S[k_3]}\Big)^2 \lesssim \|\psi\|_S^6$$

by Lemma 4.24 which finishes the analysis of B. We again leave the low-high case to the reader.

It remains the bound the contributions by the opposing high-high waves in the inner null-form. Returning to the ψ_1, ψ_2, ψ_3 notation, we may assume that $\psi_i = Q_{\le k_i}\psi_i$ for $i = 2, 3$ and that there is an angular separation of the Fourier supports of ψ_1 and ψ_2, say (since the Fourier supports of ψ_2, ψ_3 make a large angle). Hence we may bound the missing contribution to A_1 as follows, where we ignore the nullform and replace the outer $|\nabla|^{-1}$ with a weight by the usual convolution logic:

$$A_1 \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\Big(\sum_{k_0<k-5}\ \sum_{k_2=k_3+O(1)>k_0-5} 2^{-k_0}\big[\sum_{c\in\mathcal{D}_{k,k_0-k}} \|P_c\psi_1 P_{k_0}(P_{k_2}\psi_2\, P_{k_3}\psi_3)\|^2_{L^2_{t,x}}\big]^{\frac12}\Big)^2$$

We now invoke (2.30) to conclude that

$$\begin{aligned}\big[\sum_{c\in\mathcal{D}_{k,k_0-k}} \|P_c\psi_1\, P_{k_0}(P_{k_2}\psi_2\, P_{k_3}\psi_3)\|^2_{L^2_{t,x}}\big]^{\frac12} &\lesssim 2^{k_0}\big[\sum_{c\in\mathcal{D}_{k,k_0-k}} \|P_c\psi_1 P_{k_2}\psi\, P_{k_3}\psi\|^2_{L^2_t L^1_x}\big]^{\frac12}\\
&\lesssim 2^{k_0}\big[\sum_{c\in\mathcal{D}_{k,k_0-k}} \|P_c\psi_1 P_{k_2}\psi\|^2_{L^2_{t,x}} \|P_{k_3}\psi\|^2_{L^\infty_t L^2_x}\big]^{\frac12}\\
&\lesssim 2^{\frac{3k_0}{2}}\langle k_0-k\rangle\big[\sum_{c\in\mathcal{D}_{k,k_0-k}} \|P_c\psi_1\|^2_{S[k]} \|P_{k_2}\psi\|^2_{S[k_2]}\big]^{\frac12} \|P_{k_3}\psi\|_{L^\infty_t L^2_x}\\
&\lesssim 2^{\frac{3k_0}{2}}\langle k_0-k\rangle \prod_{i=1}^{3} \|P_{k_i}\psi_i\|_{S[k_i]}\end{aligned}$$

The loss of $\langle k_0 - k\rangle$ here is due to the usual issue of wave-packets which are too thick resulting in the need for Lemma 2.4. Inserting this into the bound on A_1 yields

$$A_1 \lesssim \sum_{k\in\mathbb{Z}} \Big(\sum_{k_0<k} 2^{\frac{k_0-k}{2}} \langle k_0-k\rangle \|P_k\psi_1\|_{S[k]}\Big)^2 \|\psi_2\|_S^2 \|\psi_3\|_S^3 \lesssim \prod_{i=1}^{3} \|\psi_i\|_S^2$$

as desired. The opposing high-high contributions to the other terms are similar and omitted. We still need to control the contributions from the elliptic terms χ, leading to higher order nonlinearities. This is again done in the appendix. □

We can now state the refined version of Lemma 7.6 which gives better control over the linear wave ψ_L. As in that lemma ψ are the gauged components of an admissible wave map locally on some time interval $[-T_0, T_1]$.

Corollary 7.27. *Let $\|\psi\|_S < C_0$, with ψ defined on $\mathbb{R}^{2+1}$. Given $\varepsilon_0 > 0$, there exist $M_1 = M_1(C_0, \varepsilon_0)$ many intervals I_j as in (7.2) with the following property: for each $I_j = (t_j, t_{j+1})$, there is a decomposition*

$$\psi|_{I_j} = \psi_L^{(j)} + \psi_{NL}^{(j)}, \quad \Box\psi_L^{(j)} = 0$$

which satisfies

$$\sum_{k\in\mathbb{Z}} \|P_k \psi_{NL}^{(j)}\|^2_{S[k](I_j\times\mathbb{R}^2)} < \varepsilon_0 \tag{7.49}$$

$$\|\nabla_{x,t}\psi_L^{(j)}\|_{L^\infty_t \dot{H}^{-1}} \lesssim \varepsilon_0^{-\frac14}(E+1)E \tag{7.50}$$

where the implied constant in the last inequality is universal and $E = \|\psi(t)\|_2$ is the conserved energy. In particular,

$$\|\psi_{NL}^{(j)}\|_{S(I_j\times\mathbb{R}^2)}\, \|\nabla_{x,t}\psi_L^{(j)}\|_{\dot{H}^{-1}} \ll 1$$

by choosing ε_0 small enough depending on the energy.

Proof. We first prove (7.49) and (7.50) by following the strategy of the proof of Lemma 7.6; however, we use Lemma 7.26 instead of Lemma 7.4 when the underlying time interval is small. More precisely, consider the frequency component $P_0\psi_\alpha$.

Case 1: The underlying time interval $I_0 := (-T_0, T_1)$ satisfies $|I_0| < \varepsilon_1$ with an ε_1 that is to be determined. The main property of this parameter is that it can be chosen to be an absolute constant independently of C_0. The ψ_α satisfy the system (1.12)–(1.14). Schematically, this system takes the form

$$\partial_t P_0\psi_j = \partial_j P_0\psi_0 + P_0[\psi\nabla^{-1}(\psi^2)], \quad j = 1,2$$
$$\partial_t P_0\psi_0 = \sum_{j=1}^{2} \partial_j P_0\psi_j + P_0[\psi\nabla^{-1}(\psi^2)]$$

where the nonlinearity is written schematically. Now define the linear wave $P_0\psi_L$ to be

$$P_0\psi_{L,j} := S(t)(P_0\psi_j(0), \partial_j P_0\psi_0), \quad j = 1,2$$
$$P_0\psi_{L,0} := S(t)\big(P_0\psi_0, \sum_{j=1}^{2} P_0\partial_j\psi_j(0)\big)$$

whereas $P_0\psi_{NL,\alpha} := P_0\psi_\alpha - P_0\psi_{L,\alpha}$. Thus, for $j = 1,2$,

$$P_0\psi_j(t) = P_0\psi_j(0) + \int_0^t P_0\partial_j P_0\psi_0(s)\,ds + \int_0^t P_0[\psi\nabla^{-1}(\psi^2)](s)\,ds$$
$$P_0\psi_{L,j}(t) = P_0\psi_j(0) + tP_0\partial_j\psi_0(0,\cdot) + O_{L^2_x}(t^2)$$

and similarly for ψ_0, whence for all $t \in I_0$,

$$\|P_0\psi_{NL}(t)\|_{L^2_x} \le t^2\|P_0\psi(0)\|_{L^2_x} + \Big\|\int_0^t P_0[\psi\nabla^{-1}(\psi^2)](s,\cdot)\,ds\Big\|_{L^2_x}$$
$$\lesssim t^2\|P_0\psi(0)\|_{L^2_x} + |t|^{\frac12}\big\|P_0[\psi\nabla^{-1}(\psi^2)]\big\|_{L^2_{t,x}}$$

In other words,

$$\|P_0\psi_{NL}\|_{L^\infty_t(I_0;L^2_x)} \lesssim \varepsilon_1^2\|P_0\psi(0)\|_{L^2_x} + \varepsilon_1^{\frac12}\|P_0[\psi\nabla^{-1}(\psi^2)]\|_{L^2_t(I_0;L^2_x)}$$

As in the proof of Lemma 7.6 one concludes from this that

$$\|P_0\psi_{NL}\|_{S[0](I_0\times\mathbb{R}^2)} \lesssim \varepsilon_1^2\|P_0\psi(0)\|_{L^2_x} + \varepsilon_1^{\frac12}\|P_0[\psi\nabla^{-1}(\psi^2)]\|_{L^2_t(I_0;L^2_x)}$$

Rescaling this bound to general 2^k yields the following. Suppose $|I| \le \varepsilon_1 2^{-k}$. Then

$$\|P_k\psi_{NL}\|_{S[k](I_0\times\mathbb{R}^2)} \lesssim \varepsilon_1^2\|P_k\psi(0)\|_{L^2_x} + \varepsilon_1^{\frac12}2^{-\frac{k}{2}}\|P_0[\psi\nabla^{-1}(\psi^2)]\|_{L^2_t(I_0;L^2_x)}$$

Now provided $I_0 \subset I_j$ where $\{I_j\}_{j=1}^M$, $I_j = I_j(\tilde\varepsilon_0,\psi)$, $M = M(\tilde\varepsilon_0, C_0)$ are the intervals constructed in Lemma 7.26, one concludes that

$$\sum_{k:|I_0|\le\varepsilon_1 2^{-k}} \|P_k\psi_{NL}\|^2_{S[k](I_0\times\mathbb{R}^2)} \lesssim \varepsilon_1^4\|\psi(0)\|^2_{L^2_x} + \tilde\varepsilon_0\varepsilon_1 \tag{7.51}$$

where $\tilde\varepsilon_0$ is a separate smallness parameter. We now pick $\varepsilon_1 := \varepsilon_0^{\frac14}(1+E)^{-1}$ and $\tilde\varepsilon_0 := \varepsilon_0^{\frac34}$ where $E = \|\psi(t)\|_{L^2_x}$ is the conserved energy of ψ (for this one needs to remain on the interval on which ψ equals the gauged derivative components of a wave map). This renders the right-hand side of (7.51) less than ε_0.

As already explained in the proof of Lemma 7.6, we will use this analysis also in the case of large intervals to which we now turn. However, in that case the estimates obtained here allow one to control the term $\|\psi\big|_{[-T_0,T_0]}\|_S$ in (2.73) of Section 2.5.

Case 2: The underlying time interval $I_0 = (-T_0, T_1)$ satisfies $|I_0| > \varepsilon_1$ where ε_1 is as in Case 1 (again for the P_0 frequencies). Here the analysis of Case 2 of Lemma 7.6 applies verbatim, leading to intervals $\{I'_j\}_{j=1}^{M'}$, with $M' = M'(\varepsilon_0, C_0)$ such that

$$\max_{1\le j\le M'} \sum_{k\in\mathbb{Z}} \|P_k F_\alpha\|^2_{N[k](I_j\times\mathbb{R}^2)} < \varepsilon_0$$

where $F = \sum_\alpha F_\alpha$ stands for the right-hand side of (1.14) as usual.

Now we take the intersections of the intervals I_j and I'_k which appeared in Cases 1 and 2 above. Denote this collection again by $\{I_j\}_{j=1}^{M}$ with $M = M(\tilde\varepsilon_0, \varepsilon_0, C_0)$. Fix such an I_j. Given $k \in Z$, we define $P_k\psi_L^{(j)}$ to be the free evolution of $(I\psi)[t_0]$ where $t_0 \in I_j$ is the center of I_j, whereas $P_k\psi_{NL}^{(j)}$ is everything else. By our construction,

$$\sum_{k:|I_j|\le\varepsilon_1 2^{-k}} \|P_k\psi_{NL}^{(j)}\|^2_{S[k](I_j\times\mathbb{R}^2)} \le \varepsilon_0$$

Combining this with (7.51) this bound implies (7.49). As for the linear wave $\psi_L^{(j)}$, we note that those k which belong to Case 1 yield

$$\|P_k\psi_L^{(j)}\|_{S[k](I_j\times\mathbb{R}^2)} \lesssim \|P_k\psi\|_2$$

with an absolute implicit constant, whereas (2.73) from Section 2.5 yields the bound

$$\|P_k\psi_L^{(j)}\|_{S[k](I_j\times\mathbb{R}^2)} \lesssim \varepsilon_1^{-1}\|P_k\psi\|_2$$

These estimates imply (7.50). □

Remark 7.28. Note that if we a priori work on a time interval I_j of infinite length, the statement of the Corollary may be strengthened to

$$\|\nabla_{x,t}\psi_L^{(j)}\|_{L_t^\infty \dot H^{-1}} \lesssim E$$

with universal implied constant. Indeed, in this case, the 'time averaging' around the initial data does not cost a large constant.

Later we shall need to following corollary which further specifies the Fourier support of ψ_L.

Corollary 7.29. *Let $\|\psi\|_S < C_0$. Assume that $\psi = \tilde\psi + \breve\psi$ where for some b*

$$\|\breve\psi\|_S + \|P_{(-\infty,b]^c}\tilde\psi\|_S < \delta_1$$

for some small δ_1. Then there exist intervals $\{I_j\}_{j=1}^{M_1}$ as in Corollary 7.27 so that on each I_j one has a decomposition

$$\psi|_{I_j} = \psi_L^{(j)} + \psi_{NL}^{(j)}, \quad \Box\psi_L^{(j)} = 0$$

where furthermore $\psi_L^{(j)} = \tilde\psi_L^{(j)} + \breve\psi_L^{(j)}$ and $\psi_{NL}^{(j)} = \tilde\psi_{NL}^{(j)} + \breve\psi_{NL}^{(j)}$ where both $\tilde\psi_L^{(j)}$ and $\breve\psi_L^{(j)}$ are free waves satisfying (7.50) *and both $\tilde\psi_{NL}^{(j)}$ and $\breve\psi_{NL}^{(j)}$ satisfy* (7.49)*. Furthermore,*

$$\|\breve\psi_L^{(j)}\|_S + \|P_{(-\infty,b]^c}\tilde\psi_L^{(j)}\|_S \lesssim \delta_1$$
$$\|\breve\psi_{NL}^{(j)}\|_S + \|P_{(-\infty,b]^c}\tilde\psi_{NL}^{(j)}\|_S \lesssim \delta_1$$

with an absolute implicit constant.

Proof. The proof of this statement follows the exact same lines as the proof of the previous corollary. The only difference is that each nonlinearity needs to be split into the contributions made by $\breve\psi$ and $\tilde\psi$, respectively. □

8. BMO, A_p, AND WEIGHTED COMMUTATOR ESTIMATES

In this section we develop some auxiliary tools that will be needed in the implementation of the Bahouri-Gerard theory for wave maps. More specifically, due to the lack of an imbedding from energy to L^∞ in the critical case we need to invoke methods involving BMO and the closely related A_p-classes in order to carry out Steps 1 and 2 of the program delineated in Section 1. Lemma 7.19 will play a crucial role here. Moreover, we require a weighted version of the Coifman-Meyer commutator theorem, with the weights belonging to the A_p-class. Although it does not seem to be widely known, it is an easy consequence of the standard theory and we sketch the proof for the sake of completeness. The paper [38] contains a more general form of this result. A Calderon-Zygmund kernel here is defined to be any linear operator T bounded on L^2 with the additional property that for any $f \in L^2$ with compact support and all $x \notin \mathrm{supp}(f)$,

$$Tf(x) = \int K(x,y)f(y)\,dy$$

where $|K(x,y)| \le C|x-y|^{-d}$ and for some $0 < \gamma \le 1$,

$$|K(x,y) - K(x',y)| \le C\frac{|x-x'|^\gamma}{|x-y|^{d+\gamma}} \quad \forall\, |x-y| > 2|x-x'|$$
$$|K(x,y) - K(x,y')| \le C\frac{|y-y'|^\gamma}{|x-y|^{d+\gamma}} \quad \forall\, |x-y| > 2|y-y'|$$

By the Calderon-Zygmund theorem, any such T is also bounded on $L^p(\mathbb{R}^d)$ provided $1 < p < \infty$.

Lemma 8.1. *Let $1 < p < \infty$. There exists $\delta = \delta(p) > 0$ with the following property: suppose $\phi = \phi_0 + \phi_1$ where $\|\phi_0\|_{\mathrm{BMO}(\mathbb{R}^d)} < \delta$ and $\|\phi_1\|_{L^\infty(\mathbb{R}^d)} \le A$. Then*

$$\|e^{-\phi}[T,b]e^{\phi}\|_{p\to p} \le C(d,A,T,p)\|b\|_{\mathrm{BMO}} \tag{8.1}$$

for any Calderon-Zygmund operator T and $b \in \mathrm{BMO}$. Moreover, $\inf_{p\in I}\delta(p) > 0$ and $\sup_{p\in I} C(d,A,T,p) < \infty$ for any compact $I \subset (1,\infty)$.

Proof. Since ϕ_1 contributes at most e^{2A} to the estimate, we can assume that $\phi = \phi_0$ with small BMO norm. In particular, $e^{\phi_0} \in A_p$. We will require the following inequality involving the so-called sharp maximal function $M^\sharp f$ which is defined as

$$(M^\sharp f)(x) = \sup_{Q: x\in Q}\inf_c |Q|^{-1}\int_Q \big|f(y) - c\big|\,dy$$

where c is a constant. The optimal choice of c is $c = f_Q := |Q|^{-1}\int_Q f(y)\,dy$. The estimate then reads (see Theorem 7.10 in [7])

$$\int_{\mathbb{R}^d}(Mf)^p(x)\,w(x)\,dx \le C\int_{\mathbb{R}^d}(M^\sharp f)^p(x)\,w(x)\,dx \tag{8.2}$$

for any $w \in A_\infty$ with a constant that only depends on the dimension and the constants in (7.26). To avoid trivialities like $f = const$ for which (8.2) fails, one needs to assume $Mf \in L^{p_0}(\mathbb{R}^d)$ for some $1 \le p_0 \le p$.

The proof of (8.1) combines the standard proof of the unweighted Coifman-Meyer bound with the sharp function estimate (8.2). More precisely, fix a cube Q and write

$$\begin{aligned}[T,b]f &= -(b-b_Q)Tf + T((b-b_Q)\chi_{2Q}f) + T((b-b_Q)\chi_{\mathbb{R}^d\setminus 2Q}f)\\ &=: A_Q + B_Q + C_Q\end{aligned}$$

To bound $M^\sharp([T,b]f)$, we simply note that for any $x \in Q$, and any $1 < s < \infty$,

$$|Q|^{-1}\int_Q (|A_Q(y)| + |B_Q(y)|)\,dy \le C(s,d,T)\|b\|_{\mathrm{BMO}}\big((M|Tf|^s)^{\frac1s}(x) + (M|f|^s)^{\frac1s}(x)\big)$$

Indeed, for A this follows from Hölder's inequality and the definition of BMO, whereas for B we also invoke the L^q boundedness of T for some $1 < q < s$. For C_Q we let y_Q be the center of Q and estimate for any

$y \in Q$,

$$
\begin{aligned}
|C_Q(y) - C_Q(y_Q)| &\le \int_{\mathbb{R}^d \setminus 2Q} |K(y,z) - K(y_Q,z)||(b-b_Q)(z)||f(z)|\,dz \\
&\le C \int_{\mathbb{R}^d \setminus 2Q} \frac{|y-y_Q|^\gamma}{|z-y_Q|^{d+\gamma}} |(b-b_Q)(z)||f(z)|\,dz \\
&\le C\|b\|_{\mathrm{BMO}} \inf_{x\in Q} (M|f|^s)^{\frac{1}{s}}(x)
\end{aligned}
$$

where $\gamma > 0$ is as above.
In conclusion,

$$M^\sharp([T,b]f) \le C(s,d,T)\|b\|_{\mathrm{BMO}}\big((M|Tf|^s)^{\frac{1}{s}} + (M|f|^s)^{\frac{1}{s}}\big)$$

The lemma follows from (8.2) and the weighted L^p boundedness of M and T. □

We now apply this to prove the following lemma, which will be important in the implementation of the Bahouri-Gerard decomposition for wave maps. Instead of a general Calderon-Zygmund operator, we restrict ourselves to the subclass of Mikhlin multiplier operators which are of the form $Tf = (m\hat{f})^\vee$ with $m \in C^3(\mathbb{R}^2 \setminus \{0\})$ and with

$$|D^\alpha m(\xi)| \le C(\alpha)|\xi|^{-|\alpha|} \quad \forall\, \xi \in \mathbb{R}^2 \setminus \{0\}$$

for all $|\alpha| \le 3$. For simplicity, we also limit ourselves to two dimensions.

Lemma 8.2. *Suppose $\{\varphi_n\}_{n=1}^\infty, \{\phi_n\}_{n=1}^\infty$ lie in the unit-ball of L^2. Furthermore, assume that*

$$\operatorname{supp}(\widehat{\varphi_n}),\ \operatorname{supp}(\widehat{\phi_n}) \subset \{\xi \in \mathbb{R}^2 \,:\, 2^{k_0} \le |\xi| \le 2^{k_1}\}$$

for arbitrary $k_0 < k_1 - 4$ and let $v_n := e^{(-\Delta)^{-\frac{1}{2}}\varphi_n}$. Then

$$\|P_j(v_n^{-1}T(\phi_n v_n))\|_2 \lesssim \min(2^{k_1-j}, 2^{\frac{j-k_0}{3}})$$

provided either $j \le k_0$ or $j \ge k_1$.

Proof. By Lemma 7.20, for any $1 < p < \infty$ one has $\sup_n \sup_{|t|\le 1} A_p(v_n^t) \le C(p)$. Set $R = 2^{k_1}$ and $r = 2^{k_0}$. If $j > k_1$, then

$$
\begin{aligned}
\|P_j(v_n^{-1}T(\phi_n v_n))\|_2 &\lesssim 2^{-j}\|\nabla(v_n^{-1}T(\phi_n v_n))\|_2 \\
&\lesssim 2^{-j}(\|\phi_n\|_4\|\varphi_n\|_4 + \|\nabla\phi_n\|_2) \lesssim 2^{-j}R
\end{aligned}
$$

On the other hand, if $j < k_0$, then

$$
\begin{aligned}
\|P_j(v_n^{-1}T(\phi_n v_n))\|_2 &\lesssim \int_0^1 \|P_j(v_n^{-t}[T,(-\Delta)^{-\frac{1}{2}}\varphi_n](\phi_n v_n^t))\|_2\,dt \\
&\lesssim 2^{\frac{j}{3}} \int_0^1 \|P_j(v_n^{-t}[T,(-\Delta)^{-\frac{1}{2}}\varphi_n](\phi_n v_n^t))\|_{\frac{3}{2}}\,dt \\
&\lesssim 2^{\frac{j}{3}}\|(-\Delta)^{-\frac{1}{2}}\varphi_n\|_6\|\phi_n\|_2 \lesssim 2^{\frac{j}{3}} r^{-\frac{1}{3}}
\end{aligned}
$$

In the last line, one interpolates between $\|(-\Delta)^{-\frac{1}{2}}\varphi_n\|_2 \lesssim r^{-1}$ and $\|(-\Delta)^{-\frac{1}{2}}\varphi_n\|_{\mathrm{BMO}} \lesssim 1$. □

The following result allows us to strip away weights from $T(\phi)$ provided they result from functions with frequencies which are well-separated from the Fourier support of ϕ. In what follows, we use the following terminology from [1]: Given a bounded sequence $\underline{f} := \{f_n\}_{n\ge1} \subset L^2$, and sequence $\underline{\varepsilon} := \{\varepsilon_n\}_{n\ge1} \subset R^+$, we say that $\underline{f}$ *is $\underline{\varepsilon}$-oscillatory* iff

$$\lim_{R\to\infty} \limsup_{n\to\infty} \int_{[|\xi|\varepsilon_n \in (0,\infty)\setminus(R^{-1},R)]} |f_n(\xi)|^2\,d\xi = 0$$

We say that $\underline{f}$ *is $\underline{\varepsilon}$-singular* iff

$$\limsup_{n\to\infty} \int_{[|\xi|\varepsilon_n \in (a,b)]} |f_n(\xi)|^2\,d\xi = 0$$

for all $b > a > 0$. In what follows, we shall freely use the scale selection algorithm from Section III.1 from [1], see in particular Lemma 3.1, Lemma 3.2 part (iii), and Proposition 3.4 in that section.

Lemma 8.3. *Suppose both $\{\varphi_n\}_{n=1}^\infty \subset L^2(\mathbb{R}^2)$ and $\{\phi_n\}_{n=1}^\infty \subset L^2(\mathbb{R}^2)$ are 1-oscillatory, whereas $\{\psi_n\}_{n=1}^\infty \subset L^2(\mathbb{R}^2)$ is 1-singular. Define*

$$v_n := \exp((-\Delta)^{-\frac{1}{2}}\varphi_n), \quad w_n := \exp((-\Delta)^{-\frac{1}{2}}\psi_n)$$

Then

$$(v_n w_n)^{-1} T(\phi_n\, v_n w_n) = v_n^{-1} T(\phi_n\, v_n) + o_{L^2}(1) \tag{8.3}$$

as $n \to \infty$. Moreover, $v_n^{-1} T(\phi_n\, v_n)$ is 1-oscillatory[19].

Proof. By assumption,

$$\|\phi_n\|_2 + \|\varphi_n\|_2 + \|\psi_n\|_2 \le A < \infty$$

for all $n \ge 1$. By Lemma 7.20 one has $v_n \in A_p$ and $v_n w_n \in A_p$ for all $1 < p < \infty$ with A_p constants depending only on A and p. Now fix $\varepsilon > 0$ arbitrarily small. Then there is $R > 1$ so that

$$\limsup_{n\to\infty} \int\limits_{[|\xi|<R^{-1},\ |\xi|>R]} |\hat{\varphi}_n(\xi)|^2\, d\xi < \varepsilon^2$$

Fix an $R = R(\varepsilon)$ with this property. Define $\varphi_{1n} := (\chi_{[R^{-1},R]}\widehat{\varphi_n})^\vee$, $\varphi_{2n} := \varphi_n - \varphi_{1n}$ and $\phi_{1n} := (\chi_{[R^{-1},R]}\widehat{\phi_n})^\vee$, $\phi_{2n} := \phi_n - \phi_{1n}$. Then $\|\varphi_{2n}\|_2 + \|\phi_{2n}\|_2 < \varepsilon$ for large n whence

$$\|(v_n w_n)^{-1} T(\phi_{2n}\, v_n w_n)\|_2 + \|v_n^{-1} T(\phi_{2n}\, v_n)\|_2 \le C(A,T)\varepsilon$$

as well as $\|(-\Delta)^{-\frac{1}{2}}\varphi_{2n}\|_{\mathrm{BMO}} < C\varepsilon$. Next, define

$$v_{jn} := \exp((-\Delta)^{-\frac{1}{2}}\varphi_{jn}) \qquad j = 1,2$$

By Lemmas 7.20 and 8.1,

$$\begin{aligned}
&\|(v_n w_n)^{-1} T(\phi_{1n}\, v_n w_n) - (v_{1n} w_n)^{-1} T(\phi_{1n}\, v_{1n} w_n)\|_2 \\
&\le \int_0^1 \|(w_n v_{1n} v_{2n}^t)^{-1} [T, (-\Delta)^{-\frac{1}{2}}\varphi_{2n}](\phi_{1n}\, w_n v_{1n} v_{2n}^t)\|_2\, dt \\
&\le C(A,T)\|(-\Delta)^{-\frac{1}{2}}\varphi_{2n}\|_{\mathrm{BMO}} \le C(A,T)\varepsilon
\end{aligned}$$

By the same argument,

$$\|v_n^{-1} T(\phi_{1n}\, v_n) - v_{1n}^{-1} T(\phi_{1n}\, v_{1n})\|_2 \le C(A,T)\varepsilon$$

Similarly, set

$$\psi_{2n} := (\chi_{[\rho^{-1},\rho]}\widehat{\psi_n})^\vee, \quad \psi_{1n} := \psi_n - \psi_{2n}$$

where $\rho > 1$ will be determined later. By assumption, $\|\psi_{2n}\|_2 \to 0$ as $n \to \infty$. In particular,

$$\|(-\Delta)^{-\frac{1}{2}}\psi_{2n}\|_{\mathrm{BMO}} \to 0$$

as $n \to \infty$. Applying Lemma 8.1 as before allows one to remove the weights w_{2n} from (8.3) where

$$w_{jn} := \exp((-\Delta)^{-\frac{1}{2}}\psi_{jn}) \qquad j = 1,2$$

Hence, we are reduced to establishing that

$$\|(v_{1n} w_{1n})^{-1} T(\phi_{1n}\, v_{1n} w_{1n}) - v_{1n}^{-1} T(\phi_{1n}\, v_{1n})\|_2 \le C(A,T)\varepsilon \tag{8.4}$$

for sufficiently large n. For ease of notation, we shall now drop the subscript 1 from φ_{1n} etc. with the understanding that $\widehat{\varphi_n}$ and $\widehat{\phi_n}$ are supported on $[R^{-1}, R]$ and that $\widehat{\psi_n}$ is supported off $[\rho^{-1}, \rho]$ where $\rho > 1$ is a large number depending on ε to be chosen later. Define

$$\psi_{n,\mathrm{low}} := (\chi_{(0,\rho^{-1}]}\widehat{\psi_n})^\vee, \quad \psi_{n,\mathrm{high}} := (\chi_{[\rho,\infty)}\widehat{\psi_n})^\vee$$

---

[19]Note that neither $(v_n w_n)^{-1} T(\psi_n\, v_n w_n)$ nor $v_n^{-1} T(\psi_n\, v_n)$ are in general 1-singular.

and write, correspondingly, $w_n = w_{n,\mathrm{low}} w_{n,\mathrm{high}}$. It is easy to remove $w_{n,\mathrm{high}}$:

$$\begin{aligned}
&\|(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - (v_n w_{n,\mathrm{low}})^{-1} T(\phi_n\, v_n w_{n,\mathrm{low}})\|_2 \\
&\le \int_0^1 \|(v_n w_{n,\mathrm{low}} w^t_{n,\mathrm{high}})^{-1} [T, (-\Delta)^{-\frac12}\psi_{n,\mathrm{high}}](\phi_n\, v_n w_{n,\mathrm{low}} w^t_{n,\mathrm{high}})\|_2\, dt \\
&\le C(A,T)\|(-\Delta)^{-\frac12}\psi_{n,\mathrm{high}}\|_4 \|\phi_n\|_4 \qquad (8.5)\\
&\le C(A,T,R)\rho^{-\frac12} \le \varepsilon
\end{aligned}$$

provided ρ is sufficiently large. Here we used that $v_n w_{n,\mathrm{low}} w^t_{n,\mathrm{high}}$ are A_2 weights uniformly in $0 \le t \le 1$ as well as an interpolation between L^2 and BMO to pass to the last line. For the final bound we need $\rho \gg \varepsilon^{-2}$.

To remove $w_{n,\mathrm{low}}$ we split

$$T = P_{<-\lambda} T + P_{-\lambda<\cdot<\lambda} T + P_{>\lambda} T$$

where $2^{-\lambda} R \ll \varepsilon$ and $P_{<\lambda}$ etc. denote Littlewood-Paley projections. Introducing an angular decomposition into finitely many sectors, we may assume that $|\xi_1| \ge |\xi|/10$ on the support of m. Then for large λ, and with $\mu := 2^{-\lambda}$,

$$\begin{aligned}
&\|(v_n w_{n,\mathrm{low}})^{-1} P_{>\lambda} T(\phi_n\, v_n w_{n,\mathrm{low}})\|_2 \le C\|(v_n w_{n,\mathrm{low}})^{-1} \partial_1^{-1} P_{>\lambda} T(\partial_1[\phi_n\, v_n w_{n,\mathrm{low}}])\|_2 \\
&\le C(A)\mu\big(\|\partial_1(\phi_n v_n)\|_2 + \|\phi_n \partial_1 (-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}\|_2\big) \le \varepsilon
\end{aligned}$$

For the small frequencies $P_{<-\lambda}T$ we first recall the following standard fact: with ψ a suitable Schwarz function,

$$\begin{aligned}
(P_{<-\lambda}(fg) - gP_{<-\lambda} f)(x) &= -\sum_{j=1}^2 \int_0^1 \int_{\mathbb{R}^2} \mu^2 \psi(\mu y) y_j f(x-y) \partial_j g(x-sy)\, dyds \\
&= \sum_{j=1}^2 L_{j,\lambda}(\mu^{-1} f, \partial_j g) \qquad (8.6)
\end{aligned}$$

where $L_{j,\lambda}$ in the final line denotes a multi-linear expression of the form

$$L(f,g)(x) = \int_{\mathbb{R}^4} f(x-u) g(x-v)\, \nu(du,dv) \qquad (8.7)$$

with a measure ν of mass bounded by some constant (in this case uniformly in all parameters). Using this notation, one has (since $\|v_n^{-1}\|_\infty \le C(A,R)$)

$$\begin{aligned}
\|(v_n w_{n,\mathrm{low}})^{-1} P_{<-\lambda} T(\phi_n\, v_n w_{n,\mathrm{low}})\|_2 &\le C(A,R)\|w^{-1}_{n,\mathrm{low}} T(P_{<-\lambda}(\phi_n\, v_n) w_{n,\mathrm{low}})\|_2 \\
&\quad + C(A,R)\mu^{-1} \sum_{j=1}^2 \|w^{-1}_{n,\mathrm{low}} T(L_{j,\lambda}(\phi_n v_n, w_{n,\mathrm{low}} \partial_j (-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}))\|_2 \\
&=: I_n + II_n
\end{aligned}$$

To bound I_n, note that since we may take $\mu \le R^{-1}$, one has $P_{<-\lambda}(\phi_n v_n) = P_{<-\lambda}(\phi_n(v_n-1))$. Hence, by the boundedness of T relative to the weight $w_{n,\mathrm{low}}$ and Bernstein's inequality,

$$\begin{aligned}
\|I_n\|_2 &\le \|P_{<-\lambda}(\phi_n v_n)\|_2 = \|P_{<-\lambda}(\phi_n(v_n-1))\|_2 \le C\mu \|P_{<-\lambda}(\phi_n(v_n-1)\|_1 \\
&\le C\mu \|\phi_n\|_2 \|v_n - 1\|_2 \le \varepsilon
\end{aligned}$$

for $\lambda \gg 1$. Bounding II_n requires more care, as one cannot naively remove the weights as we did in the bound for I_n. In fact, due to the translation by sy in (8.6) we are in a situation with *two different weights*, namely $w_{n,\mathrm{low}}$ and $w_{n,\mathrm{low}}(\cdot - sy)$. This means that the usual A_p weight theory applied to II_n does not simply cancel the operator T and the two weights, but rather cancels the outer weight, the operator T and the weight inside T is replaced by an expression of the form

$$w_{n,\mathrm{low}}(\cdot - sy) w^{-1}_{n,\mathrm{low}}(\cdot)$$

Thus, in view of (8.6) and these considerations, II_n is bounded by (using $\|v_n\|_\infty \le C(A,R)$)

$$C(A,R)\mu^{-2}\int_{\mathbb{R}^2}\int_0^1 \mu^2|\psi(\mu y)|\,|\mu y|\big\|\phi_n(\cdot-y)w_{n,\mathrm{low}}(\cdot-sy)w_{n,\mathrm{low}}^{-1}(\cdot)\nabla(-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}(\cdot-sy)\big\|_{L^2}\,dyds$$

$$\le C(A,R)\mu^{-2}\|\phi_n\|_2\|\nabla(-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}\|_\infty\int_{\mathbb{R}^2}\int_0^1 \mu^2|\psi(\mu y)|\,|\mu y|\big\|w_{n,\mathrm{low}}(\cdot-sy)w_{n,\mathrm{low}}^{-1}(\cdot)\big\|_{L^\infty}\,dyds$$

$$\le C(A,R)\mu^{-2}\rho^{-1}\int_{\mathbb{R}^2}\mu^2(1+\mu|y|)^{-N}\,(1+|y|\mu)^{k(A)}\,dy \le \varepsilon \tag{8.8}$$

for some constant $k(A)>0$ provided we choose ρ such that $\mu^{-2}\varepsilon^{-1}\ll\rho$. To pass to the bound in (8.8), assume first that $|y|\mu\le 1$. Then with $h_n := (-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}$ so that $w_{n,\mathrm{low}} = e^{h_n}$,

$$w_{n,\mathrm{low}}(x-y)w_{n,\mathrm{low}}^{-1}(x) \le \exp\big(|y|\|\nabla h_n\|_\infty\big) \le \exp\big(C\mu^{-1}\rho^{-1}\big) \le e^{C\varepsilon} \le 2$$

where we used that

$$\|\nabla h_n\|_\infty \le \|\nabla(-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}\|_\infty \le C\rho^{-1}$$

This implies that on scales $\le \mu^{-1}$, the weight $w_{n,\mathrm{low}}$ is essentially constant (up to multiplicative constants). Next, observe that for all cubes

$$|(h_n)_Q - (h_n)_{2^\ell Q}| \le C\|h_n\|_{\mathrm{BMO}}\ell \le C(A)\ell \qquad \forall\ \ell\ge 0$$

Hence, partitioning $\mathbb{R}^2$ into cubes of side-length μ^{-1} one obtains that

$$|h_n(y)-h_n(y')| \le C(A)\log(2+|y-y'|\mu) \tag{8.9}$$

whence

$$\sup_x w_{n,\mathrm{low}}(x-y)w_{n,\mathrm{low}}^{-1}(x) \le C(A)(1+|y|\mu)^{k(A)}$$

as claimed.

Note that the previous estimates on $P_{<-\lambda}T$ and $P_{>\lambda}T$ also prove that

$$\|v_n^{-1}P_{<-\lambda}T(\phi_n v_n)\|_2 + \|v_n^{-1}P_{>\lambda}T(\phi_n v_n)\|_2 \le \varepsilon$$

Therefore, it remains to prove that

$$\|(v_n w_{n,\mathrm{low}})^{-1}T_\lambda(\phi_n\, v_n w_{n,\mathrm{low}}) - v_n^{-1}T_\lambda(\phi_n\, v_n)\|_2 \le \varepsilon$$

where

$$T_\lambda := P_{-\lambda<\cdot<\lambda}T \tag{8.10}$$

is the operator on intermediate frequencies. Since $Tf = (m\hat f)^\vee$ with $m\in C^3(\mathbb{R}^2\setminus\{0\})$, we conclude that

$$P_{\lambda<\cdot<\lambda^{-1}}Tf(x) = \int K_\lambda(x-y)f(y)\,dy$$

with $|K_\lambda(x)| \le C(\lambda)(1+|x|)^{-3}$. Now, with $h_n = (-\Delta)^{-\frac12}\psi_{n,\mathrm{low}}$ as above, and M denoting the Hardy-Littlewood maximal operator,

$$\begin{aligned}&\|(v_n w_{n,\mathrm{low}})^{-1}T_\lambda(\phi_n\, v_n w_{n,\mathrm{low}}) - v_n^{-1}T_\lambda(\phi_n\, v_n)\|_2\\ &\le \int_0^1 \|(v_n w^t_{n,\mathrm{low}})^{-1}[T_\lambda,h_n](\phi_n\, v_n w^t_{n,\mathrm{low}})\|_2\,dt\\ &\le C(A,\lambda)\rho^{-\frac14}\int_0^1 \|(v_n w^t_{n,\mathrm{low}})^{-1}M(\phi_n\, v_n w^t_{n,\mathrm{low}})\|_2\,dt \le C(A,\lambda)\rho^{-\frac14}\end{aligned}$$

Here we used that the kernel of $[T_\lambda,h_n]$ is of the form $K_\lambda(x,y)(h_n(x)-h_n(y))$ and satisfies the bounds, cf. (8.9),

$$|K_\lambda(x,y)(h_n(x)-h_n(y))| \le C(A,\lambda)\min(\rho^{-1}|x-y|, |x-y|^{-3}\log(2+|x-y|))$$

whence

$$|[T_\lambda,h_n]f(x)| \le C(A,\lambda)\rho^{-\frac14}Mf(x)$$

Taking ρ sufficiently large (depending on ε, R, and A) finishes the proof of (8.3). Lemma 8.2 now implies that $v_n^{-1}T(\phi_n v_n)$ is 1-oscillatory. □

The following statement will be an essential technical tool for the Bahouri-Gerard method in the context of wave maps into hyperbolic space. As before, T is a Mikhlin multiplier operator.

Corollary 8.4. *Let* $\{f_n\}_{n=1}^\infty \subset L^2(\mathbb{R}^2)$ *satisfy* $\sup_{n\geq 1}\|f_n\|_2 \leq A < \infty$ *and define* $y_n = \exp\big((-\Delta)^{-\frac{1}{2}} f_n\big)$. *Let* $\Lambda_j := \big\{\lambda_{n,j}\big\}_{n=1}^\infty$ *be sequences of positive numbers for each* $1 \leq j \leq J$ *with the property that*

$$\lim_{n\to\infty}\Big\{\frac{\lambda_{n,j}}{\lambda_{n,j'}} + \frac{\lambda_{n,j'}}{\lambda_{n,j}}\Big\} \to \infty \tag{8.11}$$

for any $1 \leq j \neq j' \leq J$. *Assume further that*

$$f_n = \sum_{j=1}^J \varphi_{n,j} + \omega_n$$

where $\{\varphi_{n,j}\}_{n=1}^\infty \subset L^2(\mathbb{R}^2)$ *is* Λ_j*-oscillatory for each* $1 \leq j \leq J$, $\{\omega_n\}_{n=1}^\infty$ *is* Λ_j*-singular for every* $1 \leq j \leq J$, *and* $\sup_{n\geq 1}\|\omega_n\|_{\dot{B}^0_{2,\infty}} < \delta$.
Then $\big\{y_n^{-1}\, T(\varphi_{n,j}\, y_n)\big\}_{n=1}^\infty$ *is* Λ_j*-oscillatory,* $\big\{y_n^{-1}\, T(\omega_n\, y_n)\big\}_{n=1}^\infty$ *is* Λ_j*-singular for each* $1 \leq j \leq J$, *and*

$$\limsup_{n\to\infty}\|y_n^{-1}\, T(\omega_n y_n)\|_{\dot{B}^0_{2,\infty}} < C(A,T)\delta \tag{8.12}$$

where the constant $C(A,T)$ *only depends on* A *and* T.

Proof. Define

$$v_{n,j} := \exp((-\Delta)^{-\frac{1}{2}}\varphi_{n,j}), \qquad w_n := \exp((-\Delta)^{-\frac{1}{2}}\omega_n)$$

so that $y_n := w_n \prod_{j=1}^J v_{n,j}$. By Lemma 8.3, both $\big\{v_{n,j}^{-1}\, T(\varphi_{n,j} v_{n,j})\big\}_{n=1}^\infty$ and $\big\{y_n^{-1}\, T(\varphi_{n,j}\, y_n)\big\}_{n=1}^\infty$ are Λ_j-oscillatory. Now suppose $\{\psi_n\}_{n=1}^\infty$ is an arbitrary Λ_j-oscillatory sequence where $1 \leq j \leq J$ is fixed. Then $\widetilde{\omega}_n := y_n^{-1}\, T(\omega_n y_n)$ satisfies

$$\langle\widetilde{\omega}_n, \psi_n\rangle = \langle\omega_n, y_n T^*(\psi_n y_n^{-1})\rangle$$

By Lemma 8.3, $\{y_n T^*(\psi_n y_n^{-1})\}_{n=1}^\infty$ is Λ_j-oscillatory whence

$$\lim_{n\to\infty}\langle\widetilde{\omega}_n, \psi_n\rangle = 0$$

Therefore, $\big\{\widetilde{\omega}_n\big\}_{n=1}^\infty$ is Λ_j-singular for each $1 \leq j \leq J$.

For the proof of (8.12), we first note that passing to a subsequence if necessary, (8.11) implies that we may assume that

$$\lambda_{n,1} > \lambda_{n,2} > \ldots > \lambda_{n,J}$$

for all large n whence for any $1 \leq j \leq J-1$

$$\frac{\lambda_{n,j}}{\lambda_{n,j+1}} \to \infty \tag{8.13}$$

as $n \to \infty$. We also note that

$$\sum_{j=1}^J \|\varphi_{n,j}\|_2^2 + \|w_n\|_2^2 \leq A^2 + o(1)$$

as $n \to \infty$. Now we let $m \geq 10$ and $K \geq 10$ be integers (to be determined later) and define

$$\begin{aligned}
\widetilde{\varphi}_{n,j} &:= \varphi_{n,j}\, \chi_{[2^{-m}\leq|\nabla|\lambda_{n,j}\leq 2^m]}\\
\widetilde{y}_n &:= \prod_{j=1}^J \exp\big((-\Delta)^{-\frac{1}{2}}\widetilde{\varphi}_{n,j}\big)\\
\widetilde{\omega}_n &:= \omega_n\, \chi_{\mathbb{R}^2\setminus\bigcup_{j=1}^J[2^{-Km}\leq|\nabla|\lambda_{n,j}\leq 2^{Km}]}\\
\widetilde{w}_n &:= \exp((-\Delta)^{-\frac{1}{2}}\widetilde{\omega}_n)
\end{aligned}$$

where the multipliers involving ∇ need to be interpreted on the Fourier side. As in the proof of Lemma 8.3,

$$\limsup_{n\to\infty}\|y_n^{-1}T(w_n y_n) - \widetilde{y}_n^{-1}T(\tilde{w}_n\, \tilde{y}_n)\|_2 < C(A,T)\,\delta \tag{8.14}$$

provided m is chosen large enough and irrespective of the choice of $K \geq 1$. We will now fix m so that (8.14) holds. It therefore suffices to show that

$$\limsup_{n\to\infty} \sup_{j\in\mathbb{Z}} \|P_j\, [\tilde{y}_n^{-1} T(\tilde{w}_n\, \tilde{y}_n)]\|_2 \leq C(A,T)\,\delta \tag{8.15}$$

provided K is chosen sufficiently large. The idea behind (8.15) is that $\widetilde{w}_n$ behaves like a lacunary series, i.e., each $\widetilde{w}_n$ is the sum of functions whose Fourier supports consist of disjoint blocks which are very strongly separated. In addition, the $\widetilde{\varphi}_{n,j}$ are Fourier supported on intervals which are well separated from the Fourier support of $\widetilde{w}_n$. It will turn out that for each j – up to negligible errors as $K\to\infty$ – only one block of frequencies from $\widetilde{w}_n$ (namely the one containing 2^j) contributes to $P_j\,[\tilde{y}_n^{-1}T(\tilde{w}_n\,\tilde{y}_n)]$ and, moreover, only those $\widetilde{\varphi}_{n,j}$ with frequencies much smaller than 2^j matter. In this way, we can then essentially pass P_j onto $\tilde{w}_n$.

To establish (8.15), we introduce some more notation: set

$$\psi_n := \sum_{j=1}^{J} (-\Delta)^{-\frac12}\,\widetilde{\varphi}_{n,j}$$

and define $[T,\psi_n]^{(s)}$ iteratively via

$$[T,\psi_n]^{(1)} := [T,\psi_n], \qquad [T,\psi_n]^{(s+1)} = [[T,\psi_n]^{(s)},\psi_n]$$

Then

$$\tilde{y}_n^{-1}T(\tilde{w}_n\tilde{y}_n) = \sum_{\ell=0}^{s} \frac{1}{\ell!}[T,\psi_n]^{(\ell)}\,\tilde{w}_n + \tag{8.16}$$

$$+\frac{1}{s!}\int_0^1 (1-t)^s\, \tilde{y}_n^{-t}[T,\psi_n]^{(s+1)}(\tilde{w}_n\,\tilde{y}_n^{t})\,dt \tag{8.17}$$

Denote the remainder in (8.17) by $R_{n,s}$. To bound it in L^2, note that $\|\psi_n\|_\infty \leq CmJA$ with some absolute constant C. Therefore, placing ψ_n, $\tilde{y}_n$, and $\tilde{y}_n^{-1}$ in L^∞ yields for all $n\geq 1$

$$\|R_{n,s}\|_2 \leq \frac{e^{CmJA}}{(s+1)!}(CmJA)^{s+1}\|\tilde{w}_n\|_2 \leq e^{CmJA}\Big(\frac{CmJA}{s}\Big)^s =: \gamma_s$$

which clearly goes to zero as $s\to\infty$. In particular, $\gamma_s \leq \delta$ for large s. We now turn to the details of the analysis of the main terms in (8.16). First, one has $\tilde{w}_n = \sum_{j=0}^J \widetilde{w}_{n,j}$ where

$$\begin{aligned}
\widetilde{w}_{n,0} &:= \tilde{w}_n\,\chi_{[|\nabla|\,\lambda_{n,1}\leq 2^{-Km}]}\\
\widetilde{w}_{n,j} &:= \tilde{w}_n\,\chi_{[2^{Km}\leq|\nabla|\lambda_{n,j}\leq 2^{-Km}]} \quad \forall\, 1\leq j\leq J-1\\
\widetilde{w}_{n,J} &:= \tilde{w}_n\,\chi_{[2^{Km}\leq|\nabla|\lambda_{n,J}]}
\end{aligned}$$

with n large. Then

$$\sum_{\ell=0}^{s}\frac{1}{\ell!}[T,\psi_n]^{(\ell)}\,\tilde{w}_n = \sum_{j=0}^{J}\sum_{\ell=0}^{s}\frac{1}{\ell!}[T,\psi_n]^{(\ell)}\,\widetilde{w}_{n,j} = \sum_{j=0}^{J}\sum_{\ell=0}^{s}\frac{1}{\ell!}[T,\psi_{n,j}^{(-)}+\psi_{n,j}^{(+)}]^{(\ell)}\,\widetilde{w}_{n,j} \tag{8.18}$$

where we have set, for each $0\leq j\leq J$,

$$\psi_{n,j}^{(-)} := \sum_{1\leq k\leq j}(-\Delta)^{-\frac12}\,\widetilde{\varphi}_{n,k}, \qquad \psi_{n,j}^{(+)} := \sum_{j<k\leq J}(-\Delta)^{-\frac12}\,\widetilde{\varphi}_{n,k}$$

We shall now show that for a given $\widetilde{w}_{n,j}$ only the small frequency part of $\psi_{n,j}$, i.e., $\psi_{n,j}^{(-)}$, contributes significantly to the commutators in (8.18) (at least for very large K). To this end write

$$\begin{aligned}
[T,\psi_{n,j}^{(-)}+\psi_{n,j}^{(+)}]^{(\ell)} &= [T,\psi_{n,j}^{(-)}]^{(\ell)} + \sum_{\varepsilon}[[\cdots[T,\psi_{n,j}^{(\varepsilon_1)}],\ \psi_{n,j}^{(\varepsilon_2)}],\ \ldots,\psi_{n,j}^{(\varepsilon_\ell)}]\\
&= \mathcal{K}_{n,j,\ell}+\mathcal{R}_{n,j,\ell} \qquad (8.19)
\end{aligned}$$

where the sum here runs over ℓ-fold commutators with each $\varepsilon_k = \pm$, the choice $\varepsilon_k = -$ for all $1 \le k \le \ell$ being excluded (as it is represented by the first – and main – term on the right-hand side). Next, observe that for each

$$1 \le \ell \le s := \frac{1}{100} 2^{(K-1)m} \tag{8.20}$$

one has, for each $1 \le j \le J$ and every $k \in \mathbb{Z}$,

$$P_k\big(\mathcal{K}_{n,j,\ell}\widetilde{w}_{n,j}\big) = P_k\mathcal{K}_{n,j,\ell}P_{k-2<\cdot<k+2}\,\widetilde{w}_{n,j}$$

In fact, this vanishes unless $2^{Km-2}\lambda_{n,j}^{-1} \le 2^k \le 2^{-Km+2}\,\lambda_{n,j+1}^{-1}$. Writing

$$P_k \sum_{\ell=0}^{s} \frac{1}{\ell!}[T,\psi_n]^{(\ell)}\,\widetilde{w}_n = P_k \sum_{j=0}^{J}\sum_{\ell=0}^{s} \frac{1}{\ell!}\mathcal{K}_{n,j,\ell}P_{k-2<\cdot<k+2}\,\widetilde{w}_{n,j} \tag{8.21}$$

$$+ P_k \sum_{j=0}^{J}\sum_{\ell=0}^{s} \frac{1}{\ell!}\mathcal{R}_{n,j,\ell}P_{k-2<\cdot<k+2}\,\widetilde{w}_{n,j} \tag{8.22}$$

it follows from (8.13) that for all sufficiently large $n \ge n_0$ depending on K, m, at most one term in (8.21) can be nonzero for any choice of $k \in \mathbb{Z}$. Applying the decomposition (8.16) and (8.17) with $\psi_{n,j}^{(-)}$ instead of ψ_n to (8.21) yields

$$\begin{aligned}
&\sup_{k\in\mathbb{Z}}\Big\|P_k \sum_{j=0}^{J}\sum_{\ell=0}^{s}\frac{1}{\ell!}\mathcal{K}_{n,j,\ell}P_{k-2<\cdot<k+2}\,\widetilde{w}_{n,j}\Big\|_2 \le \sup_{k\in\mathbb{Z}}\sup_{0\le j\le J}\Big\|\sum_{\ell=0}^{s}\frac{1}{\ell!}\mathcal{K}_{n,j,\ell}P_{k-2<\cdot<k+2}\,\widetilde{w}_{n,j}\Big\|_2\\
&\le \sup_{k\in\mathbb{Z}}\sup_{0\le j\le J}\Big\|e^{-\psi_{j,n}^{(-)}}T\big(e^{\psi_{j,n}^{(-)}}P_{k-2<\cdot<k+2}\,\widetilde{w}_{n,j}\big)\Big\|_2 + \gamma_s\\
&\le C(A,T)\|w_n\|_{\dot B^0_{2,\infty}} + \gamma_s \le C(A,T)\delta
\end{aligned}$$

To pass to the final bound, we note that $\gamma_s \le \delta$ provided K is chosen sufficiently large. We also used that the weights $e^{\psi_{j,n}^{(-)}} \in A_2$ with A_2 constant $\le CA$ uniformly in j,n, cf. Lemma 8.1. As for (8.22), we make the following crude estimate for the ℓ-fold commutator as in (8.19)

$$\big\|[[\cdots[T,\psi_{n,j}^{(\varepsilon_1)}],\ \psi_{n,j}^{(\varepsilon_2)}],\ \ldots,\psi_{n,j}^{(\varepsilon_\ell)}]\widetilde{w}_{n,j}\|_2 \le C(T)(CmJA)^{\ell-1}\|\psi_{n,j}^{(+)}\|_4\|\widetilde{w}_{n,j}\|_4$$

It arises by placing one $\psi_{n,j}^{(+)}$ in L^4, all other $\psi_{n,j}^{(\varepsilon_i)}$ in L^∞, and $\tilde{w}_{n,j}$ in L^4. By Bernstein's inequality,

$$\|\widetilde{w}_{n,j}\|_4 \le C(2^{-Km}\lambda_{n,j+1})^{-\frac12}\|\widetilde{w}_{n,j}\|_2 \le CA2^{-Km/2}\lambda_{n,j+1}^{-\frac12}$$

whereas by interpolation between the L^2 and BMO bounds,

$$\|\psi_{n,j}^{(+)}\|_4 \le CA2^{m/2}\lambda_{n,j+1}^{\frac12}$$

whence

$$\big\|[[\cdots[T,\psi_{n,j}^{(\varepsilon_1)}],\ \psi_{n,j}^{(\varepsilon_2)}],\ \ldots,\psi_{n,j}^{(\varepsilon_\ell)}]\widetilde{w}_{n,j}\|_2 \le C(T,A)(CmJA)^{\ell-1}2^{(1-K)m/2}$$

Hence, the error resulting from (8.22) can be made as small as we wish by taking K large and we are done. □

In what follows, we call sequences $\Lambda_j \subset \mathbb{R}^+$ as in Lemma 8.4 pairwise *orthogonal* iff they satisfy (8.11). The following auxiliary Lemma 8.5 strengthens the result of Lemma 8.3 by replacing L^2 with $\dot B^0_{2,1}$, but under slightly different conditions. As before, T is a Mikhlin operator.

Lemma 8.5. *Suppose $\{\varphi_n\}_{n=1}^\infty, \{\phi_n\}_{n=1}^\infty, \{\psi_n\}_{n=1}^\infty$ lie in the unit-ball of L^2. Furthermore, assume that*

$$\operatorname{supp}(\widehat{\varphi_n}),\ \operatorname{supp}(\widehat{\phi_n}) \subset \{|\xi| \le 1\}, \quad \operatorname{supp}(\widehat{\psi_n}) \subset \{|\xi| \ge 1\}$$

Define

$$v_n := \exp((-\Delta)^{-\frac12}\varphi_n), \quad w_n := \exp((-\Delta)^{-\frac12}\psi_n)$$

Then given $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\big\|(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n)\big\|_{\dot B^0_{2,1}} < \varepsilon \tag{8.23}$$

$$\big\|\nabla^{-1}\big[(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n)\big]\big\|_{\infty} < \varepsilon \tag{8.24}$$

for all sufficiently large n provided

$$\limsup_{n\to\infty} \|P_{<k_0}\psi_n\|_{\dot B^0_{2,\infty}} < \delta \tag{8.25}$$

where $k_0 = k_0(T,\varepsilon)$ is some positive integer.

Proof. Since (8.23) implies (8.24) it suffices to prove the former. As before,

$$(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n) = \int_0^1 (v_n w_n^t)^{-1} [T, (-\Delta)^{-\frac12}\psi_n](\phi_n\, v_n w_n^t)\, dt \tag{8.26}$$

We now estimate the L^2-norm of this expression localized to frequency 2^j. First, we consider the case $j \ge 0$. Then, with $y_{n,t} := v_n w_n^t$, and using Bernstein's inequality, one has the bound

$$\begin{aligned}
&\|P_j\big(y_{n,t}^{-1}[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\big)\|_2 \lesssim 2^{-\frac23 j}\|P_j\nabla\big(y_{n,t}^{-1}[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\big)\|_{\frac32} \\
&\lesssim 2^{-\frac23 j}\|\nabla\big(y_{n,t}^{-1}\big)[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\|_{\frac32} + 2^{-\frac23 j}\|y_{n,t}^{-1}[T,(-\Delta)^{-\frac12}\psi_n]\nabla(\phi_n\, y_{n,t})\|_{\frac32} \\
&\quad + 2^{-\frac23 j}\|y_{n,t}^{-1}[T,\nabla(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\|_{\frac32}
\end{aligned} \tag{8.27}$$

Since uniformly in $0 \le t \le 1$,

$$y_{n,t}\nabla y_{n,t}^{-1} = -(\nabla(-\Delta)^{-\frac12}\varphi_n + t\nabla(-\Delta)^{-\frac12}\psi_n) = O_{L^2}(1)$$

we can further estimate

$$\begin{aligned}
\|\nabla\big(y_{n,t}^{-1}\big)[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\|_{\frac32} &\lesssim \|y_{n,t}^{-1}[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\|_6 \\
&\lesssim \|(-\Delta)^{-\frac12}\psi_n\|_{12}\|\phi_n\|_{12} \lesssim 1
\end{aligned}$$

To pass to the final bound, the term involving ψ_n is estimated via an L^2-BMO interpolation, whereas the ϕ_n term is controlled by Bernstein's inequality. The other two terms on the right-hand side of (8.27) are estimated similarly. As for the case $j \le 0$, Bernstein's inequality yields

$$\begin{aligned}
\|P_j\big(y_{n,t}^{-1}[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\big)\|_2 &\lesssim 2^{\frac{j}{3}}\|P_j\big(y_{n,t}^{-1}[T,(-\Delta)^{-\frac12}\psi_n](\phi_n\, y_{n,t})\big)\|_{\frac32} \\
&\lesssim 2^{\frac{j}{3}}\|(-\Delta)^{-\frac12}\psi_n\|_6\|\phi_n\|_2 \lesssim 2^{\frac{j}{3}}
\end{aligned}$$

To obtain (8.23), it suffices to show that for every $\varepsilon > 0$

$$\big\|(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n)\big\|_2 < \varepsilon^2$$

for large n. Indeed, combining this bound with the preceding then implies

$$\big\|(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n)\big\|_{\dot B^0_{2,1}} \lesssim \varepsilon^2 \log\varepsilon$$

which is more than enough. To this end, fix a large enough a, and let $w_n = w_{n,\mathrm{low}} w_{n,\mathrm{high}}$ where $w_{n,\mathrm{low}}$ corresponds to $\mathcal F^{-1}[\chi_{[|\xi|\le a]}\widehat{\psi_n}]$ and $w_{n,\mathrm{high}}$ to $\mathcal F^{-1}[\chi_{[|\xi|\ge a]}\widehat{\psi_n}]$ (with sharp cut-offs). By (8.28) and Lemma 8.1,

$$\big\|(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - (v_n w_{n,\mathrm{high}})^{-1} T(\phi_n\, v_n w_{n,\mathrm{high}})\big\|_2 \le C(T)\|\mathcal F^{-1}[\chi_{[|\xi|\le a]}\widehat{\psi_n}]\|_{\dot B^0_{2,\infty}}$$

whereas

$$\sup_n \big\|v_n^{-1} T(\phi_n\, v_n) - (v_n w_{n,\mathrm{high}})^{-1} T(\phi_n\, v_n w_{n,\mathrm{high}})\big\|_2 \le C(T)\, a^{-\frac12}$$

by the same argument as in (8.5). Choosing a so that this final bound is $< \varepsilon$ defines both $k_0(T,\varepsilon)$ and δ. □

Clearly, one has the following limiting statement.

Corollary 8.6. *Suppose* $\{\varphi_n\}_{n=1}^\infty, \{\phi_n\}_{n=1}^\infty, \{\psi_n\}_{n=1}^\infty$ *lie in the unit-ball of* L^2. *Furthermore, assume that*

$$\mathrm{supp}(\widehat{\varphi_n}),\ \mathrm{supp}(\widehat{\phi_n}) \subset \{|\xi| \le 1\}$$

and

$$\mathrm{supp}(\widehat{\psi_n}) \subset \{|\xi| \ge 1\}, \qquad \lim_{n\to\infty} \int_{|\xi|\le a} |\widehat{\psi_n}(\xi)|^2 \, d\xi = 0 \tag{8.28}$$

for each $a \ge 1$. *Define*

$$v_n := \exp((-\Delta)^{-\frac{1}{2}} \varphi_n), \quad w_n := \exp((-\Delta)^{-\frac{1}{2}} \psi_n)$$

Then

$$\big\|(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n)\big\|_{\dot{B}^0_{2,1}} \to 0$$

$$\big\|\nabla^{-1}\big[(v_n w_n)^{-1} T(\phi_n\, v_n w_n) - v_n^{-1} T(\phi_n\, v_n)\big]\big\|_\infty \to 0$$

as $n \to \infty$.

9. The Bahouri-Gerard concentration compactness method

In this section, we execute the scheme that was sketched in the introduction. We shall follow the five individual steps which we outlined there.

9.1. **The precise setup for the Bahouri-Gerard method.** As far as the concentration compactness method is concerned, our goal is to demonstrate the following main result.

Proposition 9.1. *Let* $\mathbf{u} = (\mathbf{x}, \mathbf{y}) : (-T_0, T_1) \times \mathbb{R}^2 \to \mathbb{H}^2$ *be a Schwartz class wave map. Then denoting its energy*

$$\sum_{\alpha=0,1,2} \Big(\Big\|\frac{\partial_\alpha \mathbf{x}}{\mathbf{y}}\Big\|_{L^2_x}^2 + \Big\|\frac{\partial_\alpha \mathbf{y}}{\mathbf{y}}\Big\|_{L^2_x}^2 \Big) = E < \infty,$$

there is a an increasing function $C(E) : \mathbb{R}^+ \to \mathbb{R}^+$ *with the property*

$$\|\psi\|_{S((-T_0,T_1)\times\mathbb{R}^2)} \le C(E)$$

We refer to the derivative components of $\mathbf{u}$ with respect to the standard frame $(\mathbf{y}\partial_{\mathbf{x}},\, \mathbf{y}\partial_{\mathbf{y}})$ as ϕ_α^i, $i = 1, 2$, $\alpha = 0, 1, 2$. We also use the complex notation $\phi_\alpha := \phi_\alpha^1 + i\phi_\alpha^2$. We shall refer to a wave map as *admissible*, provided its derivative components at time $t = 0$, $\phi_\alpha^i(0, \cdot)$ lie in the Schwartz class. Finally, for wave maps of Schwartz class as before, we denote the Coulomb components by

$$\psi_\alpha := \phi_\alpha\, e^{-i \sum_{k=1,2} \triangle^{-1} \partial_k \phi_k^1}$$

The energy is then given by

$$E(\mathbf{u}) = \sum_{\alpha=0,1,2} \|\phi_\alpha\|_{L^2_x}^2 = \sum_{\alpha=0,1,2} \|\psi_\alpha\|_{L^2_x}^2$$

To prove Proposition 9.1, we proceed by contradiction, assuming that the set of energy levels E for which it fails is nonempty. Then it has an infimum $E_{crit} > 0$ by the small energy result. We can then find a sequence of wave maps $\mathbf{u}^n = (\mathbf{x}^n, \mathbf{y}^n) : (-T_0^n, T_1^n) \times \mathbb{R}^2 \to \mathbb{H}^2$ with the properties

- $\lim_{n\to\infty} E(\mathbf{u}^n) = E_{crit}$ (these energies approach E_{crit} from above)
- $\lim_{n\to\infty} \|\psi^n\|_{S((-T_0^n,T_1^n)\times\mathbb{R}^2)} = \infty$.

We call such a sequence of wave maps *essentially singular*. It is now our goal to apply the Bahouri-Gerard method to the derivative components of a sequence of essentially singular data $\phi_\alpha^n(0, \cdot)$.

- In subsection 9.2, we construct decompositions of the form

$$\phi_\alpha^n = \sum_{a=1}^{A} \phi_\alpha^{na} + w_\alpha^{nA}$$

  where the ϕ_α^{na} correspond to derivative components of admissible maps which are well-frequency localized.

- In subsection 9.3, we use these decompositions to approximate the data ϕ_α^n by lower-frequency components. The goal is to inductively prove bounds on the Coulomb components of these lower-frequency approximations and finally obtain bounds on the Coulomb components ψ_α^n, unless there is only one frequency atom of maximal energy E_{crit} present.
- In subsection 9.4, the most involved, we obtain a priori bounds on the lowest frequency non-atomic components $\Psi_\alpha^{nA_0^{(0)}}$, by means of a careful induction on low-frequency approximations.
- In subsection 9.6, we construct the profile decomposition for the lowest frequency above-threshold energy frequency atoms. Here a lot of work is involved in showing that the profiles, which are obtained as weak limits of the linear covariant wave evolution associated with operators $\Box_{A^n}$, can actually be interpreted as Coulomb derivative components of actual maps, up to constant phase shifts.
- In subsection 9.7, we then complete the approximate solution which is given by the sum of the profiles and the low-frequency term to an exact solution, via a perturbative argument. This culminates in Proposition 9.30.
- Finally, in subsection 9.9 we explain how to add the remaining frequency atoms.

9.2. **Step 1: frequency decomposition of initial data.** We consider wave maps $\mathbf{u} : \mathbb{R}^{2+1} \to \mathbb{H}^2$, with Schwartz initial data. Here $\mathbb{H}^2$ stands for two-dimensional hyperbolic space which we identify with the upper half-plane. More precisely, introducing coordinates $(\mathbf{x}, \mathbf{y})$ on $\mathbb{H}^2$ in the standard model as upper half plane, and expressing u in terms of these coordinates, we assume that $\mathbf{x}$, $\mathbf{y}$, $\partial_t \mathbf{x}$, $\partial_t \mathbf{y}$ are smooth, decay toward infinity in the sense that

$$\lim_{|x|\to\infty} (\mathbf{x}(x), \mathbf{y}(x)) = (\mathbf{x}_0, \mathbf{y}_0) \in \mathbb{H}^2$$

and such that the derivative components

$$\phi_\alpha^1 = \frac{\partial_\alpha \mathbf{x}}{\mathbf{y}}, \;\; \phi_\alpha^2 = \frac{\partial_\alpha \mathbf{y}}{\mathbf{y}}, \;\; \alpha = 0, 1, 2,$$

are Schwartz, all at fixed time $t = 0$. We make the following

Definition 9.2. *We call initial data* $\{\mathbf{x}, \mathbf{y}, \partial_t \mathbf{x}, \partial_t \mathbf{y}\} : \mathbb{R}^2 \to \mathbb{H}^2 \times T\mathbb{H}^2$ admissible*, provided the derivative components* ϕ_α^k *are Schwartz functions for any* $\alpha = 0, 1, 2$ *and* $k = 1, 2$.

We note here that the property of admissibility is propagated along with the wave map flow on fixed time slices, as long as the wave map persists and is smooth. This follows from finite propagation speed, as well as the small-data well-posedness theory. We recall that the energy associated with given initial data at time $t = 0$ is given by

$$E := \int_{\mathbb{R}^2} \sum_{\alpha=0,1,2} [(\phi_\alpha^1)^2 + (\phi_\alpha^2)^2] \, dx_1 dx_2$$

We now come to the first step in the Bahouri-Gerard decomposition of a sequence of initial data, cf. [1]. More precisely, we wish to obtain a decomposition of the derivative initial data which is analogous to the one of [1]. An added feature for wave maps, which does not appear in [1], consists of the fact that the decomposition has the be performed in such a way that the individual summands in it are themselves derivatives of admissible maps. This requires some care, as the requisite condition is nonlinear, see Lemma 9.3 below. In what follows we write $\phi_\alpha := \phi_\alpha^1 + i\phi_\alpha^2$, any additional superscript referring to the index of a sequence.

Lemma 9.3. *The complex-valued Schwartz functions* ϕ_α, $\alpha = 1, 2$, *correspond to the derivative components of admissible data* $\mathbf{u} : \mathbb{R}^2 \to \mathbb{H}^2$ *iff*

$$\partial_k \phi_j - \partial_j \phi_k = \phi_k^1 \phi_j^2 - \phi_k^2 \phi_j^1, \;\; k, j = 1, 2 \tag{9.1}$$

are satisfied.

Proof. The "only if" part follows from (1.6), (1.7). For the "if" part, note first that we get

$$\partial_k \phi_j^2 - \partial_j \phi_k^2 = 0 \tag{9.2}$$

for the imaginary parts of ϕ_j and ϕ_k. This implies that

$$\phi_j^2 = \frac{\partial_j \mathbf{y}}{\mathbf{y}}, \ \ j = 1, 2$$

for a suitable positive function $\mathbf{y} : \mathbb{R}^2 \to \mathbb{R}^+$ which is unique only up to a multiplicative positive constant. We can rewrite (9.1) in the form

$$\partial_k(\mathbf{y}\phi_j^1) - \partial_j(\mathbf{y}\phi_k^1) = 0, \ \ k, j = 1, 2 \tag{9.3}$$

which in turn implies that

$$\phi_j^1 = \frac{\partial_j \mathbf{x}}{\mathbf{y}}$$

for a suitable function $\mathbf{x} : \mathbb{R}^2 \to \mathbb{R}$. To understand the behavior of $(\mathbf{x}, \mathbf{y})$ at infinity, we observe the following[20]: from (9.2),

$$\partial_2 \int_{-\infty}^{\infty} \phi_1^2(x_1, x_2)\, dx_1 = 0$$

which implies that the integral does not depend on x_2 and therefore is, in fact, zero. Similarly,

$$\int_{-\infty}^{\infty} \phi_2^2(x_1, x_2)\, dx_2 = 0 \qquad \forall\, x_1 \in \mathbb{R}$$

It follows that $\mathbf{y}$ tends to the same constant at infinity irrespective of the way in which we approach infinity. Without loss of generality, we may set this constant equal to 1. From (9.3) one further sees that

$$\int_{-\infty}^{\infty} \mathbf{y}\, \phi_1^1(x_1, x_2)\, dx_1 = \int_{-\infty}^{\infty} \mathbf{y}\, \phi_2^1(x_1, x_2)\, dx_2 = 0$$

whence $\mathbf{x}$ approaches a constant $\mathbf{x}_0$ at ∞. □

Now for the first step in the concentration compactness method, which is the Metivier-Schochet scale selection process, see [31] and Section III.1 of [1]. As already explained above, the difficulty we face here in contrast to [1] is that we need to make sure that the pieces we decompose the derivative components into are geometric, i.e., they are themselves derivative components of maps $\mathbb{R}^2 \to \mathbb{H}^2$. Section 8 provides us with the tools required for this purpose.

Proposition 9.4. *Let $\{\mathbf{x}_n, \mathbf{y}_n, \partial_t \mathbf{x}_n, \partial_t \mathbf{y}_n\}_{n\geq 1}$ be any sequence of admissible data with energy bounded by E and with associated derivative sequence $\{\phi_\alpha^n\}_{n\geq 1}$, $\alpha = 0, 1, 2$. Then up to passing to a subsequence the following holds: given $\delta > 0$, there exists a positive integer $A = A(\delta, E)$ and a decomposition*

$$\phi_\alpha^n = \sum_{a=1}^{A} \phi_\alpha^{na} + w_\alpha^{nA}$$

for $\alpha = 0, 1, 2$ and $n \geq 1$. Here the functions ϕ_α^{na}, $1 \leq a \leq A$ are derivative components of admissible maps $\mathbf{u}_n^a : \mathbb{R}^2 \to \mathbb{H}^2$, and are λ_n^a-oscillatory for a sequence of pairwise orthogonal frequency scales $\{\lambda_n^a\}_{n\geq 1}$ while the remainder w_α^{An} is λ_n^a-singular for each $1 \leq a \leq A$ and satisfies the smallness condition

$$\sup_{n\geq 1} \|w_\alpha^{nA}\|_{\dot{B}^0_{2,\infty}} < \delta$$

Finally, given any sequence $R_n \to \infty$ one has the frequency localization with $\mu_n^a := -\log \lambda_n^a$,

$$\sup_{\alpha=0,1,2} \|P_j \phi_\alpha^{na}\|_2 \leq E\, R_n^{\frac{1}{3}} 2^{-\frac{1}{3}|j-\mu_n^a|} \qquad \forall\, j \in \mathbb{Z} \tag{9.4}$$

for all $1 \leq a \leq A$ and all large n.

[20] Here the superscripts are not powers

Proof. We omit the time dependence in the notation, keeping in mind that everything takes place at initial time $t = 0$. As in Section III.1 of [1] one obtains a decomposition

$$\phi_\alpha^n = \sum_{a=1}^{A} \tilde{\phi}_\alpha^{na} + \tilde{w}_\alpha^{nA}, \quad \alpha = 0, 1, 2 \tag{9.5}$$

where the functions $\tilde{\phi}_\alpha^{na} \in L^2(\mathbb{R}^2)$ are λ_n^a-oscillatory for suitable pairwise orthogonal frequency scales $\{\lambda_n^a\}_{n\geq 1}$ for all $1 \leq a \leq A$. Moreover, there is the smallness

$$\|\tilde{w}_\alpha^{nA}\|_{\dot{B}^0_{2,\infty}} < \delta$$

We now restrict to Fourier supports of these functions. Pick a sequence $R_n \to \infty$ growing sufficiently slowly such that the intervals $[(\lambda_n^a)^{-1}R_n^{-1}, (\lambda_n^a)^{-1}R_n]$ are mutually disjoint for n large enough and different values of a. Then we replace $\tilde{w}_\alpha^{nA}$ by

$$P_{\cap_{a=1}^A [\mu_n^a - \log R_n, \mu_n^a + \log R_n]^c} \tilde{w}_\alpha^{nA} + \sum_{a=1}^{A} P_{\cap_{a'=1}^A [\mu_n^{a'} - \log R_n, \mu_n^{a'} + \log R_n]^c} \tilde{\phi}_\alpha^{na}$$

where $\mu_n^a := -\log \lambda_n^a$, while we replace each $\tilde{\phi}_\alpha^{na}$, $1 \leq a \leq A$, by

$$P_{[\mu_n^a - \log R_n, \mu_n^a + \log R_n]} \sum_{a'=1}^{A} \tilde{\phi}_\alpha^{na'} + P_{[\mu_n^a - \log R_n, \mu_n^a + \log R_n]} \tilde{w}_\alpha^{nA}$$

We need to make R_n increase sufficiently slowly so that the second term here remains λ_n^a-oscillatory. Of course the new $\tilde{\phi}_\alpha^{na}$ now also depend on the cutoff A; in order to get rid of this dependence, we may replace A by A_n where $A_n \to \infty$ suitably slowly. Then the new decomposition, which we again refer to as

$$\phi_\alpha^n = \sum_{a=1}^{A} \tilde{\phi}_\alpha^{na} + \tilde{w}_\alpha^{nA}$$

has the same properties as the original one with the added advantage of the sharp frequency localization around the scales $(\lambda_n^a)^{-1}$. In particular, since the ϕ_α^n are Schwartz functions, one concludes that the $\tilde{\phi}_\alpha^{na}$ have the same property which means that the components $\tilde{\phi}_\alpha^{na}$ are admissible, and so is $\tilde{w}_\alpha^{nA}$.

In order to prove the proposition we need to show that we can replace the components $\tilde{\phi}_\alpha^{na}$ by components ϕ_α^{na} which actually belong to admissible maps $\mathbf{u}^{na} : \mathbb{R}^2 \to \mathbb{H}^2$ up to a small error (which again can be absorbed into $\tilde{w}_\alpha^{nA}$). Note that the $\alpha = 0$ component does not present a problem here. For the $\alpha = 1, 2$ components, however, we need to ensure that the compatibility relations (9.1) hold. Continuing with the proof of the Proposition 9.4, we notice that

$$\mathbf{x}^n = \sum_{k=1,2} \triangle^{-1}\partial_k[\phi_k^{1n}\mathbf{y}^n], \quad \mathbf{y}^n = e^{\sum_{k=1,2} \triangle^{-1}\partial_k \phi_k^{2n}}$$

for the coordinate functions $(\mathbf{x}^n, \mathbf{y}^n)$; here we recall that we may impose the normalizations $\lim_{|x|\to\infty} \mathbf{x}(x) = 0$, $\lim_{|x|\to\infty} \mathbf{y}(x) = 1$. In turn, these identities imply that

$$\phi_j^{1n} = (\mathbf{y}^n)^{-1} \sum_{k=1,2} \triangle^{-1}\partial_j\partial_k[\phi_k^{1n}\mathbf{y}^n], \quad \phi_j^{2n} = \sum_{k=1,2} \triangle^{-1}\partial_j\partial_k\phi_k^{2n}$$

These relations shall allow us to replace (9.5) by a "geometric decomposition". Indeed, we simply substitute the decomposition (9.5) to obtain

$$\phi_j^{1n} = \sum_{a=1}^{A} (\mathbf{y}^n)^{-1} \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j[\tilde{\phi}_k^{1na}\mathbf{y}^n] + (\mathbf{y}^n)^{-1} \sum_{k=1,2} \triangle^{-1}\partial_j\partial_k[\tilde{w}^{1nA}\mathbf{y}^n]$$

$$\phi_j^{2n} = \sum_{a=1}^{A} \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j\tilde{\phi}_k^{2na} + \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j\tilde{w}_k^{2nA}$$

This suggests making the following choices:

$$\mathbf{x}^{na} := \sum_{k=1,2} \triangle^{-1}\partial_k[\tilde{\phi}_k^{1na}\mathbf{y}^{na}], \quad \mathbf{y}^{na} := e^{\sum_{k=1,2}\triangle^{-1}\partial_k\tilde{\phi}_k^{2na}}$$

and then defining

$$\phi_j^{1na} := (\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_j\partial_k[\tilde{\phi}^{1na}\mathbf{y}^{na}], \quad \phi_j^{2na} := \sum_{k=1,2}\triangle^{-1}\partial_j\partial_k\tilde{\phi}_k^{2na}$$
$$w_j^{1nA} := (\mathbf{y}^{n})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_j\partial_k[\tilde{w}^{1nA}\mathbf{y}^{n}], \quad w_j^{2nA} := \sum_{k=1,2}\triangle^{-1}\partial_j\partial_k\tilde{w}_k^{2nA}$$

as well as $\phi^{na} := \phi^{1na}+i\phi^{2na}$, $w^{nA} := w^{1nA}+iw^{2nA}$. Clearly the components ϕ_j^{1na}, ϕ_j^{2na} are now geometric in the sense that they derive from a map into hyperbolic space; in fact, they are associated with the maps given by the components $(\mathbf{x}^{na}, \mathbf{y}^{na})$. The proof is now concluded by appealing to Lemma 8.2, Corollary 8.4, and Lemma 8.3. For the final statement, note that by Lemma 8.2, the "geometric" components ϕ_α^{na} are also frequency localized to the interval $[\mu_n^a - \log R_n, \mu_n^a + \log R_n]$ up to exponentially decaying errors. □

As an immediate consequence of Proposition 9.4 one obtains that ϕ_j^{kna}, w_j^{knA}, $k=1,2$, are asymptotically orthogonal (where $\phi_j^{1na} = \operatorname{Re}\phi_j^{na}$ and $\phi_j^{2na} = \operatorname{Im}\phi_j^{na}$).
We now make some preparations for the second stage of the Bahouri-Gerard procedure. More specifically, we shall have to pass to the Coulomb gauge components, ψ_α, and transfer the above decomposition to the level of these components. One can split

$$\psi_\alpha^n = \phi_\alpha^n e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}} = \big[\sum_{a=1}^{A}\phi_\alpha^{na} + w_\alpha^{nA}\big]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}$$

However, the components

$$\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}$$

are not the Coulomb gauge components of a suitable wave map, and should ideally be replaced by

$$\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1na}}$$

Due to the lack of L^∞ control over the exponent, this cannot be done without further physical localizations. Nevertheless, we can state the following fact.

Lemma 9.5. *The components*

$$\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}, \qquad w_\alpha^{nA}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}$$

are λ_n^a-oscillatory and λ_n^a-singular, respectively, for each a and we have

$$\big\|w_\alpha^{nA}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big\|_{\dot{B}^0_{2,\infty}} \lesssim \delta$$

where δ is as in Proposition 9.4.

Proof. We may assume $\lambda_n^a = 1$ by scaling invariance. Given any $\varepsilon > 0$, we can choose k_0 large enough such that

$$\limsup_{n\to\infty}\|P_{[-k_0,k_0]^c}\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\|_{L^2} < \varepsilon$$

Next, for $k_1 > k_0 + C$, consider the expressions

$$P_{<-k_1}\big[P_{[-k_0,k_0]}\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big], \quad P_{>k_1}\big[P_{[-k_0,k_0]}\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big].$$

Start with the first expression, which we write as

$$\begin{aligned}&P_{<-k_1}\big[P_{[-k_0,k_0]}\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\\ &= P_{<-k_1}\big[P_{[-k_0,k_0]}\phi_\alpha^{na}\sum_{j=1,2}\triangle^{-1}\partial_j P_{[-k_0,k_0]}([-i\sum_{k=1,2}\triangle^{-1}\partial_j\partial_k\phi_k^{1n}]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}})\big]\end{aligned}$$

Using Bernstein's inequality, we can then estimate

$$\begin{aligned}&\|P_{<-k_1}\big[P_{[-k_0,k_0]}\phi_\alpha^{na}\sum_{j=1,2}\triangle^{-1}\partial_j P_{[-k_0,k_0]}([-i\sum_{k=1,2}\triangle^{-1}\partial_j\partial_k\phi_k^{1n}]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}})\big]\|_{L^2}\\&\lesssim 2^{k_0-k_1}\|P_{[-k_0,k_0]}\phi_\alpha^{na}\|_{L^2}\|\phi^{1n}\|_{L^2}<\varepsilon\end{aligned}$$

provided we choose k_1 sufficiently large in relation to k_0. The estimate for the second term is more of the same. Next, consider the "tail term" $w_\alpha^{nA}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}$. That this is λ_n^a-singular for each $1\le a\le A$ follows from the preceding via duality. It therefore remains to estimate its $\|\cdot\|_{\dot B^0_{2,\infty}}$-norm. We localize this term to fixed dyadic frequency $\sim 2^q$

$$\begin{aligned}&P_q\big[w_\alpha^{nA}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]=P_q\big[w_\alpha^{nA}P_{<q-10}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\\&+P_q\big[w_\alpha^{nA}P_{[q-10,q+10]}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]+P_q\big[w_\alpha^{nA}P_{>q+10}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\end{aligned}$$

and estimate the three terms on the right separately: first, we have

$$\begin{aligned}&\|P_q\big[w_\alpha^{nA}P_{<q-10}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\|_{L^2}=\|P_q\big[P_{[q-10,q+10]}(w_\alpha^{nA})P_{<q-10}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\|_{L^2}\\&\lesssim\|P_{[q-10,q+10]}(w_\alpha^{nA})\|_{L^2}\lesssim\|w_\alpha^{nA}\|_{\dot B^0_{2,\infty}}\lesssim\delta\end{aligned}$$

Next,

$$\begin{aligned}&\|P_q\big[w_\alpha^{nA}P_{[q-10,q+10]}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\|_{L^2}\\&=\|P_q\big[P_{<q+10}(w_\alpha^{nA})P_{[q-10,q+10]}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}}\big]\|_{L^2}\\&=\|P_q\big[P_{<q+10}w_\alpha^{nA}\sum_{j=1,2}\triangle^{-1}\partial_j P_{[q-10,q+10]}([-i\sum_{k=1,2}\triangle^{-1}\partial_j\partial_k\phi_k^{1n}]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{1n}})\big]\|_{L^2}\\&\lesssim 2^{-q}\|P_{<q+10}w_\alpha^{nA}\|_{L^\infty}\|\phi^{1n}\|_{L^2}\lesssim\|w_\alpha^{nA}\|_{\dot B^0_{2,\infty}}\lesssim\delta\end{aligned}$$

where Bernstein's inequality was used in the last step. The third term in the above Littlewood-Paley trichotomy corresponding to high-high interactions, is treated analogously and omitted. □

For later reference, it shall be important to construct "partial approximations" of the components ϕ_α^n in terms of the ϕ_α^{na}. Specifically, for $I\subset\{1,2,\dots,A\}$, we let

$$\tilde\phi_\alpha^{nI}:=\sum_{a\in I}\tilde\phi_\alpha^{na}$$

Then reasoning exactly as in the preceding, and employing the same notation as there, one obtains the following statement.

Corollary 9.6. *Let*

$$\mathbf{y}^{nI}:=e^{\sum_{k=1,2}\sum_{a\in I}\triangle^{-1}\partial_k\tilde\phi_k^{2na}},\qquad\mathbf{x}^{nI}:=\sum_{k=1,2}\sum_{a\in I}\triangle^{-1}\partial_k[\tilde\phi_k^{1na}\mathbf{y}^{nI}]$$

Then for $a\in I$

$$\phi_j^{1na}=(\mathbf{y}^{nI})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_j\partial_k[\tilde\phi_k^{1na}\mathbf{y}^{nI}]+o_{L^2}(1)$$

In particular, we have

$$\sum_{a\in I}\phi_j^{1na}=\frac{\partial_j\mathbf{x}^{nI}}{\mathbf{y}^{nI}}+o_{L^2}(1),\qquad\sum_{a\in I}\phi_j^{2na}=\frac{\partial_j\mathbf{y}^{nI}}{\mathbf{y}^{nI}}+o_{L^2}(1)$$

9.3. **Step 2: frequency localized approximations to the data.** Given an essentially singular sequence $\mathbf{u}^n$ with derivatives ϕ_α^n, Proposition 9.4 yields a new essentially singular sequence ϕ_α^n with the following property: for any $A \geq 1$ (recall the ϕ_α^{na} are defined inductively)

$$\phi_\alpha^n = \sum_{a=1}^{A} \phi_\alpha^{na} + w_\alpha^{nA}$$

Given $\delta_0 > 0$, there exists $A \geq 1$ so that $\|w_\alpha^{nA}\|_{\dot{B}^0_{2,\infty}} < \delta_0$ for large n. In what follows, we will use smallness parameters $1 \gg \varepsilon_0 \gg \delta_1 \gg \delta_0 > 0$, each of which will eventually be chosen depending only on the energy of the initial data.

Figure 5. Atoms and the Besov error

Ultimately we wish to show that there can only be a single frequency block, i.e., $A = 1$, and furthermore, that the energy of this block converges to the critical energy E_{crit} as $n \to \infty$. Thus we now use the following *dichotomy*:

- We have $A = 1$ and $\lim_{n\to\infty} \sum_{\alpha=0,1,2} \|\phi_\alpha^{na}\|_{L_x^2}^2 = E_{crit}$.
- The previous scenario does not occur. Thus, for a suitable subsequence

$$\limsup_{n\to\infty} \sum_{\alpha=0,1,2} \|\phi_\alpha^{na}\|_{L_x^2}^2 < E_{crit} - \delta_2$$

for some $\delta_2 > 0$, and all a.

If the first alternative occurs, then continue with Step 4 below. Hence we now assume that the second alternative occurs, in which case we will show that the sequence $\mathbf{u}^n$ cannot be essentially singular. We may of course assume that for each $1 \leq a \leq A$,

$$\liminf_{n\to\infty} \sum_{\alpha=0,1,2} \|\phi_\alpha^{na}\|_{L_x^2} > 0,$$

as otherwise we may pass to a subsequence for which the ϕ_α^{na} may be absorbed into the error w_α^{nA}. We may also assume that

$$\liminf_{n\to\infty} \sum_{\alpha=0,1,2} \|\phi_\alpha^{na}\|_{L_x^2}^2 = \limsup_{n\to\infty} \sum_{\alpha=0,1,2} \|\phi_\alpha^{na}\|_{L_x^2}^2$$

by passing to a subsequence. The issue now becomes how to choose the cutoff A. Due to the asymptotic orthogonality of the ϕ_α^{na} as $n \to \infty$, and for each $\alpha = 0, 1, 2$,

$$\lim_{A_0\to\infty} \sum_{a\geq A_0} \limsup_{n\to\infty} \|\phi_\alpha^{na}\|_{L_x^2}^2 = 0$$

For some absolute $\varepsilon_0 > 0$ which is small enough only depending on E_{crit}, in particular smaller than the cutoff for the small energy global well-posedness theory, we choose A_0 large enough such that

$$\sum_{a\geq A_0} \limsup_{n\to\infty} \|\phi_\alpha^{na}\|_{L_x^2}^2 < \varepsilon_0,$$

and then put $A = A_0$. Thus we now arrive at the decomposition

$$\phi_\alpha^n = \sum_{a=1}^{A_0} \phi_\alpha^{na} + w_\alpha^{nA_0}$$

We may further decompose

$$w_\alpha^{nA_0} = \sum_{a=A_0+1}^{A} \phi_\alpha^{na} + w_\alpha^{nA},$$

with the smallness property

$$\sum_{a\geq A_0+1} \sum_{\alpha=0,1,2} \limsup_{n\to\infty} \|\phi_\alpha^{na}\|_{L_x^2}^2 < \varepsilon_0$$

By adjusting A, we can further achieve

$$\limsup_{n\to\infty} \|w_\alpha^{nA}\|_{\dot{B}^0_{2,\infty}} < \delta_0$$

for any given $\delta_0 > 0$.

Re-ordering the superscripts if necessary, we may assume that the frequency scales $(\lambda_n^a)^{-1}$ of the ϕ_α^{na} are increasing with $1 \leq a \leq A_0$. The error term $w_\alpha^{nA_0}$ may be written as a sum of constituents

$$w_\alpha^{nA_0} = w_\alpha^{nA_0^{(0)}} + w_\alpha^{nA_0^{(1)}} + \ldots + w_\alpha^{nA_0^{(A_0)}} + o_{L^2}(1)$$

which satisfy the property that

$$w_\alpha^{nA_0^{(k)}} = P_{\mu_n^{k-1}+L_n<\cdot<\mu_n^k-L_n} w_\alpha^{nA_0^{(k)}} + o_{L^2}(1) \quad \text{as} \quad n\to\infty \tag{9.6}$$

with $\mu_n^a := -\log\lambda_n^a$ and a sequence $L_n \to \infty$ which increases very slowly. This can be done since $w_\alpha^{nA_0}$ is λ_n^a-singular for each $1 \leq a \leq A_0$. Thus the frequency support of $w_\alpha^{nA_0^{(k)}}$ is contained in the annulus

$$(\lambda_n^k)^{-1}e^{L_n} < |\xi| < (\lambda_n^{k+1})^{-1}e^{-L_n}, \quad (\lambda_n^0)^{-1} := 0, \quad (\lambda_n^{A_0^{(A_0)}+1})^{-1} := \infty$$

Figure 5 above is a schematic depiction of the situation $A_0 = 1$ with a unique large atom on the right, but with two smaller atoms on the left which are too large to be included in the Besov error (the three bumpy curves between the atoms). More precisely, $w_\alpha^{nA_0}$ consists of the four small curves between the atoms, and $w_\alpha^{nA_0^{(0)}}$ is the sum of the three curves to the left of the big atom *together with the two small atoms*, and $w_\alpha^{nA_0^{(1)}}$ the one to the right of the big atom.

Note that if we refine the frequency decomposition, i.e., increase A_0 to $A^{(k)} \geq A_0$, then the components $w_\alpha^{nA_0^{(k)}}$ are decomposed into

$$w_\alpha^{nA_0^{(k)}} = \sum_j \phi_\alpha^{na_j^k} + w_\alpha^{nA^{(k)}}$$

for suitable $a_j^k \in [A_0+1, A^{(k)}]$. In Figure 5 one has $j = 1,2$ for $k = 0$ corresponding to the two small atoms to the left of the large one. We may again assume that the a_j^k are increasing in j and have frequency support with increasing value of $|\xi|$, for each k. Furthermore, we have

$$\sum_{\alpha=0,1,2} \sum_j \limsup_{n\to\infty} \|\phi_\alpha^{na_j^k}\|_{L_x^2}^2 < \varepsilon_0$$

by asymptotic orthogonality and the choice of A_0. Our first goal, to be dealt with in the following section, is to control the nonlinear evolution of the minimum frequency components $w_\alpha^{nA_0^{(0)}}$. The idea behind this is as follows: due to the energy constraint

$$\limsup_{n\to\infty} \|w_\alpha^{nA_0^{(0)}}\|_2 \le E_{crit}\,,$$

we may subdivide $w_\alpha^{nA_0^{(0)}}$ into finitely many pieces be means of frequency localizations[21] $\big\{P_{J_\ell} w_\alpha^{nA_0^0}\big\}_{1\le\ell\le\frac{1000E_{crit}}{\varepsilon_0}}$ such that the dyadic intervals J_ℓ are disjoint, with $\cup_\ell J_\ell = (-\infty, (\lambda_n^1)^{-1}e^{-L_n})$, and furthermore

$$\big\|P_{J_\ell} w_\alpha^{nA_0^{(0)}}\big\|_{L_x^2} < \varepsilon_0 \quad \forall \ell$$

Recall that $(\lambda_n^1)^{-1}$ is the frequency scale of the first frequency atom ϕ_α^{n1}. In particular, this means that the frequency localized pieces $P_{J_\ell} w_\alpha^{nA_0^0}$ should be treatable via a perturbative argument. More precisely, we shall run an induction in ℓ on a sequence of approximating maps with Coulomb data essentially (up to errors which can be made arbitrarily small depending on a parameter δ_0, see Lemma 9.8 below) given by

$$\sum_{1\le j\le \ell} P_{J_j} w_\alpha^{nA_0^{(0)}}\, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{1\le j\le\ell} P_{J_j} w_k^{nA_0^{(0)}}}$$

As always, we face the issue at this point that these gauged components are not necessarily admissible, i.e., they are not given by derivative components of maps $\mathbb{R}^2 \to \mathbb{H}^2$. In order to apply the perturbative theory we shall need to show that they are close to such admissible data. This in turn follows from Lemma 8.5 provided we chose the intervals J_ℓ carefully; for this it is essential that the endpoints of these intervals do not fall onto one of the 'small' atoms ϕ_α^{na}. Otherwise, condition (8.25) would be violated. In detail, this is done as follows. Recall that the J_j are chosen to be disjoint and such that

$$w_\alpha^{nA_0^{(0)}} = \sum_j P_{J_j} w_\alpha^{nA_0^{(0)}}, \qquad \sup_j \big\|P_{J_j} w_\alpha^{nA_0^{(0)}}\big\|_{L_x^2} \lesssim \varepsilon_0$$

On the other hand, upon refining the Bahouri-Gerard frequency decomposition applied to $w_\alpha^{nA_0^{(0)}}$, we can also write

$$w_\alpha^{nA_0^{(0)}} = \sum_{j\ge1} \phi_\alpha^{na_j^{(0)}} + w_\alpha^{nA^{(0)}} \tag{9.7}$$

Here $A^{(0)} > A_0$ is chosen such that $\|w_\alpha^{nA^{(0)}}\|_{\dot B^0_{2,\infty}} \ll \delta_0$ for some constant $\delta_0 > 0$ which is to be determined, while the $a_j^{(0)}$ are certain indices in the interval $[A_0, A^{(0)}]$. Our choice of A_0 ensures that

$$\limsup_{n\to\infty} \sum_{j\ge1} \big\|\phi_\alpha^{na_j^{(0)}}\big\|_{L_x^2}^2 < \varepsilon_0$$

Now, to choose the J_j, pick for each of the $\phi_\alpha^{na_j^{(0)}}$ (which are finite in number) a frequency interval

$$\big[(\lambda_n^{a_j^{(0)}} R_j^{(0)})^{-1}, (\lambda_n^{a_j^{(0)}})^{-1} R_j^{(0)}\big]$$

with $R_j^{(0)}$ large enough such that

$$\limsup_{n\to\infty} \big\|P_{[\log(\lambda_n^{a_j^{(0)}})^{-1} - \log R_j^{(0)}, \log(\lambda_n^{a_j^{(0)}})^{-1} + \log R_j^{(0)}]^c} \phi_\alpha^{na_j^{(0)}}\big\|_{L_x^2} \ll \delta_0, \tag{9.8}$$

[21]We suppress the dependence on n of the intervals J_ℓ.

which is possible due to the frequency localization of the atoms $\phi_\alpha^{na_j^{(0)}}$. Here $\delta_0 > 0$ is a sufficiently small constant such that $\delta_0 = \delta_0(E_{crit}, \varepsilon_0)$, to be determined later. Picking n large enough, we may assume that the intervals

$$\big[(\lambda_n^{a_j^{(0)}} R_j^{(0)})^{-1}, (\lambda_n^{a_j^{(0)}})^{-1} R_j^{(0)}\big]$$

are disjoint. We can now exactly specify how to select the J_j: inductively, assume that

$$J_1 = [a_1, b_1], \dots, J_{k-1} = [a_{k-1}, b_{k-1}]$$

have been chosen. Then pick $\tilde{J}_k = [\tilde{a}_k, \tilde{b}_k]$ such that $\tilde{a}_k = b_{k-1}$ and such that the integer $\tilde{b}_k$ is maximal with the property that

$$\sum_{\alpha=0,1,2} \big\| P_{[\tilde{a}_k, \tilde{b}_k]} w_\alpha^{nA_0^{(0)}} \big\|_{L_x^2}^2 \le \varepsilon_0$$

Then if $\tilde{b}_k \in [\log(\lambda_n^{a_j^{(0)}})^{-1} - \log R_j^{(0)}, \log(\lambda_n^{a_j^{(0)}})^{-1} + \log R_j^{(0)}]$ for some j, we let

$$b_k = \log(\lambda_n^{a_j^{(0)}})^{-1} + \log R_j^{(0)}$$

Otherwise, we let $b_k = \tilde{b}_k$. The point of this construction is that if the endpoint of $\tilde{J}_k$ happens to fall on a "small atom" which may still be too large in $\dot{B}^0_{2,\infty}$ for our later purposes, we simply absorb this atom into J_k.

We can now state the approximate admissibility fact alluded to above. Recall that $\operatorname{Re} w_k^{nA_0^{(0)}} = w_k^{1nA_0^{(0)}}$. Moreover, the constant δ_0 controls the Besov norm of the tails and is kept fixed. We begin with a statement which does not involve the J_ℓ.

Lemma 9.7. *There is an admissible map $\mathbb{R}^2 \to \mathbb{H}^2$ with derivative components $\Phi_\alpha^{nA_0^{(0)}}$ such that*

$$\Big\| w_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2} \triangle^{-1}\partial_k w_k^{1nA_0^{(0)}}} - \Phi_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2} \triangle^{-1}\partial_k \Phi_k^{1nA_0^{(0)}}} \Big\|_{L_x^2} \to 0$$

as $n \to \infty$. The same applies to the difference $w_\alpha^{nA_0^{(0)}} - \Phi_\alpha^{nA_0^{(0)}}$.

Proof. Recall the relation that defines $w_j^{nA_0^{(0)}} = w_j^{1nA_0^{(0)}} + i w_j^{2nA_0^{(0)}}$:

$$w_j^{1nA_0^{(0)}} = (\mathbf{y}^n)^{-1} \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j[\tilde{w}_k^{1nA_0^{(0)}} \mathbf{y}^n], \qquad w_j^{2nA_0^{(0)}} = \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j \tilde{w}_j^{2nA_0^{(0)}}$$

We now claim that the components $w_j^{1nA_0^{(0)}}, w_j^{2nA_0^{(0)}}$ are $o_{L^2}(1)$- close to the derivative components $\Phi_j^{1,2nA_0^{(0)}}$ of a map, when $n \to \infty$. Moreover, the error satisfies $\nabla^{-1} o_{L^2}(1) = o_{L^\infty}(1)$. First, observe that by Corollary 8.6, the component $w_j^{1nA_0^{(0)}}$ is close in the above sense to

$$\Phi_j^{1nA_0^{(0)}} := (\mathbf{y}^{nA_0^{(0)}})^{-1} \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j[\tilde{w}_k^{1nA_0^{(0)}} \mathbf{y}^{nA_0^{(0)}}], \qquad \mathbf{y}^{nA_0^{(0)}} := e^{\sum_{k=1,2} \triangle^{-1}\partial_k \tilde{w}_k^{2nA_0^{(0)}}}$$

Next, introduce the auxiliary map $(\mathbf{x}^{nA_0^{(0)}}, \mathbf{y}^{nA_0^{(0)}}) : \mathbb{R}^2 \to \mathbb{H}^2$, with components defined by

$$\mathbf{x}^{nA_0^{(0)}} := \sum_{k=1,2} \triangle^{-1}\partial_k[\tilde{w}_k^{1nA_0^{(0)}} \mathbf{y}^{nA_0^{(0)}}], \qquad \mathbf{y}^{nA_0^{(0)}} = e^{\sum_{k=1,2} \triangle^{-1}\partial_k \tilde{w}_k^{2nA_0^{(0)}}}$$

Furthermore, as before we have

$$w_j^{1nA_0^{(0)}} = (\mathbf{y}^n)^{-1} \sum_{k=1,2} \triangle^{-1}\partial_k\partial_j[\tilde{w}_k^{1nA_0^{(0)}} \mathbf{y}^n], \qquad w_j^{2nA_0^{(0)}} = \sum_{k=1,2} \triangle^{-1}\partial_j\partial_k \tilde{w}_k^{2nA_0^{(0)}}$$

and we set $\Phi_j^{2nA_0^{(0)}} := w_j^{2nA_0^{(0)}}$, $w_0^{1,2nA_0^{(0)}} = \tilde{w}_0^{1,2nA_0^{(0)}}$. In view of the preceding,

$$w_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2} \triangle^{-1}\partial_k w_k^{1nA_0^{(0)}}} = \Phi_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2} \triangle^{-1}\partial_k \Phi_k^{1nA_0^{(0)}}} + o_{L^2}(1)$$

as $n \to \infty$. □

A similar result now applies to the frequency localized pieces. This time one has to use Lemma 8.5.

Lemma 9.8. *Given any $\delta_1 > 0$ one can choose $\delta_0 \ll \delta_1$ as above such that for all large n*

$$\sum_{j\leq \ell} P_{J_j} w_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\leq\ell} P_{J_j} w_k^{1nA_0^{(0)}}}, \quad \ell \geq 1$$

may be approximated within δ_1 in the energy topology by Coulomb components

$$\Psi_\alpha^{\ell n A_0^{(0)}} := \Phi_\alpha^{\ell n A_0^{(0)}} e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \Phi_k^{\ell n A_0^{(0)}}} \tag{9.9}$$

of actual maps from $\mathbb{R}^2 \to \mathbb{H}^2$, uniformly in ℓ. The same statement holds for the functions without any exponential phases.

Proof. This follows exactly along the lines of the proof of Lemma 9.7: for the components $\sum_{j\leq\ell} P_{J_j} w_\alpha^{nA_0^{(0)}}$ we use the approximating maps

$$\mathbf{x}^{\ell n A_0^{(0)}} := (\mathbf{y}^{\ell n A_0^{(0)}})^{-1} \sum_{k=1,2}\sum_{j\leq\ell} \triangle^{-1}\partial_k \big[P_{J_j}\tilde{w}_k^{1nA_0^{(0)}} \mathbf{y}^{\ell n A_0^{(0)}}\big], \qquad \mathbf{y}^{\ell n A_0^{(0)}} := e^{\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\leq\ell}\tilde{w}_k^{2nA_0^{(0)}}}$$

However, this time, the smallness of the error is contingent on the $\|\cdot\|_{\dot{B}^0_{2,\infty}}$-norm of the non-atomic part of $\tilde{w}_\alpha^{2nA_0^{(0)}}$, while the contribution of the atomic part can be made small by choosing n large enough. More precisely, (8.25) holds for all large n due to the frequency separation properties which we have imposed on the various components, see (9.8) and (9.6). These separations become effective for large n due to the orthogonality of the scales involved. □

As a general comment, we would like to remind the reader that all constructions here are not unique; moreover, they are subject to errors of the form $o_{L^2}(1)$ as $n \to \infty$.

9.4. **Step 3: Evolving the lowest-frequency nonatomic part.** As far as the evolution of $w_\alpha^{nA_0^{(0)}}$ is concerned, we claim the following result. Note that we phrase it in terms of the derivative components that we just constructed. Once we have evolved *all constituents* of the decomposition from Step 1, the perturbative theory of Section 7 will then allow us to conclude that the representation that we obtain is accurate up to a small energy error *globally in time.*

Proposition 9.9. *Let $\Phi_\alpha^{nA_0^{(0)}}$ be as in Lemma 9.7 and set*

$$\Psi_\alpha^{nA_0^{(0)}} := \Phi_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k \Phi_k^{1nA_0^{(0)}}}$$

Then provided $\varepsilon_0 \gg \delta_1 \gg \delta_0 > 0$ above are chosen sufficiently small, and provided n is large enough, the $\Phi_\alpha^{nA_0^{(0)}}$ exist globally in time as derivative components of an admissible wave map. Moreover, there is a constant $C_1(E_{crit})$ such that the solution of the gauged counterparts of these components, i.e., $\Psi_\alpha^{nA_0^{(0)}}$ satisfy the bound

$$\sup_{T_{0,1}>0} \|\Psi_\alpha^{nA_0^{(0)}}\|_{S([-T_0,T_1]\times\mathbb{R}^2)} \leq C_1(E_{crit})$$

Finally, $\Psi_\alpha^{nA_0^{(0)}}$ has essential Fourier support contained in $(0,(\lambda_n^1)^{-1})$. More precisely, for some sequence $\{R_n\}_{n=1}^\infty$ going to ∞ sufficiently slowly, one has

$$\|P_k \Psi_\alpha^{nA_0^{(0)}}\|_{S[k]} \leq R_n^{-1} e^{-\sigma|k-\mu_n^1|} \tag{9.10}$$

for all $k > \mu_n^1 = -\log\lambda_n^1$ and some absolute constant σ. As usual, all functions belong to the Schwartz class on fixed time slices.

The proof of this result will occupy this entire section. The idea is to run an induction in ℓ on a sequence of approximating maps with data $\Psi_\alpha^{\ell n A_0^{(0)}}$, see (9.9). As we start from the low frequencies, it will turn out that the differences between two consecutive such approximating components is of small energy (provided $\delta_1 \gg \delta_0$ are both sufficiently small). This allows us to pass from one approximation to the next better one by applying a perturbative argument, albeit with a linear operator involving a magnetic potential. Moreover, we need to divide the time-axis into a number of intervals which is controlled by the total energy. A key fact here which prevents energy build-up as we pass from one time interval to the next, is that the differences between these approximating components essentially preserve their energy, see Corollary 9.13. The approximate energy conservation, in turn, comes from the fact that the difference of consecutive approximating Coulomb components is essentially supported at much larger frequencies than the lower frequency approximating components. For the remainder of this section we drop the superscript $A_0^{(0)}$ from our notation since we will limit ourselves entirely to the low frequency part. We begin by showing that (still at time $t = 0$) the step from $\Psi_\alpha^{\ell-1,n}$ to $\Psi_\alpha^{\ell,n}$ amounts to adding on a term of much larger frequency, up to small errors in energy.

Lemma 9.10. *One has*

$$\Psi_\alpha^{\ell,n} - \Psi_\alpha^{\ell-1,n} = \epsilon_\alpha^{\ell,n} = P_{J_\ell}\epsilon_\alpha^{\ell,n} + \tilde{\epsilon}_\alpha^{\ell,n}$$

with $\|\tilde{\epsilon}_\alpha^{\ell,n}\|_{L_x^2} \lesssim \delta_1$. *Furthermore,*

$$\Psi_\alpha^{\ell-1,n} = P_{\cup_{j\le \ell-1} J_j}\Psi_\alpha^{\ell-1,n} + \tilde{\Psi}_\alpha^{\ell-1,n}$$

with $\|\tilde{\Psi}_\alpha^{\ell-1,n}\|_{L_x^2} \lesssim \delta_1$. *Similar statements hold on the level of the* Φ*-components.*

Proof. In view of Lemma 9.8 we may switch from $\Psi^{\ell,n}$ to the corresponding expressions involving w^n. For simplicity, write

$$\sum_{j\le\ell} P_{J_j} w_\alpha^n \, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\le\ell} P_{J_j} w_k^n} =: f_\ell e^{ig_\ell}$$

with g_ℓ real-valued. Since the Fourier support of f_ℓ is contained in $\cup_{j\le\ell} J_j = (-\infty, b_\ell]$, for any $k \ge b_\ell + 10$ one has

$$\begin{aligned}\|P_k(f_\ell\, e^{ig_\ell})\|_2 &\lesssim \|f_\ell P_{k+O(1)}\, e^{ig_\ell}\|_2 \lesssim 2^{-k}\|f_\ell\|_2 \|\nabla P_{k+O(1)} e^{ig_\ell}\|_\infty \\ &\lesssim 2^{-k}\|f_\ell\|_2 \|\Delta^{-1} D^2 f_\ell\|_\infty \lesssim 2^{-k}\|f_\ell\|_2\|\Delta^{-1}D^3 f_\ell\|_2 \lesssim 2^{b_\ell - k}\|w^n\|_2^2 \\ &\lesssim E_{\mathrm{crit}}\, 2^{b_\ell - k}\end{aligned}$$

where E_{crit} controls the total energy, and thus also the L^2-norm of w^n. By construction of w_α^n, one has for any $L > 0$

$$\limsup_{n\to\infty} \left\| P_{[b_\ell - L, b_\ell]} w_\alpha^n \right\|_2 \lesssim L\delta_0$$

Together with the preceding bound this implies that

$$\begin{aligned}&\Big\| \sum_{j\le\ell} P_{J_j} w_\alpha^n \, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\le\ell} P_{J_j} w_k^n} - P_{\cup_{j\le\ell} J_j} \sum_{j\le\ell} P_{J_j} w_\alpha^n \, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\le\ell} P_{J_j} w_k^n} \Big\|_2 \\ &\lesssim \log[(E_{\mathrm{crit}} + 1)\delta_0^{-1}]\delta_0 \ll \delta_1\end{aligned}$$

for small δ_0. Next, observe that

$$\begin{aligned}&\sum_{j\le\ell} P_{J_j} w_\alpha^n \, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\le\ell} P_{J_j} w_k^n} \\ &= \sum_{j\le\ell-1} P_{J_j} w_\alpha^n \, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\le\ell} P_{J_j} w_k^n} + P_{J_\ell} w_\alpha^n \, e^{-i\mathrm{Re}\sum_{k=1,2}\triangle^{-1}\partial_k \sum_{j\le\ell} P_{J_j} w_k^n}\end{aligned}$$

The first assertion of the lemma therefore follows from the following claims:

- The function
$$P_{J_\ell} w^n_\alpha \, e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k \sum_{j\leq \ell} P_{J_j} w^n_k}$$
has frequency support in $J_\ell = [a_\ell, b_\ell]$ up to exponentially decaying errors, and we also have
$$\limsup_{n\to\infty} \big\| P_{J^c_\ell}\big[P_{J_\ell} w^n_\alpha \, e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k \sum_{j\leq \ell} P_{J_j} w^n_k}\big]\big\|_{L^2_x} < \delta_1$$
- Furthermore, we have
$$\Big\| \sum_{j\leq \ell-1} P_{J_j} w^n_\alpha e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k \sum_{j\leq \ell} P_{J_j} w^n_k} - \sum_{j\leq \ell-1} P_{J_j} w^n_\alpha e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k \sum_{j\leq \ell-1} P_{J_j} w^n_k}\Big\|_{L^2_x} < \delta_1$$
for n large enough.

As for the first claim, note that we have already dealt with the case of frequencies larger than b_ℓ. Thus, assume that $j \leq a_\ell - 10$ and estimate
$$\begin{aligned}\|P_j(P_{J_\ell} w^n \, e^{ig_\ell})\|_2 &\lesssim 2^{\frac{j}{3}} \|P_{J_\ell} w^n \, e^{ig_\ell}\|_{\frac{3}{2}} \lesssim 2^{\frac{j}{3}} \|P_{J_\ell} w^n\|_2 \|P_{J_\ell + O(1)} e^{ig_\ell}\|_6 \\ &\lesssim 2^{\frac{j}{3}} \|P_{J_\ell} w^n\|_2 \sum_{k\in J_\ell + O(1)} 2^{-\frac{k}{3}} \|P_k \nabla e^{ig_\ell}\|_2 \lesssim E_{\mathrm{crit}} \, 2^{\frac{j-a_\ell}{3}}\end{aligned}$$
Furthermore, as before one can "fudge at the edges" meaning
$$\limsup_{n\to\infty} \big\| P_{[a_\ell, a_\ell + L]} w^n_\alpha \big\|_2 \lesssim L\delta_0$$
which concludes the first claim. For the second claim we need to show
$$\Big\| \sum_{j\leq \ell-1} P_{J_j} w^n_\alpha \, e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k \sum_{j\leq \ell-1} P_{J_j} w^n_k} \big(1 - \, e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k P_{J_\ell} w^n_k}\big)\Big\|_{L^2_x} \lesssim \delta_1$$
where the implied constant is absolute (not depending on any of the other parameters). However, this follows easily from the frequency localization up to exponentially decaying errors of
$$\sum_{j\leq \ell-1} P_{J_j} w^n_\alpha \, e^{-i\mathrm{Re}\sum_{k=1,2} \triangle^{-1}\partial_k \sum_{j\leq \ell-1} P_{J_j} w^n_k}$$
as well as the fact that
$$\limsup_{n\to\infty} \|P_{[a_\ell, a_\ell + L]\cup[b_\ell - L, b_\ell]} P_{J_\ell} w^n_k\|_{L^2_x} \lesssim L\delta_0$$
and we are done. The claim of the lemma about Φ is easier since it does not involve any phases, cf. Lemma 9.8 and Lemma 9.7. □

Our strategy now is to inductively control the nonlinear evolution of the $\Psi^{\ell,n}_\alpha$, the Coulomb components of the approximation maps, starting with $\ell = 1$. At each induction step we add a term $\epsilon^{\ell,n}_\alpha$ of energy less than ε_0. The key then is the following perturbative result. Recall that $\varepsilon_0 > 0$ is a small constant which determines the perturbative energy-cutoff (it depends on E_{crit}).

Proposition 9.11. *Let $\Psi^{\ell,n}_\alpha$, $\epsilon^{\ell,n}_\alpha$, be as before, with $1 \leq \ell \leq C_1(E_{crit}, \varepsilon_0)$. Also, let*
$$c^{(\ell-1)}_k := \max_\alpha \Big(\sum_{r\in\mathbb{Z}} 2^{-\sigma|r-k|} \|P_r P_{\cup_{j\leq \ell-1} J_j} \Psi^{\ell-1,n}_\alpha\|^2_{L^2_x}\Big)^{\frac{1}{2}}$$
for some small enough constant $\sigma > 0$ (an a priori constant). We now make the following induction hypotheses, valid for all large n: there is a decomposition $\Psi^{\ell-1,n}_\alpha = \tilde{\Psi}^{\ell-1,n}_\alpha + \breve{\Psi}^{\ell-1,n}_\alpha$ so that
$$\max_\alpha \|P_k \tilde{\Psi}^{\ell-1,n}_\alpha\|_{S[k]([-T_0,T_1]\times\mathbb{R}^2)} < C_2 \, c^{(\ell-1)}_k \tag{9.11}$$
$$\|\breve{\Psi}^{\ell-1,n}_\alpha\|_S < C_2 \, \delta_1 \tag{9.12}$$
for some positive number C_2.

Then there exists a partition $\Psi^{\ell,n}_\alpha = \tilde{\Psi}^{\ell,n}_\alpha + \breve{\Psi}^{\ell,n}_\alpha$ *so that*

$$\max_\alpha \|P_k \tilde{\Psi}^{\ell,n}_\alpha\|_{S[k]([-T_0,T_1]\times\mathbb{R}^2)} < C_3\, c_k^{(\ell)} \tag{9.13}$$

$$\|\breve{\Psi}^{\ell,n}_\alpha\|_S < C_3\, \delta_1 \tag{9.14}$$

provided $\delta_1 < \delta_1^0 = \delta_1^0(C_2, E_{crit})$, $\delta_0 \ll \delta_1$ *with* δ_0 *as in the discussion preceding Lemma 9.7, and provided* n *is sufficiently large. Here* $C_3 = C_3(C_2, E_{crit})$.

It is important to note that we iterate Proposition 9.11 $O(\frac{C_1(E_{crit}, \varepsilon_0)}{\varepsilon_0})$ many times, obtaining the induction start from the small data result of [22]. It is clear that there is some constant $\delta_{11} > 0$ (depending only on E_{crit}) such that choosing $\delta_1 < \delta_{11}$ in each step, this proposition can be applied. This $\delta_{11} > 0$ dictates our choice of $A^{(0)}$ in the decomposition

$$w^n_\alpha = \sum_j \phi_\alpha^{na_j^{(0)}} + w_\alpha^{nA^{(0)}}$$

from before, see (9.7). Another essential feature of the construction is that

$$\|\Psi^{\ell,n}\|_S \le K(E_{crit}) \tag{9.15}$$

where K is some rapidly growing function of the energy. This follows immediately from the inductive nature of the proof and the fact that the number of steps is controlled by the energy alone. However, it is crucial to the argument that we do not have to make ε_0 small depending on the function $K(E_{crit})$ as we go through the inductive process. In other words, we have to make sure that one can fix ε_0 throughout.
The idea of the proof of Proposition 9.11 is as follows: under the assumptions (9.11) and (9.12) we can find time intervals $I_1, I_2, \ldots, I_{M_1}$, $M_1 = M_1(\tilde{C}_2)$ as in Section 7, such that locally on I_j,

$$\Psi^{\ell-1,n} = \Psi_L^{\ell-1,n} + \Psi_{NL}^{\ell-1,n}$$

Here ψ_L is a linear wave and ψ_{NL} is small in a suitable sense, see Lemma 7.6 and Corollary 7.27. In order to control the evolution of $\Psi^{\ell,n}_\alpha$, we need to control the evolution of

$$\epsilon^{\ell,n}_\alpha = \Psi^{\ell,n}_\alpha - \Psi^{\ell-1,n}_\alpha$$

This we do inductively, over each interval I_j, starting with the one containing the initial time slice $t = 0$. At this point one encounters the danger that the energy of $\epsilon^{\ell,n}_\alpha$ keeps growing as we move to later (or earlier) intervals I_j, thereby effectively leaving the perturbative regime. The idea here is that we have *a priori energy conservation for the components* $\Psi^{\ell-1,n}$, $\Psi^{\ell,n}$, while at the same time, due to our assumptions on the frequency distribution of energy for $\Psi^{\ell-1,n}$, $\epsilon^{\ell,n}_\alpha$, *there cannot be much energy transfer between the latter two types of components*; more precisely, we can enforce this by choosing δ_1 small enough. This means that we have effectively *approximate energy conservation for* $\epsilon^{\ell,n}_\alpha$, whence the induction can be continued to all the I_j. We can now begin the proof in earnest.

Proof. (Proposition 9.11) We inductively control the nonlinear evolution of $\epsilon^{\ell,n}_\alpha$. For ease of notation, we set $\epsilon_\alpha := \epsilon^{\ell,n}_\alpha$ and $\psi_\alpha := \Psi^{\ell-1,n}_\alpha$ and for the most part we also ignore the α subscript. Note that while ψ exists globally in time, ϵ exists only locally in time but we will of course need to prove global existence and bounds for ϵ. But for now, any statement we make for ϵ will be locally in time on some interval I_0 around $t = 0$. Applying the divisibility statements Lemma 7.6 and Corollary 7.27 to ψ generates a decomposition of $\mathbb{R}$ into intervals $\{I_j\}_{j=1}^M$ where $M = M(\varepsilon_0, \|\psi\|_S)$. We may of course intersect these intervals with I_0 which we will tacitly assume. Fix one of these intervals, say I_1, which contains $t = 0$. It will of course be necessary for us to pass to later intervals in the temporal sense until we have exhausted the entire existence interval I_0. In other words, our induction has two direction, namely a temporal one (referring to the interval I_j), as well as a frequental one (referring to the interval J_ℓ). These two directions are indicated as vertical and horizontal ones, respectively, in Figure 6.
By construction, there is a decomposition

$$\psi = \psi_L + \psi_{NL} \tag{9.16}$$

FIGURE 6. The two directions of the induction

where $\|\psi_L\|_S \le \varepsilon_2^{-\frac{1}{4}} E_{\mathrm{crit}}^2$ and such that $\|\psi_{NL}\|_S^2 < \varepsilon_2$,

$$\|\psi_{NL}\|_S \|\psi_L\|_S < \varepsilon_2^{\frac{1}{4}} \tag{9.17}$$

Here ε_2 is small depending on E_{crit} and with $\varepsilon_0 \ll \varepsilon_2 \ll 1$. We note the following important improvement over (9.15):

$$\max_j \|\Psi^{\ell,n}\|_{S(I_j)} \lesssim \varepsilon_2^{-\frac{1}{4}} E_{\mathrm{crit}}^2 \tag{9.18}$$

Thus by restricting ourselves to one of the intervals I_j, we have essentially much reduced the nonlinear behavior of the Ψ. Proposition 9.11 will follow from a bootstrap argument, which is based on the following crucial result. Recall that J_ℓ is the Fourier support of $\epsilon(0)$ up to errors which can be made arbitrarily small in energy.

Proposition 9.12. *Let ψ satisfy the inductive assumptions* (9.11) *and* (9.12) *and let ϵ be defined as above. Suppose there is a decomposition $\epsilon = \epsilon_1 + \epsilon_2$ which satisfies the bounds*

$$\begin{aligned} \|\epsilon_2\|_{S(I_1\times\mathbb{R}^2)} &< C_2 C_4\, \delta_1 \\ \|P_k \epsilon_1\|_{S[k](I_1\times\mathbb{R}^2)} &\le C_4 d_k \quad \forall\, k \in \mathbb{Z} \end{aligned} \tag{9.19}$$

where we define

$$d_k := \Big(\sum_{r\in\mathbb{Z}} 2^{-\sigma|r-k|} \|P_r P_{J_\ell} \epsilon(0,\cdot)\|_{L_x^2}^2\Big)^{\frac{1}{2}}$$

for some $C_4 = C_4(E_{crit})$ sufficiently large, and some small absolute constant $\sigma > 0$. Then we can improve this to a similar decomposition with

$$\|\epsilon_2\|_{S(I_1\times\mathbb{R}^2)} < \frac{C_4}{2}\, C_2\, \delta_1, \qquad \|P_k \epsilon_1\|_{S[k](I_1\times\mathbb{R}^2)} \le \frac{C_4}{2} d_k \tag{9.20}$$

for all $k \in \mathbb{Z}$, provided we satisfy the smallness condition $\delta_1 < \delta_1(C_2, E_{crit})$ and $\delta_0 \ll \delta_1$ with δ_0 as in the discussion preceding Lemma 9.7.

This proposition is the key ingredient in the proof. It asserts that the frequency profile of ϵ at time $t = 0$ is essentially preserved under the evolution up to some frequency leakage, which however is controlled by the size of the *underlying Besov error.* What allows us to prevent energy of ϵ moving from high to low frequencies (which is the main difficulty here) are gains in the high-high-low interactions in the nonlinearities. Without these gains, there could indeed be this kind of energy transfer and the argument would break down. It is essential in Proposition 9.12 that C_4 is a constant that *does not* change throughout the induction, whereas C_2 *does* change.

If we accept Proposition 9.12 for now, then it is an easy matter to derive the aforementioned approximate energy conservation.

Corollary 9.13. *Under the induction hypothesis of Proposition 9.11 and assuming the validity of Proposition 9.12, one has the following: For sufficiently small δ_1 (depending on C_2 and C_4) and large n, we have*

$$\sum_{\alpha=0,1,2} \|\epsilon_\alpha\|^2_{L^\infty_t L^2_x(I_1\times\mathbb{R}^2)} < \varepsilon_0$$

where I_1 is as above.

Proof. (Corollary 9.13) Due to energy conservation for the evolution of $\psi+\epsilon$, we have

$$\sum_{\alpha=0,1,2} \|\psi_\alpha+\epsilon_\alpha\|^2_{L^2_x} = \text{constant}$$

Similarly, we have

$$\sum_{\alpha=0,1,2} \|\psi_\alpha\|^2_{L^2_x} = \text{constant}$$

The crucial observation now is that

$$\|\psi+\epsilon\|^2_{L^2_x} = \|\psi\|^2_{L^2_x} + \|\epsilon\|^2_{L^2_x} + 2\mathrm{Re}\sum_{k\in\mathbb{Z}}\int_{\mathbb{R}^2} P_k\psi\overline{P_k\epsilon}\,dx$$

on fixed time slices $t=t_0\in I_1$, and we can split

$$\sum_{k\in\mathbb{Z}}\int_{\mathbb{R}^2} P_k\psi\overline{P_k\epsilon}\,dx = \sum_{k\in\cup_{j\le l-1}J_j}\int_{\mathbb{R}^2} P_k\psi\overline{P_k\epsilon}\,dx + \sum_{k\in J_l}\int_{\mathbb{R}^2} P_k\psi\overline{P_k\epsilon}\,dx$$

Both contributions on the right are $\lesssim C_4C_2^2\delta_1$, which can be made arbitrarily small by choosing δ_1 small enough. To obtain this bound, observe that the induction hypothesis and Proposition 9.12 allow one to transfer Lemma 9.10 to all times in the interval I_1. Cauchy-Schwarz then implies the bound of $\lesssim C_4C_2^2\delta_1$. □

Corollary 9.13 allows us to keep the energy under control as we inductively pass from I_1 to its successor I_2 and so forth by restarting the procedure. Indeed, since the number of the "divisibility" intervals is bounded by $M(\varepsilon_0, E_{crit})$, we can make δ_1 in the corollary so small (depending on this number) and n so large that even the energy of the final ϵ is no bigger than $2\varepsilon_0$, say. Even though we will now work on I_1, all arguments carried out below apply to any of the later intervals $I_2, I_3, \ldots$ as well.

Proof of Proposition 9.12. We may reduce ourselves to proving the statement for frequency 2^0, i.e., $k=0$, by scaling invariance. Recall that we have chosen the intervals I_j in such fashion that (9.16) holds with the stated bounds. In order to obtain the desired estimates on ϵ, we distinguish between two cases, depending on the size of the underlying time interval. If it is short, we use the div-curl system. Otherwise we use the wave equation.

Case 1: $|I_1| < T_1$ where $T_1>0$ is some absolute small constant (to be specified). We shall use the div-curl system *linearized* around ψ, see (1.12), (1.13), which takes the schematic form

$$\partial_t\epsilon = \nabla_x\epsilon + \epsilon\nabla^{-1}(\psi^2) + \psi\nabla^{-1}(\psi\epsilon) + \epsilon\nabla^{-1}(\psi\epsilon) + \psi\nabla^{-1}(\epsilon^2) + \epsilon\nabla^{-1}(\epsilon^2)$$

The first linear term $\nabla_x\epsilon$ on the right-hand side is estimated by bootstrap, choosing T_1 smaller than some absolute constant. For each of the five nonlinear terms on the right-hand side one needs to consider two cases, depending on whether ϵ gets replaced by ϵ_1 or ϵ_2.

In light of the first part of the proof of Lemma 7.6, it suffices to prove bounds of either the form (all on the set $I_j \times \mathbb{R}^2$ and F any one of the expressions on the right above)

$$\Big(\sum_{k\in\mathbb{Z}}{}' \big[\|P_kF\|_{L_t^M \dot{H}^{\frac{1}{M}-2}+L_t^2\dot{H}^{-\frac{3}{2}}+L_t^\infty \dot{H}^{-2}}\big]^2\Big)^{\frac{1}{2}} \ll C_2C_4\delta_1$$

where the sum is over all those frequencies k such that $|I_j| < T_1 2^{-k}$, i. e. such that the re-scaled solution where k becomes 0 falls into Case 1, or else, we get

$$\|P_kF\|_{L_t^M \dot{H}^{\frac{1}{M}-2}+L_t^2\dot{H}^{-\frac{3}{2}}+L_t^\infty \dot{H}^{-2}} \ll C_4 d_k$$

We consider the frequency mode $k = 0$ and divide P_0F into pieces which get either substituted into the first or second bound above.

(a) The term $\epsilon_1\nabla^{-1}(\psi^2)$; we cannot just use Lemma 7.4 of Section 7, since smallness there can only be enforced by choosing T_1 very small, which is counter productive in Case 2, when we work on a larger interval. Hence we have to exploit the divisibility of the expression, which forces us to exploit the hidden null-structure. However, we can easily conclude from the proof of Lemma 7.4 that

$$\|P_0[\epsilon_1\nabla^{-1}P_{<-C}(\psi^2)]\|_{L_t^M L_x^2} \ll d_0$$

provided we pick $C = C(E_{crit})$ sufficiently large, and thence

$$\|\int_0^t P_0[\epsilon_1\nabla^{-1}P_{<-C}(\psi^2)]\,ds\|_{L_t^\infty L_x^2} \ll d_0$$

$$\|\int_0^t P_0[\epsilon_1\nabla^{-1}P_{<-C}(\psi^2)]\,ds\|_{L_t^2 L_x^2} \ll d_0$$

for $t \in [-T_1, T_1]$, and from there

$$\|\int_0^t P_0[\epsilon_1\nabla^{-1}P_{<-C}(\psi^2)]\,ds\|_{S[0]} \ll d_0,$$

compare (7.10) (provided $T_1 < 1$, say). Similarly, one checks that the contribution of

$$P_0[\epsilon_1\nabla^{-1}P_{>C}(\psi^2)]$$

is acceptable, and so we now need to force smallness for

$$P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(\psi^2)],$$

which we do by subdivision into small time intervals (whose number depends on $\|\psi\|_S$). First, we observe that choosing C_1 large enough depending on C and E_{crit}, we can force that

$$\|P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(Q_{>C_1}\psi\psi)]\|_{L_{t,x}^2} \ll d_0,$$

and from here one can again infer that

$$\|\int_0^t P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(Q_{>C_1}\psi\psi)]\,ds\|_{S[0]} \ll d_0$$

for $t \in [-T_1, T_1]$, $T_1 < 1$, say. The same applies to

$$P_0[Q_{>C_1}\epsilon_1\nabla^{-1}P_{[-C,C]}(\psi^2)]$$

Hence we may reduce to considering

$$P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(\psi^2)]$$

where we automatically assume that $\psi = Q_{<C_1}\psi$, $\epsilon_1 = Q_{<C_1}\epsilon_1$. Now we implement the customary Hodge decomposition

$$\psi_\nu = R_\nu\psi + \chi_\nu$$

First, substitute the gradient term for either factor ψ, which results in the expression

$$P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}\mathcal{Q}_{\nu j}(\psi,\psi)]$$

Now due to Lemma 4.17 etc that in case of high-low or low-high interactions inside $\mathcal{Q}_{\nu j}(\psi,\psi)$ we can estimate

$$\|P_{[-C,C]}\mathcal{Q}_{\nu j}(\psi,\psi)\|_{L^2_{t,x}}\lesssim\|\psi\|_S^2$$

and one may then pick time intervals I_j with the property that

$$\sum_{k\in\mathbb{Z}}\|\chi_{I_j}P_{[k-C,k+C]}\mathcal{Q}_{\nu j}(\psi,\psi)\|_{L^2_{t,x}}^2\ll 1$$

which ensures "divisibility". Thus it remains to deal with the expression

$$P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}\mathcal{Q}_{\nu j}(P_{>C_2}\psi,P_{>C_2}\psi)]$$

and indeed in light of Lemma 4.17 etc only the case when $\nu=0$ needs to be considered. We choose $C_2\gg\max\{C,C_1\}$. Note that in this case the inner null-form may have very large modulation (comparable to the frequency of the inputs), in which case we cannot take advantage of the null-structure. The idea then is to use the smoothing effect of integration over time. Specifically, we write schematically

$$\begin{aligned}
&P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}\mathcal{Q}_{\nu j}(P_{>C_2}\psi,P_{>C_2}\psi)]\\
&=P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}\partial_t(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)]\\
&-P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}\partial_j(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_0\psi)]\\
&=P_0\partial_t[\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)] &&(9.21)\\
&-P_0[\partial_t\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)] &&(9.22)\\
&-P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}\partial_j(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_0\psi)] &&(9.23)
\end{aligned}$$

Now it is straightforward to analyze the contribution of each term, keeping in mind our assumptions about hyperbolicity of each input. For the contribution of (9.21), note that we have

$$\begin{aligned}
&\int_0^t P_0\partial_s[\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)]\,ds\\
&=P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)](t,\cdot)\\
&-P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)](0,\cdot)
\end{aligned}$$

and we can then crudely bound (assuming $T_1<1$, say)

$$\begin{aligned}
&\|\chi_{[-T_1,T_1]}\big[P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)](t,\cdot)\\
&-P_0[\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)](0,\cdot)\big]\|_{L^2_{t,x}}\ll d_0
\end{aligned}$$

This again suffices for the bootstrapping.
Next, for the expression (9.22), we estimate it by

$$\begin{aligned}
&\|\chi_{[-T_1,T_1]}P_0[\partial_t\epsilon_1\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)]\|_{L^2_{t,x}}\\
&\lesssim\|\partial_t\epsilon_1\|_{L^\infty_tL^2_x}\|\nabla^{-1}P_{[-C,C]}(P_{>C_2}|\nabla|^{-1}\psi P_{>C_2}R_j\psi)\|_{L^\infty_tL^2_x}\\
&\ll d_0
\end{aligned}$$

Finally, expression (9.23) is more of the same (due to the hyperbolicity of the inputs) and omitted.
We next consider the contribution of the terms arising when the substitute an elliptic term χ_ν for ψ_ν inside $\nabla^{-1}(\psi^2)$. This leads to an expression of the schematic form

$$P_0\big[\varepsilon_1\nabla^{-1}\big(\nabla^{-1}[\psi\nabla^{-1}(\psi^2)]\psi\big)\big]$$

However, as is easily verified, we have

$$\|P_0\big[P_{k_1}\varepsilon_1\nabla^{-1}P_k\big(\nabla^{-1}P_r[\psi\nabla^{-1}(\psi^2)]\psi\big)\big]\|_{L^2_{t,x}}\lesssim 2^{-\sigma(|k_1|+|k|)}2^{-\delta(|k-k_1|)}\|P_{k_1}\varepsilon_1\|_{S[k_1]}\|P_r[\psi\nabla^{-1}(\psi^2)]\|_{L^2_t\dot{H}^{-\frac12}},$$

for suitable $\sigma > 0$, $\delta > 0$, whence we get

$$\|P_0\big[P_{k_1}\varepsilon_1\nabla^{-1}\big(\nabla^{-1}[\psi\nabla^{-1}(\psi^2)]\psi\big)\|_{L^2_{t,x}(I_j\times\mathbb{R}^2)} \lesssim 2^{-\sigma|k_1|}\|P_{k_1}\varepsilon_1\|_{S[k_1]}\big(\sum_{k\in\mathbb{Z}}\|[\psi\nabla^{-1}(\psi^2)]\|^2_{L^2_t\dot{H}^{-\frac{1}{2}}(I_j\times\mathbb{R}^2)}\big)^{\frac{1}{2}}$$

Using Lemma 7.26, we can then arrange that the right hand side is $\ll d_0$, as desired.

The corresponding estimate for $\epsilon_2\nabla^{-1}(\psi^2)$ is essentially the same, the only difference being that one square-sums over the frequencies at the end.

We recall here how one infers the desired bound on ε in the small-time case as in the proof of lemma 7.6 from the above considerations: letting $\eta(t)\in C_0^\infty(\mathbb{R})$ be a (potentially very sharp) cutoff localizing to a sufficiently close dilate of the interval I_1, and letting $\eta_1(t)$ be a cutoff localizing to an interval of length ~ 1 centered at $t=0$, we write

$$P_0\varepsilon(t,\cdot) = \eta_1(t)\big[P_0\varepsilon(0,\cdot) + \int_0^t \eta(s)\nabla_x P_0\tilde{\varepsilon}(s,\cdot)\,ds + \int_0^t \eta(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds + \dots\big]$$

where $\tilde{\varepsilon}$ is a Schwartz extension of ε satisfying the bootstrap estimate; more precisely, we can split $\tilde{\varepsilon}=\tilde{\varepsilon}_1+\tilde{\varepsilon}_2$ with each one satisfying suitable bootstrap estimates as in the Proposition. The estimate for the first time-dependent term $\eta_1(t)\int_0^t\eta(s)\nabla_x P_0\tilde{\varepsilon}(s,\cdot)\,ds$ is immediate:

$$\|\eta_1(t)\int_0^t \eta(s)\nabla_x P_0\tilde{\varepsilon}(s,\cdot)\,ds\|_{S[0](I_1\times\mathbb{R}^2)} \lesssim \|\eta\|_{L^2_t}\|\nabla_x P_0\tilde{\varepsilon}\|_{L^\infty_t L^2_x} \ll \|P_0\tilde{\varepsilon}\|_{S[0](\mathbb{R}^{2+1})}$$

Next, we can again crudely bound

$$\begin{aligned}&\|\eta_1(t)\int_0^t \eta(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds\|_{S[0](I_1\times\mathbb{R}^2)}\\ &\lesssim \|\eta_1(t)\int_0^t \eta(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds\|_{L^2_{t,x}} + \|\partial_t\big[\eta_1(t)\int_0^t \eta(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds\big]\|_{L^2_{t,x}}\end{aligned}$$

The only difference of this compared to the estimates above is the inclusion of the cutoff $\eta(s)$, which may, however, be very sharp. To deal with this, introduce a $C=C(E_{crit})$ sufficiently large, and split

$$\eta_1(t)\int_0^t \eta(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds = \eta_1(t)\int_0^t Q_{<C}(\eta)(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds + \eta_1(t)\int_0^t Q_{\geq C}(\eta)(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds$$

The second term here leads to a contribution that is bounded by

$$\lesssim \|Q_{\geq C}(\eta)\|_{L^2_t}\|P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\|_{L^\infty_t L^2_x} \ll \sup_{k\in\mathbb{Z}} 2^{-\sigma|k|}\|P_k\tilde{\varepsilon}\|_{S[k](\mathbb{R}^{2+1})}$$

For the first term above, $\eta_1(t)\int_0^t Q_{<C}(\eta)(s)P_0[\tilde{\varepsilon}\nabla^{-1}(\psi^2)]\,ds$, one implements the null-structure and performs integrations by parts exactly as explained in the first part of case (a) above. Note that when we hit the expression localized to modulation $\sim 2^j$, $j\gg 1$ with a time derivative, the definition of $S[0]$ gives us an extra weight of $2^{-\varepsilon j}$, which gives the necessary gain in $-k_2$ in the bad high-high interaction case (where $j=k_2+O(1)$).

In the remaining cases (b) and (c), we shall omit this last step (i. e. writing the Schwartz extension of $\varepsilon|_{I_1}$ explicitly), as the details are always quite similar. However, we describe it again in detail in case (d), which is slightly different.

(b) The term $\psi\nabla^{-1}(\psi\epsilon_1)$ as well as $\psi\nabla^{-1}(\psi\epsilon_2)$ both will be placed in the ϵ_2 component, meaning that we will prove that they have small S-norm. We start with ϵ_1. We claim that

$$\|P_0[\psi\nabla^{-1}(P_{k_2}\psi P_{k_3}\epsilon_1)]\|_{L^M_t L^2_x} \lesssim 2^{-\sigma_0|k_2-k_3|}\|P_{k_2}\psi\|_{S[k_2]}\|P_{k_3}\epsilon_1\|_{S[k_3]}\sup_{k_1\in\mathbb{Z}} 2^{-\sigma_0|k_1|}\|P_{k_1}\psi\|_{S[k_1]} \tag{9.24}$$

for some $\sigma_0>0$. This follows by inspecting the proof of Lemma 7.4. If $|k_2-k_3|>B|\log\delta_1|$ where B is large, one concludes from (9.24) that

$$\sum_{|k_2-k_3|>C|\log\delta_1|}\|P_0[\psi\nabla^{-1}(P_{k_2}\psi P_{k_3}\epsilon_1)]\|_{L^M_t L^2_x} \lesssim C_4\,\delta_1^{B\sigma_0}\|\psi\|_S\|\epsilon_1(0)\|_2 \sup_{k_1\in\mathbb{Z}} 2^{-\sigma_0|k_1|}\|P_{k_1}\psi\|_{S[k_1]}$$

Replacing P_0 by P_k and square summing in k yields a bound of

$$C_4\,\delta_1^{B\sigma_0}\|\psi\|_S^2\|\epsilon_1(0)\|_2 \ll C_4 C_2\,\delta_1$$

for the contribution of this case. This can be done by choosing B large depending on E_{crit}, see (9.18). On the other hand, if $|k_2-k_3|\le B|\log\delta_1|$, then we exploit that the Fourier supports of ψ and ϵ are essentially disjoint up to small errors (bounded by $\lesssim \delta_1$ in the S-norm) and exponentially decaying tails. Now we sum (9.24) over this range to obtain

$$\sum_{|k_2-k_3|\le B|\log\delta_1|}\|P_0[\psi\nabla^{-1}(P_{k_2}\psi P_{k_3}\epsilon_1)]\|_{L_t^M L_x^2} \lesssim \sum_{|k_2-k_3|\le B|\log\delta_1|}\|P_0[\psi\nabla^{-1}(P_{k_2}\breve{\psi}P_{k_3}\epsilon_1)]\|_{L_t^M L_x^2} \tag{9.25}$$

$$+\sum_{|k_2-k_3|\le B|\log\delta_1|}\|P_0[\psi\nabla^{-1}(P_{k_2}\tilde{\psi}P_{k_3}\epsilon_1)]\|_{L_t^M L_x^2} \tag{9.26}$$

For (9.25) one obtains as above

$$(9.25)\lesssim \|\breve{\psi}\|_S\|\epsilon_1(0)\|_2\sup_{k_1\in\mathbb{Z}}2^{-\sigma_0|k_1|}\|P_{k_1}\psi\|_{S[k_1]}$$

with an absolute implicit constant. Replacing P_0 with P_k and summing over all scales yields the bound

$$\lesssim \|\psi\|_S\|\breve{\psi}\|_S\|\epsilon_1(0)\|_2 \lesssim \varepsilon_2^{-\frac14}E_{\rm crit}^2\,C_2\,\delta_1\,\varepsilon_0 \ll C_2C_4\,\delta_1$$

provided we choose $\varepsilon_2^{-\frac14}E_{\rm crit}^2\varepsilon_0 \ll C_4$. Next, by the definition of the frequency envelopes c_k and d_k,

$$\begin{aligned}
(9.26) &\lesssim \sup_{k_1\in\mathbb{Z}}2^{-\sigma_0|k_1|}\|P_{k_1}\psi\|_{S[k_1]}\sum_{|k_2-k_3|\le B|\log\delta_1|}2^{-\sigma_0|k_2-k_3|}\|P_{k_2}\tilde{\psi}\|_{S[k_2]}\|P_{k_3}\epsilon_1\|_{S[k_3]}\\
&\lesssim \sup_{k_1\in\mathbb{Z}}2^{-\sigma_0|k_1|}\|P_{k_1}\psi\|_{S[k_1]}\sum_{k_2-k_3|\le B|\log\delta_1|}2^{-\sigma_0|k_2-k_3|}C_2\,c_{k_2}^{(\ell-1)}\,C_4 d_{k_3}\\
&\lesssim C_2C_4\,\delta_0\sup_{k_1\in\mathbb{Z}}2^{-\sigma_0|k_1|}\|P_{k_1}\psi\|_{S[k_1]}
\end{aligned}$$

This follows from the fact that δ_0 was chosen to control the Besov norm of $w_\alpha^{nA^{(0)}}$, as well as the fact that the intervals J_i where chosen in such a way that any of the smaller atoms contained within $w_\alpha^{nA_0^{(0)}}$ are arbitrarily far away from the endpoints of J_i as $n\to\infty$. Rescaling this bound to P_k from P_0 and square summing yields a bound of $C_2C_4\,\|\psi\|_S\delta_0 \ll C_2C_4\delta_1$ by taking δ_0 small enough, cf. (9.18).

Next, we turn to $\psi\nabla^{-1}(\psi\epsilon_2)$. Here the smallness comes from "divisibility" again as in case (a). More precisely, reasoning as in (a), we may reduce this expression to the form

$$P_0[\psi\nabla^{-1}P_{[-C,C]}(\psi\epsilon_2)]$$

where we moreover have $\psi = Q_{<C_1}\psi$, $\epsilon_2 = Q_{<C_1}\epsilon_2$. Again the argument from (a) shows that we may assume both inputs of $P_{[-C,C]}(\psi\epsilon_2)$ to have frequency $O(1)$ (implied constant depending on C, C_1, and E_{crit}). Furthermore, it is straightforward to check that if the two factors ψ have closely aligned Fourier supports, we obtain the desired smallness via Bernstein's inequality. But if the Fourier supports of the two ψ have some angular separation, interpreting the operator $\nabla^{-1}P_{[-C,C]}$ as convolution with a kernel $K(x)$ of bounded (although possibly large) L^1-mass, we may write

$$P_0[\psi\nabla^{-1}P_{[-C,C]}(\psi\epsilon_2)] = \int_{\mathbb{R}^2}P_0[\psi(\cdot,x)K(y)(\psi(\cdot,x-y)\epsilon_2)(\cdot,x-y)]\,dy$$

and then

$$\|\psi(\cdot,x)\psi(\cdot,x-y)\|_{L^2_{t,x}} \lesssim \|\psi\|_S^2,$$

which follows from our assumption about the Fourier supports, as well as the fact that both frequencies here are $< O(1)$. But then we can again force smallness by picking the I_j suitably, such that

$$\sum_{k\in\mathbb{Z}}\|\chi_{I_j}\int_{\mathbb{R}^2}P_k[P_{<k+O(1)}\psi(\cdot,x)|K(y)|P_{<k+O(1)}\psi(\cdot,x-y)]\,dy\|^2_{L_t^2\dot{H}^{-\frac12}} \ll 1$$

By replacing the output frequency 0 by k and square summing over all frequencies for which the Case 1 condition $|I_j| < T_1 2^{-k}$ is satisfied, we have then achieved that

$$\Big(\sum_k{}' \|P_k\big[\psi\nabla^{-1}(\psi\epsilon_2)\big]\|^2_{L^2_t\dot H^{-\frac32}}\Big)^{\frac12} \ll C_2C_4\delta_1$$

(c) The term $\epsilon_1\nabla^{-1}(\psi\epsilon_1)$ is easy, since it inherits the frequency profile of ϵ_1. More precisely, using the same type of trilinear estimates as in (a) and (b) one obtains

$$\|P_k(\epsilon_1\nabla^{-1}(\psi\epsilon_1))\|_{S[k](I_1\times\mathbb{R}^2)} \lesssim C_4\, d_k\|\psi\|_{L^\infty_t L^2_x}\|\epsilon_1\|_{L^\infty_t L^2_x} \ll C_4\, d_k$$

using (9.18) and the fact that $\|\epsilon_1\|_{L^\infty_t L^2_x} \le 2\varepsilon_0$ (taking δ_1 small). The other cases are easier due to the presence of δ_1 coming from ϵ_2.

(d) The term $\psi\nabla^{-1}(\epsilon^2)$ splits into the terms $\psi\nabla^{-1}(\epsilon_1^2)$, $\psi\nabla^{-1}(\epsilon_1\,\epsilon_2)$, and $\psi\nabla^{-1}(\epsilon_2^2)$. The last two are easier due to the smallness of ϵ_2. The first one is harder, as it inherits the frequency profile of ψ and therefore needs to be incorporated in ϵ_2. This means that we need to gain the very small δ_0, which is only possible if there are high-high gains in the inner term of $\psi\nabla^{-1}(\epsilon_1^2)$ resulting from ϵ_1. Of course, this requires that we expand this inner expression into a null-form via the usual Hodge decomposition.

(i): High-High-Low interactions in $\nabla^{-1}(\epsilon^2)$. This is the following (schematic) type of term:

$$\sum_{k,\,k_{1,2,3},\,k\ll k_2} P_0[P_{k_1}\psi\nabla^{-1}P_k(P_{k_2}\epsilon P_{k_3}\epsilon)].$$

It is straightforward to see that we may assume $|k| < \sigma_3 k_2$ for some $\sigma_3 > 0$ (absolute constant independent of the other smallness parameters), and furthermore $k_2 = k_3 + O(1) > B|\log\delta_1|$, since otherwise the desired smallness follows as in the preceding Case (b). We may thus essentially assume $k_1 = O(1)$, $k = O(1)$, and reduce to the simplified expression

$$\sum_{k_1=O(1)=k,\,k_2>B|\log\delta_1|} P_0[P_{k_1}\psi\nabla^{-1}P_k(P_{k_2}\epsilon P_{k_3}\epsilon)]$$

Suppressing the frequency localizations for now, we use the schematic relation

$$\begin{aligned} P_0\big[\psi\nabla^{-1}(\epsilon^2)\big] =& P_0\big[\psi\nabla^{-1}(R_\nu\epsilon^1R_j\epsilon^2 - R_j\epsilon^1R_\nu\epsilon^2) + \psi\nabla^{-1}(\nabla^{-1}(\epsilon\nabla^{-1}(\epsilon^2))R_\nu\epsilon) + \dots \\ &+ \psi\nabla^{-1}(\nabla^{-1}\big([\epsilon\nabla^{-1}(\epsilon^2)]^2\big)\big] + \dots \end{aligned}$$

where we omit the remaining quintilinear and septilinear terms. More precisely, we shall use this provided both inputs ϵ have relatively small modulation, i.e., are of hyperbolic type. In the immediately following we shall be a bit careless about the order in which we apply space-time frequency localizations and apply the Hodge decomposition. Due to the fact that the functions ε are a priori only defined locally in time, this is a potential technical issue (which did not come up when we applied the Hodge decomposition to the ψ's, as these are a priori defined globally in time). We shall explain how to del with this difficulty further below, when we explain how to construct the contribution to the actual Schwartz extension of ε from the present case. Thus for $k_2 = k_3 + O(1) > B|\log\delta_1|$, we write

$$P_0\big[P_{k_1}\psi\nabla^{-1}(P_{k_2}\epsilon P_{k_3}\epsilon)\big] = P_0\big[P_{k_1}\psi\nabla^{-1}(P_{k_2}Q_{>k_2}\epsilon P_{k_3}\epsilon)\big] + P_0\big[P_{k_1}\psi\nabla^{-1}(P_{k_2}Q_{<k_2}\epsilon P_{k_3}Q_{>k_3}\epsilon)\big] \tag{9.27}$$

$$+ P_0\big[P_{k_1}\psi\nabla^{-1}(R_\nu P_{k_2}Q_{<k_2}\epsilon R_jP_{k_3}Q_{<k_3}\epsilon - R_jP_{k_2}Q_{<k_2}\epsilon R_\nu P_{k_3}Q_{<k_3}\epsilon) \tag{9.28}$$

$$+ P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\epsilon\nabla^{-1}(\epsilon^2))R_\nu P_{k_3}Q_{<k_3}\epsilon) \tag{9.29}$$

$$+ P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\psi\nabla^{-1}(\epsilon^2))R_\nu P_{k_3}Q_{<k_3}\epsilon) \tag{9.30}$$

$$+ P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\epsilon\nabla^{-1}(\psi\epsilon))R_\nu P_{k_3}Q_{<k_3}\epsilon) \tag{9.31}$$

$$+ P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\epsilon\nabla^{-1}(\psi^2))R_\nu P_{k_3}Q_{<k_3}\epsilon) \tag{9.32}$$

$$+ P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\psi\nabla^{-1}(\psi\epsilon))R_\nu P_{k_3}Q_{<k_3}\epsilon) \tag{9.33}$$

$$+ P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}[\epsilon\nabla^{-1}(\epsilon^2)]\nabla^{-1}P_{k_3}Q_{<k_3}[\epsilon\nabla^{-1}(\epsilon^2)])\big] + \dots \tag{9.34}$$

where $\dots$ denotes the remaining septilinear terms containing mixed ψ-ϵ-interactions. Again we may substitute ϵ_1 everywhere for ϵ, the contributions from ϵ_2 leading to much smaller contributions. The first two terms on the right are straightforward to estimate: using Bernstein's inequality, one obtains for (9.27) the bound

$$\|P_0\big[P_{k_1}\psi\nabla^{-1}(P_{k_2}Q_{>k_2}\epsilon_1 P_{k_3}\epsilon_1)\big]\|_{L^2_{t,x}} \lesssim \min\{\|P_{k_1}\psi\|_{L_t^\infty L_x^2}, \|P_{k_1}\psi\|_{L_t^\infty L_x^\infty}\}\|P_{k_2}Q_{>k_2}\epsilon_1\|_{L^2_{t,x}}\|P_{k_3}\epsilon_1\|_{L_t^\infty L_x^2}$$
$$\lesssim 2^{-\frac{k_2}{2}}\min\{\|P_{k_1}\psi_1\|_{L_t^\infty L_x^2}, \|P_{k_1}\psi_1\|_{L_t^\infty L_x^\infty}\}\|P_{k_2}\epsilon_1\|_{S[k_2]}\|P_{k_3}\epsilon_1\|_{S[k_3]}$$

Keep in mind here we assume $k_1 = O(1)$. Then by an argument similar to the one used to estimate (9.26), replacing the output frequency by 2^k and square summing over $k = k_1 + O(1)$ while also summing over $|k_1 - k_2| > B|\log\delta_1|$, one can bound this contribution by $\lesssim C_2 C_4^2 \varepsilon_0^2 \delta_0$, which is enough to incorporate this term into ϵ_2. The second term in the expansion is of course handled identically, and so we now turn to the third term (9.28), which is the most delicate one. The potential difficulty comes when $\nu = 0$, as the $Q_{\nu j}$-null-form allows us to pull out one derivative otherwise; indeed, assume first that $\{\nu, j\} = \{1, 2\}$. Then using the identity (and omitting the subscript from ϵ for simplicity)

$$R_1\epsilon^1 R_2\epsilon^2 - R_1\epsilon^2 R_2\epsilon^1 = \partial_1[\nabla^{-1}\epsilon^1 R_2\epsilon^2] - \partial_2[\nabla^{-1}\epsilon^1 R_1\epsilon^2],$$

we can estimate (always under the assumption $k_1 = O(1) = k$)

$$\|P_0\big[P_{k_1}\psi\nabla^{-1}P_k(R_1P_{k_2}Q_{<k_2}\epsilon R_2P_{k_3}Q_{<k_3}\epsilon - R_2P_{k_2}Q_{<k_2}\epsilon R_1P_{k_3}Q_{<k_3}\epsilon)\|_{L^2_{t,x}}$$
$$\lesssim \|P_{k_1}\psi\|_{L_t^\infty L_x^2}\|P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_{1,2}P_{k_3}Q_{<k_3}\epsilon]\|_{L_t^2 L_x^\infty}$$

In order to estimate the right-hand factor, we use the improved Strichartz estimates: we have

$$P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_{1,2}P_{k_3}Q_{<k_3}\epsilon] = \sum_{\substack{c_{1,2}\in\mathcal{D}_{k_2,-k_2}\\ \mathrm{dist}(c_1,-c_2)=O(1)}} P_k[\nabla^{-1}P_{c_1}Q_{<k_2}\epsilon R_{1,2}P_{c_2}Q_{<k_3}\epsilon]$$

whence we get

$$\|P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_{1,2}P_{k_3}Q_{<k_3}\epsilon]\|_{L_t^2L_x^\infty} \lesssim 2^{-k_2}\Big(\sum_{c\in\mathcal{D}_{k_2,-k_2}}\|P_{k_2}Q_{<k_2}\epsilon\|^2_{L_t^4L_x^\infty}\Big)^{\frac{1}{2}}\Big(\sum_{c\in\mathcal{D}_{k_3,-k_3}}\|P_{k_3}Q_{<k_3}\epsilon\|^2_{L_t^4L_x^\infty}\Big)^{\frac{1}{2}}$$
$$\lesssim 2^{-\frac{k_2}{2+}}\prod_{j=2,3}\|P_{k_j}\epsilon\|_{S[k_j]},$$

whence we now have

$$\|P_0\big[P_{k_1}\psi\nabla^{-1}P_k(R_1P_{k_2}Q_{<k_2}\epsilon R_2P_{k_3}Q_{<k_3}\epsilon - R_2P_{k_2}Q_{<k_2}\epsilon R_1P_{k_3}Q_{<k_3}\epsilon)\|_{L^2_{t,x}}$$
$$\lesssim \|P_{k_1}\psi\|_{L_t^\infty L_x^2}2^{-\frac{k_2}{2+}}\prod_{j=2,3}\|P_{k_j}\epsilon\|_{S[k_j]}$$

From here one can again conclude as in case (b).
Hence we now consider the more difficult case where $\nu = 0$. First, it is straightforward to check that we may reduce the first input $P_{k_1}\psi$ to modulation $< 2^{\sigma_4 k_2}$, where for example we may put $\sigma_4 = \frac{1}{2}$. Then we use the schematic representation

$$P_0\big[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k(R_0P_{k_2}Q_{<k_2}\epsilon R_1P_{k_3}Q_{<k_3}\epsilon - R_1P_{k_2}Q_{<k_2}\epsilon R_0P_{k_3}Q_{<k_3}\epsilon)$$
$$= P_0\partial_t[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1P_{k_3}Q_{<k_3}\epsilon]] - P_0[P_{k_1}Q_{<\frac{k_2}{2}}\partial_t\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1P_{k_3}Q_{<k_3}\epsilon]]$$
$$- P_0[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_kR_1[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_0P_{k_3}Q_{<k_3}\epsilon]]$$

If one then integrates the transport equation for ϵ, the contribution from the above terms is

$$\begin{aligned}
&P_0[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1 P_{k_3}Q_{<k_3}\epsilon]](t,\cdot)\\
&-P_0[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1 P_{k_3}Q_{<k_3}\epsilon]](0,\cdot)\\
&-\int_0^t P_0[P_{k_1}Q_{<\frac{k_2}{2}}\partial_t\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1 P_{k_3}Q_{<k_3}\epsilon]](s,\cdot)\,ds\\
&-\int_0^t P_0[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k R_1[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_0 P_{k_3}Q_{<k_3}\epsilon]](s,\cdot)\,ds
\end{aligned}$$

But under our current assumption $k_1 = O(1)$, $k = O(1)$, we have the estimate (using Bernstein's inequality)

$$\begin{aligned}
&\|P_0[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1 P_{k_3}Q_{<k_3}\epsilon]](t,\cdot)\\
&-P_0[P_{k_1}Q_{<\frac{k_2}{2}}\psi\nabla^{-1}P_k[\nabla^{-1}P_{k_2}Q_{<k_2}\epsilon R_1 P_{k_3}Q_{<k_3}\epsilon]](0,\cdot)\|_{L_t^\infty L_x^2}\\
&\lesssim 2^{-k_2}\|P_{k_1}\psi\|_{L_t^\infty L_x^2}\|P_{k_2}\epsilon\|_{L_t^\infty L_x^2}\|P_{k_3}\epsilon\|_{L_t^\infty L_x^2}
\end{aligned}$$

and the remaining integral expressions on the right also easily lead to exponential gains in $-k_2$ due to the extra ∇^{-1} applied to $P_{k_2}Q_{<k_2}\epsilon$. Our assumption $k_2 > B|\log\delta_1|$ then allows us to incorporate the contribution of all these source terms into ϵ_2. Note that the cutoff $Q_{<\frac{k_2}{2}}$ in front of $\partial_t\psi$ allows us to control the effect of the ∂_t.

We explain here how to deal with the construction of the actual Schwartz extension of ε for the contribution of the preceding terms, since this is a bit more complicated than in case (a); thus as at the end of case (a) consider

$$\eta_1(t)\int_0^t \eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds$$

where we have a high-high interaction inside $\nabla^{-1}(\varepsilon^2)$ but all other frequencies are $O(1)$, as discussed in the preceding. In particular, we have $\varepsilon = P_{k_{2,3}}\varepsilon$ with $k_2 = k_3 + O(1) \geq B|\log\delta_1|$. We first observe that we are done provided $|I_1| \lesssim 2^{-\gamma k_2}$ for some small $\gamma > 0$, since then we get

$$\begin{aligned}
&\|\eta_1(t)\int_0^t \eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds\|_{S[0](I_1\times\mathbb{R}^2)} \lesssim \|\eta\|_{L_t^2}\|P_0[\psi\nabla^{-1}(\varepsilon^2)]\|_{L_t^\infty L_x^2}\\
&\lesssim 2^{-\frac{\gamma}{2}k_2}\big(\sup_{k\in Z}2^{-\sigma|k|}\|P_k\psi\|_{S[k]}\big) \leq \delta_1^{\frac{\gamma}{2}B_1}\big(\sup_{k\in Z}2^{-\sigma|k|}\|P_k\psi\|_{S[k]}\big),
\end{aligned}$$

which is more than enough for inclusion of this contribution into the ε_2-part. Next, fixing some $1 \gg \gamma' \gg \gamma$ and letting ϕ_1 be a smooth cutoff localizing to I_1 and which equals 1 for all t at distance $\geq 2^{-\gamma' k_2}$ from the endpoint of I_1, we write

$$\eta_1(t)\int_0^t \eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds = \eta_1(t)\int_0^t \phi_1\eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds + \eta_1(t)\int_0^t \phi_2\eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds,$$

with $\phi_2 = 1 - \phi_1$. Then as before we get

$$\|\eta_1(t)\int_0^t \phi_2\eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds\|_{S[0](I_1\times\mathbb{R}^2)} \lesssim \delta_1^{\frac{\gamma'}{2}B_1}\big(\sup_{k\in Z}2^{-\sigma|k|}\|P_k\psi\|_{S[k]}\big)$$

which is again more than enough to include this term into ε_2. Next, we decompose for some $1 \gg \gamma'' \gg \gamma'$

$$\begin{aligned}
&\eta_1(t)\int_0^t \phi_1\eta(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds\\
&= \eta_1(t)\int_0^t Q_{<\gamma'' k_2}(\phi_1\eta)(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds + \eta_1(t)\int_0^t Q_{\geq\gamma'' k_2}(\phi_1\eta)(s)P_0[\psi\nabla^{-1}(\varepsilon^2)]\,ds
\end{aligned}$$

The second term on the right is again small since $\|Q_{\geq\gamma'' k_2}(\phi_1\eta)\|_{L_t^2} \lesssim \delta_1^{\frac{\gamma''}{2}k_2}$. For the first term on the right, we note that

$$Q_{<\gamma'' k_2}(\phi_1\eta) = \eta_1 Q_{<\gamma'' k_2}(\phi_1\eta) + O(2^{-N\gamma'' k_2}),$$

whence up to errors which can again be immediately absorbed into ε_2, we can perform the Hodge-type decomposition for the factors in $\nabla^{-1}(\varepsilon^2)$ and continue the calculations as after (9.27). Note that the localization due to the factor $\phi_1\eta$ also allows us to reduce the high-frequency inputs $P_{k_{2,3}}\varepsilon$ by their hyperbolic reductions $P_{k_{2,3}}Q_{<k_{2,3}}\varepsilon$ and still be able to perform the Hodge decomposition up to negligible errors. This is because

$$\eta\phi_1 P_{k_2}Q_{<k_2}\varepsilon = \eta\phi_1 P_{k_2}Q_{<k_2}(\eta\varepsilon) + O(2^{-Nk_2})$$

and we have $\eta\varepsilon = \eta(R_\nu\varepsilon + \chi_\nu)$. Of course, inclusion of the cutoff η destroys the frequency localization again, but we have

$$\eta\phi_1 P_{k_2}Q_{<k_2}(\eta R_\nu\varepsilon) = \eta\phi_1 P_{k_2}Q_{<k_2}(R_\nu\varepsilon) - \eta\phi_1 P_{k_2}Q_{<k_2}([1-\eta]R_\nu\varepsilon)$$

and $\eta\phi_1 P_{k_2}Q_{<k_2}([1-\eta]R_\nu\varepsilon) = O_{L^2_{t,x}}(2^{-Nk_2})$. As at the end of case (a), we observe that if the expression at modulation $\sim 2^j$, $j \gg 1$, is hit by a time derivative ∂_t, the definition of $S[0]$ gives us a gain of $2^{-\varepsilon j}$, which translates into a gain of $2^{-\varepsilon k_2}$.

The remaining terms (9.29)-(9.34) no longer require an integration by parts trick and can be directly placed into $L^2_{t,x}$ with the requisite gain in k_2. We treat here the term (9.30) given by

$$P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\psi\nabla^{-1}(\epsilon^2))R_\nu P_{k_3}Q_{<k_3}\epsilon)$$

where we always keep in mind the localizations $k_1 = O(1) = k$, $k_2 = k_3 + O(1) > B|\log\delta_1|$. The key here is as before the improved Strichartz estimates. Write

$$P_{k_2}Q_{<k_2}(\psi\nabla^{-1}(\epsilon^2)) = P_{k_2}Q_{<k_2}(\psi\nabla^{-1}P_{<0}(\epsilon^2)) + \sum_{s\geq 0} P_{k_2}Q_{<k_2}(\psi\nabla^{-1}P_s(\epsilon^2))$$

We treat here the contribution of the second term on the right, the first being treated in the same vein. Now if $s < k_2 - 10$, we get

$$\Big(\sum_{c\in\mathcal{D}_{k_2,s-k_2}} \|P_c Q_{<k_2}(\psi\nabla^{-1}P_s(\epsilon^2))\|^2_{L^4_t L^1_x}\Big)^{\frac{1}{2}} \lesssim 2^{\frac{3k_2}{4}} 2^{\frac{s-k_2}{2+}} 2^{-s}\|P_{k_2}\psi\|_{S[k_2]}\|\epsilon\|^2_{L^\infty_t L^2_x}$$

Thus in the case $s < k_2 - 10$ from Bernstein's inequality we get

$$\begin{aligned}
&\|P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{k_2}Q_{<k_2}(\psi\nabla^{-1}P_s(\epsilon^2))R_\nu P_{k_3}Q_{<k_3}\epsilon)\|_{L^2_{t,x}}\\
&= \sum_{\substack{c_{1,2}\in\mathcal{D}_{k_2,s-k_2}\\ \mathrm{dist}(c_1,-c_2)\lesssim 2^s}} \|P_{k_1}\psi\nabla^{-1}(\nabla^{-1}P_{c_1}Q_{<k_2}(\psi\nabla^{-1}P_s(\epsilon^2))R_\nu P_{c_2}Q_{<k_3}\epsilon)\|_{L^2_{t,x}}\\
&\lesssim \|P_{k_1}\psi\|_{L^\infty_t L^2_x}\Big(\sum_{c\in\mathcal{D}_{k_2,s-k_2}} \|P_{c_1}Q_{<k_2}(\psi\nabla^{-1}P_s(\epsilon^2))\|^2_{L^4_t L^1_x}\Big)^{\frac{1}{2}}\Big(\sum_{c\in\mathcal{D}_{k_2,s-k_2}} \|R_\nu P_{c_2}Q_{<k_3}\epsilon\|^2_{L^4_t L^\infty_x}\Big)^{\frac{1}{2}}\\
&\lesssim 2^{-k_2}2^{\frac{3k_2}{2}}2^{2(\frac{s-k_2}{2+})}2^{-s}\|P_{k_1}\psi_1\|_{S[k_1]}\|P_{k_2}\psi\|_{S[k_2]}\|\epsilon\|^2_{L^\infty_t L^2_x}
\end{aligned}$$

Summing over $0 < s < k_2$ results in the bound

$$\lesssim 2^{-\frac{k_2}{2+}}\|P_{k_1}\psi_1\|_{S[k_1]}\|P_{k_2}\psi\|_{S[k_2]}\|\epsilon\|^2_{L^\infty_t L^2_x}$$

On the other hand, when $s \geq k_2 - 10$, we simply bound

$$\|P_{k_2}Q_{<k_2}(\psi\nabla^{-1}P_s(\epsilon^2))\|_{L^4_t L^1_x} \lesssim 2^{-\frac{k_1}{4}}\|\psi\|_S\|\epsilon\|^2_{L^\infty_t L^2_x}$$

and from here one estimates the $L^2_{t,x}$-norm of the output as before but without using the improved Strichartz, just the standard $L^4_t L^\infty_x$-bound. The remaining terms (9.31) are handled similarly.

(ii) : High-Low/ Low-High interactions within $\nabla^{-1}(\epsilon^2)$ In this case one gains exponentially in the maximum frequency occurring among the two factors ϵ, provided this is much larger than 1. In this case one can argue as in case (b) to include this contribution into ϵ_2.

(e) The cubic term $\epsilon\nabla^{-1}(\epsilon^2)$ is easy, and can be treated as in (a) and (b) above. Here the smallness comes simply from the size of ϵ.

The bootstrap argument for ϵ in the small time case is now completed as in the proof of Lemma 7.6, cf. (7.10).

Case 2: $|I_1| \geq T_1$, where $T_1 > 0$ is a small constant depending on E_{crit}. Here we have to work with the wave equation satisfied by $P_0\epsilon$. We start by recording this equation schematically in its original trilinear form, to which we apply various Hodge type decompositions as well as localizations in frequency space. The goal is to write the equation in the form of a nonlinear wave equation with a low-frequency magnetic potential term, which we will treat as part of the linear operator. To begin with, we have the schematic equation (here we suppress the fact that ϵ really stands for the system of variables $\{\epsilon_\alpha\}, \alpha = 0, 1, 2$)

(9.35)

$$\begin{aligned}\Box P_0\epsilon &= \nabla_{x,t}P_0\big[(\psi+\epsilon)\nabla^{-1}([\psi+\epsilon]^2)\big] - \nabla_{x,t}P_0\big[(\psi)\nabla^{-1}(\psi^2)\big]\\ &= P_0\nabla_{x,t}[\epsilon\nabla^{-1}(\psi^2)] + P_0\nabla_{x,t}[\psi\nabla^{-1}(\psi\epsilon)] + P_0\nabla_{x,t}[\epsilon\nabla^{-1}(\psi\epsilon)] + P_0\nabla_{x,t}[\psi\nabla^{-1}(\epsilon^2)] + P_0\nabla_{x,t}[\epsilon\nabla^{-1}(\epsilon^2)]\end{aligned}$$

More precisely, the terms on the right-hand side of (9.35) are exactly those given by (1.14). It is precisely the first term on the last line which causes technical difficulties for the bootstrap argument, and we shall have to include parts of it into the linear operator. However, this will only be made specific once we have localized the terms suitably in frequency space. To begin with, note that we will implement a bootstrap argument in order to deduce bounds on ϵ. For this we substitute Schwartz extensions $\tilde{\epsilon}_\alpha$ for each ϵ_α on the right-hand side (these extensions agreeing with ϵ_α on the time interval $I_1 \times \mathbb{R}^2$ we are working on), and then solve the inhomogeneous wave equation for ϵ_α, improving the bounds we used for $\tilde{\epsilon}_\alpha$. Denoting the right-hand source term above — with $\tilde{\epsilon}_\alpha$ instead of ϵ_α — by $\tilde{F}_\alpha$, what we really do is solving the problem

$$\Box P_0\epsilon_\alpha = P_0\tilde{F}_\alpha$$

In order to deduce the S-bounds on $P_0\epsilon_\alpha$, we split this variable into two parts

$$P_0\epsilon_\alpha = P_0Q_{\geq D}\epsilon_\alpha + P_0Q_{<D}\epsilon_\alpha$$

Here the parameter D is chosen sufficiently large depending on T_1 from Case 1 and thus depends on E_{crit} (but is independent of the induction stage). Then we solve the preceding wave equation by setting

$$\begin{aligned}P_0Q_{\geq D}\epsilon_\alpha &= \Box^{-1}Q_{\geq D}P_0\tilde{F}_\alpha\\ P_0Q_{<D}\epsilon_\alpha &= S(t)\big(P_0Q_{<D}\epsilon_\alpha\big)[0] + \int_0^t U(t-s)P_0Q_{<D}\tilde{F}_\alpha(s)\,ds\end{aligned}$$

In other words, $P_0Q_{<D}\epsilon_\alpha$ solves the following inhomogeneous wave equation:

$$\Box P_0Q_{<D}\epsilon = P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}(\psi^2) + \psi\nabla^{-1}(\psi\tilde{\epsilon}) + \tilde{\epsilon}\nabla^{-1}(\psi\tilde{\epsilon}) + \psi\nabla^{-1}(\tilde{\epsilon}^2) + \tilde{\epsilon}\nabla^{-1}(\tilde{\epsilon}^2)\big] \tag{9.36}$$

First, we identify the terms which can be included in the right-hand side as source terms since they gain smallness, which is achieved in part by introducing suitable Fourier localizations. To begin with, recall that the basic version of the wave maps equation at the level of the Coulomb gauge is of the schematic form

$$\Box\psi_\alpha = i\partial^\beta[\psi_\alpha A_\beta] - i\partial^\beta[\psi_\beta A_\alpha] + i\partial_\alpha[\psi^\nu A_\nu]$$

The estimates of Section 5 will be seen to imply that the middle term here can be included entirely in the right-hand side, and the immediately ensuing discussion is only applied to the first and third terms. Split the first term on the right in (9.36) (which is understood to be of the first or third type) into

$$P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}(\psi^2)\big] = P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi^2)\big] + P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{\geq -D_1}(\psi^2)\big]$$

Here D_1 is a large constant depending like D on the energy in a "mild" way, i.e., independently of the stage of the induction we are at, as will be seen shortly. Recalling that on $I_1 \times \mathbb{R}^2$ we have the decomposition

$$\psi = \psi_L + \psi_{NL},$$

we further decompose (schematically)

$$\begin{aligned}&P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi^2)\big]\\ &= P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big] + P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_{NL}^2)\big] + P_0Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L\psi_{NL})\big]\end{aligned}$$

Due to the smallness of ψ_{NL} and (9.17), it is only the first term on the right which we need to incorporate in part into the linear operator. Of course this requires replacing $\tilde{\epsilon}$ by ϵ, which requires some care due to the non-local operator $Q_{<D}$ interfering with our aim. First, write

$$P_0 Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big] = P_0\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big] - P_0 Q_{\geq D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big]$$

Since we only need to solve the equation on $I_1\times\mathbb{R}^2$, where $\tilde{\epsilon}$ and ϵ agree, we may replace the right-hand side by

$$\begin{aligned}&P_0\nabla_{x,t}\big[\epsilon\nabla^{-1}P_{<-D_1}(\psi_L^2)\big] - P_0 Q_{\geq D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big]\\ &= P_0\nabla_{x,t}\big[Q_{<D}\epsilon\nabla^{-1}P_{<-D_1}(\psi_L^2)\big] + P_0\nabla_{x,t}\big[Q_{\geq D}\epsilon\nabla^{-1}P_{<-D_1}(\psi_L^2)\big] - P_0 Q_{\geq D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big]\end{aligned}$$

Now we introduce null-structure by performing Hodge decompositions as in Section 3, for all the trilinear terms. In particular, the preceding discussion yields that we replace the schematic term

$$P_0 Q_{<D}\nabla_{x,t}\big[\tilde{\epsilon}\nabla^{-1}P_{<-D_1}(\psi_L^2)\big]$$

by

$$\begin{aligned}&\sum_{j=1,3}F_\alpha^{3j}(P_0Q_{<D}\epsilon; P_{<-D_1};\psi_L,\psi_L) + P_0Q_{<D}F_\alpha^{32}(\tilde{\epsilon};P_{<-D_1};\psi_L,\psi_L)\\ &+\Big[\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L) - \sum_{j=1,3}F_\alpha^{3j}(P_0Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L)\Big]\\ &+\Big[\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{\geq D}\epsilon;P_{<-D_1};\psi_L,\psi_L) - \sum_{j=1,3}P_0Q_{\geq D}F_\alpha^{3j}(\tilde{\epsilon};P_{<-D_1};\psi_L,\psi_L)\Big]\\ &+\sum_{k=2}^{5}P_0Q_{<D}F_\alpha^{2k+1}(\tilde{\epsilon};P_{<-D_1};\psi_L,\psi_L)\end{aligned}$$

We can now write the wave equation that we use to solve for $P_0Q_{<D}\epsilon$ as follows:

$$\begin{aligned}\Box(P_0Q_{<D}\epsilon) = &\sum_{j=1,3}F_\alpha^{3j}(P_0Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L) + P_0Q_{<D}F_\alpha^{32}(\tilde{\epsilon};P_{<-D_1};\psi_L,\psi_L)\\ &+\Big[\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L) - \sum_{j=1,3}F_\alpha^{3j}(P_0Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L)\Big]\\ &+\Big[\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{\geq D}\epsilon;P_{<-D_1};\psi_L,\psi_L) - \sum_{j=1,3}P_0Q_{\geq D}F_\alpha^{3j}(\tilde{\epsilon};P_{<-D_1};\psi_L,\psi_L)\Big]\\ &+\sum_{k=2}^{5}\big[P_0Q_{<D}F_\alpha^{2k+1}(\psi+\tilde{\epsilon},(\psi+\tilde{\epsilon}),(\psi+\tilde{\epsilon})) - P_0Q_{<D}F_\alpha^{2k+1}(\psi,\psi,\psi)\big]\\ &+P_0Q_{<D}F_\alpha^3(\tilde{\epsilon};P_{<-D_1};\psi_{NL},\psi_L) + P_0Q_{<D}F_\alpha^3(\tilde{\epsilon};P_{<-D_1};\psi_L,\psi_{NL})\\ &+P_0Q_{<D}F_\alpha^3(\tilde{\epsilon},\psi_{NL},\psi_{NL}) + P_0Q_{<D}F_\alpha^3(\tilde{\epsilon};P_{\geq -D_1};\psi_L,\psi_L)\\ &+P_0Q_{<D}F_\alpha^3(\psi,\tilde{\epsilon},\psi) + P_0Q_{<D}F_\alpha^3(\psi,\psi,\tilde{\epsilon}) + P_0Q_{<D}F_\alpha^3(\tilde{\epsilon},\tilde{\epsilon},\psi) + P_0Q_{<D}F_\alpha^3(\psi,\tilde{\epsilon},\tilde{\epsilon})\\ &+P_0Q_{<D}F_\alpha^3(\tilde{\epsilon},\tilde{\epsilon},\tilde{\epsilon})\end{aligned}\tag{9.37}$$

The significance of the first term on the right, i.e., the expression

$$\sum_{j=1,3}F_\alpha^{3j}(P_0Q_{<D}\epsilon,\psi_L,\psi_L),$$

is that it implicitly contains a magnetic potential interaction term, see the discussion at the end of Section 3. In order to deduce estimates, we shall re-arrange terms and move the magnetic interaction term contained in the above term

$$2i\partial^\beta(P_0Q_{<D}\epsilon)A_\beta,\qquad A_\beta := -P_{<-D_1}\partial_j^{-1}I\mathcal{Q}_{\beta j}(\psi_L,\psi_L)\tag{9.38}$$

to the left, thereby obtaining an equation of the schematic type

$$\Box(P_0 Q_{<D}\epsilon) + 2i\partial^\beta (P_0 Q_{<D}\epsilon) A_\beta = F \tag{9.39}$$

The next issue occupying us is the derivation of a priori estimates for this type of equation, at first treating F as a function with good Fourier localization properties and bounded with respect to $\|\cdot\|_N$.

9.4.1. *Solving the wave equation with a magnetic potential in the Coulomb gauge.* For simplicity's sake, replace $(P_0 Q_{<D}\epsilon)$ at the end of the preceding section by ϵ for this subsection. The key fact that is proven here is the following:

Proposition 9.14. *Assume that F is a function at frequency ~ 1, and $\|\psi_L\| \lesssim E_{crit}$. Also, assume the solution to (9.39) with data $(\epsilon(0,\cdot), \partial_t\epsilon(0,\cdot)) = (f,g)$, all supported at frequency ~ 1, to be supported at frequency ~ 1 and modulation $\lesssim 1$. Finally, assume that*

$$D_1 > D_1(E_{\text{crit}}).$$

Then ϵ satisfies the bound

$$\|\epsilon\|_{S[0]} \lesssim \|F\|_{N[0]} + \|(f,g)\|_{L^2_x \times \dot H^{-1}}$$

with implied constant only depending on E_{crit}. Furthermore, there is approximate energy conservation:

$$\|\partial_t \epsilon(t,\cdot)\|^2_{L^2_x} + \|\nabla_x \epsilon(t,\cdot)\|_{L^2_x} = \|\partial_t \epsilon(0,\cdot)\|^2_{L^2_x} + \|\nabla_x \epsilon(0,\cdot)\|_{L^2_x} + c(D_1) + O(\|F\|_{N[0]}\|\varepsilon\|_S)$$

with $c(D_1) \to 0$ as $D_1 \to \infty$, independently of t.

Proof. Recall that $0 \le \beta \le 2$,

$$A_\beta = -\triangle^{-1} \sum_{j=1,2} \partial_j P_{<-D_1} I[R_\beta \psi_L^1 R_j \psi_L^2 - R_\beta \psi_L^2 R_j \psi_L^1]$$

and observe that these functions are real-valued and Schwartz for fixed times. The key difficulty comes from the fact that there appears no obvious way to obtain smallness for the linear interaction term $2i\partial^\beta \epsilon A_\beta$, even when restricting to small time intervals. The easiest way out of this impasse is to use an approximate a priori bound resulting from energy conservation. This will allow us to split the bad interaction term into two, one of which is small due to angular alignment of the inputs, the other of which is controlled due to the a priori bound. Moreover, we note that we may always move parts of the expression $2i\partial^\beta \epsilon A_\beta$ with additional smallness properties, such as extreme frequency discrepancies inside A_β or special angular alignments, to the right-hand side, since we gain smallness for them as shown in Section 5. More precisely, let us pick a cap size $|\kappa| = |\kappa|(E_C)$, and write the underlying equation (9.39) in localized form as

$$\Box P_\kappa \epsilon + 2i\partial^\beta P_\kappa \epsilon A_\beta = P_{2\kappa}\tilde F, \quad P_\kappa\epsilon[0] = (P_\kappa f, P_\kappa g) \tag{9.40}$$

Here we also assume that D_1 above is large enough in relation to $|\kappa|(E_C)$. We next pick a cap size $|\kappa_1| = |\kappa_1(E_C)| \ll |\kappa(E_C)$, but such that $D_1 \gg |\kappa_1|^{-100}$, say. Note that a computation similar to (8.19) reveals that

$$\|P_{2\kappa}\tilde F - P_{2\kappa} F\|_{N[0]} \le c_6 \|P_\kappa \epsilon\|_{S[0]}$$

where $c_6 = c_6(D_1, E_C)$ can be made arbitrarily small in relation to $|\kappa|(E_C)$. Now make the following

Apriori Bound Assumption: There exist constants $c_7 = c_7(E_C, |\kappa|)$, $C_7 = C_7(E_C)[\|F\|_{N[0]} + \|f\|_{L^2_x} + \|g\|_{\dot H^{-1}}]$, such that

$$\sup_{\omega \notin \pm 2\kappa} \|P_\kappa Q^{\pm}_{<2\log|\kappa_1|}\epsilon\|_{L^\infty_{t_\omega} L^2_{x_\omega}} < c_7 \|\epsilon\|_{S[0]} + C_7$$

We first show that this assumption, together with a standard bootstrap procedure, implies the bound of the proposition. Then we establish the a priori bound. Observe that the localization to caps of size $|\kappa|$ is important in the first step, while we need to pass to finer caps κ_1 in order to establish the a priori bound. Thus return to the original equation, which we write in the form

$$\Box\epsilon = F - \sum_{|\kappa| = |\kappa|(E_C)} 2i\partial^\beta P_\kappa \epsilon A_\beta, \quad \epsilon[0] = (f,g) \tag{9.41}$$

Decompose the term on the right hand side into

$$F - \sum_{|\kappa|=|\kappa|(E_C)} 2i\partial^\beta P_\kappa \epsilon A_\beta = F - \sum_{|\kappa|=|\kappa|(E_C),\pm} 2i\partial^\beta P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon \tilde{A}_\beta - \sum_{|\kappa|=|\kappa|(E_C),\pm} 2i\partial^\beta P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon A^\dagger_\beta - \sum_{|\kappa|=|\kappa|(E_C)} Q_{\geq 2\log|\kappa_1|}\partial^\beta P_\kappa \epsilon A_\beta$$

Here we define

$$A^\dagger_\beta := -\sum_{\kappa_{1,2}\in K_{\log|\kappa|(E_{crit})}, \mathrm{dist}(\pm\kappa,\kappa_{1,2})<10|\kappa|(E_{crit})} \sum_{\substack{\max\{k_{1,2,3}\}\leq \min\{k_{1,2,3}\}+C_8(E_{crit})\\ k_1<-D_1}} I\triangle^{-1}\sum_{j=1,2}\partial_j P_{k_1+O(1)}[R_\beta P_{k_2,\kappa_1}\psi^1_L R_j P_{k_3,\kappa_2}\psi^2_L - R_\beta P_{k_3,\kappa_2}\psi^2_L R_j P_{k_2,\kappa_1}\psi^1_L] - \sum_{\substack{\max\{k_{1,2,3}\}>\min\{k_{1,2,3}\}+C_8(E_{crit})\\ k_1<-D_1}} I\triangle^{-1}\sum_{j=1,2}\partial_j P_{k_1}[R_\beta P_{k_2}\psi^1_L R_j P_{k_3}\psi^2_L - R_\beta P_{k_2}\psi^2_L R_j P_{k_3}\psi^1_L]$$

and furthermore $\tilde{A}_\beta := A_\beta - A^\dagger_\beta$. By choosing $|\kappa|(E_C)$ small enough in relation to E_C, and further using Corollary 5.2 as well as Lemma 5.5 as well as their improvements in the small angle case, see section 5.3, we infer that

$$\|\sum_\pm 2i\partial^\beta P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon A^\dagger_\beta\|_{N[0]} \lesssim |\kappa|^{\delta_{10}}\|P_\kappa\epsilon\|_{S[0]},$$

and furthermore, exploiting the alignment of the inputs in the definition of $A^\dagger_\beta$ as well as Cauchy-Schwarz, we get

$$\|\sum_{|\kappa|=|\kappa|(E_C),\pm} 2i\partial^\beta P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon A^\dagger_\beta\|_{N[0]} \lesssim |\kappa|^{(\delta_{10})-}\|\epsilon\|_{S[0]}$$

where the implied constant is universal. Next, consider

$$\sum_{|\kappa|=|\kappa|(E_C)} Q_{\geq 2\log|\kappa_1|}\partial^\beta P_\kappa \epsilon A_\beta$$

Here, we estimate

$$\|Q_{<O(1)}\big[Q_{\geq 2\log|\kappa_1|}\partial^\beta P_\kappa\epsilon A_\beta\big]\|_{N[0]} \leq \|Q_{O(1)>\cdot\geq 2\log|\kappa_1|}\partial^\beta P_\kappa\epsilon\|_{L^2_t L^2_x}\|A_\beta\|_{L^2_t L^\infty_x} < c_8(E_C,|\kappa|)\|P_\kappa\epsilon\|_{S[0]}$$

provided D_1 is large enough, where we pick $c_8(E_C,|\kappa|)$ small enough in relation to the indicated quantities. Similarly, we get

$$\|Q_{>O(1)}\big[Q_{\geq 2\log|\kappa_1|}P_\kappa\epsilon A_\beta\big]\|_{N[0]} \leq \|Q_{O(1)>\cdot\geq 2\log|\kappa_1|}P_\kappa\epsilon\|_{L^2_t L^2_x}\|A_\beta\|_{L^\infty_{t,x}} < c_8(E_C,|\kappa|)\|P_\kappa\epsilon\|_{S[0]}$$

Finally, consider the most delicate term above, $\sum_{|\kappa|=|\kappa|(E_C),\pm} 2i\partial^\beta P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon\tilde{A}_\beta$. By definition of $\tilde{A}_\beta$, at least one input (both inputs being free waves) has some angular separation from $\pm\kappa$ for its Fourier support. On the other hand, the frequencies of the inputs are approximately equal to the frequency of the output $\tilde{A}_\beta$. Now using the "a priori bound assumption" from above, we obtain (with implied constants only depending on E_C)

$$\|\sum_{|\kappa|=|\kappa|(E_C),\pm} 2i\partial^\beta P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon\tilde{A}_\beta\|_{N[0]} \lesssim \sum_{|\kappa|=|\kappa|(E_C),\pm}\big[\sup_{\omega\notin\pm 2\kappa}\|P_\kappa Q^\pm_{<2\log|\kappa_1|}\epsilon\|_{L^\infty_{t_\omega}L^2_{x_\omega}}\big] \leq \sum_{|\kappa|=|\kappa|(E_C),\pm}[c_7\|\epsilon\|_{S[0]} + C_7]$$

By picking c_7 small enough in relation to $|\kappa|$, we can bound the preceding by

$$< c_8\|\epsilon\|_{S[0]} + C_9$$

with $C_9 = \tilde{C}_9(E_C)[\|F\|_{N[0]} + \|f\|_{L^2} + \|g\|_{\dot{H}^{-1}}]$. Recalling (9.41) as well as picking c_8 small enough, it is now straightforward to deduce the bound

$$\|\epsilon\|_{S[0]} \lesssim \|F\|_{N[0]} + \|(f,g)\|_{L^2_x \times \dot{H}^{-1}}$$

We now turn to the proof of the aforementioned a priori bound: in effect, we will first prove a conceptually somewhat simpler standard energy bound where we replace the null-frame energy by the standard energy; this, together with the a priori bound, will in particular imply approximate energy conservation as $D_1 \to \infty$, as specified in the proposition.
To begin with, pick localizers P_κ, $|\kappa| = |\kappa|(E_C, D_1)$ with $|\kappa| \to 0$ sufficiently slowly as $D_1 \to \infty$, such that if $\chi_\kappa(\xi)$ is the corresponding cutoff on the Fourier side, we have

$$\sum_\kappa \chi_\kappa^2(\xi) = 1$$

Consider the inhomogeneous problem

$$\Box\epsilon + 2i\partial^\nu \epsilon A_\nu = F,\ \epsilon[0] = (f,g);$$

Under the assumptions of the proposition, we intend to show approximate energy conservation as in the statement of the proposition. We localize this equation as before

$$\Box P_\kappa \epsilon + 2i\partial^\nu P_\kappa \epsilon \tilde{A}_\nu = -2i\partial^\nu P_\kappa \epsilon A_\nu^\dagger - P_\kappa\big[2i\partial^\nu \epsilon A_\nu\big] + \big[2i\partial^\nu P_\kappa \epsilon A_\nu\big] + P_{2\kappa}F_\kappa =: \tilde{F}_\kappa \tag{9.42}$$

where $A_\nu^\dagger$ is defined as above but with $C_8 = C_8(E_C, D_1)$ and $C_8 \to \infty$ sufficiently slowly as $D_1 \to \infty$. Note that we may arrange that

$$\sum_\kappa (-\log|\kappa|)^l \| - P_{2\kappa}F_\kappa + \tilde{F}_\kappa\|_{N[0]} \to 0$$

as $D_1 \to \infty$, for any l. Finally, we shall also assume that $\epsilon = P_0 Q_{<D_2}\epsilon$, where $D_2 = D_2(D_1) \to \infty$ as $D_1 \to \infty$. Indeed, one may apply such an operator to the equation and move the errors on the right-hand side, as they can be iterated away. We leave these technical details to the reader. Now consider the *covariant energy density*

$$\sum_\kappa \frac{1}{2}[|\partial_t P_\kappa \epsilon + i\tilde{A}_0 P_\kappa \epsilon|^2 + \sum_{j=1,2} |\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon|^2] \tag{9.43}$$

Compute

$$\begin{aligned}
&\partial_t\big[\frac{1}{2}|\partial_t P_\kappa \epsilon + i\tilde{A}_0 P_\kappa \epsilon|^2 + \sum_{j=1,2}\frac{1}{2}|\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon|^2\big] \\
&= \mathrm{Re}\big[(\partial_t P_\kappa \epsilon + i\tilde{A}_0 P_\kappa \epsilon)\overline{(\partial_t + i\tilde{A}_0)^2 P_\kappa \epsilon} + \sum_{j=1,2}(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)\overline{\partial_t(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)}\big]
\end{aligned}$$

The second term on the right satisfies

$$\begin{aligned}
\mathrm{Re}[(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)\overline{\partial_t(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)}] &= \mathrm{Re}[(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)\overline{(\partial_{x_j} + i\tilde{A}_j)(\partial_t + i\tilde{A}_0)P_\kappa \epsilon}] \\
&\quad + \mathrm{Re}[(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)\overline{i(\partial_t \tilde{A}_j - \partial_{x_j}\tilde{A}_0)P_\kappa \epsilon}] \\
&= \partial_{x_j}\mathrm{Re}[(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)\overline{(\partial_t + i\tilde{A}_0)P_\kappa \epsilon}] \\
&\quad - \mathrm{Re}[(\partial_{x_j} + i\tilde{A}_j)^2 P_\kappa \epsilon\,\overline{(\partial_t + i\tilde{A}_0)P_\kappa \epsilon}] \\
&\quad + \mathrm{Re}[(\partial_{x_j} P_\kappa \epsilon + i\tilde{A}_j P_\kappa \epsilon)\overline{i(\partial_t \tilde{A}_j - \partial_{x_j}\tilde{A}_0)P_\kappa \epsilon}]
\end{aligned}$$

In summary, one obtains the following *local form of energy conservation*:

(9.44)

$$\partial_t\sum_\kappa\Big[\frac{1}{2}|\partial_t P_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon|^2+\sum_{j=1,2}\frac{1}{2}|\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon|^2\Big]-\sum_\kappa\sum_{j=1}^{2}\partial_{x_j}\mathrm{Re}\big[(\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon)\overline{(\partial_t+i\tilde{A}_0)P_\kappa\epsilon}\big]$$
$$=\sum_\kappa\big[\mathrm{Re}\big[\big[(\partial_t+i\tilde{A}_0)^2-\sum_{j=1,2}(\partial_{x_j}+i\tilde{A}_j)^2\big]P_\kappa\epsilon\,\overline{(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)}\big]+\mathrm{Re}[(\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon)\overline{i(\partial_t\tilde{A}_j-\partial_{x_j}\tilde{A}_0)P_\kappa\epsilon}]\big]$$

We furthermore observe that any solution of $\Box\varepsilon+2iA^\alpha\,\partial_\alpha\varepsilon=F$ satisfies

$$\big[(\partial_t+iA_0)^2-\sum_{j=1,2}(\partial_{x_j}+iA_j)^2\big]\epsilon=F+i(\partial_tA_0-\sum_{j=1,2}\partial_{x_j}A_j)\epsilon+(\sum_{j=1,2}A_j^2-A_0^2)\epsilon$$

We now integrate the above relation over a time slice $[0,t_0]\times\mathbb{R}^2$, which gives

$$\sum_\kappa\int_{\mathbb{R}^2}\big[\frac{1}{2}|\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon|^2+\sum_{j=1,2}\frac{1}{2}|\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon|^2\big](t_0,x)\,dx$$
$$=\sum_\kappa\int_{\mathbb{R}^2}\big[\frac{1}{2}|\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon|^2+\sum_{j=1,2}\frac{1}{2}|\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon|^2\big](0,x)\,dx$$
$$+\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}\big[\big(\tilde{F}_\kappa+i(\partial_t\tilde{A}_0-\sum_{j=1,2}\partial_{x_j}\tilde{A}_j)P_\kappa\epsilon+(\sum_{j=1,2}\tilde{A}_j^2-\tilde{A}_0^2)\big)P_\kappa\epsilon\,\overline{(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)}\big]\,dtdx$$
$$+\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}[(\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon)\overline{i(\partial_t\tilde{A}_j-\partial_{x_j}\tilde{A}_0)P_\kappa\epsilon}]\,dtdx$$
$$+\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}\big[\,[(\partial_t+iA_0)^2-\sum_{j=1,2}(\partial_{x_j}+iA_j)^2,P_\kappa]\overline{(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)}\big]\,dtdx$$

We now estimate the three last integrals,

$$\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}\big[(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)\overline{\big(\tilde{F}_\kappa+i(\partial_t\tilde{A}_0-\sum_{j=1,2}\partial_{x_j}\tilde{A}_j)P_\kappa\epsilon+(\sum_{j=1,2}\tilde{A}_j^2-\tilde{A}_0^2)P_\kappa\epsilon\big)}\big]\,dtdx$$
$$+\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}[(\partial_{x_j}P_\kappa\epsilon+i\tilde{A}_jP_\kappa\epsilon)\overline{i(\partial_t\tilde{A}_j-\partial_{x_j}\tilde{A}_0)P_\kappa\epsilon}]\,dtdx$$
$$+\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}\big[\,[(\partial_t+iA_0)^2-\sum_{j=1,2}(\partial_{x_j}+iA_j)^2,P_\kappa]\epsilon\overline{(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)}\big]\,dtdx$$

One can classify four types of terms.

(1) The term $\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}\big[(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)\overline{\tilde{F}_\kappa}\big]\,dtdx$. Here one uses the duality of N and S, Lemma 2.19, as well as the space-time frequency localization of ϵ:

$$\Big|\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\mathrm{Re}\big[(\partial_tP_\kappa\epsilon+i\tilde{A}_0P_\kappa\epsilon)\overline{\tilde{F}_\kappa}\big]\,dtdx\Big|\lesssim[\sum_\kappa\|\tilde{F}_\kappa\|_{N[0]}]\|\epsilon\|_S$$

Application of Lemma 2.19 is justified due to our assumptions on the modulation of ϵ, which in turn restrict the modulation of $\tilde{F}$ to the hyperbolic regime via the equation.

(2) The terms of the form $\sum_\kappa\int_{[0,t_0]\times\mathbb{R}^2}\nabla_{x,t}\tilde{A}\;P_\kappa\epsilon\nabla_{x,t}P_\kappa\epsilon\,dtdx$. These are controlled due to the angular separation inherent in the definition of $\tilde{A}_\beta$. Note the schematic identity

$$\nabla_{x,t}\tilde{A}_\beta=\sum_{k_{1,2}<-D_1}\nabla_{x,t}\nabla^{-1}P_{k_1}[P_{k_2}\psi_LP_{k_3}\psi_L]$$

Here our reductions for $\tilde{A}_\beta$ imply $k_1=k_2+O(1)=k_3+O(1)$ (where the implied constant may be quite large depending on E_{crit}, D_1) and furthermore the inputs $P_{k_{2,3}}\psi_L$ have some angular separation between their Fourier supports and $\pm\kappa$. But from this one infers that

$$\Big|\int_{[0,t_0]\times\mathbb{R}^2}\nabla_{x,t}P_{<k}\tilde{A}\;P_\kappa\epsilon\overline{\nabla_{x,t}P_\kappa\epsilon}\,dtdx\Big|\lesssim 2^k\|P_\kappa\epsilon\|_{S[0]}^2$$

since one may pair each factor ψ_L against a factor $P_\kappa\epsilon$.
(3) The terms

$$\int_{[0,t_0]\times\mathbb{R}^2}A^2\;P_\kappa\epsilon\overline{\nabla_{x,t}P_\kappa\epsilon}\,dtdx$$

are easier to handle. Here, one may use that

$$\|[P_kA]^2\|_{L_t^{\frac{4}{3}}L_x^2}\lesssim E_{\text{crit}}^4\,2^{\frac{k}{4}},$$

which follows from the usual Strichartz estimates, cf. Lemma 2.17:

$$\|\nabla^{-1}P_k(\psi_L^2)\|_{L_t^{\frac{8}{3}}L_x^4}\lesssim 2^{-k}\|\psi_L\|_{L_t^{\frac{16}{3}}L_x^8}^2\lesssim 2^{\frac{k}{8}}\|\psi_L\|_2^2$$

One may then use the $L_t^8L_x^4$-control for ϵ to get

$$\sum_\kappa\Big|\int_{[0,t_0]\times\mathbb{R}^2}P_{<k}A^2\;P_\kappa\epsilon\overline{\nabla_{x,t}P_\kappa\epsilon}\,dtdx\Big|\lesssim 2^{\frac{k}{4}}\|\epsilon\|_{S[0]}^2,$$

this bound of course being sub-optimal.
(4) The terms $\int_{[0,t_0]\times\mathbb{R}^2}[(\nabla_{t,x}+iA)^2,P_\kappa]\epsilon\overline{\nabla_{t,x}\epsilon}\,dtdx$. Here, using the observation that $P_\kappa(fg)=gP_\kappa f+\Lambda(f,\nabla g)$ provided f is supported at frequency ~ 1, while g is supported at frequency $\ll\log|\kappa|$, and further Λ represents a convolution operator of bounded L^1-mass, we reduce this case to either case (2) or (3) in the immediately preceding.

Summation over small $k<-D_1$ in (2), (3), (4) now yields the desired smallness provided D_1 is large.

In view of the preceding, we may conclude that

$$\|\nabla_{x,t}\epsilon(t,\cdot)\|_{L_x^2}^2=\|\nabla_{x,t}\epsilon(0,\cdot)\|_{L_x^2}^2+O(\gamma^2\|\epsilon\|_{S[0]}^2+\|F\|_{N[0]}\|\epsilon\|_{S[0]}),\tag{9.45}$$

where γ may be made arbitrarily small by choosing $D_1(E_{crit})$ in the statement of the proposition large enough. Note that we eliminated the magnetic potential here from the covariant energy by means of the estimate $\|A_\beta\|_{L_t^\infty L_x^\infty}\lesssim\gamma\ll 1$. Almost energy conservation claimed in the proposition follows from this if we assume a priori control over $\|\epsilon\|_{S[0]}$.
To achieve the latter, all that remains is to establish the "a priori bound" above over one of the null-frame ingredients of $\|\cdot\|_{S[0]}$. This will follows by a very slight modification of the above argument; indeed, the only difference will be that now we are integrating (9.44) over a region $A^\omega_{t,c}:=[0,t]\times\mathbb{R}^2\cap\{t_\omega>c\}$ for arbitrary c, with $\omega\in S^1$ being a fixed direction. Recall equation (9.41). Also, recall that we introduced a smaller scale $|\kappa_1|$ immediately before the "a priori bound assumption" above. This extra scale now becomes important: we may localize (9.41) further to obtain

$$\Box P_\kappa\epsilon=P_\kappa F-\sum_{|\kappa_1|=|\kappa_1|(E_C,|\kappa|),\,\kappa_1\subset\frac{3}{2}\kappa}P_\kappa\big(2i\partial^\nu P_{\kappa_1}\epsilon A_\nu\big)\tag{9.46}$$

provided D_1 is chosen large enough. We further localize this to scale $|\kappa_1|$ to obtain for $\kappa_1\subset\frac{3}{2}\kappa$

$$\Box P_{\kappa_1}Q^\pm_{<2\log|\kappa_1|}\epsilon+2i\partial^\nu P_{\kappa_1}Q^\pm_{<2\log|\kappa_1|}\epsilon\,\tilde{A}_\nu=P_{2\kappa_1}\tilde{F}^\pm_{\kappa_1},$$

where we construct $\tilde{A}_\nu$ as in the first part of the proof, but with inputs of angular separation from $\pm\kappa_1$ now comparable to $|\kappa_1|$, as well as the quotient of all frequencies involved in this definition. On the other hand, $P_{2\kappa_1}\tilde{F}_{\kappa_1}$ incorporates all errors generated, in particular those involving $A^\dagger_\nu$. In the sequel, we shall omit the additional localizer $Q^\pm_{<2\log|\kappa_1|}$ but keep in mind that $P_{\kappa_1}\epsilon$ has this additional localization property. Now we integrate the corresponding divergence identity (9.44) over $A^\omega_{t,c}$ for fixed c and $\omega\notin\pm 2\kappa$.

This yields (for fixed κ_1)

$$\begin{aligned}
&\int_{A^\omega_{t,c}} \mathrm{Re}\big[(\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon)\overline{\big[(\partial_t + i\tilde{A}_0)^2 - \sum_{j=1,2}(\partial_{x_j} + i\tilde{A}_j)^2\big]P_{\kappa_1}\epsilon}\big] \\
&\qquad\qquad + \mathrm{Re}[(\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon)\overline{i(\partial_t A_j - \partial_{x_j} A_0)P_{\kappa_1}\epsilon}]\, dtdx \\
&= \int_{0\times\mathbb{R}^2\cap A^\omega_{t,c}} \big[\frac{1}{2}|\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon|^2 + \sum_{j=1,2}\frac{1}{2}|\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon|^2\big]\, dtdx \\
&- \int_{t\times\mathbb{R}^2\cap A^\omega_{t,c}} \big[\frac{1}{2}|\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon|^2 + \sum_{j=1,2}\frac{1}{2}|\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon|^2\big]\, dtdx \\
&+ \int_{\{t_\omega = c\}\cap A^\omega_{t,c}} \big[\frac{1}{2}|\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon|^2 \\
&\qquad + \sum_{j=1,2}\big(\frac{1}{2}|\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon|^2 - \omega_j \mathrm{Re}[(\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon)\overline{(\partial_t + i\tilde{A}_0)P_{\kappa_1}\epsilon}]\big)\big]\, dx_\omega
\end{aligned} \tag{9.47}$$

It is the latter integral expression that gives us the additional information we need: Indeed, use the decomposition

$$\sum_{j=1,2} |\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon|^2 = \big|\sum_{j=1,2}(\omega_j \partial_{x_j} P_{\kappa_1}\epsilon + i\omega_j \tilde{A}_j P_{\kappa_1}\epsilon)\big|^2 + \big|\sum_{j=1,2}(\omega_j^\perp \partial_{x_j} P_{\kappa_1}\epsilon + i\omega_j^\perp \tilde{A}_j P_{\kappa_1}\epsilon)\big|^2$$

Recalling that $\omega \notin 2\kappa$, we can conclude that

$$\sup_c \sup_t \int_{\{t_\omega = c\}\cap A^\omega_{t,c}} \big|\sum_{j=1,2}(\omega_j^\perp \partial_{x_j} P_{\kappa_1}\epsilon + i\omega_j^\perp \tilde{A}_j P_{\kappa_1}\epsilon)\big|^2 \gtrsim \|P_{\kappa_1}\epsilon\|^2_{L^\infty_{t_\omega} L^2_{x_\omega}},$$

since the magnetic potential is small in $L^\infty_t L^\infty_x$. Here the implicit constant depends on $|\kappa|$, but not $|\kappa_1|$ (which we recall was chosen $\ll |\kappa|$); this will be important since we can compensate a loss in this implicit constant by picking $|\kappa_1|$ small enough. Next, observe that

$$\frac{1}{2}\big|\sum_{j=1,2}(\omega_j \partial_{x_j} P_{\kappa_1}\epsilon + i\omega_j \tilde{A}_j P_{\kappa_1}\epsilon)\big|^2 + \frac{1}{2}|\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon|^2 - \sum_{j=1,2}\omega_j \mathrm{Re}[(\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon)\overline{(\partial_t + i\tilde{A}_0)P_{\kappa_1}\epsilon}] \geq 0,$$

and also that, due to the additional localization coming from the (suppressed) $Q^\pm_{<2\log|\kappa_1|}$ applied to $P_{\kappa_1}\epsilon$, we have

$$\sum_{\kappa_1 \subset \frac{3}{2}\kappa} \|P_{\kappa_1}\epsilon\|^2_{NF[\kappa]*} \gtrsim \|P_\kappa \epsilon\|^2_{NF[\kappa]*}$$

In order to derive the desired "a priori bound assumption", we need to estimate the first three integral expressions in (9.47). To begin with, the argument given above for the standard energy conservation implies that

$$\begin{aligned}
&\sum_{\kappa_1 \subset \frac{3}{2}\kappa} \big[\int_{0\times\mathbb{R}^2\cap A^\omega_{t,c}} \big[\frac{1}{2}|\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon|^2 + \sum_{j=1,2}\frac{1}{2}|\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon|^2\big]\, dtdx \\
&\qquad - \int_{t\times\mathbb{R}^2\cap A^\omega_{t,c}} \big[\frac{1}{2}|\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon|^2 + \sum_{j=1,2}\frac{1}{2}|\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon|^2\big]\, dtdx\big] \\
&\qquad \lesssim \|\nabla_{x,t}\epsilon(0,\cdot)\|^2_{\dot{H}^{-1}} + \|F\|^2_{N[0]} + \nu(D_1)\|\epsilon\|^2_{S[0]} + \|F\|_{N[0]}\|\epsilon\|_{S[0]}
\end{aligned} \tag{9.48}$$

where we have $\nu(D_1) \to 0$ as $D_1 \to \infty$, and the implicit constant is absolute.
It remains to control

$$\begin{aligned}
&\sum_{\kappa_1 \subset \frac{3}{2}\kappa} \int_{A^\omega_{t,c}} \mathrm{Re}\big[(\partial_t P_{\kappa_1}\epsilon + i\tilde{A}_0 P_{\kappa_1}\epsilon)\overline{\big[(\partial_t + i\tilde{A}_0)^2 - \sum_{j=1,2}(\partial_{x_j} + i\tilde{A}_j)^2\big]P_{\kappa_1}\epsilon}\big] \\
&\qquad\qquad + \mathrm{Re}[(\partial_{x_j} P_{\kappa_1}\epsilon + i\tilde{A}_j P_{\kappa_1}\epsilon)\overline{i(\partial_t A_j - \partial_{x_j} A_0)P_{\kappa_1}\epsilon}]\, dtdx,
\end{aligned}$$

which we do by essentially following the steps in (1), (2), (3) of the standard energy estimate above: proceeding by exact analogy, we need to estimate the following expressions:
(1'): The terms

$$\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{A^{\omega}_{t,c}}\mathrm{Re}\big[(\partial_t P_{\kappa_1}\epsilon+i\tilde{A}_0P_{\kappa_1}\epsilon)\overline{P_{2\kappa_1}\tilde{F}^{\pm}_{\kappa_1}}\big]\,dtdx$$

Here we recall that we suppress the additional localization operator $Q^{\pm}_{<2\log|\kappa_1|}$ in front of ϵ. To estimate this term, write it as

$$\begin{aligned}&\mathrm{Re}\big[\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{\mathbb{R}^{2+1}}(\partial_t P_{\kappa_1}\epsilon+i\tilde{A}_0P_{\kappa_1}\epsilon)\overline{\chi_{A^{\omega}_{t,c}}(t,x)P_{2\kappa_1}\tilde{F}^{\pm}_{\kappa_1}}\,dtdx\big]\\&=\mathrm{Re}\big[\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{\mathbb{R}^{2+1}}(\partial_t P_{\kappa_1}\epsilon+i\tilde{A}_0P_{\kappa_1}\epsilon)\overline{\chi_{A^{\omega}_{t,c}}(t,x)P_{2\kappa_1}\tilde{F}^{\pm}_{\kappa_1}}\,dtdx\big]\\&=\mathrm{Re}\big[\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{\mathbb{R}^{2+1}}(\partial_t P_{\kappa_1}\epsilon+i\tilde{A}_0P_{\kappa_1}\epsilon)\overline{P_{\kappa_1}Q^{\pm}_{<2\log|\kappa_1|}\big(\chi_{A^{\omega}_{t,c}}(t,x)P_{2\kappa_1}\tilde{F}^{\pm}_{\kappa_1}\big)}\,dtdx\big]\end{aligned}$$

where we again exploited the suppressed localization of ϵ close to the light cone. Now we use Lemma 2.20, together with Lemma 2.11, as well as Cauchy-Schwarz and the suppressed localization of ϵ to bound the preceding by

$$\begin{aligned}&\big|\mathrm{Re}\big[\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{\mathbb{R}^{2+1}}(\partial_t P_{\kappa_1}\epsilon+i\tilde{A}_0P_{\kappa_1}\epsilon)\overline{P_{\kappa_1}Q^{\pm}_{<2\log|\kappa_1|}\big(\chi_{A^{\omega}_{t,c}}(t,x)P_{2\kappa_1}\tilde{F}^{\pm}_{\kappa_1}\big)}\,dtdx\big]\big|\\&\lesssim\|P_{\frac{3}{2}\kappa}\epsilon\|_{S[0]}[\nu\|P_{\frac{3}{2}\kappa}\epsilon\|_{S[0]}+\|P_{\frac{3}{2}\kappa}F\|_{N[0]}],\end{aligned}$$

where again $\nu=\nu(D_1,|\kappa_1|)$ can be made small as $D_1,|\kappa_1|^{-1}\to\infty$. Note that the implicit constant in the preceding inequality is depending on $|\kappa|$.
(2'), (3'), (4'): The terms

$$\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{A^{\omega}_{t,c}}\nabla_{x,t}\tilde{A}P_{\kappa_1}\epsilon\overline{\nabla_{x,t}P_{\kappa_1}\epsilon}\,dxdt,$$

$$\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{A^{\omega}_{t,c}}\tilde{A}^2P_{\kappa_1}\epsilon\overline{\nabla_{x,t}P_{\kappa_1}\epsilon}\,dxdt$$

$$\sum_{\kappa_1\subset\frac{3}{2}\kappa}\int_{A^{\omega}_{t,c}}[(\nabla_{t,x}+iA)^2,P_{\kappa_1}]\epsilon\overline{\nabla_{t,x}\epsilon}\,dtdx$$

These are estimated exactly as in (2), (3), (4), exploiting the angular separation of the inherent factors, as well as the fact that we may let $D_1\to\infty$ independently of $|\kappa_1|$. The "a priori bound assumption" now follows from (1')-(3') by picking D_1 and $|\kappa_1|^{-1}$ large enough such that all terms in the above bounds involving $\|\epsilon\|_{S[0]}$ come with a factor of at most size $c_7(E_C,|\kappa|)$, the latter as in the "a priori bound assumption" above. This completes the proof of Proposition 9.14. □

Due to frequency leakage coming from the magnetic term we shall also require energy estimates that take $N\to S$, or alternatively, preservation of frequency envelope. For the following lemma, we allow more general frequency support of A_α. Hence consider the following equation[22]

$$\Box u+2i\partial^{\alpha}[u\,A_\alpha]=F,\quad u[0]=(f,g)\tag{9.49}$$

where F has the property that $F=F_1+F_2$ where $\|F_1\|_N:=\Big(\sum_{k\in\mathbb{Z}}\|F\|^2_{N[k]}\Big)^{\frac{1}{2}}$ is finite and with F_2 controlled by a frequency envelope, i.e., $\|P_kF_2\|_{N[k]}\le c_k$ and $\{c_k\}_{k\in\mathbb{Z}}$ is sufficiently flat (as defined above).

[22]Note that here we have to take the derivative outside the product, since otherwise we cannot control high-high interactions. In step 4 we will revert to the original form of the operator as before, without making a priori modulation or frequency restrictions; this does not lead to problems there since we work with a function u instead of ϵ with a different scaling.

Furthermore, $(f,g) = (f_1, g_1) + (f_2, g_2)$ with $\|(f_1, g_1)\|_{L^2\times \dot H^{-1}}$ finite and $\|P_k(f_2, g_2)\|_{L^2\times \dot H^{-1}} \le d_k$ where d_k is again a sufficiently flat envelope. Finally,

$$A_\alpha = \Delta^{-1} \sum_{j=1,2} \partial_j [R_\beta \psi_L^1 R_j \psi_L^2 - R_\beta \psi_L^2 R_j \psi_L^1]$$

is more general than in the previous proposition. Here ψ_L^1 and ψ_L^2 are finite energy free waves (with energy bounded by E_{crit}). Now one has the following result.

Lemma 9.15. *Let u be a solution of (9.49) with F and f, g as above. Then $u = u_1 + u_2 + u_3$ where $\|u_1\|_S \lesssim \|F_1\|_N + \|(f_1, g_1)\|_{L^2\times\dot H^{-1}}$, and $\|P_k u_2\|_{S[k]} \lesssim c_k$, $\|P_k u_3\|_{S[k]} \lesssim d_k$. The implied constants only depend on the energy of $\psi_L^{1,2}$.*

Proof. We restrict ourselves to $P_j u$. By scaling $j = 0$. Now split $A_\alpha = \sum_{i=1}^3 A_\alpha^{(i)}$ where

$$A_\alpha^{(1)} = \sum_{k<-C} P_k A_\alpha, \quad A_\alpha^{(2)} = \sum_{-C<k<C} P_k A_\alpha \tag{9.50}$$

The constant C in (9.50) is chosen such that the proof of Proposition 9.14 applies to the low frequency part of A_α. Then we write

$$\begin{aligned} \Box P_0 u + 2i\partial^\alpha P_0 u\, A_\alpha^{(1)} &= P_0 F - 2iP_0[\partial^\alpha u\, A_\alpha^{(2)} + \partial^\alpha u\, A_\alpha^{(3)}] + L(\partial^\alpha u, \nabla A_\alpha^{(1)}) \\ P_0 u[0] &= P_0(f,g) \end{aligned}$$

Here $L(\cdot,\cdot)$ stands for the commutator in $P_0(uv) = vP_0u + L(u, \nabla v)$. We now divide $\mathbb{R} = \bigcup_{\ell=1}^M I_\ell$ into disjoint intervals with the property that

$$\max_\ell \|P_0[\partial^\alpha u\, A_\alpha^{(2)} + \partial^\alpha u\, A_\alpha^{(3)}]\|_{N[0](I_\ell\times\mathbb{R}^2)} \le \gamma\, \|\psi_L\|_S^2 \sum_{k\in\mathbb{Z}} 2^{-\sigma_0|k|} \|P_k u\|_{S[k]}$$

where $\gamma > 0$ can be made arbitrarily small, and $M = M(E_{crit}, \gamma)$. To see this, one argues as in several previous instances. First consider $A^{(1)}$. In case of angular alignment of the Fourier supports of any two of the inputs, one obtains a gain as shown in Section 5. On the other hand, in case of angular separation the desired smallness is achieved by a careful choice of the I_ℓ, see Section 7. Finally, for $A^{(2)}$ one uses the high-high-low gains in the trilinear estimates (see the form of the weights $w(j_1, j_2, j_3)$ in Section 5 when $\max(j_2, j_3) > C$). The commutator terms satisfies the bound

$$\|L(\partial^\alpha u, \nabla A_\alpha^{(1)})\|_{N[0]} \lesssim 2^{-C}\, \|\psi_L\|_S^2 \sum_{k\in\mathbb{Z}} 2^{-\sigma_0|k|} \|P_k u\|_{S[k]}$$

since $\nabla A_\alpha^{(1)}$ gains a factor of 2^{-C}. We now apply the covariant energy bound of Proposition 9.14 to conclude that (with P_j instead of P_0)

$$\|P_j u\|_{S[j]} \le C(E_{crit})(\|P_j(f,g)\|_{L^2\times\dot H^{-1}} + (\gamma + 2^{-C}) \sum_{k\in\mathbb{Z}} 2^{-\sigma_0|k-j|} \|P_k u\|_{S[k]} + \|P_j F\|_{N[j]})$$

The lemma now follows from this estimate provided the frequency envelopes are flat enough compared to σ_0. □

9.4.2. *Controlling the error terms.* In this section, we complete the proof of Proposition 9.12. This amounts to bounding each of the terms on the right-hand side of (9.37) one by one using the covariant energy estimate of the previous section.

We begin with the first term in (9.37), i.e., $\sum_{j=1,3} F_\alpha^{3j}(P_0 Q_{<D}\epsilon; P_{<-D_1}; \psi_L, \psi_L)$ from which we which we have subtracted the magnetic potential term. Thus, we claim that we can decompose, with A_β as in (9.38),

$$\sum_{j=1,3} F_\alpha^{3j}(P_0 Q_{<D}\epsilon; P_{<-D_1}; \psi_L, \psi_L) - 2i\partial^\beta P_0 Q_{<D}\epsilon\, A_\beta$$

into the sum of two terms, one of which has controlled frequency envelope and the other small S-norm as in (9.20). By (3.14), this difference equals

$$iP_0Q_{<D}\epsilon_\alpha\, I\partial^\beta\partial_j^{-1}P_{<-D_1}\mathcal{Q}_{\beta j}(\psi_L,\psi_L)+iP_0Q_{<D}R^\beta\epsilon\;\partial_j^{-1}IP_{<-D_1}\partial_\alpha\mathcal{Q}_{\beta j}(\psi_L,\psi_L)$$
$$+\,iP_0Q_{<D}\partial^\beta\big[\epsilon_\alpha\, I^c\partial_j^{-1}P_{<-D_1}\mathcal{Q}_{\beta j}(\psi_L,\psi_L)\big]+iP_0Q_{<D}\partial_\alpha\big[\epsilon^\beta\, I^c\partial_j^{-1}P_{<-D_1}\mathcal{Q}_{\beta j}(\psi_L,\psi_L)\big]$$

Denoting these terms by $term_{11}$ - $term_{14}$, respectively, we now proceed to estimate them by means of Section 5.2. Let us now assume that ϵ is of the envelope type, see ϵ_1 in (9.19). Then by (5.88)

$$\|term_{11}\|_{N[0]}\lesssim 2^{-\sigma D_1}\|P_0\epsilon_\alpha\|_{S[0]}\|\psi_L\|_S^2\ll C_4\,d_0$$

for D_1 large depending on E_{crit}. The contribution of ϵ_2 is estimated similarly. For $term_{12}$ one uses that by Lemma 5.7

$$\|P_0Q_{<D}R^\beta\epsilon\;\partial_j^{-1}IP_k\partial_\alpha\mathcal{Q}_{\beta j}(\psi_L,\psi_L)\|_{N[0]}\lesssim 2^{\sigma k}\|P_0\epsilon\|_{S[0]}\|\psi_L\|_S^2$$

which is sufficient for both $\epsilon_{1,2}$ since $k\le -D_1$.
For $term_{13}$, we use the last part of Corollary 5.4; thus if we choose the implicit constant in the definition of I^c sufficiently large (depending only on E_{crit}), we get

$$\|iP_0Q_{<D}\partial^\beta\big[\epsilon_\alpha\, I^c\partial_j^{-1}P_{<-D_1}\mathcal{Q}_{\beta j}(\psi_L,\psi_L)\big]\|_{N[0]}\ll\|P_0\epsilon_\alpha\|_{S[0]}\|\psi_L\|_S^2$$

which is again enough for both the contribution of $\epsilon_{1,2}$. The last term $term_{14}$ is of course handled analogously.

The second term in (9.37) is bounded by (see (3.16))

$$\|\partial^\beta P_0Q_{<D}\big[\tilde\epsilon\,\partial_j^{-1}P_{<-D_1}I^c\mathcal{Q}_{\alpha j}(\psi_L,\psi_L)\|_{N[0]}+\|P_0Q_{<D}\big[R_\beta\tilde\varepsilon\,\partial_j^{-1}P_{<-D_1}I\partial^\beta\mathcal{Q}_{\alpha j}(\psi_L,\psi_L)\|_{N[0]}\tag{9.51}$$

We can bound the first by

$$\|\partial^\beta P_0Q_{<D}\big[\tilde\epsilon\,\partial_j^{-1}P_{<-D_1}I^c\mathcal{Q}_{\alpha j}(\psi_L,\psi_L)\|_{N[0]}\ll\|P_0\tilde\epsilon\|_{S[0]}\|\psi_L\|_S^2,$$

using Corollary 5.4 as well as choosing the constant in I^c large enough, while we can bound

$$\|P_0Q_{<D}\big[R_\beta\tilde\varepsilon\,\partial_j^{-1}P_{<-D_1}I\partial^\beta\mathcal{Q}_{\alpha j}(\psi_L,\psi_L)\|_{N[0]}\lesssim 2^{-\sigma D_1}\|\tilde\varepsilon\|_{S[0]}\|\psi_L\|_S^2,$$

having used Lemma 5.7.

The third term in (9.37) is the commutator

$$\begin{aligned}&\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L)-\sum_{j=1,3}F_\alpha^{3j}(P_0Q_{<D}\epsilon;P_{<-D_1};\psi_L,\psi_L)\\&=\tilde P_0\partial^\beta[Q_{<D}\epsilon_\alpha\,\nabla A_\beta]+\tilde P_0\partial^\alpha[Q_{<D}\epsilon^\beta\,\nabla A_\beta]\end{aligned}$$

where the second line is schematic. Hence, the smallness for this term is obtained just as in the preceding term via Lemma 5.7.

Next, as the fourth term we face the commutator

$$\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{\ge D}\epsilon;P_{<-D_1};\psi_L,\psi_L)-\sum_{j=1,3}P_0Q_{\ge D}F_\alpha^{3j}(\tilde\epsilon;P_{<-D_1};\psi_L,\psi_L)$$
$$=\sum_{j=1,3}P_0F_\alpha^{3j}(Q_{\ge D}\epsilon;P_{<-D_1};\psi_L,\psi_L)-\sum_{j=1,3}P_0Q_{\ge D}F_\alpha^{3j}(\epsilon;P_{<-D_1};\psi_L,\psi_L)\tag{9.52}$$
$$+\sum_{j=1,3}P_0Q_{\ge D}F_\alpha^{3j}(\epsilon;P_{<-D_1};\psi_L,\psi_L)-\sum_{j=1,3}P_0Q_{\ge D}F_\alpha^{3j}(\tilde\epsilon;P_{<-D_1};\psi_L,\psi_L)\tag{9.53}$$

First, (9.52) is bounded by

$$2^{-\sigma D_1}\|P_0\epsilon\|_{S[0]}\|\psi_L\|_S^2,$$

by the same commutator logic as before and Lemma 5.1. Second, since $\epsilon = \tilde{\epsilon}$ on I_1, the length of which is bounded below by an absolute constant by Case 1 above, one obtains that, upon letting $\chi_{\tilde{I}_1}$ be a smooth cutoff localizing to an interval containing I_1 obtained by adding an interval of length $2^{-\frac{D}{2}}$ to the top of I_1

$$\begin{aligned}
&\Big\| \sum_{j=1,3} P_0 Q_{\geq D} F_\alpha^{3j}(\epsilon - \tilde{\epsilon}; P_{<-D_1}; \psi_L, \psi_L)\Big\|_{N[0](I_1\times\mathbb{R}^2)} \\
&\lesssim \Big\| \sum_{j=1,3} P_0 Q_{\geq D}\big[\chi_{\tilde{I}_1} F_\alpha^{3j}(\epsilon - \tilde{\epsilon}; P_{<-D_1}; \psi_L, \psi_L)\big]\Big\|_{L_t^1(I_1; L_x^2)} \\
&+ \Big\| \sum_{j=1,3} P_0 Q_{\geq D}\big[(1-\chi_{\tilde{I}_1}) F_\alpha^{3j}(\epsilon - \tilde{\epsilon}; P_{<-D_1}; \psi_L, \psi_L)\big]\Big\|_{N[0](I_1\times\mathbb{R}^2)} \\
&\lesssim 2^{-D}\|P_0(\epsilon - \tilde{\epsilon})\|_{L_t^\infty L_x^2}\|P_{<-D_1}\nabla^{-1}(\psi_L^2)\|_{L_t^\infty L_x^\infty} + 2^{-ND}\|P_0(\epsilon-\tilde{\epsilon})\|_{S[0]}\|\psi_L\|_S^2 \\
&\lesssim (2^{-D-D_1} + 2^{-ND})\|P_0(\epsilon - \tilde{\epsilon})\|_{L_t^\infty L_x^2}\|\psi_L\|_S^2
\end{aligned}$$

where we used Bernstein's inequality in the last line.

The fifth term is a collection of quintilinear and higher order terms, and we deal with it in the appendix.

The terms six through eight are easy by Lemma 5.1, Lemma 5.5, Lemma 5.7 and (9.17). More precisely, they inherit the frequency profile of $\tilde{\epsilon}$ times a factor of $\varepsilon_2^{\frac{1}{4}}$; this is good enough to bootstrap both ϵ_1 and ϵ_2.

The ninth term in (9.37) is split as follows, see (3.14):

$$\begin{aligned}
&P_0 Q_{<D} F_\alpha^3(\tilde{\epsilon}; P_{\geq -D_1}; \psi_L, \psi_L) \\
&= i\partial^\beta P_0 Q_{<D}[\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{\geq -D_1}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] - iP_0 Q_{<D}[P_0 R_\beta \tilde{\epsilon}\; \partial_j^{-1} I\partial^\beta P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\alpha j}(\psi_L, \psi_L)] \\
&- i\partial^\beta P_0 Q_{<D}[P_{>0} R_\beta \tilde{\epsilon}\, \partial_j^{-1} I P_{>0}\mathcal{Q}_{\alpha j}(\psi_L, \psi_L)] - i\partial^\beta P_0 Q_{<D}[P_{<-5} R_\beta \tilde{\epsilon}\, \partial_j^{-1} I P_0 \mathcal{Q}_{\alpha j}(\psi_L, \psi_L)] \\
&+ iP_0 Q_{<D}[P_0 \partial^\beta \tilde{\epsilon}_\alpha\, \partial_j^{-1} I P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] + iP_0 Q_{<D}[P_0 R^\beta \tilde{\epsilon}\; \partial_j^{-1} I\partial_\alpha P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] \\
&+ i\partial_\alpha P_0 Q_{<D}[P_{>0} R^\beta \tilde{\epsilon}\, \partial_j^{-1} I P_{>0}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] + i\partial_\alpha P_0 Q_{<D}[P_{<-5} R^\beta \tilde{\epsilon}\, \partial_j^{-1} I P_0 \mathcal{Q}_{\beta j}(\psi_L, \psi_L)] \\
&+ i\partial^\beta[\tilde{\epsilon}_\alpha\, I^c \partial_j^{-1} P_{\geq -D_1}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] - i\partial^\beta[\tilde{\epsilon}_\beta\, I^c \partial_j^{-1} P_{\geq -D_1}\mathcal{Q}_{\alpha j}(\psi_L, \psi_L)]
\end{aligned}$$

Denote these terms in this order by $term_{91}$ through $term_{100}$. First, we rewrite $term_{91}$ in terms of the usual trichotomy:

$$\begin{aligned}
i\partial^\beta P_0 Q_{<D}[\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{\geq -D_1}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] &= i\partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] \\
+ i\partial^\beta P_0 Q_{<D}[P_{<-5}\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_0\mathcal{Q}_{\beta j}(\psi_L, \psi_L)] &+ i\partial^\beta P_0 Q_{<D}[P_{>0}\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{>0}\mathcal{Q}_{\beta j}(\psi_L, \psi_L)]
\end{aligned} \tag{9.54}$$

The first term in (9.54) is rewritten as the sum

$$\begin{aligned}
&\partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\tilde{\psi}_L, \tilde{\psi}_L)] + \partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\tilde{\psi}_L, \breve{\psi}_L)] \\
&+ \partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\breve{\psi}_L, \tilde{\psi}_L)] + \partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\breve{\psi}_L, \breve{\psi}_L)]
\end{aligned}$$

where we followed the notation of Corollary 7.29. Each of the terms containing $\breve{\psi}$ is bootstrapped easily, using the smallness of $\tilde{\epsilon}_\alpha$ and Lemma 5.5. Rescaling and square-summing these contributions are placed in ϵ_2; alternatively, one can recover the frequency envelope using the smallness of δ_1 for the bootstrap. For the first term, we proceed as in (b) of Case 1. More precisely, using the smallness of δ_0 (and the fact that the Besov smallness of ψ at the edges of the intervals J_ℓ inherits itself to $\tilde{\psi}$) as well as the frequency evacuation property for large n, one obtains that

$$\|\partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1} P_{-D_1\leq\cdot<-5}\mathcal{Q}_{\beta j}(\tilde{\psi}_L, \tilde{\psi}_L)]\|_{N[0]} \lesssim C_4^3 C_2^2 \delta_0 \sum_{r\in[-D_1,-5]} \|P_r\tilde{\psi}_L\|_{S[k]}$$

As usual, replacing the output frequency 0 by k and square summing, this gets turned into an S bound, leading to an ϵ_2 contribution. If $\tilde{\epsilon} = \tilde{\epsilon}_2$, then it again suffices to consider $\tilde{\psi}$. In this case, one needs to gain extra smallness by partitioning I_1 further; however, the number of intervals needed for this partition only depends on the energy in an absolute way (i.e., not on the stage of the induction). First, we may assume that there is angular separation between the Fourier supports of the two $\tilde{\psi}_L$ inputs due to the bound

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{-m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\lesssim 2^{-m_0}}} \|\partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1}P_{-D_1\le\cdot<-5}\mathcal{Q}_{\beta j}(P_{\kappa_1}\tilde{\psi}_L, P_{\kappa_2}\tilde{\psi}_L)]\|_{N[0]} \ll \|P_0\tilde{\epsilon}\|_{S[0]}$$

see Section 5.3. Here we used that $\|\psi_L\|_S^2$ is bounded by the energy in an absolute way, which allows us to chose m_0 in the same fashion. On the other hand, the remaining term

$$\sum_{\substack{\kappa_1,\kappa_2\in\mathcal{C}_{-m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)> 2^{-m_0}}} \|\partial^\beta P_0 Q_{<D}[P_0\tilde{\epsilon}_\alpha\, I\partial_j^{-1}P_{-D_1\le\cdot<-5}\mathcal{Q}_{\beta j}(P_{\kappa_1}\tilde{\psi}_L, P_{\kappa_2}\tilde{\psi}_L)]\|_{N[0]}$$

is estimated by placing $\mathcal{Q}_{\beta j}(P_{\kappa_1}\tilde{\psi}_L, P_{\kappa_2}\tilde{\psi}_L)$ into $L^2_{t,x}$, see the reasoning leading up to (7.16), followed by a decomposition of the interval of integration. Here is important to note that D_1 only depends on the energy.

For the second term in (9.54) consider first $\tilde{\epsilon}_1$; then the frequency envelope of $\tilde{\epsilon}_1$ is inherited by this expression. More precisely, for $P_k\tilde{\epsilon}$ one gains a weight $2^{-\sigma k}$ from Lemma 5.5 which is sufficient for the bootstrap provided k is sufficiently large and negative; if not, then one applies the same divisibility as for the previous terms. the same reasoning applies to $\tilde{\epsilon}_2$.

Finally, for the third term in (9.54) consider first the contribution by $\tilde{\epsilon}_1$. In that case one has

$$\|\partial_\alpha P_0 Q_{<D}[P_{>0}R^\beta\tilde{\epsilon}\,\partial_j^{-1}IP_{>0}\mathcal{Q}_{\beta j}(\psi_L,\psi_L)]\|_{N[0]} \lesssim \sup_{k_1,k_2>0} 2^{-\sigma_0|k_1-k_2|}\|P_{k_1}\tilde{\epsilon}\|_{S[k_1]}\|P_{k_2}\psi_L\|_{S[k_2]} \tag{9.55}$$

which can be made $\ll C_4C_2\,\delta_1$ by choosing δ_0 small and n large. On the other hand, if $\tilde{\epsilon}=\tilde{\epsilon}_2$, then one gains smallness in two ways: if any one of $\tilde{\epsilon}$, or the two ψ_L inputs has large frequency, then ones gains smallness from the weight w in Lemma 5.5. If the three inputs have frequency of size $O(1)$, then one gains smallness by divisibility as before.

Next, we note that $term_{92}$ is treated in the same fashion as the first term on the right-hand side of (9.54).

The terms $term_{93}$ and $term_{97}$ are of the high-high type and are estimated exactly as in (9.55), and $term_{94}$, $term_{98}$ are essentially the same as the low-high term on the righ-hand side of (9.54). To bound $term_{95}$ and $term_{96}$ one applies the same divisibility considerations as in the high-low case of $term_{91}$. Finally, the terms $term_{99}$ and $term_{100}$ can be handled similarly, first reducing

$$I^c\partial_j^{-1}P_{\ge -D_1}\mathcal{Q}_{\beta j}(\psi_L,\psi_L), I^c\partial_j^{-1}P_{\ge -D_1}\mathcal{Q}_{\alpha j}(\psi_L,\psi_L)$$

to frequency $O(1)$ via Lemma 5.1, and then using the divisibility property for $\|I^c\partial_j^{-1}P_{O(1)}\mathcal{Q}_{\beta j}(\psi_L,\psi_L)\|_{\dot{X}_0^{0,-\epsilon,2}}$ etc.

The tenth and eleventh terms in (9.37) are essentially the same so it suffices to estimate the former. Since the details are quite similar to the preceding arguments, we will proceed schematically. Beginning with $\tilde{\epsilon}=\tilde{\epsilon}_1$, we split

$$\begin{aligned} P_0Q_{<D}F^3_\alpha(\psi,\psi,\tilde{\epsilon}) &= P_0Q_{<D}F^3_\alpha(\psi_L,\psi_L,\tilde{\epsilon}) + P_0Q_{<D}F^3_\alpha(\psi_L,\psi_{NL},\tilde{\epsilon}) \\ &\quad + P_0Q_{<D}F^3_\alpha(\psi_{NL},\psi_L,\tilde{\epsilon}) + P_0Q_{<D}F^3_\alpha(\psi_{NL},\psi_{NL},\tilde{\epsilon}) \end{aligned} \tag{9.56}$$

and furthermore, using Corollary 7.29,

$$\begin{aligned} P_0Q_{<D}F^3_\alpha(\psi_L,\psi_L,\tilde{\epsilon}) &= P_0Q_{<D}F^3_\alpha(\tilde{\psi}_L,\tilde{\psi}_L,\tilde{\epsilon}) + P_0Q_{<D}F^3_\alpha(\tilde{\psi}_L,\breve{\psi}_L,\tilde{\epsilon}) \\ &\quad + P_0Q_{<D}F^3_\alpha(\breve{\psi}_L,\tilde{\psi}_L,\tilde{\epsilon}) + P_0Q_{<D}F^3_\alpha(\breve{\psi}_L,\breve{\psi}_L,\tilde{\epsilon}) \end{aligned} \tag{9.57}$$

All terms here are going to be placed into the S error ϵ_2 since they inherit the frequency envelope of ψ. The trilinear estimates of Lemma 5.5 allows for this, with the required smallness for the terms containing ψ_{NL} is gained by the smallness of $\|\psi_L\|_S\|\psi_{NL}\|_S$. Furthermore, the terms containing $\breve{\psi}$ are easy due to the smallness $\|\breve{\psi}\|_S < C_2\,\delta_1$ and the bootstrap assumption on $\tilde{\epsilon}$ (one then chooses ε_0 small enough). The most interesting term here is $P_0Q_{<D}F^3_\alpha(\tilde{\psi}_L,\tilde{\psi}_L,\tilde{\epsilon})$. To place it in ϵ_2 one uses the same small Besov/frequency evacuation logic that we have used several times before.

The twelfth term in (9.37) is easy since it inherits the frequency envelope of $\tilde{\epsilon}$ and basically bootstraps itself.

The thirteenth term has to be placed entirely into the S-error ϵ_2. This can be done using the high-high gain in Lemma 5.5 and in (5.41) as demonstrated several times before.

Finally, the fourteenth term is the cubic one which is again easy. This concludes obtaining the bootstrap for $Q_{<D}\epsilon$. We now need to do the same thing for

$$Q_{\geq D}\epsilon$$

Since this is a technical repetition of similar reasoning, we again defer this to the appendix.
Finally, to complete the bootstrap, we of course also need to take into account the contribution of the linear evolution of the data for $P_0Q_{<D}\epsilon$, corresponding to the first term on the right in the parametrix (2.72) (with $I\psi$ corresponding to $P_0Q_{<D}\epsilon$). But according to the considerations following (2.72), we can write

$$P_0Q_{<D}\epsilon = P_0Q_{<D}\big(\chi_{[-T_1,T_1]}\epsilon\big) + P_0Q_{<D}\big((1-\chi_{[-T_1,T_1]})\epsilon\big)$$

where T_1 is as in Case 1 above, and χ smoothly truncates to this interval and equals 1 on a smaller sub-interval of length $\gg D^{-1}$. Then using the bootstrap already accomplished in Case 1, we can split

$$P_0Q_{<D}\big(\chi_{[-T_1,T_1]}\epsilon\big)[0] = f_0 + g_0$$

where we have $\|f_0\|_{L^2\times\dot{H}^{-1}} \ll C_4 d_0$ while for g_0, replacing 0 by general $k\in\mathbb{Z}$, we get $\sum_{k\in\mathbb{Z}}\|g_k\|^2_{L^2\times\dot{H}^{-1}} \ll C_4C_2\delta_1$. On the other hand, we get

$$\|P_0Q_{<D}\big((1-\chi_{[-T_1,T_1]})\epsilon\big)[0]\|_{L^2\times\dot{H}^{-1}} \ll \|\varepsilon\|_{S[0]}$$

by choosing D large enough, and so this also leads to an acceptable contribution This concludes the proof of Proposition 9.12. □

It is now easy to conclude the proof of 9.11. More precisely, as indicated in Figure 6, one proceeds in the direction of increasing time by passing from I_1 to I_2 and so on. Writing $I_2 = [a_2, b_2]$, one introduces the frequency envelope

$$\tilde{d}_k = \big(\sum_{r\in\mathbb{Z}} 2^{-\sigma|r-k|}\|P_r\|\varepsilon_1(a_2,\dot{)}\|^2_{L^2_x}\big)^{\frac{1}{2}}$$

and by the preceding argument, using the fact that $\|\varepsilon(a_2,0)\|_E < \varepsilon_0$, we obtain

$$\varepsilon|_{I_2} = \varepsilon_1 + \varepsilon_2$$

with

$$\|\varepsilon_2\|_{S(I_2\times\mathbb{R}^2)} \lesssim \delta_1,\ \|P_k\varepsilon_1\| \lesssim \tilde{d}_k$$

with implied constant only depending on E_{crit}. But by the preceding step, we also have

$$\tilde{d}_k \lesssim d_k$$

and we can then again conclude via Corollary 9.13 that the energy of $\varepsilon|_{I_2} < \varepsilon_0$ provided n and δ_1^{-1} are sufficiently large. Even though ϵ is initially only defined locally, Proposition 7.2 and the $\|\cdot\|_S$-norm bound of Proposition 9.12 imply that ϵ exists globally with the bounds stated in Proposition 9.11, see (9.13) and (9.14). □

Proof of Proposition 9.9. This follows simply by iterating Proposition 9.12, i.e., by passing from J_1 to J_2 and so forth in Figure 6. Even though the constant C_2 increases with ℓ, in the end one obtains a bound of the form (9.15). The final statement (9.10) is a consequence of our proof of Proposition 9.12 due to the frequency evacuation of the first Besov error from the atom ϕ_n^a. In fact, our estimates are based on control of the frequency envelope which therefore implies (9.10) at all stages of the induction. □

We include here two Corollaries of the proof just given: note from the proof of Proposition 9.12 that even if the perturbation $\epsilon[0]$ at time zero is controlled by a frequency envelope c_k which is concentrated at much higher frequencies than Ψ, nonetheless the evolution of ϵ will also involve a component that is no longer well frequency-localized (which forces us to implement the splitting $\epsilon = \epsilon_1 + \epsilon_2$). The reason for this are interactions of the schematic form such as $\nabla_{t,x}[\psi\nabla^{-1}(\epsilon^2)]$, which 'inherit' the frequency envelope of ψ. This issue is moot provided there is no distinction between ψ and ϵ:

Corollary 9.16. *Assume that ψ_α are the Coulomb components of an admissible wave map from $(-T_0, T_1)\times\mathbb{R}^2$ into $\mathbb{H}^2$, and that we have $\|P_k\psi_\alpha\|_{L^2} \leq c_k$ for a frequency envelope with $c_l 2^{-\sigma|k-l|} \leq c_k \leq c_l 2^{\sigma|k-l|}$ $\forall l, k$, with $\sigma > 0$ small enough. Then assuming a bound*

$$\|\psi\|_{S([-T_0,T_1]\times\mathbb{R}^2)} < C_0,$$

we can conclude

$$\|P_k\psi\|_{S([-T_0,T_1]\times\mathbb{R}^2)} < M(C_0)c_k$$

Proof. One proceeds just as in the preceding proof, but with ψ taking the role of ϵ; thus write schematically

$$\Box P_0\psi = \nabla_{t,x}P_0[\psi_H\nabla^{-1}(\psi^2)_L] + \nabla_{t,x}P_0[\psi_H\nabla^{-1}(\psi^2)_H] + \nabla_{t,x}P_0[\psi_L\nabla^{-1}(\psi^2)_H]$$

where we set

$$f_H g_L = \sum_{k\in Z} P_k f P_{<k-C_1} g,\ f_H g_H = \sum_{k\in Z} P_k f P_{[k-C_1.k_1+C_1]} g$$

with C_1 a constant depending on C_0 above. Then arguing exactly as in the preceding proof (in particular, one has to implement the null-form expansion in the nonlinearity) one finds a collection of time intervals I_j, $j = 1, 2, \ldots, N(C_0)$ such that

$$\|\nabla_{t,x}P_0[\psi_H\nabla^{-1}(\psi^2)_H]\|_{N[0](I_j\times\mathbb{R}^2)} + \|\nabla_{t,x}P_0[\psi_L\nabla^{-1}(\psi^2)_H]\|_{N[0](I_j\times\mathbb{R}^2)} \ll \sup_{l\in\mathbb{Z}} 2^{-\sigma|k-l|}\|P_l\psi\|_{S[l](I_j\times\mathbb{R}^2)}$$

provided $C_1, N(C_0)$ is chosen large enough in relation to C_0. But then using Proposition 9.14, one obtains inductively in j bounds of the form

$$\|P_k\psi\|_{S[k](I_j\times\mathbb{R}^2)} \leq M_j(C_0)c_k$$

via a bootstrap argument on $I_j \times \mathbb{R}^2$. □

Corollary 9.17. *Let ψ_α be as in the preceding corollary, with $\|\psi\|_{S([-T_0,T_1]\times\mathbb{R}^2)} < C_0$. Then there exists $\delta_1 = \delta_1(C_0) > 0$ such that if $\psi_\alpha + \varepsilon_\alpha$ are Coulomb components of an admissible map wave map with*

$$\|\varepsilon(0,\cdot)\|_{L^2_x} < \delta_1,$$

then the wave maps evolution of data $\psi_\alpha + \varepsilon_\alpha$ exists on $[-T_0, T_1]\times\mathbb{R}^2$, and we have

$$\|\varepsilon\|_{S([-T_0,T_1]\times\mathbb{R}^2)} \lesssim \delta_1$$

with implied constant depending on C_0.

Proof. This is exactly as in the proof of Proposition 9.12, with $\varepsilon_1 = 0$, and $\varepsilon_2 = \varepsilon$. □

9.5. **Completion of the proofs of Lemma 7.10 and Proposition 7.11.** We commence by proving the final assertion of Lemma 7.10. This follows immediately from Corollary 9.16.
Next, the proof of Proposition 7.11 follows from Corollary 9.17.

9.6. **Step 4: Adding the first large atomic component; preparing the second stage of Bahouri Gerard.** Recall from Section 9.2 that we wrote the data ϕ^n_α of the essentially singular sequence (at time $t = 0$) in the form

$$\phi^n_\alpha = \sum_{a=1}^{A_0}\phi^{na}_\alpha + w^{nA_0}_\alpha,$$

where A_0 was chosen such that the sum

$$\limsup_{n\to\infty}\sum_{a\geq A_0+1}\|\phi^{na}_\alpha\|^2_{L^2_x} \ll \varepsilon_0$$

As before $\varepsilon_0(E_{crit}) > 0$ is an absolute constant that depends only on the energy. Then recall from Section 9.3 that the atoms ϕ^{na}_α "split" the error term $w^{nA_0}_\alpha$ into finitely many pieces $w^{nA_0^{(i)}}_\alpha$, $0 \leq i \leq A_0$, ordered by the size of $|\xi|$ in their Fourier support. Of course our eventual goal is to describe the evolution of the Coulomb components (with $\phi_\alpha^n = \phi_\alpha^{n1} + i\phi_\alpha^{n2}$)

$$\psi_\alpha^n = \phi_\alpha^n e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\phi_k^{n1}}$$

Our strategy then is to construct "intermediate wave maps" bootstrapping the bounds from one to the next, starting with the low frequency ones to the higher frequency ones. In the previous section, we have shown that we can derive an a priori bound

$$\|\Psi_\alpha^{nA_0^{(0)}}\|_S = \left\|\Phi_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{nA_0^{(0)}1}}\right\|_S < C_{10}(E_{crit})$$

provided we choose A_0 above large enough and also pick n large enough. Moreover, we can then prove frequency localized bounds of the form

$$\left\|P_k\left[\Phi_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{nA_0^{(0)}1}}\right]\right\|_{S[k](\mathbb{R}^{2+1})} \leq C_{11}(E_{crit})c_k$$

for a suitable frequency envelope c_k with $\sum_{k\in\mathbb{Z}} c_k^2 \leq 1$, say, and c_k rapidly decaying for $k \notin (-\infty, \log(\lambda_n^1)^{-1})$, where the frequency scales of the ϕ^{na} are given by $(\lambda_n^a)^{-1}$.

We now pass to the next approximating map, with data given by

$$[w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1}]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1} + \phi_k^{n1}]} + o_{L^2}(1) = w_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1} + \phi_k^{n1}]}$$
$$+ \phi_\alpha^{n1} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1} + \phi_k^{n1}]} + o_{L^2}(1)$$

Here the first component satisfies

$$w_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1} + \phi_k^{n1}]} = w_\alpha^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k w_k^{nA_0^{(0)}1}} + o_{L^2}(1)$$

as $n \to \infty$ since $w_\alpha^{nA_0^{(0)}}$ is singular with respect to the scale of $\phi_k^{nA_0^{(0)}1}$. Technically speaking, this follows by means of the usual trichotomy considerations. We now need to understand the lack of compactness of the large added term

$$\tilde{\psi}_\alpha^{na} = \tilde{\psi}_\alpha^{n1} := \phi_\alpha^{n1} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1} + \phi_k^{n1}]},$$

which is where the second phase of Bahouri-Gerard needs to come in.

We now normalize via re-scaling to $\lambda_n^1 = 1$. This means now that the frequency support of $\tilde{\psi}_\alpha^{na}$ with $a = 1$ is uniformly concentrated around frequency $|\xi| \sim 1$. Observe that here *we cannot get rid of the phase* $e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k w_k^{nA_0^{(0)}1}}$, which may indeed "twist" the Coulomb components additionally. This will have a negligible effect, however, since the ψ-system (1.12)–(1.14) is invariant with respect to the modulation symmetry $\psi \mapsto e^{i\gamma}\psi$.

For technical reasons[23], we now apply a Hodge type decomposition to the components $\psi_{1,2}^{na}$ (here $1, 2$ refer to the derivatives on $\mathbb{R}^2$ with respect to the two coordinate directions), as well as for $\tilde{\psi}_{1,2}^{na}$. Thus write

$$\phi_1^{na} = \partial_1\tilde{\phi}^{na} + \partial_2\mathring{\phi}^{na} \tag{9.58}$$

$$\phi_2^{na} = \partial_2\tilde{\phi}^{na} - \partial_1\mathring{\phi}^{na} \tag{9.59}$$

$$\tilde{\psi}_1^{na} = \partial_1\zeta^{na} + \partial_2\eta^{na} \tag{9.60}$$

$$\tilde{\psi}_2^{na} = \partial_2\zeta^{na} - \partial_1\eta^{na} \tag{9.61}$$

More precisely, we define the components $\tilde{\phi}^{na}, \mathring{\phi}^{na}$, ζ^{na}, η^{na} using the preceding relations, imposing a vanishing condition at spatial infinity. All of this is at time $t = 0$, of course. Now following the procedure of the preceding section, using the bound

$$\|\Psi_\alpha^{nA_0^{(0)}}\|_S < C_{10}(E_{crit}),$$

[23]This has to do with the fact that the energy of the free wave equation involves a derivative.

we can select finitely many intervals I_j (whose number depends on $C_{10}(E_{crit}))$ such that

$$\Psi^{nA_0^{(0)}}|_{I_j} = \Psi_{jL}^{nA_0^{(0)}} + \Psi_{jNL}^{nA_0^{(0)}} \tag{9.62}$$

for each interval j, see Corollary 7.27. Moreover, it is straightforward to verify that our normalization $\lambda_n^1 = 1$ implies that $|I_j| \to \infty$ as $n \to \infty$; indeed, this follows from L^∞-bounds.
Next, pick the interval I_1 containing the initial time slice $t = 0$. Consider the magnetic potential (note that we do not use the Hodge decomposition here)

$$A_\nu^n := \sum_{j=1,2} \triangle^{-1}\partial_j[\Psi_\nu^{1nA_0^{(0)}}\Psi_j^{2nA_0^{(0)}} - \Psi_j^{1nA_0^{(0)}}\Psi_\nu^{2nA_0^{(0)}}]$$

Here we restrict everything to a non-resonant situation, i.e., we shall replace the above by

$$\begin{aligned} A_\nu^n = &\sum_{\substack{\kappa_{1,2}\in K_{-\Lambda_n}\\ \mathrm{dist}(\kappa_{1,2})\gtrsim 2^{-\Lambda_n}}} \sum_{|k-k_1|<\Lambda_n, |k_1-k_2|<\Lambda_n} \\ &\sum_{j=1,2} \triangle^{-1}\partial_j P_k I^{(n)}[P_{k_1,\kappa_1}\Psi_\nu^{1nA_0^{(0)}} P_{k_2,\kappa_2}\Psi_j^{2nA_0^{(0)}} - P_{k_1,\kappa_1}\Psi_j^{1nA_0^{(0)}} P_{k_2,\kappa_2}\Psi_\nu^{2nA_0^{(0)}}], \end{aligned} \tag{9.63}$$

where we have introduced the modulation cutoff

$$I^{(n)} := \sum_{k\in Z} P_k Q_{<k+\Lambda_n}$$

Here we shall let $\Lambda_n \to \infty$ as $n \to \infty$ sufficiently slowly. The errors thereby generated shall be treatable as perturbative errors. *This time we use the full Ψ, and not just the free wave part.* Our notation is somewhat inconsistent, since we do not include $A_0^{(0)}$. Since we keep this parameter fixed throughout this section, this omission will be inconsequential. From now on we shall denote $\Psi_\nu^{1nA_0^{(0)}} = \Psi_\nu^{1n}$ etc. to simplify the notation. We shall tacitly assume that the Ψ_ν^n allow the usual Hodge type decompositions as in the preceding.

Definition 9.18. *The covariant wave operator $\Box_{A^n}$ is defined via*

$$\Box_{A^n} u := \Box u + 2i\partial^\nu u A_\nu^n$$

The fundamental fact about this operator is that solutions obeying $\Box_{A^n} u = 0$ preserve the energy in the limit $n \to \infty$. This will allow us to modify the second stage of the Bahouri-Gerard method to the covariant d'Alembertian instead of the "flat" d'Alembertian. We state this rigorously as follows.

Lemma 9.19. *Assume that u is essentially supported at frequency 1, and that A_ν is essentially supported at frequencies $\ll 1$. By this we mean that*

$$\lim_{R\to\infty} \|P_{[-R,R]^c} u[0]\|_{L_x^2} = 0 \tag{9.64}$$

as well as

$$\lim_{n\to\infty} \|P_{>-R}\Psi^{1n}\|_S = 0 \tag{9.65}$$

for any $R > 0$. If u solves

$$\Box_{A^n} u = 0,\ u[0] = (\partial_t u, \nabla_x u) = (u_0, u_1) \in L^2 \times L^2$$

then one obtains a global bound (uniformly in the implicit Λ_n)

$$\|u\|_{\tilde{S}(\mathbb{R}^{2+1})} \lesssim \|u[0]\|_{L_x^2}$$

with implied constant depending on E_{crit} as well as $\sup_n \|\Psi^n\|_S$ (which control A^n), and we can conclude that

$$\|\partial_t u(t,\cdot)\|_{L_x^2}^2 + \|\nabla_x u(t,\cdot)\|_{L_x^2}^2 = \|u_0(t,\cdot)\|_{L_x^2}^2 + \|u_1(t,\cdot)\|_{L_x^2}^2 + o_{L^2}(1)$$

as $n \to \infty$, uniformly in $t \in \mathbb{R}$, provided $\Lambda_n \to \infty$ sufficiently slowly. We have introduced the slightly altered norm

$$\|u\|_{\tilde{S}(\mathbb{R}^{2+1})} = \|\nabla_x u\|_{S(\mathbb{R}^{2+1})} + \big(\sum_{k\in Z} \nabla_{x,t} P_k Q_{\geq k} u\|^2_{\dot{X}^{0,\frac{1}{2},\infty}}\big)^{\frac{1}{2}}$$

We also have

$$\lim_{R\to\infty} \|P_{[-R,R]^c} u\|_{\tilde{S}(\mathbb{R}^{2+1})} = 0$$

Proof. This follows by the same argument that we used to prove Proposition 9.14. In the latter proof, we assumed that the Coulomb potential was defined in terms of free waves. In order to get the present conclusion, one needs to invoke the decomposition from Corollary 7.27 on suitable time intervals. The additional contributions can be handled just as in the proof of Proposition 9.14. We have modified the norm to $\|\cdot\|_{\tilde{S}}$ to accommodate the different scaling and to strengthen it in the large modulation regime. This is possible since we don't have a time derivative hitting the term $2i\partial^\nu u A^n_\nu$, see e. g. the estimates in [56] in the simple case of large modulation for the output. These considerations easily furnish a bound of the form

$$\|u\|_{\tilde{S}} \lesssim \|u[0]\|_{L^2_x}$$

with implicit constant depending on $\|\Psi\|_S$. Similarly, one obtains the frequency localized bounds

$$\|P_k u\|_{\tilde{S}} \lesssim c_k$$

where we put

$$c_k = \big(\sum_{l\in\mathbb{Z}} 2^{-\sigma|l-k|} \|P_l u[0]\|^2_{L^2_x}\big)^{\frac{1}{2}}$$

with $\sigma > 0$ sufficiently small, just as in the proof of Proposition 9.12. In order to get the asymptotic (in n) energy conservation, one writes

$$\begin{aligned} &\Box P_{[-R,R]} Q_{<2R} u + 2i\partial^\nu P_{[-R,R]} Q_{<2R} u P_{<-10R} A^n_\nu \\ &= -P_{[-R,R]} Q_{<2R}\big[2i\partial^\nu u A^n_\nu\big] + 2i\partial^\nu P_{[-R,R]} Q_{<2R} u P_{<-10R} A^n_\nu \\ &= -P_{[-R,R]}\big[2i\partial^\nu Q_{\geq 2R} u A^n_\nu\big] + P_{[-R,R]} Q_{\geq 2R}\big[2i\partial^\nu u A^n_\nu\big] \\ &+ \big[2i\partial^\nu P_{[-R,R]} Q_{<2R} u A^n_\nu\big] - P_{[-R,R]}\big[2i\partial^\nu Q_{<2R} u A^n_\nu\big] =: F^n_R \end{aligned} \tag{9.66}$$

But then for fixed $R \gg 1$ we have

$$\lim_{n\to\infty} \|F^n_R\|_{\tilde{N}} = 0,$$

where $\|F\|_{\tilde{N}} := \big(\sum_{k\in\mathbb{Z}} \|\nabla_x P_k F\|^2_{N_1[k]}\big)^{\frac{1}{2}}$, and $\|\cdot\|_{N_1[k]}$ is defined as in Definition 2.9 but with $\|\cdot\|_{\dot{X}_k^{-\frac{1}{2}+\varepsilon,-1-\varepsilon,2}}$ replaced by $\|\cdot\|_{\dot{X}_k^{-1,-\frac{1}{2},1}}$. The latter modification is again a consequence of the fact that in the large modulation case, one gets a better estimate for the expressions $2iP_{[-R,R]} u A^n_\nu$ as there is no outer time derivative, and simple modifications of the proofs in section 2.3 (see also the proof of the energy estimate in the high modulation case in [56]) yield that this modifcation of $\|\cdot\|_N$ suffices to recover $\|\cdot\|_{\tilde{S}}$. Similarly, we have

$$\Box P_{[-R,R]} Q_{\geq 2R} u + P_{[-R,R]} Q_{\geq 2R}\big[\partial^\nu u A^n_\nu\big] = 0$$

and one checks readily that

$$\lim_{n\to\infty} \|P_{[-R,R]} Q_{\geq 2R}\big[\partial^\nu u A^n_\nu\big]\|_{\tilde{N}} = 0$$

But then the argument of the proof of Proposition 9.14 yields that

$$\|\nabla_{t,x} P_{[-R,R]} Q_{<2R} u(t,\cdot)\|_{L^2} = \|\nabla_{t,x} P_{[-R,R]} Q_{<2R} u(0,\cdot)\|_{L^2} + o(1)$$

and further $\forall t$ (where $o(1)$ indicates the behavior as $n \to \infty$)

$$\lim_{R\to\infty} \|\nabla_{t,x} P_{[-R,R]} Q_{<2R} u(t,\cdot)\|_{L^2} = \|\nabla_{t,x} u(t,\cdot)\|_{L^2} + o(1)$$

which gives the desired asymptotic energy conservation. □

In our applications of Lemma 9.19, (9.64) will hold due to the frequency localization inherent in our construction of the atoms; in other words, u will be 1-oscillatory after rescaling. The other condition (9.65) will hold due to (9.10), at least at the first stage of the construction (i.e., when adding the first atom as we are doing here). For $a = 2$ etc. we will use the exact same frequency evacuation property which gave rise to (9.10) in the first place.

9.6.1. *Dispersion for the covariant wave equation.* In this section we prove a weak form of dispersion for the initial value problem

$$\Box_{A^n} u = 0, \quad u[0] := (f, g) \tag{9.67}$$

where $\Box_{A^n}$ is as in Definition 9.18. For simplicity, we first consider the case where A^n_ν is defined as in (9.63) but with *free waves* Ψ^n_L. We shall assume that (f, g), whence also u by Lemma 9.15, are essentially supported at frequency 1, see Lemma 9.19. Generally speaking, u depends on n away from the time $t_0 = 0$, but the above limit is uniform in n and holds on any time-slice. We assume that the free waves ψ^n_L satisfy

$$\lim_{n\to\infty} \|P_{>-R}\psi^n_L\|_{L^2_x} = 0 \tag{9.68}$$

for any $R > 0$. We now claim the following main result of this subsection for the covariant wave equation (9.67). For simplicity, we drop n as a superscript.

Proposition 9.20. *Let u be a solution of (9.67), with $(f, g) \in \dot H^1 \times L^2$. Given $\gamma > 0$, there exists a decomposition*

$$u = u_1 + u_2$$

with the following properties:

- *$u_{1,2}$ satisfy the same a priori estimates which were proved for u in Lemma 9.19*
- *$\|u_2\|_{\tilde S} < \gamma$*
- *there exists $t_0 = t_0(\gamma, f, g, E_{crit})$ (but t_0 does not depend otherwise on ψ_L) such that for $|t| > t_0$ one has that*

$$\|u_1(t, \cdot)\|_{L^\infty_x} < \gamma, \tag{9.69}$$

uniformly for large enough n.

The proof of this result will be split into several pieces. The idea is to first obtain a "parametrix" for u, which is established by restricting to suitable time intervals (this is done via "divisibility"). Once we have such a parametrix (more precisely, a representation of u as a sum of Volterra iterates starting with the free wave), we can use the dispersion of the wave equation to prove the desired result. First, we follow Tao to establish the following divisibility lemma[24] .

Lemma 9.21. *For any $\varepsilon_1 > 0$ there exist a partition of $\mathbb{R}$ into intervals $\{I_j\}_{j=1}^M$ where $M \lesssim (E_{crit}\,\varepsilon_1^{-1})^C$ for some absolute constant C with the property[25] that for any u*

$$\max_{1\le j\le M} \|\partial^\alpha u\, A_\alpha\|_{\nabla_x^{-1} N(I_j\times\mathbb{R}^2)} \le \varepsilon_1 \|u\|_S$$

Note that the intervals depend on ψ_L (but not on u), but their number does not (other than through the energy).

Proof. According to the trilinear estimates of Section 5, we may assume that there is angular separation between $\hat u$ and the waves in A_α. Otherwise there is the desired gain. The amount of angular separation is very small and depends on E_{crit} and ε_1. We shall now implicitly assume that $\partial^\alpha u\, A_\alpha$ respects this type of angular separation. Note that we may restrict ourselves to the case of high-low interactions between u and A_α, since for the other cases, the divisibility follows by using the same argument as in the proof of Lemma 9.15. Also, since by the preceding lemma we have that u is concentrated along frequency ~ 1, we may reduce to considering the zero frequency mode $P_0 u$.

[24]It appears likely that an alternative approach to the pointwise decay is to use commuting vector fields. This would force us to strengthen the norm assumptions on Ψ^n_ν even more, however, and so we opted for the present approach

[25]We define $\|F\|_{\nabla_x^{-1}N} := \|\nabla_x F\|_N$

By (2.29),

$$\|\partial^\alpha P_0 u\, A_\alpha\|_{\nabla_x^{-1} N} \leq C(E_{crit}\,, \varepsilon_1) \sum_{k_1 < -C} 2^{-\frac{k_1}{2}} \|P_0 u\, P_{k_1}\psi_L\|_{L^2_{t,x}} \|P_{k_2}\psi_L\|_S$$

$$\leq C(E_{crit}\,, \varepsilon_1) \Big(\sum_{k_1 < -C} 2^{-k_1} \|P_0 u\, P_{k_1}\psi_L\|^2_{L^2_{t,x}} \Big)^{\frac{1}{2}} \|\psi_L\|_S \tag{9.70}$$

Next, by Theorem 1.11 of [50], assuming u to be a free wave, for each $k \in \mathbb{Z}$ there exists a collection $\mathcal{T}_k$ of tubes τ_k^i of size $\infty \times 2^k \times 2^k$ centered along a light-ray and aligned with the Fourier support of u such that $\#\mathcal{T}_k \leq (E_{crit}\, \varepsilon^{-1})^C$ and so that, where $\varepsilon > 0$ is small and will be determined,

$$\|P_0 u\, P_{k_1}\psi_L\|_{L^2_{t,x} \backslash \Omega_{k_1}} \leq \varepsilon\, 2^{\frac{k_1}{2}} \|u\|_2 \|P_{k_1}\psi_L\|_2 \tag{9.71}$$

where $\Omega_k := \bigcup_{\tau \in \mathcal{T}_{k_1}} \tau$. In our case u is of course not a free wave; however, by Remark 5.12 as well as Remark 6.6 in conjunction with Lemma 2.22, we conclude that we can write

$$u = u_1 + u_2$$

where

$$\|u_2\|_{\tilde{S}} < \varepsilon_2 \|u\|_{\tilde{S}}$$

while

$$u_1 = \int f_a u_a\, \nu(da)$$

is a superposition of free waves u_a with the same frequency support properties as u and

$$\int \|f_a u_a\|_{L^2_x}\, \nu(da) \leq C(\varepsilon_2) \|u\|_{\tilde{S}}, \quad f_a \in L^\infty_{t,x}, \quad \|f_a\|_{L^\infty_{t,x}} \leq C$$

Thus for u in the original sense, choosing ε in (9.71) of the form $C(E_{crit}\,, \varepsilon_2)^{-1}\varepsilon_2$, we get

$$\|P_0 u\, P_{k_1}\psi_L\|_{L^2_{t,x} \backslash \Omega_{k_1}} \leq \varepsilon_2\, 2^{\frac{k_1}{2}} \|u\|_{S[0]} \|P_{k_1}\psi_L\|_2$$

Inserting this bound in (9.70) yields

$$\begin{aligned}
\|\partial^\alpha P_0 u\, A_\alpha\|_{\nabla_x^{-1} N} &\leq C(E_{crit}\,, \varepsilon_1) \Big(\sum_{k_1 < -C} 2^{-k_1} \|P_0 u\, P_{k_1}\psi_L\|^2_{L^2_{t,x} \backslash \Omega_{k_1}} \Big)^{\frac{1}{2}} \|\psi_L\|_S \\
&\quad + C(E_{crit}\,, \varepsilon_1) \Big(\sum_{k_1 < -C} 2^{-k_1} \sum_{\tau_k^i \in \mathcal{T}_k} \|\chi_{\tau_k^i} P_0 u\, P_{k_1}\psi_L\|^2_{L^2_{t,x}} \Big)^{\frac{1}{2}} \|\psi_L\|_S \\
&\leq C(E_{crit}\,, \varepsilon_1)\varepsilon_2 \|u\|_{S[0]} \|\psi_L\|^2_S \\
&\quad + C(E_{crit}\,, \varepsilon_1) \Big(\sum_{k_1 < -C} 2^{-k_1} \sum_{\tau_{k_1}^i \in \mathcal{T}_{k_1}} \|P_0 u\|^2_{L^\infty_t L^2_x} \|\chi_{\tau_k^i}\, P_{k_1}\psi_L\|^2_{L^2_t L^\infty_x} \Big)^{\frac{1}{2}} \|\psi_L\|_S
\end{aligned}$$

By picking ε_2 small enough in relation to ε_1, we can achieve the desired smallness gain for the first expression on the right. Next, by a standard TT^* estimate, and for all $k_1 \in \mathbb{Z}$,

$$\|\chi_{\tau_k^i}\, P_{k_1}\psi_L\|_{L^2_t L^\infty_x} \lesssim 2^{\frac{k_1}{2}} \|P_{k_1}\psi_L\|_2$$

whence

$$\Big(\sum_{k_1 \in \mathbb{Z}} 2^{-k_1} \sum_{\tau_{k_1}^i \in \mathcal{T}_{k_1}} \|\chi_{\tau_k^i}\, P_{k_1}\psi_L\|^2_{L^2_t L^\infty_x} \Big)^{\frac{1}{2}} \lesssim (E_{crit}\, \varepsilon^{-1})^C \|\psi_L\|_2$$

Therefore, the exist intervals $\{I_j\}_{j=1}^M$ as claimed. Since the constants $C(\varepsilon_2)$ and $C(E_{crit}\,, \varepsilon_1)$ depend polynomially on the parameters, we are done. □

We can now prove Proposition 9.20. We will assume that the energy of the data (f, g) is also controlled by E_{crit} although this is only a notational convenience.

Proof of Proposition 9.20. With $\{I_j\}_{1\le j\le M}$ as in the lemma, we relabel them as follows: with initial time $0\in I_{j_0}$, we set $J_0 := I_{j_0}$. At the next step, we define $J_1 = I_{j_1}$ and $J_{-1} := I_{j_2}$ where I_{j_1} is the successor of I_{j_0} (with respect to positive orientation of time), whereas I_{j_2} is the predecessor. In this fashion one obtains a sequence J_i with $0\le i\le M'$ and $M'\le (E_{crit}\,\varepsilon_1^{-1})^C$ as in Lemma 9.21 where ε_1 is small depending only on E_{crit}. Next, let u be the solution of

$$\Box u + 2i\partial^\alpha u\, A_\alpha = 0, \quad u[0] = (f,g)$$

We claim that $u^{(0)} := u\big|_{J_0}$ can be written as an infinite Duhamel expansion in the form

$$u^{(0)} := \sum_{\ell=0}^\infty u^{(J_0,\ell)}, \quad u^{(J_0,0)}(t) := S(t)u[0],$$
$$u^{(J_0,\ell)} := -2i\int_0^t U(t-s)\partial^\alpha u^{(J_0,\ell-1)}\,A_\alpha(s)\,ds$$

where $S(t) = (U,V)(t)$ is the free wave evolution, and $U(t) = \frac{\sin(t|\nabla|)}{|\nabla|}$, $V(t) = \cos(t|\nabla|)$. Of course, $t\in J_0$ in this equation. Due to the energy estimate of Section 2.3 and Lemma 9.21, this series converges with respect to the S-norm. In a similar fashion, we can pass to later times: $u^{(i)} := u\big|_{J_i}$ satisfies

$$u^{(i)} := \sum_{\ell=0}^\infty u^{(J_i,\ell)}, \quad u^{(J_i,0)}(t) := S(t-t_i)u^{(i-1)}[t_i],$$
$$u^{(J_i,\ell)} = -2i\int_{t_i}^t U(t-s)\partial^\alpha u^{(J_i,\ell-1)}\,A_\alpha(s)\,ds \tag{9.72}$$

where $t\in J_i$ and $t_i := \max J_{i-1} = \min J_i$ for $i\ge 1$ and $t_0 := 0$. Observe that

$$\begin{aligned} u^{(J_i,0)}(t) &:= S(t-t_{i-1})u^{(i-1)}[t_{i-1}] \\ &\quad - 2i\sum_{\ell=1}^\infty \int_{t_{i-1}}^{t_i} U(t-s)\chi_{J_{i-1}}(s)\partial^\alpha u^{(J_{i-1},\ell)}(s)\,A_\alpha(s)\,ds \end{aligned} \tag{9.73}$$

for all $t\in J_i$. If $i\ge 2$, we expand further to obtain

$$\begin{aligned} S(t-t_{i-1})u^{(i-1)}[t_{i-1}] &:= S(t-t_{i-2})u^{(i-2)}[t_{i-2}] \\ &\quad - 2i\sum_{\ell=1}^\infty \int_{t_{i-2}}^{t_{i-1}} U(t-s)\chi_{J_{i-2}}(s)\partial^\alpha u^{(J_{i-2},\ell)}(s)\,A_\alpha(s)\,ds \end{aligned}$$

This procedure can be continued all the way back to $t_0 = 0$ and yields

$$u^{(J_i,0)}(t) := S(t)(f,g) - \sum_{k=0}^{i-1} 2i\sum_{\ell=1}^\infty \int_{t_k}^{t_{k+1}} U(t-s)\chi_{J_k}(s)\partial^\alpha u^{(J_k,\ell)}(s)\,A_\alpha(s)\,ds \tag{9.74}$$

for all $t\in J_i$. Inductively, one passes from this term to $u^{(J_i,\ell)}$ for all $\ell\ge 0$ by means of (9.72). We next claim that for each i, the functions $u^{(J_i,\ell)}$ become small with respect to $\|\cdot\|_{\tilde S}$ provided ℓ is large enough. This is a direct consequence of applying Lemma 9.21 to the above iterative definition of $u^{(J_i,\ell)}$ as well as the basic energy estimate.

Now fix a number $\gamma > 0$. We will show that there exist $t_0 = t_0(\gamma)$ and $n_0(\gamma)$ with the property that if $|t| > t_0(\gamma)$ and $n > n_0(\gamma)$, then we can write

$$u = u_1 + u_2$$

where

$$\|u_2\|_{\tilde S} < \gamma$$

and

$$|u_1(t,x)| < \gamma$$

for $|t| > t_0$, uniformly in $n > n_0(\gamma)$. We start by reducing ourselves to a double light cone. Indeed, pick a large enough disc D_γ in the time slice $\{0\} \times \mathbb{R}^2$ with the property that

$$\|\chi_{D_\gamma^c} u[0]\|_{L_x^2} \ll \gamma$$

Here $\chi_{D_\gamma^c}$ is a smooth cutoff localizing to a large dilate of D_γ. If we denote the covariant propagation of $\chi_{D_\gamma^c} u[0]$ by $\tilde{u}_2$, then we can achieve that

$$\|\tilde{u}_2\|_{\tilde{S}} \ll \gamma$$

by means of Lemma 9.15. We are thus reduced to estimating $\tilde{u}_1 = u - \tilde{u}_2$, which by construction is supported in a (large) double cone whose base depends only on γ. We can then expand $\tilde{u}_1$ in terms of Volterra iterates just as before, and there exists ℓ_γ with the property that

$$\sum_i \sum_{\ell > \ell_\gamma} \|\tilde{u}_1^{J_i,\ell}\|_{\tilde{S}} \ll \gamma \tag{9.75}$$

Furthermore, note that all the iterates $\tilde{u}_1^{J_i,\ell}$ are supported in the same double light cone with base D_γ. We now show that $\tilde{u}_1 = u_1 + u_2^\dagger$ where $\|u_2^\dagger\|_{\tilde{S}} \ll \gamma$ and u_1 has the desired dispersive property. Setting $u_2 := \tilde{u}_2 + u_2^\dagger$ then concludes the argument. First, in view of (9.75) and the fact that the total number of J_i is controlled by the energy, we may include the contributions of $\ell > \ell_\gamma$ in $u_2^\dagger$.
By Huyghens principle, $\tilde{u}_1 = \chi(t,x)\tilde{u}_1$ where for the remainder of the proof $\chi(t,x)$ is a smooth cut-off to the region $|x| \le |t| + \rho$ with ρ being the radius of D_γ. Then we can write

$$\begin{aligned} \tilde{u}_1^{(J_i,\ell)}(t) &= -2i\chi(t,x) \int_{t_i}^t U(t-s) P_{[-k_0 < \cdot < k_0]} \big[\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)} A_\alpha(s)\big]\, ds \\ &\quad - 2i\chi(t,x) \int_{t_i}^t U(t-s) P_{[-k_0 < \cdot < k_0]^c} \big[\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)} A_\alpha(s)\big]\, ds \end{aligned} \tag{9.76}$$

We now show that the second integral splits into a term of small $L_t^\infty L_x^\infty$-norm, and one of small $\tilde{S}$ norm. First, consider $P_{[<-k_0]}$. Then by Bernstein's inequality, and the energy estimate

$$\begin{aligned} &\Big\|\chi(t,x) \int_{t_i}^t U(t-s) P_{[<-k_0]} \big[\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)} A_\alpha(s)\big]\, ds\Big\|_{L_x^\infty} \\ &\lesssim 2^{-k_0} \Big\| \int_{t_i}^t U(t-s) P_{[<-k_0]} \big[\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)} A_\alpha(s)\big]\, ds\Big\|_{L_t^\infty L_x^2} && (9.77) \\ &\lesssim 2^{-k_0} E_{crit}\, \|\tilde{u}_1^{(J_i,\ell-1)}\|_{\tilde{S}} \le C(E_{crit}) 2^{-k_0} && (9.78) \end{aligned}$$

whereas for $P_{[>k_0]}$ one can essentially (up to tails which are handled by Lemma 7.23, for example) remove the exterior χ since the interior $\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)}$ obeys that very localization. In conclusion, the resulting term is placed in $u_2^\dagger$. Now consider the main term (9.76). Decompose S^1 into caps κ of size $c(E_{crit}, \gamma)$ which is a small constant. Denote the corresponding decomposition of the double light-cone $\{|x| \le |t| + \rho\}$ into angular sectors by $\{S_\kappa\}_\kappa$. Associated with the S_κ there is a smooth partition of unity $\sum_\kappa \chi_\kappa = \chi$. Write (9.76) as the sum

$$2i \sum_\kappa \chi_\kappa(t,x) \int_{t_i}^t U(t-s) P_{[-k_0 < \cdot < k_0]} P_{[\hat{\xi} \in \mp 2\kappa]} \big[\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)} A_\alpha(s)\big]\, ds \tag{9.79}$$

$$+ 2i \sum_\kappa \chi_\kappa(t,x) \int_{t_i}^t U(t-s) P_{[-k_0 < \cdot < k_0]} P_{[\hat{\xi} \notin \mp 2\kappa]} \big[\partial^\alpha \tilde{u}_1^{(J_i,\ell-1)} A_\alpha(s)\big]\, ds \tag{9.80}$$

Here $\hat{\xi} = \frac{\xi}{|\xi|}$ and the sign is selected according to the decomposition into incoming and outgoing propagator:

$$U(t) = \frac{1}{2i|\nabla|}[e^{it|\nabla|} - e^{-it|\nabla|}]$$

By Bernstein's inequality the first term (9.79) satisfies

$$\|(9.79)\|_{L_t^\infty L_x^\infty} \le C(E_{crit}) |\kappa|^{\frac{1}{2}} \|\tilde{u}_1^{(J_i,\ell-1)}\|_S$$

which can be made small for small κ. Also note that

$$|t|\xi| \pm x\cdot\xi| \gtrsim |t| \quad \forall\,(t,x,\xi) \text{ such that } \chi_\kappa(t,x)\neq 0,\ 2^{-k_0}<|\xi|<2^{k_0},\quad \hat{\xi}=\frac{\xi}{|\xi|}\notin \mp 2\kappa$$

where the choice of $\pm$ depends on whether the propagator U is incoming or outgoing. Now we make the inductive assumption (relative to ℓ and i) that

$$\|[\chi_\kappa P_{[\pm\hat{\xi}\notin -2\kappa]}\tilde{u}_1^{(J_i,\ell-1)}](t)\|_{L^\infty_x} + \|\chi_{[||t|-|x||\gtrsim |t|]}\tilde{u}_1^{(J_i,\ell-1)}(t,x)\|_{L^\infty_x} \le C_N(i,\ell,E_{crit}\,,\gamma)|t|^{-N} \tag{9.81}$$

where the $\pm$ sign is according to whether the function has space-time Fourier support in the upper or lower half-spaces, i.e., whether $\tau>0$ or $\tau<0$. Strictly speaking, the cap size here depends on (i,ℓ) with the size $c(E_{crit}\,,\gamma)$ from above being the size at the end of the induction (recall that there are only finitely many choices for these parameters). But for simplicity of notation, we suppress this dependence from the notation. Note that we only have finitely many values of ℓ, i. Now to estimate the second integral term (9.80), we distinguish between a number of cases: first if $|s|\ll|t|$ (where the implicit small constant depends on $|\kappa|$), due to the a priori support conditions satisfied by $\tilde{u}_1^{(J_i,\ell-1)}$ which forces $|y|<|s|+\rho$, we obtain the desired gain in t by integrating by parts with respect to $|\xi|$ in the Fourier integral representation of U.. Next, assume that $|s|\sim|t|$ (where the implicit small constant again depends on $|\kappa|$ - this will be tacitly understood for the remainder of the proof). Then we first reduce to $||s|-|y||\ll|s|$. For this consider the term

$$2i\sum_\kappa \chi_\kappa(t,x)\int_{t_i}^t U(t-s)P_{[-k_0<\cdot<k_0]}P_{[\hat{\xi}\notin\mp 2\kappa]}\big[\chi_{[||s|-|y||\gtrsim |s|]}\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)}\,A_\alpha(s)\big]\,ds$$

Since we assume $|s|\sim|t|$, the desired gain t^{-N} here follows by using the induction hypothesis. Hence we now reduce to estimating

$$2i\sum_\kappa \chi_\kappa(t,x)\int_{t_i}^t U(t-s)P_{[-k_0<\cdot<k_0]}P_{[\hat{\xi}\notin\mp 2\kappa]}\big[\chi_{[||s|-|y||\ll |s|]}\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)}\,A_\alpha(s)\big]\,ds$$

Here we apply a further decomposition

$$\begin{aligned}\chi_{[||s|-|y||\ll|s|]}\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)} &= \chi_{[||s|-|y||\ll|s|]}\sum_{\kappa'}\chi_{\kappa'}(s,y)\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)}\\ &= \chi_{[||s|-|y||\ll|s|]}\sum_{\kappa'}\chi_{\kappa'}(s,y)P_{[\pm\hat{\xi}\in-2\kappa']}\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)}\\ &\quad+\chi_{[||s|-|y||\ll|s|]}\sum_{\kappa'}\chi_{\kappa'}(s,y)P_{[\pm\hat{\xi}\in-(2\kappa')^c]}\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)}\end{aligned}$$

The contribution of the second term here is again rapidly decaying due to the induction assumption. Hence we have now reduced to estimating

$$\begin{aligned}2i\sum_\kappa \chi_\kappa(t,x)\int_{t_i}^t U(t-s)P_{[-k_0<\cdot<k_0]}P_{[\hat{\xi}\notin\mp 2\kappa]}\big[\chi_{[||s|-|y||\ll|s|]}\\ \sum_{\kappa'}\chi_{\kappa'}(s,y)P_{[\pm\hat{\xi}\in-2\kappa']}\partial^\alpha\tilde{u}_1^{(J_i,\ell-1)}\,A_\alpha(s)\big]\,ds\end{aligned}$$

Now writing out the free wave parametrix, we see that on the support of the resulting integral in the variables ξ, y, s, we have that

$$|\pm|\xi|s+y\cdot\xi|\ll|t|,$$

and choosing κ' as well as the implied constant in $||s|-|y||\ll|s|$ suitably small, we can ensure that

$$|\pm t|\xi|+x\cdot\xi|\sim|t|\gg|\pm|\xi|s+y\cdot\xi|$$

on the support of the integrand. Integrations by parts in $|\xi|$ yield the desired rapid decay with respect to $|t|$. This recovers the first part of the inductive assumption, and the second follows identically, since if $||t|-|x||\gtrsim|t|$, then we necessarily have

$$|\pm t|\xi|+x\cdot\xi|\gtrsim|t|$$

The inductive procedure is now completed by means of (9.74) which takes account of the changes in the level i. $\square$

Recall that we restricted Ψ to be a free wave in (9.63). In order to treat the general case, we apply the usual decomposition (9.62). As usual, the smallness of the Ψ_{NL} allows one to iterate these terms away. Furthermore, the proof of Proposition 9.20 applies to these terms equally well since we do not rely on any specific structure of the $u_1^{(J_i,\ell-1)}$ other than the inductive assumption (9.81), and the formalism of the Volterra iteration by which we represented these solutions. In this way, the same dispersive property may be proved globally, i. e. not just on I_1 where (9.62) holds.

9.6.2. *The second stage of Bahouri Gerard, applied to the first large atomic component.* Recall that we are considering only $a = 1$. Nevertheless, we keep the parameter "a" in our notation general. We now need to quantify the lack of compactness for the functions $\tilde{\phi}^{na}$, $\mathring{\phi}^{na}$, ζ^{na}, η^{na}, all at time $t = 0$. We evolve each of these using the covariant wave flow from before and select a number of concentration profiles. The method for this follows exactly the Bahouri-Gerard template, but using Lemma 9.19 instead of standard energy conservation for the free wave flow. In order to define the temporal flow for each component, we need to impose time derivatives at time $t = 0$. We do this by defining

$$\partial_t\tilde{\phi}^{na}(0,\cdot) := \phi_0^{na}(0,\cdot), \quad \partial_t\mathring{\phi}(0,\cdot) := 0$$
$$\partial_t\zeta^{na}(0,\cdot) := \tilde{\psi}_0^{na}, \quad \partial_t\eta^{na}(0,\cdot) := 0$$

Introduce the following terminology:

Definition 9.22. *Given data $u[0] = (u_0, u_1)$ at time $t = 0$, we denote by*

$$S_{A^n}(t)(u[0])$$

the solution of $\Box_{A^n}(u) = 0$ with the given data, evaluated at time t.

We now describe the important process of *extraction of concentration profiles:* Consider $S_{A^n}(\zeta^{na}[0])$, with $\zeta^{na}[0] = (\tilde{\psi}_0^{na}, \zeta^{na})$. Following [1] introduce the family $\mathcal{V}_{A^n}(\underline{\zeta^a})$, consisting of all functions on $V_\zeta(t,x) \in L^2_{t,\mathrm{loc}}H^1_x \cap C^1L^2_x$ such that

$$\big(S_{A^n}(\zeta^{na}[0])\big)(t + t_n, x + x_n) \rightharpoonup V_\zeta(t,x)$$

as $n \to \infty$ for some sequence $\{(t_n,x_n)\}_{n=1}^\infty \in \mathbb{R}\times\mathbb{R}^2$. Here, the weak limit is in the sense of $L^2_{t,\mathrm{loc}}H^1_x$. Observe that such a function $V_\zeta(t,x)$ solves $\Box V_\zeta = 0$ in the sense of distributions. Thus it makes sense to introduce the quantity

$$\eta_{A^n}(\underline{\zeta^a}) := \sup\{E(V_\zeta),\, V_\zeta \in \mathcal{V}_{A^n}(\underline{\zeta^a})\},$$

where

$$E(V_\zeta) := \int_{\mathbb{R}^2} |\nabla_{x,t}V_\zeta|^2 dx$$

We can now state the following lemma that is at the core of the second stage of the Bahouri-Gerard process for wave maps. Recall that $a = 1$ here.

Lemma 9.23. *There exists a collection of sequences $\{(t_n^{ab}, x_n^{ab})\} \subset \mathbb{R}\times\mathbb{R}^2$, $b \geq 1$, as well as a family of concentration profiles $V_\zeta^{ab}[0] := (V_{\zeta 0}^{ab}(x), V_{\zeta 1}^{ab}(x)) \in L^2(\mathbb{R}^2)\times\dot{H}^1(\mathbb{R}^2)$, with the following properties: introducing the shifted gauge potentials*

$$\tilde{A}^{nab} := A^n(t + t_n^{ab}, x + x_n^{ab}), \tag{9.82}$$

one has

- *For any $B \geq 1$, one can write*

$$\big(S_{A^n}(\zeta^{na}[0])\big)(t,x) = \sum_{b=1}^B \big(S_{\tilde{A}^{nab}}(V_\zeta^{ab}[0])\big)(t - t_n^{ab}, x - x_n^{ab}) + W_\zeta^{naB}(t,x)$$

Here each function $\big(S_{\tilde{A}^{nab}}(V_\zeta^{ab}[0])\big)(t-t_n^{ab},x-x_n^{ab})$*,* $W_\zeta^{naB}(t,x)$*, solves the equation* $\Box_{A^n}u=0$*, and we have*

$$\lim_{B\to\infty}\big[\limsup_{n\to\infty}\eta_{A^n}(\underline{W}^{aB})\big]=0 \tag{9.83}$$

- *One has the divergence relations*

$$\lim_{n\to\infty}[|t_n^{ab}-t_n^{ab'}|+|x_n^{ab}-x_n^{ab'}|]=\infty$$

for $b\neq b'$*.*

- *There is the asymptotic orthogonality relation*

$$E(\zeta^{na}[0])=\sum_{b=1}^{B}E(V_\zeta^{ab}[0])+E(W_\zeta^{naB}(t,\cdot))+o(1)$$

Here E *refers to the standard (flat) energy and the o-term satisfies* $\lim_{B\to\infty}\limsup_{n\to\infty}o(1)=0$*.*

- *All* $V_\zeta^{ab}[0]$*, as well as their evolutions* $S_{\tilde{A}^{nab}}(V_\zeta^{ab}[0])$ *and the* W_ζ^{naB} *are* 1*-oscillatory.*

Proof. We follow [1]: There is nothing to do provided $\eta_{A^n}(\underline{\zeta^a})=0$. Hence assume this quantity is >0. Then pick a profile $V_\zeta^{a1}(t,x)\in L^2_{t,\mathrm{loc}}H^1_x\cap C^1L^2_x$ and associated sequence $\{(t_n^{a1},x_n^{a1})\}_{n\geq1}$ such that

$$\big(S_{A^n}(\zeta^{na}[0])\big)(t+t_n^{a1},x+x_n^{a1})\rightharpoonup V_\zeta^{a1}(t,x)$$

with

$$E(V_\zeta^{a1})>\frac12\eta_{A^n}(\underline{\zeta^a})$$

Using the notation of the lemma, consider then

$$\begin{aligned}&\big(S_{A^n}(\zeta^{na}[0])\big)(t+t_n^{a1},x+x_n^{a1})-\big[S_{A^n}\big(S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot-x_n^{a1})\big)\big](t+t_n^{a1},x+x_n^{a1})\\&=\big(S_{A^n}(\zeta^{na}[0])\big)(t+t_n^{a1},x+x_n^{a1})-\big(S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])\big)(t,x)\end{aligned}$$

But by our construction, this expression converges weakly to 0.
Furthermore, due to Lemma 9.19, we have that

$$E(S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot-x_n^{a1}))=E(V_\zeta^{a1}[0])+o_{L^2}(1)$$

Now we repeat the preceding step, but replace $\zeta^{na}[0]$ by

$$\zeta^{na}[0]-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot-x_n^{a1})$$

Thus select a sequence $\{(t_n^{a2},x_n^{a2})\}_{n\geq1}$ and a concentration profile $V_\zeta^{a2}(t,x)$ such that

$$E(V_\zeta^{a2})\geq\frac12\eta(\zeta^{na}-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot-x_n^{a1}))$$

and furthermore

$$\big[S_{A^n}\big(\zeta^{na}-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot-x_n^{a1})\big)\big](t+t_n^{a2},x+x_n^{a2})\rightharpoonup V_\zeta^{a2}(t,x)$$

We obtain that necessarily

$$\lim_{n\to\infty}|t_n^{a1}-t_n^{a2}|+|x_n^{a1}-x_n^{a2}|=\infty$$

Furthermore, we claim that

$$E(V_\zeta^{a1}[0])+E(\zeta^{na}[0]-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}])=E(\zeta^{na}[0])+o_{L^2}(1)$$

This follows again just as in the free case, using Lemma 9.19: We need to show that

$$\int_{\mathbb{R}^2}\nabla_{x,t}S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot)\cdot\nabla_{x,t}[\zeta^{na}[0]-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])(0-t_n^{a1},\cdot)]\,dx=o_{L^2}(1)$$

Due to Lemma 9.19, up to $o_{L^2}(1)$, the left-hand side equals

$$\begin{aligned}\int_0^1\int_{\mathbb{R}^2}&\nabla_{x,t}S_{A^n}\big(S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}]\big)(t+t_n^{a1},\cdot+x_n^{a1})\cdot\\&\cdot\nabla_{x,t}S_{A^n}\big([\zeta^{na}[0]-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}]]\big)(t+t_n^{a1},\cdot+x_n^{a1})\,dx\end{aligned}$$

But here we can again use that

$$S_{A^n}\big(S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}]\big)(t+t_n^{a1},\cdot+x_n^{a1}) = V_\zeta^{a1}(t,\cdot)+o_{L^2}(1)$$

provided $t\in[0,1]$, while by construction

$$S_{A^n}\big([\zeta^{na}[0]-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}]]\big)(t+t_n^{a1},\cdot+x_n^{a1}) \rightharpoonup 0$$

The conclusion is that

$$\begin{aligned}\int_0^1\int_{\mathbb{R}^2}\nabla_{x,t}S_{A^n}\big(S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}]\big)(t+t_n^{a1},\cdot+x_n^{a1})\cdot\\ \cdot\nabla_{x,t}S_{A^n}\big([\zeta^{na}[0]-S_{\tilde{A}^{na1}}(V_\zeta^{a1}[0])[0-t_n^{a1}]]\big)(t+t_n^{a1},\cdot+x_n^{a1})\,dx\\ =o_{L^2}(1),\end{aligned}$$

from which the asymptotic orthogonality follows. All assertions of the lemma now follow by applying the preceding considerations inductively B times. $\square$

FIGURE 7. The dependence domains of various concentration profiles

Figure 7 depicts various concentration profiles. More precisely, one should view these profiles as being well-localized in physical space centered at their cores in space-time. The figure then shows the approximate support of the wave evolutions of these profiles.

Generally speaking, a will always refer to a frequency atom, whereas b refers to the concentration profile generated by a frequency atom. We shall now apply Lemma 9.23 to the covariant evolution of $\zeta^{na}[0]$, as well as the remaining components η^{na}, $\tilde{\psi}^{na}$, $\mathring{\psi}^{na}$.

9.6.3. *Selecting geometric concentration profiles.* At this stage, we face the same issue as in Step 1 above: we have a sequence of component functions V_α^{ab} associated with the essentially singular sequence ϕ_α^n, but in order to apply the "energy induction hypothesis", i.e., the assumption that E_{crit} is the minimal energy for which uniform control fails, we need to show that the V_α^{ab} can be assembled to form the Coulomb components of actual maps from $\mathbb{R}^2\to\mathbb{H}^2$. We now address this task. To begin with, we may assume that $\Psi^{nA_0^{(0)}}\neq o_{L^2}(1)$ since otherwise $\Psi^{nA_0^{(0)}}$ is a perturbative error.

To summarize our construction of the concentration profiles: we started with the Hodge decomposition (all at time $t = 0$)

$$\begin{aligned}
\phi_1^{na} &= \partial_1 \tilde{\phi}^{na} + \partial_2 \mathring{\phi}^{na},\ \partial_t \tilde{\phi}^{na} := \phi_0^{na}, \quad \partial_t \mathring{\phi}^{na} = 0\\
\phi_2^{na} &= \partial_2 \tilde{\phi}^{na} - \partial_1 \mathring{\phi}^{na}\\
\tilde{\psi}_1^{na} &= \partial_1 \zeta^{na} + \partial_2 \eta^{na}, \quad \partial_t \zeta^{na} := \tilde{\psi}_0^{na}, \quad \partial_t \eta^{na} = 0\\
\tilde{\psi}_2^{na} &= \partial_2 \zeta^{na} - \partial_1 \eta^{na}
\end{aligned}$$

From here it is immediate that

$$E(\phi^{na}) = \sum_{\alpha=0}^{2} \|\phi_\alpha^{na}\|_{L_x^2}^2 = \sum_{\alpha=0}^{2} \|\partial_\alpha \tilde{\phi}^{na}\|_{L_x^2}^2 + \sum_{\alpha=0}^{2} \|\partial_\alpha \mathring{\phi}^{na}\|_{L_x^2}^2 = \sum_{\alpha=0}^{2} \|\partial_\alpha \zeta^{na}\|_{L_x^2}^2 + \sum_{\alpha=0}^{2} \|\partial_\alpha \eta^{na}\|_{L_x^2}^2$$

Now, we evolve each of the $\tilde{\phi}^{na}$ etc. in time using the covariant flow, and apply Lemma 9.23. Changing the notation from that lemma, one obtains the decompositions (with $\tilde{A}$ as in (9.82))

$$\nabla_{x,t} \tilde{\phi}^{na} = \sum_{b=1}^{B} \nabla_{x,t} \big[S_{\tilde{A}^{nab}}(\tilde{V}_1^{ab}[0])\big](0 - t^{nab}, x - x^{nab}) + \nabla_{x,t} \tilde{W}_1^{naB}$$

$$\nabla_{x,t} \mathring{\phi}^{na} = \sum_{b=1}^{B} \nabla_{x,t} \big[S_{\tilde{A}^{nab}}(\tilde{V}_2^{ab}[0])\big](0 - t^{nab}, x - x^{nab}) + \nabla_{x,t} \tilde{W}_2^{naB}$$

$$\nabla_{x,t} \zeta^{na} = \sum_{b=1}^{B} \nabla_{x,t} \big[S_{\tilde{A}^{nab}}(V_1^{ab}[0])\big](0 - t^{nab}, x - x^{nab}) + \nabla_{x,t} W_1^{naB}$$

$$\nabla_{x,t} \eta^{na} = \sum_{b=1}^{B} \nabla_{x,t} \big[S_{\tilde{A}^{nab}}(V_2^{ab}[0])\big](0 - t^{nab}, x - x^{nab}) + \nabla_{x,t} W_2^{naB}$$

where the W errors are small in the η-sense when B is large, see (9.83). Here we use the same sequences t^{nab}, x^{nab} for all decompositions, which of course we can by passing to suitable subsequences. Note that we are working with *both* the ϕ and ψ components here, which is needed for the following result.

Proposition 9.24. *For any $1 \le b \le B$, and any $\delta_2 > 0$, there exists an admissible (derivative components are Schwartz) map from $\mathbb{R}^2$ into $\mathbb{H}^2$, with derivative components $\phi_{j\delta_2}^{nab}$, $j = 1, 2$, and a number $\gamma_{\delta_2 nab} \in \mathbb{R}$, such that*

$$\Big\| \Big(\partial_1 \big[S_{\tilde{A}^{nab}}(V_1^{ab}[0])\big](0 - t^{nab}, x) + \partial_2 \big[S_{\tilde{A}^{nab}}(V_2^{ab}[0])\big](0 - t^{nab}, x),\ \partial_2 \big[S_{\tilde{A}^{nab}}(V_1^{ab}[0])\big](0 - t^{nab}, x)$$
$$- \partial_1 \big[S_{\tilde{A}^{nab}}(V_2^{ab}[0])\big](0 - t^{nab}, x)\Big) - e^{i\gamma_{\delta_2 nab}} \big(\phi_{1\delta_2}^{nab} e^{-i\sum_{k=1,2} \triangle^{-1}\partial_k \phi_{k\delta_2}^{nab}}, \phi_{2\delta_2}^{nab} e^{-i\sum_{k=1,2} \triangle^{-1}\partial_k \phi_{k\delta_2}^{nab}}\big) \Big\|_{L_x^2} < \delta_2$$

for large n.

Proof. Due to the asymptotic orthogonality relation of Lemma 9.23, given δ_2 there exists B_0 so that for all $b > B_0$ one can simply take the derivative components to equal zero. In other words, it suffices to consider $1 \le b \le B_0$.

Fix a b, we shall pick B larger, if necessary, and also pick n large enough later. For simplicity introduce the notation

$$V_3^{nab} := \partial_1 S_{\tilde{A}^{nab}}(V_1^{ab}[0]) + \partial_2 S_{\tilde{A}^{nab}}(V_2^{ab}[0])$$

$$V_4^{nab} := \partial_2 S_{\tilde{A}^{nab}}(V_1^{ab}[0]) - \partial_1 S_{\tilde{A}^{nab}}(V_2^{ab}[0])$$

and similarly for $W_{3,4}^{nab}$. Note that we here introduce dependence on n again.
Thus at time $t = 0$, we have the identities

$$\phi_1^{na} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1}+\phi_k^{na1}]} = \sum_{b=1}^{B} V_3^{nab}(0 - t^{nab}, x - x^{nab}) + W_3^{naB}$$

$$\phi_2^{na} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}1}+\phi_k^{na1}]} = \sum_{b=1}^{B} V_4^{nab}(0 - t^{nab}, x - x^{nab}) + W_4^{naB}$$

where the W's satisfy the smallness property (9.83). Then we distinguish between the following two cases:

(A): $V_{3,4}^{nab}(\cdot - t^{nab}, x - x^{nab})$ *is of temporally bounded type.* By this we mean that

$$\liminf_{n\to\infty} |t^{nab}| < \infty$$

By passing to a subsequence, we may then assume that

$$\limsup_{n\to\infty} |t^{nab}| < \infty$$

or in fact, that $\lim_{n\to\infty} t^{nab}$ exists.

(B): $V_{3,4}^{nab}(\cdot - t^{nab}, x - x^{nab})$ *of temporally unbounded type.* By this we mean that

$$\lim_{n\to\infty} |t^{nab}| = \infty$$

Observe that in this latter case, due to Proposition 9.20, we can conclude that

$$V_{3,4}^{nab}(\cdot - t^{nab}, x - x^{nab}) = o_{L^\infty}(1) + o_{L^2}(1)$$

as $n \to \infty$.

We treat these cases separately, commencing with the temporally bounded Case A. We need to show that $V_{3,4}^{nab}(\cdot - t^{nab}, x - x^{nab})$ can be approximated arbitrarily well by the Coulomb components of admissible maps. We shall do this by physical localization: Note that for $b' \neq b$, we have either

$$\lim_{n\to\infty} |t^{nab'}| = \infty$$

or else

$$\lim_{n\to\infty} |x^{nab} - x^{nab'}| = \infty$$

We conclude that if χ_R^{nab} is a smooth spatial cutoff localizing to a disc of radius R, $R < \infty$, centered at x^{nab}, then we have

$$\lim_{n\to\infty} \|\chi_R^{nab} V_{3,4}^{nab'}(0 - t^{nab'}, x - x^{nab'})\|_{L_x^2} = 0,$$

using Proposition 9.20. We also claim

Lemma 9.25. *Choosing $B = B(\delta_2, R)$ large enough, we get (here $\delta_2 > 0$ can be prescribed arbitrarily)*

$$\limsup_{n\to\infty} \|\chi_R^{nab} W_{3,4}^{naB}\|_{L_x^2} \ll \delta_2$$

for all $1 \leq b \leq B_0$.

Proof. Recall that

$$W_{3,4}^{naB} = \partial_{1,2} W_1^{naB} \pm \partial_{2,1} W_2^{naB},$$

where $W_{1,2}^{naB}$ solve the covariant wave equation $\Box_{A^n} u = 0$ (where as before A^n is defined using $\Psi^{nA_0^{(0)}}$). But then it is straightforward to check that the space-time Fourier support of u is contained in a small neighborhood of the light cone intersected with the set $|\xi| \sim 1$, up to arbitrarily small errors. One can then reason exactly as in [1], see Lemma 3.8 in that paper. $\square$

Therefore, given $\delta_2 > 0$, we can pick $R = R(\delta_2, V^{ab}_{1,2})$ with the property that

$$\limsup_{n\to\infty}\Big\|\chi_R^{nab}\big(\phi_1^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{1na}]},\phi_2^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{1na}]}\big) - \big(V_3^{ab}(0-t^{nab},x-x^{nab}),V_4^{ab}(0-t^{nab},x-x^{nab})\big)\Big\|_{L^2_x} \ll \delta_2$$

We now need to show that the components

$$\big(\chi_R^{nab}\phi_1^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{1na}]},\chi_R^{nab}\phi_2^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{1na}]}\big)$$

are close to the Coulomb components of an admissible map, up to a constant phase shift. To achieve this, we insert the profile decomposition we obtained above for $\phi^{na}_{1,2}$, i.e., write

$$\chi_R^{nab}\phi_1^{na} = \chi_R^{nab}[\sum_{b'=1}^{B}\sum_{j=1,2}\partial_j\big(S_{\tilde{A}^{nab'}}(\tilde{V}_j^{ab'}[0])(0-t^{nab'},x-x^{nab'})+\tilde{W}_j^{naB}\big)]$$

$$\chi_R^{nab}\phi_2^{na} = \chi_R^{nab}[\sum_{b'=1}^{B}\sum_{j=1,2}(-1)^{j+1}\partial_{j+1}\big(S_{\tilde{A}^{nab'}}(\tilde{V}_j^{ab'}[0])(0-t^{nab'},x-x^{nab'})+\tilde{W}_j^{naB}\big)]$$

where ∂_{j+1} has to be interpreted modulo 2. Now if we choose B large enough (depending on R, chosen further above), and then choose n large enough, we can ensure that

$$\|\chi_R^{nab}\phi_1^{na} - \chi_R^{nab}\sum_{j=1,2}\partial_j\big(S_{\tilde{A}^{nab}}(\tilde{V}_j^{ab}[0])(0-t^{nab},x-x^{nab})\|_{L^2_x} \ll \delta_2$$

$$\|\chi_R^{nab}\phi_2^{na} - \chi_R^{nab}\sum_{j=1,2}(-1)^{j+1}\partial_{j+1}\big(S_{\tilde{A}^{nab}}(\tilde{V}_j^{ab}[0])(0-t^{nab},x-x^{nab})\|_{L^2_x} \ll \delta_2$$

We continue by approximating the truncated components $\chi_R^{nab}\phi_1^{na}$ by the derivative components of an admissible map $(\tilde{\mathbf{x}}^{nab},\tilde{\mathbf{y}}^{nab}):\mathbb{R}^2\to\mathbb{H}^2$.
For this purpose we recall the identity

$$\phi_j^{1na} = (\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_k\partial_j[\phi_k^{1na}\mathbf{y}^{na}],\quad j=1,2$$

Inserting the above decomposition for the ϕ_k^{1na}, we obtain

$$\begin{aligned}\phi_j^{1na} =&(\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_1\partial_j[[\sum_{b'=1}^{B}\sum_{\tilde{j}=1,2}\partial_{\tilde{j}}\big(\mathrm{Re}S_{\tilde{A}^{nab'}}(\tilde{V}_{\tilde{j}}^{ab'}[0])(0-t^{nab'},x-x^{nab'})+\mathrm{Re}\tilde{W}_{\tilde{j}}^{naB}\big)]\mathbf{y}^{na}]\\ &+(\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_2\partial_j[[\sum_{b'=1}^{B}\sum_{\tilde{j}=1,2}(-1)^{\tilde{j}+1}\partial_{\tilde{j}+1}\big(\mathrm{Re}S_{\tilde{A}^{nab'}}(\tilde{V}_{\tilde{j}}^{ab'}[0])(0-t^{nab'},x-x^{nab'})+\mathrm{Re}\tilde{W}_{\tilde{j}}^{naB}\big)]\mathbf{y}^{na}]\end{aligned}$$

Using the frequency localization of all functions involved, and increasing R if necessary (independently of B), we can then achieve that for n large enough

$$\begin{aligned}&\big\|\chi_R^{nab}\phi_j^{1na} - (\mathbf{y}^{na})^{-1}\sum_{k=1,2}\sum_{b'}\triangle^{-1}\partial_1\partial_j[\chi_R^{nab}[\sum_{\tilde{j}=1,2}\partial_{\tilde{j}}\big(\mathrm{Re}S_{\tilde{A}^{nab'}}(\tilde{V}_{\tilde{j}}^{ab'}[0])(0-t^{nab'},x-x^{nab'})+\mathrm{Re}\tilde{W}_{\tilde{j}}^{naB}\big)]\mathbf{y}^{na}]\\ &+(\mathbf{y}^{na})^{-1}\sum_{k=1,2}\sum_{b'}\triangle^{-1}\partial_2\partial_j[\chi_R^{nab}[\sum_{\tilde{j}=1,2}(-1)^{\tilde{j}+1}\partial_{\tilde{j}+1}\big(\mathrm{Re}S_{\tilde{A}^{nab'}}(\tilde{V}_{\tilde{j}}^{ab'}[0])(0-t^{nab'},x-x^{nab'})+\mathrm{Re}\tilde{W}_{\tilde{j}}^{naB}\big)]\mathbf{y}^{na}]\big\|_{L^2_x} \ll \delta_2\end{aligned}$$

From here we infer that

$$\limsup_{n\to\infty}\big\|\chi_R^{nab}\phi_j^{1na} - (\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_k\partial_j[\chi_R^{nab}\phi_k^{1na}\mathbf{y}^{na}]\big\|_{L^2_x} \ll \delta_2$$

Now modify $\mathbf{y}$ to a function $\tilde{\mathbf{y}}^{nab}$ by picking numbers R'', R' with

$$R \ll R'' \ll R',$$

to be specified shortly, and setting

$$\tilde{\mathbf{y}}^{nab} = e^{\sum_{k=1,2}\chi_{R'}^{nab}\triangle^{-1}\partial_k P_{[-R'',R'']}\phi_k^{2na}}$$

whence

$$\frac{\partial_j \tilde{\mathbf{y}}^{nab}}{\tilde{\mathbf{y}}^{nab}} = \partial_j \sum_{k=1,2}\chi_{R'}^{nab}\triangle^{-1}\partial_k P_{[-R'',R'']}\phi_k^{2na} = \chi_{R'}^{nab}\phi_j^{2na} + \text{error},$$

where we can achieve that $\|\text{error}\|_{L_x^2} \ll \delta_2$ by choosing R'' large enough depending on δ_2 and the localization of ϕ^{na} in frequency space, and then R' large enough in relation to R''. Increasing B if necessary and then choosing n large enough, we can then also achieve that

$$\Big\|\chi_{R'}^{nab}\phi_2^{na} - \chi_{R'}^{nab}\sum_{j=1,2}(-1)^{j+1}\partial_{j+1}S_{\tilde{A}^{nab}}(\tilde{V}_j^{ab}[0])(0-t^{nab}, x-x^{nab})\Big\|_{L_x^2} \ll \delta_2$$

and then

$$\|\chi_{R'}^{nab}\phi_2^{na} - \chi_R^{nab}\phi_2^{na}\|_{L_x^2} \ll \delta_2$$

We next show that the expression

$$(\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_k\partial_j[\chi_R^{nab}\phi_k^{1na}\mathbf{y}^{na}]$$

is well approximated by

$$(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_k\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}]$$

To see this, write

$$(\mathbf{y}^{na})^{-1}\sum_{k=1,2}\triangle^{-1}\partial_k\partial_j[\chi_R^{nab}\phi_k^{1na}\mathbf{y}^{na}] = e^{-\sum_{k=1,2}\chi_{R'}^{nab}\triangle^{-1}\partial_k\phi_k^{2na}}\sum_{k=1,2}\triangle^{-1}\partial_k\partial_j[\chi_R^{nab}\phi_k^{1na}e^{\sum_{k=1,2}\chi_{R'}^{nab}\triangle^{-1}\partial_k\phi_k^{2na}}] + \text{error},$$

where we can achieve $\|\text{error}\|_{L_x^2} \ll \delta_2$ by choosing R' large enough in relation to R and the intrinsic Fourier localization properties of ϕ^{na}.

Split the phase into the product

$$e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}\phi_k^{2na}} = e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{[-R'',R'']}\phi_k^{2na}}e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{[-R'',R'']^c}\phi_k^{2na}}$$

We need to show that we can eliminate the factor $e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{[-R'',R'']^c}\phi_k^{2na}}$. Using similar arguments as in Step 1, choosing R'' large enough in relation to R, it is straightforward to show that, with $\partial_k^{-1} := \Delta^{-1}\partial_k$,

$$\|e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}\phi_k^{2na}}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}e^{\sum_{k=1,2}\chi_{R'}^{nab}P_{(-\infty,R'']}\partial_k^{-1}\phi_k^{2na}}(e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{>R''}\phi_k^{2na}}-1)]\|_{L_x^2}$$
$$\ll \delta_2$$

$$\|(e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{>R''}\phi_k^{2na}}-1)e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{(-\infty,R'']}\phi_k^{2na}}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}e^{\sum_{k=1,2}\chi_{R'}^{nab}\triangle^{-1}P_{(-\infty,R'']}\partial_k\phi_k^{2na}}]\|_{L_x^2}$$
$$\ll \delta_2$$

We next show that we can also eliminate $e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}$. Indeed, proceeding as in the first section, write

$$e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}]$$
$$= \sum_{l\geq 2}[\chi_{lR}^{nab} - \chi_{(l-1)R}^{nab}]e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}]$$
$$+ \chi_R^{nab}e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}]$$

Here the cutoff χ_{lR}^{nab} localizes to a disc of radius lR around x^{nab}. Then pick a point p_{lnab} in this disc, for each l, and write for fixed $l \geq 2$

$$e^{-\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}]$$
$$=\Big(\frac{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}}{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}(p_{lnab})}}\Big)^{-1}(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}\frac{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}}{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}(p_{lnab})}}]$$

But then we can estimate

$$[\chi_{lR}^{nab}-\chi_{(l-1)R}^{nab}]\Big(\frac{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}}{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}(p_{lnab})}}-1\Big)=O(\frac{lR}{R''}),$$

and then using the machinery from Step 1 (which yields a l^{-N} gain), and choosing $R'' \gg R$, we can achieve that

$$\Big\|[\Big(\frac{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}}{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}(p_{lnab})}}\Big)^{-1}-1](\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}\frac{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}}{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}(p_{lnab})}}]\Big\|_{L_x^2}$$
$$\ll \delta_2 l^{-N}$$

Similarly, one can eliminate the second instance of

$$\frac{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}}}{e^{\sum_{k=1,2}\chi_{R'}^{nab}\partial_k^{-1}P_{<-R''}\phi_k^{2na}(p_{lnab})}}$$

Hence we have now shown that for B and then n large enough, we have that the functions

$$\tilde{\phi}_j^{1nab}:=(\tilde{\mathbf{y}}^{nab})^{-1}\sum_{k=1,2}\partial_k^{-1}\partial_j[\chi_R^{nab}\phi_k^{1na}\tilde{\mathbf{y}}^{nab}],\ \tilde{\phi}_j^{2nab}:=\frac{\partial_j\tilde{\mathbf{y}}^{nab}}{\tilde{\mathbf{y}}^{nab}},$$

which of course are the derivative components of admissible maps, satisfy the inequalities

$$\|\tilde{\phi}_j^{1nab}-\chi_R^{nab}\phi_j^{1na}\|_{L_x^2}\ll\delta_2,\quad \|\tilde{\phi}_j^{2nab}-\chi_R^{nab}\phi_j^{2na}\|_{L_x^2}\ll\delta_2$$

Our next task is to approximate the *Coulomb components*. For this consider

$$\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}[w_k^{nA_0^{(0)}}+\phi_k^{1na}]},\quad \tilde{\phi}_j^{nab}=\tilde{\phi}_j^{1nab}+i\tilde{\phi}_j^{2nab}$$

From the preceding, we can arrange that

$$\|\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}[w_k^{nA_0^{(0)}}+\phi_k^{1na}]}-\chi_R^{nab}\phi_j^{na}e^{-i\sum_{k=1,2}\partial_k^{-1}[w_k^{nA_0^{(0)}}+\phi_k^{1na}]}\|_{L_x^2}\ll\delta_2$$

We need to show that we can also arrange (i.e., upon choosing B, n large enough) that

$$\|\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}[w_k^{nA_0^{(0)}}+\phi_k^{1na}]}-\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\tilde{\phi}_k^{1nab}}e^{i\gamma_{nab}}\|_{L_x^2}\ll\delta_2$$

for a suitable constant γ_{nab}.

We first get rid of the phase $e^{-i\sum_{k=1,2}\partial_k^{-1}w_k^{nA_0^{(0)}}}$: simply pick a point p_{nab} in the support of χ_R^{nab}, and replace $e^{-i\sum_{k=1,2}\partial_k^{-1}w_k^{nA_0^{(0)}}}$ by $e^{-i\sum_{k=1,2}\partial_k^{-1}w_k^{nA_0^{(0)}}(p_{nab})}$.

Next, we need to show that $\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{1na}}$ is close to $\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\tilde{\phi}_k^{1nab}}$, up to a constant phase shift. First, pick $R_1 \gg R$ such that

$$\|\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{1na}}-\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{1na}}\|_{L_x^2}\ll\delta_2$$

Next, pick $R' \gg R_1$ such that

$$\|\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{1na}}-e^{i\gamma_{1nab}}\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}\phi_k^{1na}}\|_{L_x^2}\ll\delta_2$$

for suitable γ_{1nab}. Next, we claim that picking $R_2 \gg R'$, we can arrange that

$$\|\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}\phi_k^{1na}}-\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}[\chi_{R_2}^{nab}\tilde{\phi}_k^{1nab}]}\|_{L_x^2}\ll\delta_2$$

This is a consequence of the fact that (for R_2 large and then n sufficiently large)

$$\|\chi_{R_1}^{nab}\big[e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}[\chi_{R_2}^{nab}\tilde{\phi}_k^{1nab}-\phi_k^{1na}]}-1\big]\|_{L_x^\infty}\ll\delta_2$$

Finally, we claim that

$$\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}[\chi_{R_2}^{nab}\tilde{\phi}_k^{1nab}]}$$

is very close to $\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\tilde{\phi}_k^{1nab}}$, which is what we need to finish case **(A)**. To see this, note that by choosing R_2 large enough in relation to R', we get from Bernstein's inequality

$$\|\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}[\chi_{R_2}^{nab}\tilde{\phi}_k^{1nab}]-\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}\tilde{\phi}_k^{1nab}\|_{L_x^\infty}\ll\delta_2$$

This immediately implies

$$\|\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}[\chi_{R_2}^{nab}\tilde{\phi}_k^{1nab}]}-\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}\tilde{\phi}_k^{1nab}}\|_{L_x^2}\ll\delta_2$$

To conclude, picking R' large enough in relation to R_1 allows us to find a phase $e^{i\gamma_{2nab}}$ such that

$$\|\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}P_{[-R',R']}\tilde{\phi}_k^{1nab}}-\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\tilde{\phi}_k^{1nab}}e^{i\gamma_{2nab}}\|_{L_x^2}\ll\delta_2$$

Since we also have, as mentioned before, that

$$\|\chi_{R_1}^{nab}\tilde{\phi}_j^{nab}-\tilde{\phi}_j^{nab}\|_{L_x^2}\ll\delta_2$$

Combining all of the preceding steps, we infer the existence of a phase $e^{i\gamma_{nab}}$ such that

$$\|\chi_R^{nab}\phi_j^{na}e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{1na}}-\tilde{\phi}_j^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\tilde{\phi}_k^{1nab}}e^{i\gamma_{nab}}\|_{L_x^2}\ll\delta_2$$

We then get for suitable γ'_{nab}

$$\|e^{i\gamma'_{nab}}\tilde{\phi}_{1,2}^{nab}e^{-i\sum_{k=1,2}\partial_k^{-1}\tilde{\phi}_k^{1nab}}e^{i\gamma_{nab}}-V_{3,4}^{ab}(0-t^{nab},x-x^{nab})\|_{L_x^2}\ll\delta_2$$

This finally concludes case **(A)**, i.e., the temporally bounded case.

(B): *temporally unbounded case.* Here we have $\lim_{n\to\infty}|t^{nab}|=\infty$, whence using Proposition 9.20, we get that

$$V_{3,4}^{nab}(0-t^{nab},x-x^{nab})=o_{L^\infty}(1)+o_{L^2}(1),$$

where we recall the notation

$$V_3^{nab}=\partial_1S_{\tilde{A}^{nab}}(V_1^{ab}[0])+\partial_2S_{\tilde{A}^{nab}}(V_2^{ab}[0])$$
$$V_4^{nab}:=\partial_2S_{\tilde{A}^{nab}}(V_1^{ab}[0])-\partial_1S_{\tilde{A}^{nab}}(V_2^{ab}[0])$$

We now make the following

Claim: *Choosing n large enough, we have*

$$\|\nabla_{x,t}S_{\tilde{A}^{nab}}(V_2^{ab}[0])\|_{L_x^2}\ll\delta_3$$

for any given $\delta_3>0$. In particular, $V_2^{ab}[0]=0$ for all b of temporally unbounded type.

Thus in Case (B), the components $V_{3,4}^{nab}(0-t^{nab},x-x^{nab})$ are approximately given by the gradient of a suitable (complex valued) function. Once the Claim is established, Case (B) will be straightforward to conclude.

In order to prove the Claim, we shall use the curl equations satisfied by the components ϕ_j^{na}. To begin with, pick R large enough such that

$$\|P_{[-R,R]}(\phi_j^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}-\sum_{b=1}^{B}(V_{j+2}^{nab}(0-t^{nab},x-x^{nab})+W_{j+2}^{naB})\|_{L_x^2}\ll\delta_2,\quad j=1,2$$

Then using the Littlewood-Paley trichotomy, and choosing R larger if necessary, we can arrange that

$$\|P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_j^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big) - \sum_{b=1}^{B}(V_{j+2}^{nab}(0-t^{nab},x-x^{nab}) + W_{j+2}^{naB})\|_{L_x^2} \ll \delta_2, \quad j=1,2$$

Now fix a cutoff χ^{nab} which localizes to a large annulus of radius $|t^{nab}|$ around x^{nab} and thickness R_n large enough, such that

$$\limsup_{n\to\infty}\|\chi^{nab}V_{3,4}^{nab}(0-t^{nab},x-x^{nab}) - V_{3,4}^{nab}(0-t^{nab},x-x^{nab})\|_{L_x^2} \ll \delta_2$$

By removing finitely many 'holes' from this annulus and adjusting χ^{nab} correspondingly, we can ensure that

$$\lim_{n\to\infty}\|\chi^{nab}V_{3,4}^{nab'}(0-t^{nab'},x-x^{nab'})\|_{L_x^2} = 0, \quad b\neq b',\ 1\leq b'\leq B$$

for all b' of temporally bounded type. We cannot simply arrange that

$$\lim_{n\to\infty}\|\chi^{nab}W_{3,4}^{naB}\|_{L_x^2} = 0, \quad \lim_{n\to\infty}\|\chi^{nab}V_{3,4}^{nab'}(0-t^{nab'},x-x^{nab'})\| = 0$$

where $V_{3,4}^{nab'}(0-t^{nab'},x-x^{nab'})$ is of unbounded type, and it will be more complicated to disentangle $W_{3,4}^{naB}$ and temporally unbounded $V_{3,4}^{nab'}(0-t^{nab'},x-x^{nab'})$ from $V_{3,4}^{nab}$. From the preceding, choosing R and then n large enough, we can arrange that (here the sum $\sum'$ is over temporally unbounded profiles)

$$\begin{aligned}&\|\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_j^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\\ &-\chi^{nab}[V_{j+2}(0-t^{nab},x-x^{nab}) + \sum_{b\neq b'}{}' V_{j+2}^{nab'}(0-t^{nab'},x-x^{nab'}) + W_{j+2}^{naB}]\|_{L_x^2} \ll \delta_2, \quad j=1,2\end{aligned}$$

Here R only depends on the frequency concentration of ϕ^{na}. We now analyze the curl expression

$$\begin{aligned}&\nabla^{-1}\partial_1\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_2^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\\ &-\nabla^{-1}\partial_2\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\end{aligned}$$

We shall show that this expression becomes arbitrarily small when n is sufficiently large. Decompose the above expression into

$$\begin{aligned}&\nabla^{-1}\big[\partial_1\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_2^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\\ &-\nabla^{-1}\big[\partial_2\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\\ &+\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\partial_1\phi_2^{na}-\partial_2\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\\ &+\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_2^{na})\partial_1(e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}})\big)\big]\\ &-\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_1^{na})\partial_2(e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}})\big)\big]\end{aligned}$$

For the first two terms, choosing the cutoff χ^{nab} suitably, it is clear that for n large enough we have

$$\begin{aligned}&\|\nabla^{-1}\big[\partial_1\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_2^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\\ &-\nabla^{-1}\big[\partial_2\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2}\\ &\ll \delta_2\end{aligned}$$

For the third term, we use the schematic curl relation $\partial_1\phi_2^{na} - \partial_2\phi_1^{na} = $ "$(\phi^{na})^2$". Note that by including a suitable cutoff $\tilde{\chi}^{nab}$ having similar characteristics as χ^{nab}, we get

$$\begin{aligned}&\|\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\partial_1\phi_2^{na}-\partial_2\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\\ &-\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\tilde{\chi}^{nab}\text{"}(\phi^{na})^2\text{"})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2} \ll \delta_2\end{aligned}$$

Now we insert the decomposition

$$\phi_j^{na} = \sum_{b'=1}^{B} \tilde{V}_{j+2}^{nab'}(0 - t^{nab'}, x - x^{nab'}) + \tilde{W}_{j+2}^{naB}, \quad j = 1, 2$$

For any chosen B, by picking n large enough, we can achieve that *for all temporally bounded* b'

$$\Big\|\tilde{\chi}^{nab} \sum_{b'=1, \quad b'\neq b}^{B} \tilde{V}_{j+2}^{nab}(0 - t^{nab}, x - x^{nab})\Big\|_{L_x^2} \ll \delta_2,$$

and hence we reduce to estimating (where now the sum $\sum'$ only involves temporally unbounded profiles)

$$\|\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\tilde{\chi}^{nab}[\tilde{V}_{3,4}^{nab} + {\sum_{b'\neq b}}' \tilde{V}_{3,4}^{nab'} + \tilde{W}_{3,4}^{naB}]^2)e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2}$$

Now recall from Proposition 9.20 that for all temporally unbounded b'

$$\tilde{\chi}^{nab}\tilde{V}_{3,4}^{nab'} = o_{L^\infty}(1) + o_{L^2}(1)$$

Hence we obtain

$$\|\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\tilde{\chi}^{nab}[\tilde{V}_{3,4}^{nab} + {\sum_{b'\neq b}}' \tilde{V}_{3,4}^{nab'}]^2)e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2}$$
$$+ \|\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\tilde{\chi}^{nab}[\tilde{V}_{3,4}^{nab} + {\sum_{b'\neq b}}' \tilde{V}_{3,4}^{nab'}]\tilde{W}_{3,4}^{naB})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2} \ll \delta_2$$

for n large enough. Finally, consider the term

$$\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\tilde{\chi}^{nab}[\tilde{W}_{3,4}^{naB}]^2)e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]$$

Here we split

$$\tilde{W}_{3,4}^{naB} = P_{[-R_1,R_1]^c}\tilde{W}_{3,4}^{naB} + P_{[-R_1,R_1]}\tilde{W}_{3,4}^{naB}$$

Then if B is chosen large enough in relation to R_1, we obtain both

$$\|P_{[-R_1,R_1]^c}\tilde{W}_{3,4}^{naB}\|_{L_x^2} \ll \delta_2, \|P_{[-R_1,R_1]}\tilde{W}_{3,4}^{naB}\|_{L_x^\infty} \ll \delta_2$$

Here the first inequality holds of course uniformly in n, B due to the frequency localization. From here we infer that for B and then n large enough, we get

$$\|\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\tilde{\chi}^{nab}[\tilde{W}_{3,4}^{naB}]^2)e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2} \ll \delta_2$$

The argument for showing

$$\|\nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_2^{na})\partial_1(e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}})\big)\big]$$
$$- \nabla^{-1}\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_1^{na})\partial_2(e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}})\big)\big]\|_{L_x^2} \ll \delta_2$$

of course proceeds in identical fashion.

Summarizing what we have achieved thus far in Case (B), we have shown that for n large enough, we get

$$\|\nabla^{-1}\partial_1\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_2^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]$$
$$- \nabla^{-1}\partial_2\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]\|_{L_x^2} \ll \delta_2$$

In light of the fact pointed out earlier that $(j = 1, 2)$

$$\big[\chi^{nab}P_{[-10R,10R]}\big(P_{[-R,R]}(\phi_j^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}\phi_k^{na}}\big)\big]$$

is well approximated by[26]

$$\chi^{nab}[V_{j+2}^{nab}+\sum_{b'\neq b}{}' V_{j+2}^{nab'}+W_{j+2}^{naB}],$$

we then infer that (recalling the definition of $V_{3,4}$, $W_{3,4}$)

$$\begin{aligned}&\|\nabla^{-1}\partial_1\big[\chi^{nab}\partial_2[((S_{\tilde{A}^{nab}}(V_1^{ab}[0]))(0-t^{nab},x-x^{nab})+\sum_{b'\neq b}{}' S_{\tilde{A}^{nab'}}(V_1^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+W_1^{naB})]\\&-\partial_1[(S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})+\sum_{b'\neq b}{}' S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+W_2^{naB}]]\\&-\nabla^{-1}\partial_2\big[\chi^{nab}\partial_1[((S_{\tilde{A}^{nab}}(V_1^{ab}[0]))(0-t^{nab},x-x^{nab})+\sum_{b'\neq b}{}' S_{\tilde{A}^{nab'}}(V_1^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+W_1^{naB})]\\&+\partial_2[(S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})+\sum_{b'\neq b}{}' S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+W_2^{naB}]]\|_{L_x^2}\\&\ll\delta_2\end{aligned}$$

But then choosing the cutoff χ^{nab} as above and picking n large enough, we conclude (noting cancelations in the preceding expression) that

$$\|\nabla^{-1}\triangle[(S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})+\chi^{nab}\sum_{b'\neq b}{}' S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+\chi^{nab}W_2^{naB}]\|_{L_x^2}\ll\delta_2$$

This inequality, together with the approximate orthogonality of the two summands involved, then gives the smallness of either summand separately: recall from Lemma 9.23 and its proof that we have (for sufficiently large n)

$$\Big|\int_{\mathbb{R}^2}\nabla_{x,t}(S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})\cdot\overline{\nabla_{x,t}W_2^{naB}(0,\cdot)}\,dx\Big|\ll\delta_2$$

$$\Big|\int_{\mathbb{R}^2}\nabla_{x,t}(S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})\cdot\overline{\nabla_{x,t}(S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})}\,dx\Big|\ll\delta_2,\quad b'\neq b$$

Now recall the vanishing condition at time $t=0$

$$\sum_{b'=1}^{B}\partial_t S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+\partial_t W_2^{naB}=0$$

which we used to define the linear covariant evolution of η^{na}. Applying the cutoff χ^{nab}, and choosing n large enough, we get that

$$\|\partial_t S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})+\chi^{nab}\sum_{b'\neq b}{}'\partial_t S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})+\chi^{nab}\partial_t W_2^{naB}\|_{L_x^2}\ll\delta_2$$

However, this inequality, together with the two preceding ones, implies that

$$\begin{aligned}&\|\nabla_{x,t}(S_{\tilde{A}^{nab}}(V_2^{ab}[0]))(0-t^{nab},x-x^{nab})\|_{L_x^2}+\|\chi^{nab}\nabla_{x,t}W_2^{naB}\|_{L_x^2}\\&\qquad+\sum_{b'\neq b}\|\chi^{nab}\nabla_{x,t}(S_{\tilde{A}^{nab'}}(V_2^{ab'}[0]))(0-t^{nab'},x-x^{nab'})\|_{L_x^2}\ll\delta_2\end{aligned}$$

Summarizing the state of affairs in Case (B), we have shown thus far that the Claim holds. But this then says that the 'diluted concentration profile' given by

$$\begin{aligned}V_3^{nab}&=[\partial_1 S_{\tilde{A}^{nab}}(V_1^{ab}[0])+\partial_2 S_{\tilde{A}^{nab}}(V_2^{ab}[0])](0-t^{nab},x-x^{nab})\\V_4^{nab}&:=[\partial_2 S_{\tilde{A}^{nab}}(V_1^{ab}[0])-\partial_1 S_{\tilde{A}^{nab}}(V_2^{ab}[0])](0-t^{nab},x-x^{nab})\end{aligned}$$

[26]Here $\sum'$ again only involves temporally unbounded profiles

is given, up to an L^2-error of size δ_2, by the *pure gradient term*

$$V_3^{nab} = \partial_1 S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab}, x-x^{nab})$$

$$V_4^{nab} = \partial_2 S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab}, x-x^{nab})$$

We shall now use this to construct a map from $\mathbb{R}^2 \to \mathbb{H}^2$ whose Coulomb derivative components are close to $V_{3,4}^{nab}$.

Indeed, picking R large enough and then n sufficiently large depending on R, it is straightforward to check that

$$\partial_j P_{[-R,R]}\big(S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab}, x-x^{nab})\big)$$
$$= \partial_j P_{[-R,R]}\big(S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab}, x-x^{nab})\big)e^{-i\sum_{k=1,2}\partial_k^{-1}\partial_k P_{[-R,R]}\left(S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab},x-x^{nab})\right)} + \text{error},$$

where we have $\|\text{error}\|_{L_x^2} \ll \delta_2$. Then we define a map $(\mathbf{x},\mathbf{y}) : \mathbb{R}^2 \to \mathbb{H}^2$ (here we abuse notation heavily, this map of course depends on n, a, b) via

$$\mathbf{x} := \text{Re}P_{[-R,R]}\big(S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab}, x-x^{nab})\big), \quad \mathbf{y} := e^{\text{Im}P_{[-R,R]}\left(S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab},x-x^{nab})\right)},$$

These then satisfy

$$\Big\|\frac{\partial_j\mathbf{x}}{\mathbf{y}} + i\frac{\partial_j\mathbf{y}}{\mathbf{y}} - \partial_j P_{[-R,R]}\big(S_{\tilde{A}^{nab}}(V_1^{ab}[0])](0-t^{nab}, x-x^{nab})\big)\Big\|_{L_x^2} \ll \delta_2,$$

and the associated Coulomb derivative components are the desired approximations. This concludes the proof of Proposition 9.24. □

Summary thus far, for both (A), (B): *we have shown that we have the "covariant Bahouri Gerard decompositions"*

$$\phi_1^{na} e^{-i\sum_{k=1,2}\partial_k^{-1}[w_k^{nA_0^{(0)}} + \phi_k^{1na}]} = \sum_{b=1}^{B} V_3^{nab}(0-t^{nab}, x-x^{nab}) + W_3^{nab}$$

$$\phi_2^{na} e^{-i\sum_{k=1,2}\partial_k^{-1}[w_k^{nA_0^{(0)}} + \phi_k^{1na}]} = \sum_{b=1}^{B} V_4^{nab}(0-t^{nab}, x-x^{nab}) + W_4^{nab},$$

where we have

$$V_3^{nab} := \partial_1 S_{\tilde{A}^{nab}}(V_1^{ab}[0]) + \partial_2 S_{\tilde{A}^{nab}}(V_2^{ab}[0])$$

$$V_4^{nab} := \partial_2 S_{\tilde{A}^{nab}}(V_1^{ab}[0]) - \partial_1 S_{\tilde{A}^{nab}}(V_2^{ab}[0])$$

and similarly for $W_{3,4}$. Furthermore, for n large enough and any given $\delta_2 > 0$, we can find maps $(\mathbf{x}^{\delta_2 nab}, \mathbf{y}^{\delta_2 nab}) : \mathbb{R}^2 \to \mathbb{H}^2$, with the property that their (spatial) Coulomb derivative components are δ_2 close (within the L^2-metric) to constant phase shifts of the $V_{3,4}^{nab}(0-t^{nab}, x-x^{nab})$.

We shall now refine the information we have by proving the following

Lemma 9.26. *Given $\delta_2 > 0$, we can pick B and then n large enough such that*

$$\|\nabla_{x,t} W_2^{naB}\|_{L_x^2} \ll \delta_2$$

Remark 9.27. Recalling the identities

$$W_3^{naB} = \partial_1 W_1^{naB} + \partial_2 W_2^{naB}$$

$$W_4^{naB} = \partial_2 W_1^{naB} - \partial_1 W_2^{naB}$$

We see that this says that $W_{3,4}^{naB}$ are essentially pure gradient terms, like in Case (B).

Proof. (Lemma 9.26) The proof is quite similar to the Case (B) above. Given $\delta_2>0$, first choose an index B_1 such that we have

$$\limsup_{n\to\infty}\|\sum_{b=B_1}^{B}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})\|_{L_x^2}\ll\delta_2,$$

for any $B\geq B_1$. Further, pick $R=R(\delta_2)$ with the property that

$$\limsup_{n\to\infty}\|P_{[-R,R]^c}(\phi_j^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\|_{L_x^2}\ll\delta_2,\quad j=1,2$$

Increasing R if necessary, we can then also achieve that (for n large enough)

$$\|P_{[-10R,10R]}\big[P_{[-R,R]}(\phi_{1,2}^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\big]-\sum_{b=1}^{B_1}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})-W_{3,4}^{naB}\|_{L_x^2}\ll\delta_2$$

Here we will choose B sufficiently large in relation to B_1,δ_2. Now pick a cutoff χ which localizes to the union of large discs covering most of the support (in the L^2-sense) of the atoms $V_{3,4}^{nab}(0-t^{nab},x-x^{nab})$ of bounded type, i.e., for which $\limsup|t^{nab}|<\infty$, $1\leq b\leq B_1$. Of course χ then depends on a,B_1,n, but we suppress this dependence here. Picking χ suitably and then choosing n large enough, we can then ensure that

$$\|(1-\chi)[\sum_{b=1}^{B_1}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})-\sum_{b=1}^{\prime B_1}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})]\|_{L_x^2}\ll\delta_2,$$

where $\sum_{b=1}^{\prime B_1}$ indicates that we only sum over the atoms of "unbounded type". Summarizing the above steps, we now have

$$\|(1-\chi)P_{[-10R,10R]}\big[P_{[-R,R]}(\phi_{1,2}^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\big]-\sum_{b=1}^{\prime B_1}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})-(1-\chi)W_{3,4}^{naB}\|_{L_x^2}\ll\delta_2$$

By picking B large enough (recall that we can do so independently of B_1), we may also assume that

$$\|(1-\chi)W_{3,4}^{naB}-W_{3,4}^{naB}\|_{L_x^2}\ll\delta_2$$

Here we use Lemma 9.25.

Next, we calculate the curl of the Coulomb components, localized as above, and with an extra cutoff $(1-\chi)$. Thus we want to estimate the expression

$$\partial_2\big((1-\chi)P_{[-10R,10R]}\big[P_{[-R,R]}(\phi_1^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\big]\big)$$
$$-\partial_1\big((1-\chi)P_{[-10R,10R]}\big[P_{[-R,R]}(\phi_2^{na})e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\big]\big)$$

This we can estimate as in Case (B): Of course the case when a derivative falls on $(1-\chi)$ is negligible. Then repeating the arguments in Case (B) above, we need to estimate the schematic expression

$$\big((1-\chi)P_{[-10R,10R]}\big[P_{[-R,R]}([\phi^{na}]^2)e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\big]\big)$$

Here we use the Bahouri Gerard decomposition of the ϕ^{na}, i.e.,

$$\phi_{1,2}^{na}=\sum_{b=1}^{B}V_{3,4}(0-t^{nab},x-x^{nab})+W_{3,4}^{naB}$$

It is clear that we then reduce to estimating

$$\big((1-\chi)P_{[-10R,10R]}\big[P_{[-R,R]}((1-\tilde\chi)[\sum_{b=1}^{\prime B_1}V_{3,4}(0-t^{nab},x-x^{nab})+W_{3,4}^{naB}]^2)e^{-i\sum_{k=1,2}\partial_k^{-1}[\phi_k^{1na}+w_k^{1nA_0^{(0)}}]}\big]\big)$$

But the contribution of the terms $V_{3,4}(0-t^{nab},x-x^{nab})$, $1\leq b\leq B_1$ can be made arbitrarily small by choosing n large enough, while the contribution of $W_{3,4}^{naB}$ is handled by placing one factor into L_x^2 and the other into L^∞_x.

Summarizing, we have now shown that

$$\Big\|\nabla^{-1}\partial_2\big[\sum_{b=1}^{'B_1}V_3^{nab}(0-t^{nab},x-x^{nab})+W_3^{naB}\big]-\nabla^{-1}\partial_2\big[\sum_{b=1}^{'B_1}V_4^{nab}(0-t^{nab},x-x^{nab})+W_4^{naB}\big]\Big\|_{L^2_x}\ll\delta_2$$

But then recalling the defining relations for $V_{3,4}^{nab}, W_{3,4}^{naB}$, we can repeat the argument from part (B) in the preceding proof to conclude that for B and then n large enough, we have

$$\|\nabla_{x,t}W_2^{naB}\|_{L^2_x}\ll\delta_2,$$

as desired. □

Proposition 9.24 together with Lemma 9.26 are key technical tools we shall use in the next section when bounding the wave maps with data

$$w^{nA_0^0}+\phi^{na},$$

where $a=1$.

9.7. **Step 4: Adding the first large atomic component and invoking the induction hypothesis.** In Step 3, we constructed a wave map with data corresponding to the lowest frequency "non-atomic" part, whose Coulomb components are

$$\Phi_\alpha^{nA_0^{(0)}}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{nA_0^{(0)}}}=w_\alpha^{nA_0^{(0)}}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k w_k^{nA_0^{(0)}}}+o_{L^2}(1)$$

Our next step now is to prove bounds for the wave map whose Coulomb components are given by

$$\psi_\alpha^{n(<1)}:=[w_\alpha^{nA_0^{(0)}}+\phi_\alpha^{n1}]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{n1}]}+o_{L^2}(1),$$

provided we make the following key

Energy Assumption: *All concentration profiles have energy $<E_{crit}$. Thus*

$$E(V^{ab})<E_{crit}\ \forall\, b \tag{9.84}$$

As before, in order to avoid confusion, we shall denote the superscript 1 here instead by a, it being understood that $a=1$. Thus we now intend to prove global bounds for the evolution of the Coulomb data

$$\begin{aligned}\psi_\alpha^{n(<a)}&:=[w_\alpha^{nA_0^{(0)}}+\phi_\alpha^{na}]e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{na}]}+o_{L^2}(1)\\&=w_\alpha^{nA_0^{(0)}}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k w_k^{nA_0^{(0)}}}+\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{na}]}+o_{L^2}(1)\end{aligned}$$

¿From the preceding section, we obtain a decomposition of the added term

$$\tilde{\psi}_\alpha^{na}:=\phi_\alpha^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{na}]}$$

as a sum of concentration profiles *at time* $t=0$. Note that this time in principle plays no distinguished role, other than that we are guaranteed existence of the evolution of the wave maps with above data on some small time interval centered at $t=0$. Recall the decompositions (for any $B\geq 1$)

$$\phi_1^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{na}]}=\sum_{b=1}^{B}V_3^{nab}(0-t^{nab},x-x^{nab})+W_3^{naB}$$

$$\phi_2^{na}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{na}]}=\sum_{b=1}^{B}V_4^{nab}(0-t^{nab},x-x^{nab})+W_4^{naB}$$

where

$$\begin{aligned}V_3^{nab}&:=\partial_1 S_{\tilde{A}^{nab}}(V_1^{ab}[0])+\partial_2 S_{\tilde{A}^{nab}}(V_2^{ab}[0])\\V_4^{nab}&:=\partial_2 S_{\tilde{A}^{nab}}(V_1^{ab}[0])-\partial_1 S_{\tilde{A}^{nab}}(V_2^{ab}[0])\end{aligned}$$

and similarly for $W_{3,4}$, while we also have

$$\phi_0^{na} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k[w_k^{nA_0^{(0)}}+\phi_k^{na}]} = \sum_{b=1}^{B}\partial_t S_{\tilde{A}^{nab}}(V_1^{ab}[0])(0-t^{nab}, x-x^{nab}) + \partial_t W_1^{naB},$$

$$\sum_{b=1}^{B}\partial_t S_{\tilde{A}^{nab}}(V_2^{ab}[0])(0-t^{nab}, x-x^{nab}) + \partial_t W_2^{naB} = 0$$

These decompositions are understood to hold at time $t=0$, of course. Now the fact that for B large enough (and then n large enough) we can arrange that $\|\nabla_{x,t}W_2^{naB}\|_{L_x^2} \ll \delta_2$ implies that

$$\|\sum_{b=1}^{B}\partial_t S_{\tilde{A}^{nab}}(V_2^{ab}[0])(0-t^{nab}, x-x^{nab})\|_{L_x^2} \ll \delta_2.$$

Recall that we have *temporally bounded concentration profiles*, as well as *temporally unbounded ones.* Then it is intuitively clear that the evolution of $\tilde{\psi}^{na}$ (this is not well-defined strictly speaking, we can only evolve Coulomb components of actual maps; however, we can think of $\tilde{\psi}^{na}$ as the difference between the components of maps) will be dominated for a large time interval around $t=0$ by the evolution of the temporally bounded concentration profiles, which will exhibit nonlinear behavior, while the temporally unbounded ones will behave like free waves for a long time. In order to make things precise, we introduce a *hierarchy of temporal scales*, which means we order the times t^{nab} according to whether they are positive or negative and then whether

$$\lim_{n\to\infty}(t^{nab}-t^{nab'}) = \pm\infty$$

Assume that this way, we arrive at the list of representative time scales, $M=M(B)$,

$$0 = t^{nab_1}, t^{nab_2}, \ldots, t^{nab_M}$$

where we have $t^{nab_i}>0$, say, and

$$\lim_{n\to\infty}(t^{nab_j}-t^{nab_{j-1}}) = \infty,$$

and furthermore for each $b\in\{1,2,\ldots,B\}$, we have $t^{nab}=t^{nab_j}$ for some j as above. Note that we have chosen to equate those times here that do not diverge from each other. This can of course be done by passing to a subsequence such that the difference of these times converges, and the redefining the concentration profiles accordingly.

We then implement an inductive procedure, controlling the evolution of $\psi_\alpha^{n(<a)}$ on the interval $[0, t^{nab_2}-C]$ for some huge C (such that we are guaranteed that all the concentration profiles focussing at times t^{nab_j}, $j\geq 2$, will not display any nonlinear behavior there yet), while the temporally bounded ones start to disperse and behave linearly around time $t^{nab_2}-C$, for sufficiently large n. *This then guarantees that there is essentially no nonlinear interactions going on between evolutions of concentration profiles at different time scales.*

9.7.1. *Proving a priori bounds for the evolution of $\psi_\alpha^{n(<a)}$; the lowest time scale.* Here we prove a priori bounds on the (wave map) evolution of the Coulomb components $\psi_\alpha^{n(<a)}$. Recall that at time $t=0$, we have the decomposition

$$\psi_{1,2}^{n(<a)}(0,\cdot) = w_{1,2}^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k w_k^{nA_0^{(0)}}} + \sum_{b=1}^{B} V_{3,4}^{nab}(0-t^{nab}, x-x^{nab}) + W_{3,4}^{naB} + o_{L^2}(1)$$

$$\psi_0^{n(<a)}(0,\cdot) = w_0^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k w_k^{nA_0^{(0)}}} + \sum_{b=1}^{B}\partial_t S_{\tilde{A}^{nab}}(V_1^{ab}[0])(0-t^{nab}, x-x^{nab}) + \partial_t W_1^{naB}$$

The "tail ends" $W_{3,4}^{naB}$, $\partial_t W_1^{naB}$ here satisfy the smallness condition

$$W_{3,4}^{naB},\ \partial_t W_1^{naB} = o_{L^2}(1) + o_{L^\infty}(1)$$

where $o(\cdot)$ here is meant in case $B, n \to \infty$. Observe from Lemma 9.26 that we actually have

$$W_{3,4}^{naB} = \partial_{1,2} W_1^{naB} + \text{error}, \quad \|\text{error}\|_{L_x^2} \ll \delta_2,$$

provided we choose B and then n large enough. Furthermore, the proof of Proposition 9.24, case (B), reveals that for concentration profiles which are *temporally unbounded*, we have

$$V_{3,4}^{nab}(0 - t^{nab}, x - x^{nab}) = \partial_{1,2} S_{\tilde{A}^{nab}}(V_1^{ab}[0])(0 - t^{nab}, x - x^{nab})$$

We shall now build the evolution of $\psi_\alpha^{n(<a)}$ as the sum of well-known pieces, namely the evolutions of the atomic profiles, plus an error term, which we will show will remain small. To make things precise, we now use the following construction: We shall use $\delta_2 > 0$ as a smallness parameter which will ultimately hinge on intrinsic properties of the concentration profiles as well as the S-bound on the already constructed low frequency part $\Psi_\alpha^{nA_0^(0)}$, and be specified at the end of the construction. Thinking of $\delta_2 > 0$ as fixed for now, we first pick a large cutoff B_1 with the property that

$$\limsup_{n\to\infty} \big[\| \sum_{B \geq b \geq B_1} V_{4,3}^{nab} \|_{L_x^2} + \| \sum_{B \geq b \geq B_1} \partial_t S_{\tilde{A}^{nab}}(V_1^{ab}[0]) \|_{L_x^2} + \|\partial_\alpha W_1^{naB}\|_{L_{t,x}^\infty + L_t^\infty L_x^2}\big] \ll \delta_2$$

for any $B \geq B_1$. Then we evolve the concentration profiles corresponding to a $b \in \{1, 2, \ldots, B_1\}$ as follows:

(I): Evolution of temporally bounded concentration profile.

Here, by passing to a subsequence, we may assume that t^{nab} converges as $n \to \infty$, and we may then set $t^{nab} = 0$ by time translation. Also, it is apparent that then

$$\begin{aligned} V_3^{nab} &= \partial_1 V_1^{ab}(0, \cdot)) + \partial_2 V_2^{ab}(0, \cdot) + o_{L^2}(1) \\ V_4^{nab} &= \partial_2 V_1^{ab}(0, \cdot)) - \partial_1 V_2^{ab}(0, \cdot) + o_{L^2}(1), \partial_t S_{\tilde{A}^{nab}}(V_1^{ab}[0]) \\ &= \partial_t V_1^{ab}(0, \cdot) + o_{L^2}(1) \end{aligned}$$

are all essentially independent of n. Now according to Proposition 9.24, we can find, for each $\delta_2 > 0$, a constant phase $\gamma_{\delta_2 ab}$ and an admissible map from $\mathbb{R}^2 \to \mathbb{H}^2$ whose Coulomb components $\psi_\alpha^{ab\delta_2}$ satisfy

$$\|e^{i\gamma_{\delta_2 ab}} \psi_{1,2}^{ab\delta_2} - V_{3,4}^{nab}\|_{L_x^2} \ll \delta_3, \quad \|e^{i\gamma_{\delta_2 ab}} \psi_{1,2}^{ab\delta_2} - \partial_t V_1^{ab}(0, \cdot)\|_{L_x^2} \ll \delta_3,$$

For the sake of simplicity, we now refer to the Coulomb components of such a map, which we choose for δ_3 extremely small (depending on B_1 etc. and to be specified later), simply as ψ_α^{ab}.
First, we evolve the components of ψ_α^{ab} on a large time interval I^{ab} centered at $t = 0$, using the wave maps flow for the Coulomb components. This yields an a priori bound

$$\|\psi^{ab}\|_S < C_{ab}$$

due to our energy assumption (9.84). Furthermore, due to Corollary 7.27 as well as Remark 7.28, given $\delta_2 > 0$, one can then choose time intervals

$$I_1, I_2, \ldots, I_{M_{ab}(\delta_2)},$$

where the final one is of the form $[t_1^{ab\delta_2}, \infty)$, say, such that

$$\psi^{ab}|_{I_j} = \psi_{jL}^{ab} + \psi_{jNL}^{ab}$$

with[27]

$$\|\psi_{jNL}^{ab}\|_{S(I_j \times \mathbb{R}^2)} \ll \delta_2, \quad \|\nabla_{x,t} \psi_{jL}^{ab}\|_{L_t^\infty \dot{H}^{-1}} \lesssim E_{crit}$$

Here of course $\Box \psi_{jL}^{ab} = 0$. Note that the intervals I_j here only depend on a, b as well as the smallness parameter δ_2. By the Huyghen's principle, one may assume that the support of ψ_{jL}^{ab} is contained in the set $|x| \leq |t| + D_{ab}(\delta_2)$ for some (possibly very large number $D_{ab}(\delta_2)$). But then by choosing a much larger time $T^{ab\delta_2}$, we can arrange that

$$\|\psi_{M_{ab}L}^{ab}([T^{ab\delta_2}, \infty), \cdot)\|_{L_{t,x}^\infty + L_t^\infty L_x^2} \ll \delta_2$$

[27]The implied constant in the second inequality here is universal, independent of δ_2.

These considerations reveal that pursuing the wave maps evolution of the components ψ^{ab} long enough, we eventually find that

$$\psi^{ab}(t,\cdot) = o_{L^2}(1) + o_{L^\infty}(1),$$

where $o(\cdot)$ is in the sense as $|t| \to \infty$.

This conclusion is of critical importance: note that thus far we have not taken the low frequency contribution from $w^{nA_0^{(0)}}$ (from Step 3) into account, which starts to play an important role for extremely large times. The above asymptotic description allows us to incorporate this low-frequency effect by adjusting the linear evolution of $\psi^{ab}_{M_{ab}L}$ from flat to covariant. In more precise terms, we now make the following choice of an extension $\tilde{\psi}^{ab}_\alpha$ of the data ψ^{ab}_α:

- On the interval $[0, T^{ab\delta_2}]$, we let $\tilde{\psi}^{ab} = \psi^{ab}$.
- On the interval $[T^{ab\delta_2}, \infty)$, we let $\tilde{\psi}^{ab}$ be the covariant extension of $\psi^{ab}[T^{ab\delta_2}]$, i.e., we have

$$\Box_{A^n}\tilde{\psi}^{ab} = 0$$

on $[T^{ab\delta_2}, \infty)$, where A^n is defined with inputs $\Phi^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{1nA_0^{(0)}}}$. More precisely, we apply a Hodge decomposition to the data $\tilde{\psi}^{ab}_\alpha$ as in (9.60), (9.61), and evolve these components as in Step 3. In order to avoid a "kink" at the juncture of these two regimes, we define

$$\tilde{\psi}^{ab} = \chi_{(-\infty,T^{ab\delta_2}+10)}(t)\psi^{ab} + (1-\chi_{(-\infty,T^{ab\delta_2}+10)})(t)S_{A^n}\psi^{ab}[T^{ab\delta_2}] \tag{9.85}$$

where the notation for the second term is schematic, and $\chi_{(-\infty,T^{ab\delta_2}+10)}(t)$ smoothly localizes to the indicated interval and satisfies

$$\chi_{(-\infty,T^{ab\delta_2}+10)}|_{[0,T^{ab\delta_2}]} = 1$$

With these definitions, one can prove the following bound.

Proposition 9.28. *We have a bound of the form*

$$\|\tilde{\psi}^{ab}\|_S < C(C_{ab}),$$

where we recall the assumption $\|\psi^{ab}\|_S < C_{ab}$ from above. Furthermore, denoting by c_k, $k \in \mathbb{Z}$, a frequency envelope controlling the data at time $t = 0$, i.e.,

$$c_k = \Big(\sum_{l\in\mathbb{Z}} 2^{-\sigma|l-k|}\|P_l\psi^{ab}(0,\cdot)\|^2_{L^2_x}\Big)^{\frac{1}{2}}$$

for sufficiently small a priori constant $\sigma > 0$, one has

$$\|P_k\tilde{\psi}^{ab}\|_{S[k]} \le C(C_{ab})c_k$$

Proof. The proof of this follows from Proposition 9.14, as well as Lemma 7.23 and its proof. □

The idea now in Case (I) is to use $\tilde{\psi}^{ab}$ as approximate evolution of the data ψ^{ab} globally in time, for n large enough. Thus $\tilde{\psi}^{ab}|_{[0,T^{ab\delta_2}]}$ is the actual wave maps flow, while beyond time $T^{ab\delta_2}$, we use the covariant linear evolution.

(II): Evolution of temporally unbounded concentration profile.

Here we have $\lim_{n\to\infty}|t^{nab}| = \infty$, and as before, $1 \le b \le B_1$, where we have chosen B_1 above. In this case, using the argument from Case (B) in the proof of Proposition 9.24 and arguing as at the beginning of the preceding Case (I) (we again write ψ^{nab} instead of $\psi^{nab\delta_2}$),

$$\psi^{nab}_\alpha = \partial_\alpha S_{\tilde{A}^{nab}}(V_1^{ab}[0])(0 - t^{nab}, x - x^{nab}) + \text{error}, \quad \alpha = 0,1,2,$$

with $\|\text{error}\|_{L^2_x} \ll \delta_2$. In this case we set

$$\tilde{\psi}^{nab}_\alpha = \partial_\alpha S_{\tilde{A}^{nab}}(V_1^{ab}[0])(t - t^{nab}, x - x^{nab}),$$

the covariant linear evolution. Of course this becomes inaccurate when $t \to t^{nab}$ and the nonlinear effects start to become relevant, but we recall that we are on the lowest time scale in this subsection, i.e., $t \ll t^{nab_2}$. Then we have the following bound.

Proposition 9.29. *There is a bound of the form*

$$\|\tilde{\psi}^{nab}\|_S < C(E_{crit}),$$

Furthermore, denoting by c_k, $k \in \mathbb{Z}$, a frequency envelope controlling the data at time $t = 0$, i.e.,

$$c_k = (\sum_{l\in\mathbb{Z}} 2^{-\sigma|l-k|}\|P_l\psi^{nab}(0,\cdot)\|_{L^2_x}^2)^{\frac{1}{2}}$$

for sufficiently small a priori constant $\sigma > 0$,

$$\|P_k\tilde{\psi}^{ab}\|_{S[k]} \leq C(E_{crit})c_k$$

(III): Evolution of the weakly small error.

These are the components $W_{3,4}^{naB}, \partial_t W_1^{naB}$. From Lemma 9.26, we know that

$$W_{3,4}^{naB} = \partial_{1,2}W_1^{naB} + \text{error},$$

where we can force $\|\text{error}\|_{L^2_x} \ll \delta_2$ by choosing B and then n large enough. We then evolve W_1^{naB} using the covariant linear evolution, i.e.,

$$\Box_{A^n}W_1^{naB}(t,x) = 0, \quad W_1^{naB}[0] = (W_1^{naB}, \partial_t W_1^{naB}),$$

and then define $W_{3,4}^{naB}(t,x) = \partial_{1,2}W_1^{naB}(t,x)$.

We have now defined the evolutions of all the ingredients of $\tilde{\psi}_\alpha^{na}$. We claim that by choosing δ_2 small enough and then B and n large enough, the sum of all these constituents gives the correct evolution of $\tilde{\psi}_\alpha^{na}$ up to a small error. This is clarified the following *Core Proposition for Bahouri Gerard II* which ties it all together.

Proposition 9.30. *There is a cutoff $\delta_2 > 0$ sufficiently small, depending on the profiles $V_{1,2}^{ab}[0]$, $1 \leq b \leq B_1$, as well as the a priori bound we have established for $\Psi^{nA_0^{(0)}}$, such that the following holds: picking B_1 and then n large enough, we can write (with B_1 chosen as above) on $[0, t^{nab_2} - C] \times \mathbb{R}^2$ for C sufficiently large and depending on the $\tilde{\psi}_\alpha^{nab}$ of unbounded type, $b = 1, 2, \ldots, B_1$,*

$$\psi_\alpha^{n(<a)}(t,x) = \Phi_\alpha^{nA_0^{(0)}}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{nA_0^{(0)}}}(t,x) + \sum_{b=1}^{B_1}\tilde{\psi}_\alpha^{nab}(t,x) + \partial_\alpha W_1^{naB_1}(t,x) + \epsilon_\alpha(t,x), \quad \alpha = 0,1,2$$

where the components $\tilde{\psi}_\alpha^{nab}(t,x)$, $\partial_\alpha W_1^{naB_1}(t,x)$, are constructed as in (I)–(III) above, and with

$$\|\epsilon\|_{S([0,t^{nab_2}-C]\times\mathbb{R}^2)} \ll \delta_2$$

Moreover, $\|P_k\epsilon\|_{S([0,t^{nab_2}-C]\times\mathbb{R}^2)}$ is exponentially decaying for frequencies $k > -\log(\lambda_n^a)$ Thus the inequality above implies uniform smallness of $\epsilon(t,x)$ for $t \in [0, t^{nab_2} - C]$.

Remark 9.31. There appears to be circular reasoning in the statement of this result: we need to choose $\delta_{2,3} > 0$ extremely small depending on the profiles $V_{1,2}^{ab}[0]$, $1 \leq b \leq B_1$, but here B_1 itself was defined based on δ_2. This is clarified by noting that all the profiles $V_{1,2}^{ab}[0]$ are small (more precisely, the square sum of their energies is small) for b sufficiently large, and this implies that enlarging B_1 past a certain cutoff will not affect the condition on δ_2; for more clarification see the "important technical observation" below.

Proof. (Proposition 9.30) We will prove the inequality for $P_k\epsilon$ using a bootstrap argument. The challenge consists in careful book-keeping of all the possible interactions. The idea is to essentially replicate the proof of Proposition 9.12 with $\epsilon = \epsilon_2$. The main novel feature here is that we now have to deal with a large number of additional source terms stemming from the nonlinear interactions of the various constituents in the decomposition of $\psi_\alpha^{n(<a)}$. To begin with, we split the (large) time interval $[0, t^{nab_2} - C]$ into finitely many intervals

$$[0, t^{nab_2} - C] = \cup_{j=1}^{M_1} I_j,$$

where we have a decomposition (with $\Psi_\alpha^{nA_0^{(0)}} = \Phi_\alpha^{nA_0^{(0)}}e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{1nA_0^{(0)}}}$)

$$\Psi^{nA_0^{(0)}}|_{I_j} = \Psi_{jL}^{nA_0^{(0)}} + \Psi_{jNL}^{nA_0^{(0)}}$$

with, see Corollary 7.27,

$$\|\Psi_{jNL}^{nA_0^{(0)}}\|_{S(I_j\times\mathbb{R}^2)}<\varepsilon_2,\quad \|\nabla_{x,t}\psi_{jL}^{nA_0^{(0)}}\|_{L_t^\infty\dot{H}^{-1}}\lesssim \varepsilon_2^{-\frac{1}{2}}E_{\rm crit}^2$$

We then run a bootstrap argument inductively on each of these intervals, where of course $M_1=M_1(E_{crit})$ is not too large. We shall now work on the interval I_1, say. This enables us to use the covariant energy estimate from Step 3.

Clearly, the evolved concentration profiles also interact with ϵ; we then further subdivide the intervals I_j into smaller ones, which by abuse of notation we again label as I_j, such that

$$\Big(\sum_{b\in\{1,2,\dots B_1\}}\tilde{\psi}^{nab}\Big)|_{I_j}=\sum_{b\in\{1,2,\dots B_1\}}\tilde{\psi}_{jL}^{nab}+\sum_{b\in\{1,2,\dots B_1\}}\tilde{\psi}_{jNL}^{nab}$$

Note that now the number of intervals is of the form $M_1=M_1(E_{crit},\{V_{1,2}^{ab}[0]\}_{b\in\{1,2,\dots,B_1\}})$. Furthermore, one has

$$\|\Big(\sum_{b\in\{1,2,\dots B_1\}}\tilde{\psi}^{nab}\Big)_{jNL}\|_{S(I_j\times\mathbb{R}^2)}<\varepsilon_2$$

while also

$$\|\Big(\sum_{b\in\{1,2,\dots B_1\}}\nabla_{x,t}\tilde{\psi}^{nab}\Big)_{jL}\|_{L_t^\infty\dot{H}^{-1}}\lesssim\varepsilon_2^{-\frac{1}{2}}E_{\rm crit}^2$$

where ε_2 is a universal constant depending only on E_{crit}. The fact that we get the last inequality with universal implied constant hinges on the approximate orthogonality of the $\tilde{\psi}^{nab}$ for n large enough. One may object at this point that the choice of B_1 was dictated by δ_2, and hence may be extremely large, which in turn means that the number M_1 of intervals above depends on δ_2 and may also become extremely large. The following observation, however, shows that M_1 only depends on a *fixed number* of concentration profiles independent of δ_2:

Important technical observation:

Here we note that M_1 really only depends on $\{V_{1,2}^{ab}[0]\}_{b\in\{1,2,\dots,B_0\}}$, for some B_0 with the property that

$$\sum_{b\geq B_0}\|V_{1,2}^{ab}[0]\|_{L_x^2}^2<\epsilon_0$$

where ϵ_0 is the small-energy global well-posedness cutoff. Thus we can make δ_2 small without increasing M_1 concurrently. To see this, write

$$\{V_{1,2}^{ab}[0]\}_{b\in\{B_0,B_0+1,\dots,B_1\}}=\{V_{1,2}^{ab}[0]\}_{b\in\Lambda_1}\cup\{V_{1,2}^{ab}[0]\}_{b\in\Lambda_2},\quad \Lambda_1\cup\Lambda_2=\{B_0,B_0+1,\dots,B_1\},$$

so that $\{S_{\tilde{A}^{nab}}(V_{1,2}^{ab}[0])(0-t^{nab},x-x^{nab})\}_{b\in\Lambda_1}$ is the collection of temporally bounded concentration profiles with $b\in\{B_0,B_0+1,\dots,B_1\}$. Then the argument that was used for Case (A) in the proof of Proposition 9.24 reveals that we can approximate

$$\sum_{b\in\Lambda_2}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})$$

up to a constant phase shift arbitrarily well by the Coulomb components of an admissible map, and then Proposition 9.14 allows us to evolve the data

$$\sum_{b\in\Lambda_2}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})+o_{L^2}(1)$$

using the covariant linear flow on $[0,t^{nab_1}]$. This leads to bounds that are uniform in B_0,n only involving ϵ_0. Handling the contribution of

$$\sum_{b\in\Lambda_1}V_{3,4}^{nab}(0-t^{nab},x-x^{nab})+o_{L^2}(1)$$

i.e., the "tail" of bounded concentration profiles, is more complicated since we may no longer necessarily approximate this sum by Coulomb components of admissible maps, but only the individual summands $V_{3,4}^{nab}(0-t^{nab},x-x^{nab})$. Thus the correct evolution of this term has to consist of the evolution of the individual ingredients, and one then needs to bound the S-norm of this (very large) sum in terms of an a priori bound, provided n is large enough. In this regard we have the following result.

Lemma 9.32. *For each $b \in \Lambda_1$ and $t \in [0, t^{nab_1}]$, denote by $V_{3,4}^{nab}(t-t^{nab}, x-x^{nab})+o_{L^2}(1)$ the (nonlinear) wave maps evolution of the Coulomb components of an admissible sufficiently good approximation to the data $V_{3,4}^{nab}(t-t^{nab}, x-x^{nab})$, as in the preceding discussion. Then for n large enough, we have*

$$\|\sum_{b\in\Lambda_1} V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\|_{S[0,t^{nab_1}]} \lesssim \varepsilon_0$$

for a suitable universal implied constant.

Proof. (Lemma 9.32) For each $V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)$, pick an interval $[0, \tilde{t}^{ab}]$ with the property that we can write

$$\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab})+o_{L^2}(1)\big]|_{[0,\tilde{t}^{ab}]^c} = \big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab})+o_{L^2}(1)\big]_L+\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab})+o_{L^2}(1)\big]_{NL}$$

where we impose the condition

$$\|\nabla_{x,t}\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]_L\|_{L_t^\infty \dot{H}_x^{-1}} \lesssim \|V_{3,4}^{nab}(0-t^{nab}, x-x^{nab})\|_{L_x^2}$$

$$\|\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]_{NL}\|_{S([0,\tilde{t}^{ab}]^c\times\mathbb{R}^2)} \ll \frac{\varepsilon_0}{B_1}$$

where the implied constant in the first inequality is universal. That this is possible follows from Corollary 7.27 and Remark 7.28. Choosing n large enough and exploiting essential disjointness of the supports at time t^{nab_1}, we can arrange that

$$\big(\sum_{b\in\Lambda_1}\|\nabla_{x,t}\big[V_{3,4}^{nab}(t^{nab_1}-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]_L\|_{\dot{H}_x^{-1}}^2\big)^{\frac{1}{2}} \lesssim \varepsilon_0$$

which then implies (for large enough n)

$$\|\sum_{b\in\Lambda_1}\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]_L\|_{S([0,\max_{b\in\Lambda_1}\tilde{t}^{ab}]^c\times\mathbb{R}^2)} \lesssim \varepsilon_0$$

In order to complete the proof of the lemma, we need to also control

$$\|\sum_{b\in\Lambda_1}\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]\|_{S([0,\max_{b\in\Lambda_1}\tilde{t}^{ab}]\times\mathbb{R}^2)}$$

Here we exploit the fact that for n large, the functions $\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]$ are supported on disjoint light cones up to small errors with respect to S. One then concludes that $\sum_{b\in\Lambda_1}\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]$ are the spatial Coulomb components of a solution to the wave maps problem up to arbitrarily small error (as $n\to\infty$), with energy $\lesssim \varepsilon_0$ at time $t=0$. The small energy well-posedness then implies

$$\|\sum_{b\in\Lambda_1}\big[V_{3,4}^{nab}(t-t^{nab}, x-x^{nab}) + o_{L^2}(1)\big]\|_{S([0,\tilde{t}^{ab}]\times\mathbb{R}^2)} \lesssim \varepsilon_0$$

where the implied constant is universal. □

Now assume the bound[28]

$$\|P_k\epsilon\|_S \leq C_5\delta_2$$

We show that provided we choose $C_5 = C_5(E_{crit})$ large enough, we can bootstrap C_5 to $\frac{C_5}{2}$, whence we get the bound on all of I_1. Then we continue the argument to I_2 etc. Note that by choosing δ_2 small enough in relation to M_1 as well as the other a priori data E_{crit}, $V_{1,2}^{ab}[0]$, $b = 1, 2, \ldots, B_0$, the error term will then remain small.

By scaling invariance, it suffices to bootstrap the estimate for $P_0\epsilon$. We now bootstrap the bound for $P_0\epsilon$. Here we essentially proceed as in step (3), the a priori bound for the first non-atomic component $\psi^{nA_0^{(0)}} = \Phi^{nA_0^{(0)}} e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k\Phi_k^{nA_0^{(0)}}}$. Thus we distinguish as there between the small time case, when

[28]Here ϵ stands for the vector with components ε_α, $\alpha = 0, 1, 2$

the div-curl system suffices, and the large time case, when the wave equations are important: we shall work here on the interval I_1 containing the initial time slice $t = 0$.

(i): small time case $|I_1| < T_1$. Here T_1 is a sufficiently small absolute constant. Write the equation for ϵ, using the div-curl system, schematically as follows:

$$\begin{aligned}\partial_t P_0\epsilon =&\nabla_x P_0\epsilon + P_0\big[\epsilon\nabla^{-1}([\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{a(>B_0)} + W^{naB}]^2)\big]\\ &+ P_0\big[[\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{a(>B_0)} + W^{naB}]\nabla^{-1}(\epsilon[\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{a(>B_0)} + W^{naB}])\big]\\ &+ P_0\big[\epsilon\nabla^{-1}(\epsilon[\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{na(>B_0)} + W^{naB}])\big] + P_0\big[[\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{a(>B_0)} + W^{naB}]\nabla^{-1}(\epsilon^2)\big]\\ &+ P_0[\epsilon\nabla^{-1}(\epsilon^2)] + \text{interactions terms}\end{aligned}$$

"Interactions terms" here refers to all possible expressions which do not involve the radiation term ϵ such as

$$P_0[\psi^{nA_0^{(0)}}\nabla^{-1}[(\tilde{\psi}^{nab})^2]]$$

Indeed, the complete list of the error terms included under this heading is complicated, due to our construction of the evolutions $\tilde{\psi}^{nab}$ in (I)-(III) above. Recall that for the *temporally bounded type components*, we use the nonlinear wave maps flow on a large time interval $T^{ab\delta_2}$, but we then use the covariant linear evolution past that time. This means that on $[0, T^{ab\delta_2}]$, we generate error interaction terms like the preceding one coming from the interactions with the low frequency part $\psi^{nA_0^{(0)}}$, while on the interval $[T^{ab\delta_2}, t^{nab_2} - C]$ generate errors due to the *nonlinear self-interactions* of $\tilde{\psi}^{nab}$.

On the other hand, for the *temporally unbounded type components*, we use the linear covariant evolution on $[0, t^{nab_2} - C]$, which means that we generate errors due to the nonlinear self-interactions.

In addition to all these, we generate errors due to different concentration profiles interacting with each other, as well with the small frequency component $\psi^{nA_0^{(0}}$, or the weakly small error, and the latter also generates nonlinear errors due to interactions with itself. We will deal with this rather large collection of errors later, showing that we can make its $N[0]$-norm arbitrarily small by choosing B_1 large enough, and then n large enough.

We also use the notation $\tilde{\psi}^{a(>B_0)}$ for the evolution of

$$\sum_{b\in\Lambda_1\cup\Lambda_2} V_{3,4}^{nab}(0 - t^{nab}, x - x^{nab}) + o_{L^2}(1), \sum_{b\in\Lambda_1\cup\Lambda_2} \partial_t S_{\tilde{A}^{nab}}(V_1^{ab}[0])(0 - t^{nab}, x - x^{nab}) + o_{L^2}(1),$$

as explained in the "important technical observation" above.

We first deal with the terms involving ϵ. Our task is to gain a smallness constant that allows us to improve the a priori bound we are assuming about ϵ.

(i.1): Terms involving ϵ. These can be handled exactly as Case 1 in the proof of Proposition 9.12, in light of the bound

$$\|[\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{a(>B_0)} + W^{naB}]\|_S \leq C(\psi^{nA_0^{(0)}}, \{\tilde{\psi}^{nab}\}_{b=1}^{B_0}, E_{crit})$$

Thus for example paralleling Case 1 (a) in the proof of Proposition 9.12, one obtains a bound

$$\sum_{k\in\mathbb{Z}}\|\chi_{I_j}P_k\big[\epsilon\nabla^{-1}([\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0}\tilde{\psi}^{nab} + \tilde{\psi}^{a(>B_0)]^2})\|_{L_t^2\dot{H}^{-\frac{1}{2}}}^2 \ll \|\epsilon\|_S^2$$

provided we choose the time cutoffs suitably (such that the number M_1 of such time intervals is as above).

(i.2) Errors due to nonlinear (self)interactions of the $\tilde{\psi}^{nab}$, $\psi^{nA_0^{(0)}}$, W^{naB}. Note that these errors serve as source terms for ϵ, and hence we need to show that they are extremely small (of order controlled by δ_2).

The mechanism for this is first choosing B_1 sufficiently large (for the contributions involving W^{naB}), and then choosing n large enough. As these estimates are analogous to those in Case 1 of Proposition 9.12, we explain here only the mechanism for generating arbitrary smallness (as $n \to \infty$).

(i.2.a) Errors generated by the temporally bounded type $\tilde{\psi}^{nab}$. If $\tilde{\psi}^{nab}$ is the evolution of a temporally bounded concentration profile, then recall that we let $\tilde{\psi}^{nab}$ be the wave maps evolution on the interval $[0, T^{ab\delta_2}]$, *provided $\tilde{\psi}^{nab}$ is supported at frequency scale* ~ 1. Now we want to track the evolution of an arbitrary frequency mode $P_k\epsilon$, which we have scaled to $k = 0$. But then we have also re-scaled all the source terms. Now the source terms generated by $\tilde{\psi}^{nab}$ itself come from a number of sources: first, the "gluing definition" of (9.85) implies that we generate errors of the form (before frequency localization)

$$\chi'_{(-\infty, T^{ab\delta_2}+10)}(t)\psi^{ab} - \chi'_{(-\infty, T^{ab\delta_2}+10)}(t)S_{A^n}\psi^{ab}[T^{ab\delta_2}]$$

The only way for this term to contribute in the Case (i) for a fixed frequency (which we assume equals one after scaling) is when the original frequency (which gets scaled to one) is extremely large. But this contribution is then easily seen to be very small in $L_t^\infty L_x^2$, say, due to the frequency localization of $\tilde{\psi}^{ab}$. Next, the self-interaction errors generated from the usual div-curl system are (schematically)

$$P_0[\partial_t\tilde{\psi}^{nab} - \nabla_x\tilde{\psi}^{nab} - \tilde{\psi}^{nab}\nabla^{-1}[(\tilde{\psi}^{nab})^2]],$$

which vanishes provided the I_1 fits into the *re-scaled interval* $[0, T^{ab\delta_2}]$. Otherwise, one obtains a contribution of the above form on the complement of the *re-scaled interval* $[0, T^{ab\delta_2}]$ inside I_1, and which is of the above form. We need to show that picking n large enough, this can be made arbitrarily small. For this purpose we use the following observation.

Lemma 9.33. *Let $\tilde{\psi}^{nab}$ be the evolved Coulomb components of a temporally bounded type concentration profile, concentrated at frequency ~ 1. Then letting $T^{ab\delta_2}$ be the time indicating transition from nonlinear to linear evolution (as explained in the preceding discussion), we have*

$$\tilde{\psi}_\alpha^{nab}(T^{ab\delta_2}, \cdot) = \partial_\alpha\tilde{\psi}^{nab}(T^{ab\delta_2}, \cdot) + \mathit{error},$$

where

$$\|\mathit{error}\|_{L_x^2} \to 0$$

as $\delta_2 \to 0$, and furthermore

$$\tilde{\psi}^{nab} = \sum_{k=1,2} \triangle^{-1}\partial_k\tilde{\psi}_k^{nab}$$

The proof of this lemma follows exactly as in the proof of Case (B) of Proposition 9.24. It then follows that in case we are on the complement of the re-scaled interval $[0, T^{ab\delta_2}]$ inside I_1, we generate errors of the form

$$P_0[\tilde{\psi}^{nab}\nabla^{-1}[(\tilde{\psi}^{nab})^2]] + \text{error}$$

with error as in the preceding lemma, in addition to errors stemming from interactions of $\tilde{\psi}^{nab}$ with the other components $\psi^{nA_0^{(0)}}$ etc. to be considered later. But then, using the $L_{t,x}^\infty$-dispersion for the $\tilde{\psi}^{nab}(t, \cdot)$ as $|t| \to \infty$, it is seen that

$$\|P_0[\tilde{\psi}^{nab}\nabla^{-1}[(\tilde{\psi}^{nab})^2]]\|_{L_t^\infty L_x^2(I_1\cap[0,T^{ab\delta_2}]^c\times\mathbb{R}^2)} \ll \delta_2$$

if we choose $T^{ab\delta_2}$ large enough in relation to δ_2. Next, we need to analyze the errors generated by $\tilde{\psi}^{nab}$ through interaction with the other ingredients $\psi^{nA_0^{(0)}}$, $\tilde{\psi}^{a(>B_0)}$, and W^{naB}. We begin with the interactions between two *distinct* terms $\tilde{\psi}^{nab}$, $b = 1, 2, \ldots, B_1$. Thus we are considering

$$P_0[\tilde{\psi}^{nab_1}\nabla^{-1}(\tilde{\psi}^{nab_2}\tilde{\psi}^{nab_3})]$$

where $b_i \neq b_j$ for some i, j. By the frequency localization of all these factors, we may assume that, up to negligible errors, each of them satisfies $\tilde{\psi}^{nab_j} = P_{[-C_6, C_6]}\tilde{\psi}^{nab_j}$ where C_6 is a potentially extremely large constant depending on the frequency localizations (i.e., how well-localized the factors are in frequency space), as well as δ_2 and B_1, and that $\log\left[(\lambda_n^a)^{-1}\right] \in [-C_6, C_6]$. Now assume first that $I_1 \subset [0, T^{ab_j\delta_2}]$ for all j, i.e., our time interval is such that we are in the "nonlinear regime" for each of these factors. But then choosing n large enough, we can force

$$\|P_0[\tilde{\psi}^{nab_1}\nabla^{-1}(\tilde{\psi}^{nab_2}\tilde{\psi}^{nab_3})]\|_{L_t^\infty L_x^2} \ll \delta_2 \frac{1}{C_6^{100}}$$

by the essential disjointness of the supports of the factors, and this suffices to handle Case 1, see the proof of Proposition 9.12. Indeed, the expression

$$[\tilde{\psi}^{nab_1}\nabla^{-1}(\tilde{\psi}^{nab_2}\tilde{\psi}^{nab_3})]$$

is essentially supported in a frequency interval $[-10C_6, 10C_6]$, and repeating the above estimate for each of these frequencies and square summing easily yields the bound

$$\sum_k \|\chi_{[-T_1,T_1]}(2^k t)P_k[\tilde{\psi}^{nab_1}\nabla^{-1}(\tilde{\psi}^{nab_2}\tilde{\psi}^{nab_3})]\|_{L_t^2\dot{H}^{-\frac{1}{2}}} \ll \delta_2$$

If, on the other hand, at least one of the factors $\tilde{\psi}^{nab_j}$ is in the "linear regime" (i.e., satisfies the covariant wave equation), then smallness follows from the L^∞-decay.

Next we consider the term

$$P_0[\tilde{\psi}^{nab}\nabla^{-1}[(\psi^{nA_0^{(0)}})^2]]$$

Here of course it is essential that we are in Case (i) and so it suffices to estimate the $L_t^\infty L_x^2$ or also $L_{t,x}^2$ norm of this term, see Case 1 of the proof of Proposition 9.12. Due to the essential disjointness of the Fourier supports of $\tilde{\psi}^{nab}$ and $\psi^{nA_0^{(0)}}$, see Proposition 9.9, we may assume that the first input $\tilde{\psi}^{nab}$ has frequency of size one, while the second input has extremely small frequency (controlled by picking n large enough). But then we may estimate this contribution by placing $\nabla^{-1}[(\psi^{nA_0^{(0)}})^2]$ into $L_{t,x}^\infty$, and re-scaling and square-summing over the output frequencies results in the desired small bound.
Finally, the interactions of temporally bounded $\tilde{\psi}^{nab}$ with the remaining weakly small errors W^{naB_1} are handled similarly by exploiting the smallness of the latter with respect to $L_{t,x}^\infty$. Here the "Important Technical Observation" from before becomes important again.

(i.2.b) Errors generated by temporally unbounded $\tilde{\psi}^{nab}$. Again the errors generated are of the form

$$P_0[\partial_t\tilde{\psi}^{nab} - \nabla_x\tilde{\psi}^{nab} - \tilde{\psi}^{nab}\nabla^{-1}[(\tilde{\psi}^{nab})^2]],$$

as well as terms involving interactions of $\tilde{\psi}^{nab}$ with $\psi^{nA_0^{(0)}}$, $\tilde{\psi}^{a(>B_0)}$, as well as W^{naB}. From Part (B) of the proof of Proposition 9.24, we know that $\tilde{\psi}^{nab}$ is of gradient form up to an error which can be made arbitrarily small. Hence the above simplifies, up to a negligible error, to the nonlinear term

$$-P_0[\tilde{\psi}^{nab}\nabla^{-1}[(\tilde{\psi}^{nab})^2]],$$

To estimate this, we can first reduce this to

$$-P_0[\tilde{\psi}^{nab}\nabla^{-1}P_{[-C_6,C_6]}[(P_{[-C_6,C_6]}\tilde{\psi}^{nab})^2]],$$

arguing as in Case 1 of the proof of Proposition 9.12, and then by using the $L_{t,x}^\infty$-dispersion, i.e., Lemma 9.20, to write

$$P_{[-C_6,C_6]}\tilde{\psi}^{nab}(0,\cdot) = o_{L^\infty}(1)$$

from which the desired smallness follows easily. The interaction terms of temporally unbounded $\tilde{\psi}^{nab}$ with the remaining components $\psi^{nA_0^{(0)}}$, $\tilde{\psi}^{a(>B_0)}$, W^{naB}, are handled as before and are omitted.

(i.2.c) Errors generated by the weakly small remainder W^{naB}. Again recalling Part (B) of the proof of Proposition 9.24, and Lemma 9.26, we know that W_α^{naB} is of pure gradient form up to a negligible error (provided B and n are large enough). The conclusion is that the error of the form

$$P_0[\partial_t W^{naB} - \nabla_x W^{naB} - W^{naB}\nabla^{-1}([W^{naB}]^2)$$

reduces up to a negligible error to the nonlinear self-interaction term

$$P_0[-W^{naB}\nabla^{-1}([W^{naB}]^2)],$$

which can be estimated as in the preceding case, using the smallness of $\|W^{naB}\|_{L^\infty_{t,x}}$ after reducing to frequencies of size $O(1)$.

(ii) $|I_1| > T_1$, T_1 as in Case 1. Proceeding as in Case 2 of the proof of Proposition 9.12, we decompose $P_0\epsilon$ into

$$P_0\epsilon = P_0 Q_{<D}\epsilon + P_0 Q_{\geq D}\epsilon,$$

where $D = D(E_{crit})$ is a sufficiently large constant. Then arguing as in the proof of Proposition 9.12, we obtain two equations

$$\Box_A P_0 Q_{<D}\epsilon_\alpha = F^1_\alpha$$
$$\Box P_0 Q_{\geq D}\epsilon = F^2_\alpha$$

Here the magnetic potential A in the first equation is defined as in the proof of Proposition 9.12 but with ψ_L replaced by

$$\psi^{nA_0^{(0)}} + \sum_{b=1}^{B_0} \tilde{\psi}^{nab}_L + W^{naB}$$

The source terms F^1_α are obtained as in Section 3, and here we of course linearize around the above expression. Then we re-iterate the estimates in the proof of Proposition 9.12, with ϵ replacing ϵ_2 and $\epsilon_1 = 0$. As in Case 1 above, the only new feature are the source terms coming from nonlinear interactions between the various $\tilde{\psi}^{nab}$, W^{naB_1}. Fortunately, the fact that each of these functions is essentially frequency localized to the same interval, the mechanisms that force smallness reduce as before to either physical separation or dispersive decay. We explain here how to obtain smallness for the trilinear null-form source terms, which we write schematically in the form $\nabla_{x,t}[\rho_1\nabla^{-1}\mathcal{Q}_{\nu j}(\rho_2,\rho_3)]$, were ρ represents one of the functions $\psi^{nA_0^{(0)}}$, $\sum_{b=1}^{B_0}\tilde{\psi}^{nab}_L$, W^{naB}. We consider the following cases:

(ii.0) *Errors due to the gluing construction* (9.85). These errors are of the form

$$\chi''_{(-\infty,T^{ab\delta_2}+10)}(t)\psi^{ab} - \chi''_{(-\infty,T^{ab\delta_2}+10)}(t)S_{A^n}\psi^{ab}[T^{ab\delta_2}]$$
$$\chi'_{(-\infty,T^{ab\delta_2}+10)}(t)\partial_t\psi^{ab} - \chi'_{(-\infty,T^{ab\delta_2}+10)}(t)\partial_t S_{A^n}\psi^{ab}[T^{ab\delta_2}]$$

To show the smallness of these, note that ψ^{ab} solves the schematic div-curl system

$$\nabla_t\psi - \nabla_x\psi = \psi\nabla^{-1}(\psi^2)$$

Now since we have $\psi^{ab} = o_{L^\infty}(1) + o_{L^2}(1)$, choosing $T^{ab\delta_2}$ large enough, we see that (with $o(1)$ in case $T\to\infty$)

$$\chi_{[T,T+10]}\big[\nabla_t\psi^{ab} - \nabla_x\psi^{ab}\big] = o_{L^M_t\dot{H}^{-(1-\frac{1}{M})}}(1)$$

Similarly, by construction, the extensions $S_{A^n}\psi^{ab}[T^{ab\delta_2}]$ also satisfy the (schematic) relations

$$\chi_{[T,T+10]}\big[\partial_t\big(S_{A^n}\psi^{ab}[T^{ab\delta_2}]\big) - \nabla_x\big(S_{A^n}\psi^{ab}[T^{ab\delta_2}]\big)\big] = o_{L^\infty_T L^2_x}(1),$$

see Lemma 9.33. But then it easily follows that

$$\|\chi''_{(-\infty,T^{ab\delta_2}+10)}(t)\psi^{ab} - \chi''_{(-\infty,T^{ab\delta_2}+10)}(t)S_{A^n}\psi^{ab}[T^{ab\delta_2}]\|_N \ll \delta_2$$
$$\|\chi'_{(-\infty,T^{ab\delta_2}+10)}(t)\partial_t\psi^{ab} - \chi'_{(-\infty,T^{ab\delta_2}+10)}(t)\partial_t S_{A^n}\psi^{ab}[T^{ab\delta_2}]\|_N \ll \delta_2$$

(ii.1) *Self-interactions of temporally bounded* $\tilde{\psi}^{nab}$. These only occur provided $I_1\cap[0,T^{ab\delta_2}]^c \neq \emptyset$. Thus assume the latter is the case, and consider

$$\nabla_{x,t}[\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab},\tilde{\psi}^{nab})]$$

Now the estimates of Section 5.3 imply that we obtain

$$\|\nabla_{x,t}[\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab},\tilde{\psi}^{nab})]\|_{N(I_1\times\mathbb{R}^2)} \ll \delta_2$$

provided at least two of the inputs have Fourier support with very close angular alignment, depending on $\|\psi^{nab}\|_S$, δ_2. Thus we may assume that these inputs have Fourier supports with some amount (albeit very small) of angular separation. Similarly, localizing the Fourier support to frequency ~ 1, say, we may reduce to the expression

$$\sum_{k_{1,2,3}=O(1)} \nabla_{x,t} P_0[P_{k_1}\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(P_{k_2}\tilde{\psi}^{nab}, P_{k_3}\tilde{\psi}^{nab})],$$

where the implied constant $O(1)$ is of course potentially extremely large, depending on $\|\psi^{nab}\|_S$, δ_2. We may similarly assume that all the modulations present are of size $O(1)$ at most (which may again be quite large, depending on $\|\psi^{nab}\|_S$, δ_2). But then the assumed angular separation between all factors allows us to bound this expression (for fixed frequencies) by

$$\|\nabla_{x,t} P_0[P_{k_1}\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(P_{k_2}\tilde{\psi}^{nab}, P_{k_3}\tilde{\psi}^{nab})]\|_{N[0]} \lesssim \|P_{k_1}\tilde{\psi}^{nab}\|_{S[k_1]}\|\nabla^{-1}\mathcal{Q}_{\nu j}(P_{k_2}\tilde{\psi}^{nab}, P_{k_3}\tilde{\psi}^{nab})\|_{L^2_{t,x}}$$

But then the desired smallness follows by interpolating the improved bilinear Strichartz type bound

$$\|\nabla^{-1}\mathcal{Q}_{\nu j}(P_{k_2}\tilde{\psi}^{nab}, P_{k_3}\tilde{\psi}^{nab})\|_{L^p_{t,x}} \lesssim \prod_{j=1,2}\|P_{k_j}\tilde{\psi}^{nab}\|_{S[k_j]}$$

for some $p < 2$ following from a result due to Bourgain[29] [2] as well as Lemma 2.22, and the smallness bound

$$\|\nabla^{-1}\mathcal{Q}_{\nu j}(P_{k_2}\tilde{\psi}^{nab}, P_{k_3}\tilde{\psi}^{nab})\|_{L^\infty_{t,x}} \ll 1$$

which we obtain by letting $T^{ab\delta_2}$ be large enough in relation to δ_2. Replacing 0 by k and square summing over the output frequencies, the desired bound follows easily.

(ii.2): *Interactions of two different temporally bounded* $\tilde{\psi}^{nab}$. Here the mechanism at work is the physical separation of the centers of mass for n large. Thus consider

$$\nabla_{x,t} P_0[P_{k_1}\tilde{\psi}^{nab_1}\nabla^{-1}\mathcal{Q}_{\nu j}(P_{k_2}\tilde{\psi}^{nab_2}, P_{k_3}\tilde{\psi}^{nab_3})]$$

where we have $b_i \neq b_j$ for at least one pair i, j. Now if we have $I_1 \subset [0, T^{ab_j\delta_2}]$ for both i, j, then $\tilde{\psi}^{nab_{i,j}}$ are essentially supported in disjoint light cones for n large enough. Specifically, due to Lemma 7.22 as well as Lemma 7.23, given $\delta_2 > 0$, we may write

$$\tilde{\psi}^{nab_i} = \tilde{\psi}^{nab_i}_{\text{cone}} + \tilde{\psi}^{nab_i}_{\text{cone}^c}$$
$$\tilde{\psi}^{nab_j} = \tilde{\psi}^{nab_j}_{\text{cone}} + \tilde{\psi}^{nab_j}_{\text{cone}^c}$$

where we have

$$\|\tilde{\psi}^{nab_{i,j}}_{\text{cone}^c}\|_S \ll \delta_2$$

while the functions $\tilde{\psi}^{nab_{i,j}}_{\text{cone}}$ are supported in disjoint double cones while still satisfying

$$\|\tilde{\psi}^{nab_{i,j}}_{\text{cone}}\|_S \lesssim C(\tilde{\psi}^{nab})$$

It is then straightforward to conclude that by choosing n large enough, we may force

$$\|\nabla_{x,t}[\tilde{\psi}^{nab_1}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab_2}, \tilde{\psi}^{nab_3})]\|_N \ll \delta_2$$

If on the other hand we have $I_1^c \cap [0, T^{ab_j\delta_2}] \neq \emptyset$ for at least one j, then we use L^∞ dispersion on this intersection to get smallness as before.

(ii.3) *Interactions of temporally bounded* $\tilde{\psi}^{nab}$ *and* $\psi^{nA_0^{(0)}}$. We distinguish between $I_1 \subset [0, T^{ab\delta_2}]$ and $I_1 \cap [0, T^{ab\delta_2}]^c \neq \emptyset$. In the former case, where $\tilde{\psi}^{nab}$ is given by the actual wave map propagation, we generate error terms of the form

$$\nabla_{x,t}[\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(\psi^{nA_0^{(0)}}, \psi^{nA_0^{(0)}})]$$

As in case (i.2.a) above, localizing the output to frequency ~ 1, we may reduce to the case when $\nabla^{-1}\mathcal{Q}_{\nu j}(\psi^{nA_0^{(0)}}, \psi^{nA_0^{(0)}})$ has extremely small frequency. But then one obtains

$$\|\nabla_{x,t} P_0[\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(\psi^{nA_0^{(0)}}, \psi^{nA_0^{(0)}})]\|_{L^1_t \dot{H}^{-1}} \ll \|P_{[-5,5]}[\tilde{\psi}^{nab}\|_{L^\infty_t L^2_x}$$

[29] Of course one also has the optimal results due to Wolff and Tao, but those are not really needed here.

and re-scaling and square summing over the output frequencies, we can force an upper bound $\ll \delta_2$ by choosing n large enough. We further generate interaction terms of the form

$$\nabla_{x,t}[\psi^{nA_0^{(0)}}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab},\psi^{nA_0^{(0)}})],\quad \nabla_{x,t}[\psi^{nA_0^{(0)}}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab},\tilde{\psi}^{nab})],$$

However, the trilinear estimates in Section 5 in addition to the frequency support properties of these inputs reveal that choosing n large enough, we can force

$$\|\nabla_{x,t}[\psi^{nA_0^{(0)}}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab},\psi^{nA_0^{(0)}})]\|_N \ll \delta_2,\quad \|\nabla_{x,t}[\psi^{nA_0^{(0)}}\nabla^{-1}\mathcal{Q}_{\nu j}(\tilde{\psi}^{nab},\tilde{\psi}^{nab})]\|_N \ll \delta_2$$

Note that in the second situation above, i.e., $I_1\cap[0,T^{ab\delta_2}]^c\neq\emptyset$, we *essentially* no longer generate errors of the form

$$\nabla_{x,t}[\tilde{\psi}^{nab}\nabla^{-1}\mathcal{Q}_{\nu j}(\psi^{nA_0^{(0)}},\psi^{nA_0^{(0)}})], \tag{9.86}$$

as well as similar higher order terms arising from the Hodge expansion of the potential term

$$-\sum_{j=1,2}\triangle^{-1}\partial_j\big[\Psi_\nu^{1nA_0^{(0)}}\Psi_j^{2nA_0^{(0)}}-\Psi_\nu^{2nA_0^{(0)}}\Psi_j^{1nA_0^{(0)}}\big]$$

within $I_1\cap[0,T^{ab\delta_2}]^c$ since now $\tilde{\psi}^{nab}$ is given by the linear covariant evolution. This is made precise as follows: by construction, on the latter intersection, we can write

$$\tilde{\psi}_1^{nab}=\partial_1\zeta^{nab}+\partial_2\eta^{nab},\ \tilde{\psi}_2^{nab}=\partial_2\zeta^{nab}-\partial_1\eta^{nab},\ \tilde{\psi}_0^{nab}=\partial_t\zeta^{nab},$$

where the functions ζ^{nab},η^{nab} solve the covariant wave equations

$$\Box_{A^n}\zeta^{nab}=\Box_{A^n}\eta^{nab}=0$$

Here the potential term A^n is defined as in (9.63), which is to be contrasted with the 'true potential'

$$A^n_{*\nu}:=-\sum_{j=1,2}\triangle^{-1}\partial_j\big[\Psi_\nu^{1nA_0^{(0)}}\Psi_j^{2nA_0^{(0)}}-\Psi_\nu^{2nA_0^{(0)}}\Psi_j^{1nA_0^{(0)}}\big]$$

Consider the terms $\partial_\alpha\zeta^{nab}=:\partial_\alpha\zeta$, $\alpha=0,1,2$. We make the **Claim** that the expression

$$\Box\partial_\alpha\zeta+i\partial^\beta\big[\partial_\alpha\zeta A^n_\beta\big]+i\partial_\alpha\big[\partial^\beta\zeta I^c A^n_\beta\big]+i\sum_k P_k\big[\partial^\beta\partial_\alpha\zeta P_{<k-5}IA^n_\beta\big] \tag{9.87}$$

is negligible in that its $\|\cdot\|_N$-norm converges to zero as $n\to\infty$. In light of the decompositions of the nonlinearity in section 3, one then easily concludes that terms of the form (9.86) as well as similar higher order terms are indeed accounted for by the covariant wave evolution $\Box_{A^n}$.
To see the above Claim, we use the notation

$$fg=f_Hg_H+f_Hg_L+f_Lg_H$$

where we put

$$f_Hg_H=\sum_k P_kfP_{[k-5,k+5]}g,\ f_Hg_L=\sum_k P_kfP_{<k-5}g,\ f_Lg_H=\sum_k P_kfP_{>k+5}g$$

Then write

$$i\partial^\beta[\partial_\alpha\zeta A^n_\beta]=i\partial^\beta[\partial_\alpha\zeta_H(A^n_\beta)_H]+i\partial^\beta[\partial_\alpha\zeta_L(A^n_\beta)_H]+i\partial^\beta[\partial_\alpha\zeta_H(A^n_\beta)_L]$$

Then the trilinear estimates of section 5 as well as our assumptions on the frequency localization of ζ, $\psi^{nA_0^{(0)}}$ imply that

$$\|i\partial^\beta[\partial_\alpha\zeta_H(A^n_\beta)_H]\|_N\longrightarrow 0,\ \|i\partial^\beta[\partial_\alpha\zeta_L(A^n_\beta)_H]\|_N\longrightarrow 0$$

as $n\to\infty$, and so it suffices to replace $i\partial^\beta[\partial_\alpha\zeta A^n_\beta]$ by $i\partial^\beta[\partial_\alpha\zeta_H(A^n_\beta)_L]$. Due to Corollary 5.4, Remark 5.6 and Lemma 5.7 we can pick a sequence $\Lambda_n\to\infty$ sufficiently slowly and such that if we put $I^n=\sum_k P_kQ_{<k+\Lambda_n}$, $I^{nc}=\sum_k P_kQ_{\geq k+\Lambda_n}$, then we have

$$\|i\partial^\beta\big[\partial_\alpha\zeta_HI^{nc}(A^n_\beta)_L\big]\|_N\longrightarrow 0,\ \|i\big[\partial_\alpha\zeta_HI^n\partial^\beta(A^n_\beta)_L\big]\|_N\longrightarrow 0,\ \|i[\partial^\beta\zeta_H\partial_\alpha(I^nA^n_\beta)_L]\|_N\longrightarrow 0$$

$$\|i\partial_\alpha[\partial^\beta\zeta_H(I^{nc}A^n_\beta)_L]\|_N\longrightarrow 0$$

We then replace $i\partial^\beta[\partial_\alpha\zeta_H(A^n_\beta)_L]$ by $i\partial^\beta[\partial_\alpha\zeta_H(I^nA^n_\beta)_L]$, up to asymptotically vanishing error. Then write

$$\begin{aligned} i\partial^\beta[\partial_\alpha\zeta_H(I^nA^n_\beta)_L] &= i[\partial^\beta\partial_\alpha\zeta_H(I^nA^n_\beta)_L] + i[\partial_\alpha\zeta_H\partial^\beta(I^nA^n_\beta)_L] \\ &= i\partial_\alpha[\partial^\beta\zeta_H(I^nA^n_\beta)_L] + i[\partial_\alpha\zeta_H\partial^\beta(I^nA^n_\beta)_L] \\ &\quad - i[\partial^\beta\zeta_H\partial_\alpha(I^nA^n_\beta)_L] \\ &= i\partial_\alpha[\partial^\beta\zeta_H(A^n_\beta)_L] + o_N(1) \end{aligned}$$

Proceeding similarly for the remaining terms of (9.87), it then follows that

$$\begin{aligned} &\Box\partial_\alpha\zeta + i\partial^\beta\left[\partial_\alpha\zeta A^n_\beta\right] + i\partial_\alpha\left[\partial^\beta\zeta I^c A^n_\beta\right] + i\sum_k P_k\left[\partial^\beta\partial_\alpha\zeta P_{<k-5}IA^n_\beta\right] \\ &= \partial_\alpha\left[\Box\zeta + 2i\partial^\beta\zeta_H(A^n_\beta)_L\right] + o_N(1) = \partial_\alpha\left[\Box\zeta + 2i\partial^\beta\zeta(A^n_\beta)\right] + o_N(1) \end{aligned}$$

where in the last step we have again used (a slight variation[30] of) Lemma 5.7. The **Claim** above follows from this.

(ii.4) *The remaining interactions the $\tilde{\psi}^{nab}$ of bounded or unbounded type, $\psi^{nA_0^{(0)}}$, as well as $\partial_\alpha W^{naB_1}$.* These offer nothing new: note that both the components $\tilde{\psi}^{nab}$ of unbounded type as well as the covariant linear waves W^{naB_1} have extremely small $L^\infty_{t,x}$-norm, but enjoy the same frequency localization properties as $\tilde{\psi}^{nab}$; indeed, for unbounded type $\tilde{\psi}^{nab}$, this follows by choosing the C in the interval we work on $[0, t^{nab_2} - C]$ sufficiently large. Thus any trilinear interactions involving them can be handled as in case (ii.1) in the asymptotic regime. Also, not that interactions of $\partial_\alpha W^{naB_1}$ with $\psi^{nA_0^{(0)}}$ are of schematic type

$$\nabla_{x,t}[\psi^{nA_0^{(0)}}\nabla^{-1}\mathcal{Q}_{\nu j}(\partial_\alpha W^{naB_1}, \psi^{nA_0^{(0)}})]$$

$$\nabla_{x,t}[\psi^{nA_0^{(0)}}\nabla^{-1}\mathcal{Q}_{\nu j}(\partial_\alpha W^{naB_1}, \partial_\alpha W^{naB_1})]$$

$$\nabla_{x,t}[\partial_\alpha W^{naB_1}\nabla^{-1}\mathcal{Q}_{\nu j}(\psi^{nA_0^{(0)}}, \partial_\alpha W^{naB_1})],$$

and hence can be made arbitrarily small with respect to $\|\cdot\|_N$ by choosing n large enough.

We omit the treatment of the higher order interactions between the $\tilde{\psi}^{nab}$ as this offers nothing qualitatively new. Applying the arguments from the proof of Proposition 9.12, we now conclude the proof of Proposition 9.30. □

Proposition 9.30 allows us to extend the Coulomb components $\psi^{n(<a)}_\alpha$ to the interval $[0, t^{nab_2} - C]$. But now the profiles $\tilde{\psi}^{nab}$ which were temporally bounded *with respect to* $t = 0$ become temporally unbounded with respect to the new starting time $t^{nab_2} - C$ as $n \to \infty$. Now by repeating the arguments in Section 9.6.3, we see that for those concentration profiles $\tilde{\psi}^{nab}$ for which (see the discussion in Section 9.6.3) $\limsup_{n\to\infty}|t^{nab_2} - t^{nab}| < \infty$, i.e., they concentrate at time t^{nab_2} or alternatively time $t^{nab_2} - C$, the exact same arguments as in that subsection imply that they can be approximated arbitrarily well in the L^2-sense by Coulomb components of admissible maps (but for this we have to know that the Components $\psi^{n(<a)}_\alpha$ and the associated wave maps actually extend to time $t^{nab_2} - C$). But then we have an exact analogue for Proposition 9.30 on the interval $[t^{nab_2} - C, t^{nab_3} - \tilde{C}]$. Repeating this process finitely many times, we extend $\psi^{n(<a)}_\alpha$ to $\mathbb{R}^{2+1}$, and obtain an a priori bound

$$\|\psi^{n(<a)}_\alpha\|_S < C_a$$

as well as exponential decay of the $\|P_k\psi^{n(<a)}_\alpha\|_{S[k]}$ for $k \gg \log[(\lambda^a_n)^{-1}]$.

[30] Recall that we have stronger estimates for ζ in the regime of large modulations, viz. Lemma 9.19

9.8. **Completion of the proof of Proposition 7.15 as well as of Corollary 7.16.** Both of these can be deduced by a simpler version of the proof of Proposition 9.30. For Proposition 7.15, one makes the ansatz

$$\psi_\alpha = \partial_\alpha\big(S(0-t_0)(\partial_t V, V)\big) + \epsilon_\alpha$$

and performs a bootstrap argument for $\|\epsilon_\alpha\|_{S((-\infty,0]\times\mathbb{R}^2)}$ for t_0 large enough. This is as in the proof of Proposition 9.30 where the free linear evolution of $\partial_\alpha\big(S(0-t_0)(\partial_t V, V)\big)$ replaces one of the temporally unbounded $\tilde{\phi}^{nab}$, say, while all the other components $\tilde{\phi}^{nab}$, $\psi^{nA_0^{(0)}}$, $\partial_\alpha W^{naB_1}$ vanish. If we pick t_0 large enough, all the error terms due to nonlinear self-interactions of $\partial_\alpha\big(S(t-t_0)(\partial_t V, V)\big)$ become arbitrarily small due to the reasoning in case (ii.1) of the proof of Proposition 9.30. As there, one then obtains the estimates for ϵ via the technique used in the proof of Proposition 9.12. We conclude that for given $\delta_3 > 0$, if t_0 is chosen large enough, we obtain the a priori bound

$$\|\epsilon_\alpha\|_{S((-\infty,0]\times\mathbb{R}^2)} \ll \delta_3$$

and from here the smoothness of the solution follows, see Proposition 7.3.

Next, we prove Corollary 7.16: from Proposition 7.15, we know that we can construct admissible Coulomb components of the form

$$\psi_\alpha^n := \partial_\alpha\big(S(t-t_n)(\partial_t V, V)\big) + \epsilon_\alpha$$

for $t \in (-\infty, t_n - C]$ for some large enough absolute constant C, with

$$\limsup_{n\to\infty} \|\epsilon_\alpha\|_{S((-\infty,t_n-C]\times\mathbb{R}^2)} \ll 1.$$

Now we claim that the functions $\psi_\alpha^n(t_n - 10C, \cdot)$ form a Cauchy sequence in the L_x^2-sense. To see this, note that for $n > m$

$$\psi_\alpha^n(t_n - t_m, \cdot) = \psi_\alpha^m(0, \cdot) + o_{L^2}(1)$$

as $n, m \to \infty$, whence by Proposition 7.11 one has

$$\psi_\alpha^n(t_n - 10C, \cdot) = \psi_\alpha^m(t_m - 10C, \cdot) + o_{L^2}(1)$$

But then also

$$\psi_\alpha^n(t + t_n, \cdot) = \psi_\alpha^m(t + t_m, \cdot) + o_{L^2}(1), \quad t \in (-\infty, -10C)$$

again by Proposition 7.11 , and furthermore, due to the uniform bounds

$$\limsup_{n\to\infty} \|\psi_\alpha^n\|_{S((-\infty,t_n-C]\times\mathbb{R}^2)} < M < \infty$$

for suitable $M \in \mathbb{R}$, we conclude upon denoting

$$\Psi_\alpha^\infty(t, \cdot) := \lim_n \psi_\alpha^n(t + t_n, \cdot)$$

that

$$\|\Psi_\alpha^\infty\|_{S((-\infty,-\tilde{C}]\times\mathbb{R}^2)} \le M$$

for any $\tilde{C} > 10C$, as desired.

9.9. **Step 5 of the Bahouri Gerard process; adding all atoms.** In the preceding subsection we derived a priori bounds for the wave maps evolution of the (admissible) Coulomb components

$$(w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1})e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k(w_k^{nA_0^{(0)}} + \phi_k^{n1})} + o_{L^2}(1)$$

under the assumption that either

$$\liminf_{n\to\infty} \|w^{nA_0^{(0)}}\|_{L_x^2} > 0$$

or else, applying the second stage Bahouri Gerard decomposition to the large atom ϕ^{n1}, that all the concentration profiles have energy $< E_{crit}$. We shall henceforth make this assumption. Now we continue the process by extending the data at time $t = 0$ for the Coulomb components to

$$(w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1} + w_\alpha^{nA_0^{(1)}})e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k(w_k^{nA_0^{(0)}} + \phi_k^{n1} + w_k^{nA_0^{(1)}})} + o_{L^2}(1),$$

where we recall that the error term $o_{L^2}(1)$ is necessary in order to ensure that the data correspond to exact Coulomb components of an admissible map. Denote the wave maps evolution of

$$(w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1} + w_\alpha^{nA_0^{(1)}})e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k(w_k^{nA_0^{(0)}}+\phi_k^{n1}+w_k^{nA_0^{(1)}})} + o_{L^2}(1),$$

which is defined at time $t=0$, by the same symbol. We state the result:

Proposition 9.34. *Under the preceding assumptions, the evolution of the preceding Coulomb components exists globally in time. For n large enough, we have an a priori bound*

$$\|(w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1} + w_\alpha^{nA_0^{(1)}})e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k(w_k^{nA_0^{(0)}}+\phi_k^{n1}+w_k^{nA_0^{(1)}})} + o_{L^2}(1)\|_{S(\mathbb{R}^{2+1})} < \infty$$

The bound here depends on E_{crit} as well as the a priori bounds for the evolution of the concentration profiles extracted by adding ϕ^{n1}. Furthermore, we have the same bounds as in Proposition 9.11 (applied to the union of all J_j), where the implied constants depend on $E_{\rm crit}$ as well as the a priori bounds for the evolution of the concentration profiles extracted by adding ϕ^{n1}.

The proof of this is a precise replica of the one given in Step 3. The difference consists in the fact that in the decomposition (see Step 2)

$$w^{nA_0^{(1)}} = \sum_j \phi^{na_j^k} + w^{nA^{(1)}}$$

we now need to ensure that $\|w^{nA^{(1)}}\|_{\dot{B}^0_{2,\infty}}$ is small enough depending on both E_{crit} as well as the a priori bounds for the concentration profiles from Step 4.

Next, one extends the data at time $t=0$ to

$$(w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1} + w_\alpha^{nA_0^{(1)}} + \phi_\alpha^{n2})e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k(w_k^{nA_0^{(0)}}+\phi_k^{n1}+w_k^{nA_0^{(1)}}+\phi_k^{n2})} + o_{L^2}(1)$$

Repeating the procedure of Step 4 but with magnetic potential defined in terms of the ψ-evolution of

$$(w_\alpha^{nA_0^{(0)}} + \phi_\alpha^{n1} + w_\alpha^{nA_0^{(1)}})e^{-i\sum_{k=1,2}\triangle^{-1}\partial_k(w_k^{nA_0^{(0)}}+\phi_k^{n1}+w_k^{nA_0^{(1)}})} + o_{L^2}(1),$$

one again derives the same types of bounds as in Proposition 9.30 and the process continues A_0 many times, as we recall from the discussion at the beginning of Step 2. We have finally arrived at the following grand conclusion to this section.

Theorem 9.35. *Let ψ^n be a sequence of gauged derivative components of admissible wave maps $\mathbf{u}^n : [-T_0^n, T_1^n]\times\mathbb{R}^2 \to \mathbb{H}^2$. The hypothesis*

$$\lim_{n\to\infty}\|\psi^n\|_{S([-T_0^n,T_1^n]\times\mathbb{R}^2)} = \infty,\ \lim_{n\to\infty}\|\psi_n\|_E = E_{crit} \tag{9.88}$$

implies that two possible cases occur: up to rescaling and spatial translations, either we have

$$\psi_\alpha^n(0,\cdot) = V_\alpha + o_L^2(1)$$

for some fixed L^2-profile V_α, or else we have for some sequence $t^n\to\infty$ (or $t^n\to-\infty$) and suitable $(\partial_t V, V)\in L^2\times\dot{H}^1$,

$$\psi_\alpha^n(0,\cdot) = \partial_\alpha\big(S(0-t^n)[\partial_t V, V]\big) + o_L^2(1),$$

where $S(t)$ refers to the standard free wave propagator. In the former case

$$\sum_{\alpha=0}^2 \|V_\alpha\|_{L^2}^2 = E_{crit}\,,$$

while in the latter case, one has

$$\sum_{\alpha=0}^2 \|\partial_\alpha V\|_{L^2}^2 = E_{crit} \tag{9.89}$$

Note that due to Lemma 7.10, in the first case, there exist $T_0 > 0$, $T_1 > 0$ with the property that

$$\sup_n \|\psi^n\|_{S([-T_0,T_1]\times\mathbb{R}^2)} < \infty,$$

and we can then define

$$\lim_{n\to\infty} \psi^n_\alpha(t,x) =: \Psi^\infty_\alpha(t,x) \tag{9.90}$$

where the limit is in the sense of $L^\infty_{\text{loc}}([-T_0,T_1];L^2(\mathbb{R}^2))$. Similarly, in the second case, due to Corollary 7.16, we have the corresponding statements on some semi-infinite interval $I = (-\infty,T_0)$ respectively (T_0,∞). We call the maximal such open interval $(-T_0,T_1)$ (respectively $(-\infty,T_0)$ or (T_0,∞)) the *lifespan* of the asymptotic object $\Psi^\infty_\alpha(t,x)$. Finally, in order to apply the Kenig-Merle type argument, we need the following essential *compactness property*:

Corollary 9.36. *There exist continuous functions $\bar{x} : I \to \mathbb{R}^2$ and $\lambda : I \to \mathbb{R}^+$ so that the family of functions $\{\lambda(t)^{-1}\,\Psi^\infty_\alpha(t,(\cdot - \bar{x}(t))\lambda(t)^{-1})\}_{t\in I} \subset L^2_x$ is pre-compact.*

Proof. We may assume that

$$\sup_{0<T_2<T_1} \|\Psi^\infty_\alpha\|_{S([0,T_2)\times\mathbb{R}^2)} = \infty \tag{9.91}$$

see Lemma 7.17. The proof follows [13], [14] and amounts to an argument by contradiction. More precisely, we begin by showing that one can find functions $\lambda(t)$, $\bar{x}(t)$ not necessarily continuous with the desired compactness property. Suppose this fails. Then there exists $\varepsilon > 0$ and a sequence of times $\{t_n\} \subset I$ so that

$$\inf_{\lambda>0,\bar{x}\in\mathbb{R}^2} \|\lambda^{-1}\,\Psi^\infty_\alpha(t_n,(\cdot-\bar{x})\lambda^{-1}) - \Psi^\infty_\alpha(t_m,\cdot)\|_2 \geq \varepsilon \tag{9.92}$$

for any $n \neq m$. Necessarily $t_n \to T_1$. Now apply Theorem 9.35 to the sequence $\{\Psi^\infty_\alpha(t_n,\cdot)\}_{n=1}^\infty$, which satisfies (9.88), but on a shifted time-interval. Note that $\Psi^\infty_\alpha(t_n,\cdot)$ are not admissible in the sense that they are not necessarily given as the Coulomb derivative components of admissible wave maps. However, by approximation by the original sequence ψ^n_α (up to symmetries) one concludes that either for some $V_\alpha \in L^2(\mathbb{R}^2)$,

$$\Psi^\infty_\alpha(t_n,x) = \lambda_n^{-1} V_\alpha((x-x_n)\lambda_n^{-1}) + o^2_L(1) \tag{9.93}$$

for some sequence λ_n, x_n, or that for some $s^n \to \infty$ or $s_n \to -\infty$,

$$\Psi^\infty_\alpha(t_n,x) = \lambda_n^{-1}\partial_\alpha\big(S(-s^n)[\partial_t V,V]\big)((x-x_n)\lambda_n^{-1}) + o^2_L(1), \tag{9.94}$$

where V is as in (9.89). Clearly, (9.93) contradicts (9.92). For (9.94), we first show that $\{s_n\}_{n=1}^\infty$ has to be bounded. Assume that $s_n \to -\infty$. Then Proposition 7.15 implies for large n that Ψ^∞_α exists on $[0,\infty)\times\mathbb{R}^2$ and

$$\|\Psi^\infty_\alpha\|_{S([0,\infty)\times\mathbb{R}^2)} < \infty$$

which contradicts our assumption (9.91). If on the other hand $s_n \to \infty$, then this implies by the same proposition that

$$\sup_n \|\Psi^\infty_\alpha\|_{S((-\infty,t_n])\times\mathbb{R}^2)} < \infty$$

This again contradicts our assumption (9.91) and we are done. As in [13] one proves by approximation that λ and $\bar{x}$ can be taken to be continuous. □

10. The proof of the main theorem

For the purposes of this section, it is sometimes preferable pass to the *extrinsic point of view.* Specifically, let $\mathcal{S}$ be a compact Riemann surface of the hyperbolic type, i.e., it is uniformized by the hyperbolic plane. Given a covering map $\pi : \mathbb{H}^2 \to \mathcal{S}$, we obtain a Riemannian structure on $\mathcal{S}$ which makes π a local isometry. By Nash's theorem, we may isometrically embed $\mathcal{S} \hookrightarrow \mathbb{R}^N$ into an ambient Euclidean space. Now denote the compositions

$$U^n := \pi \circ \mathbf{u}^n \;:\; I \times \mathbb{R}^2 \to \mathcal{S}$$

defined on $I\times\mathbb{R}^2$, see the above discussion. We can express these maps in terms of the ambient coordinates. Our first task is to identify an actual *map* U from $I\times\mathbb{R}^2$ into $\mathcal{S}\hookrightarrow\mathbb{R}^N$ which in some sense corresponds to the limiting object $\Psi_\alpha^\infty(t,x)$. The fact that this can be done follows again from the compactness property of the $\Psi_\alpha^\infty(t,x)$. We have the following

Proposition 10.1. *Under the above assumptions, there exists a subsequence of $\{U^n,\,\phi^n,\,\psi^n\}$ which we denote in the same fashion as well as a function $U(t,\cdot)\in C^0(I;\dot H^1)\cap C^1(I;L^2)$, such that*

$$\lim_{n\to\infty}U^n(t,x)=:U(t,x),\qquad \lim_{n\to\infty}\nabla_{x,t}U^n(t,x)=\nabla_{x,t}U(t,x)$$

where the former limit is the a.e. pointwise sense and the latter limit is in the L^2_x-sense on fixed time intervals. The map U is a weak wave map (in the distributional sense). Also, the second limit is uniform on compact intervals $J\subset I$. Finally, the family of functions

$$\{\nabla_{x,t}U(t,\,\cdot)\}_{t\in I}\subset L^2_x$$

is compact up to rescaling and translational symmetries (which may depend on time).

Proof. We may assume that for times $t\in I$ we have

$$\psi_\alpha^n(t,\,\cdot)=\Psi_\alpha^\infty(t,\,\cdot)+o_{L^2}(1)$$

But then it follows that for each such $t\in I$, there is a subsequence (depending on t) such that also $\phi_\alpha^n(t,\,\cdot)$ converges in the L^2-sense. To see this, note that

$$\phi_\alpha^n(t,\cdot)=(\Psi_\alpha^\infty(t,\,\cdot)+o_{L^2}(1))e^{i\Delta^{-1}\sum_{j=1}^2\phi_j^n}$$

inherits both the physical L^2-localization coming from $\Psi_\alpha^\infty(t,\cdot)$ as well as the Fourier localization of this profile[31] whence it is compact and a subsequence converges as claimed. Picking a dense subset of times $\{t_i\}_{i=1}^\infty\subset I$ and using the Cantor diagonal argument, one obtains a subsequence which we again denote by ψ^n etc. such that $\phi^n(t_i,\,\cdot)$ converges for each i in the L^2 sense. By Corollary 9.36, it then follows that $\phi^n(t,\,\cdot)$ converges in the L^2 sense, uniformly on compact sub-intervals of I. In particular, the limit ϕ^∞ satisfies $\phi^\infty\in C^0(I;L^2(\mathbb{R}^2))$. We now use this to infer the existence of $U(t,x)$. First, introduce a global frame $\{e_{1,2}\}$ on the pull-back bundle of $T\mathcal{S}$ under the wave map U^n by projecting down the standard frame $\{\mathbf{e_1},\,\mathbf{e_2}\}$, i.e., $e_j(t,x):=\pi_*(\mathbf{e}_j)(\mathbf{u}^n(t,x))$. Thus

$$\partial_\alpha U^n(t,x)=\sum_{k=1,2}e_k^n(t,x)\phi_\alpha^{kn}(t,x)\tag{10.1}$$

Fix some $I'\subset I$ which is compactly contained in I. We now use that the pull-back frame is bounded. By the preceding, given $\varepsilon>0$ there exists R so large that

$$\limsup_{n\to\infty}\big\|\nabla_{t,x}U^n\chi_{[|x|>R]}\big\|_{L^\infty(I';L^2(|x|>R))}<\varepsilon$$

On the other hand, it is clear that

$$\limsup_{n\to\infty}\big\|\nabla_{t,x}U^n\|_{L^\infty(I';L^2)}<\infty$$

By Rellich's theorem we now conclude that up to passing to a subsequence, $\partial_\alpha U^n\rightharpoonup X_\alpha$ in $L^\infty(I';L^2)$ (in the weak-* sense), as well as $U^n\to U$ in $L^\infty_{\rm loc}(I;L^2)$ strongly. Necessarily then $U\in L^\infty(I',\dot H^1(\mathbb{R}^2))$, see (10.1) as well as $X_\alpha=\partial_\alpha U$. One immediately obtains the stronger statement that $U\in C^0(I,L^2)$ by integrating in time. One in fact has stronger convergence: first note that

$$\partial_\alpha e_k^n(t,x)=d(\pi_*)(d\mathbf{e}_j)(\mathbf{u}^n(t,x))\partial_\alpha\mathbf{u}^n(t,x)$$

which implies that $\{e_k^n\}_{n=1}^\infty$ is compact in $\dot H^1(\mathbb{R}^2)$. It now follows from (10.1) and Rellich's theorem as before that up to a subsequence one has

$$\partial_\alpha U^n(t_i,\cdot)\to\partial_\alpha U(t_i,\cdot)$$

[31] This follows as usual from a Littlewood-Paley trichotomy argument and the energy conservation of the ϕ^n.

strongly in L^2. By compactness, one therefore also has strongly in L^2

$$\partial_\alpha U^n(t,\cdot) \to \partial_\alpha U(t,\cdot)$$

uniformly on compact subsets of I. This implies all the convergence and regularity statements of the proposition. The fact that U is a weak wave map follows from this, as well as from [9]. □

Note that we do not claim that we have uniqueness for the limiting object U, and indeed we only have a well-posedness theory at the level of the ψ_α. Thus we cannot purely work at the level of wave maps with compact target $\mathcal{S}$. Nevertheless, the latter will play an important role when ruling out certain pathological behaviors, or also to formulate the conservation laws.

For example, we have the following

Corollary 10.2. *Let U be the weak wave map as in Proposition 10.1. Then one has the following conservation laws: with $|\cdot|^2 = \langle\cdot,\cdot\rangle$ being the metric on $\mathcal{S}$,*

- $\frac{d}{dt}\sum_{\alpha=0}^2 \int_{\mathbb{R}^2} |\partial_\alpha U(t,x)|^2\,dx = 0$
- $\frac{d}{dt}\int_{\mathbb{R}^2} \langle \partial_t U(t,x), \partial_i U(t,x)\rangle\,dx = 0 \quad i = 1,2$
- $\frac{d}{dt}\sum_{i=1}^2 \int_{\mathbb{R}^2} x_i\phi(x/R)\langle \partial_t U(t,x), \partial_i U(t,x)\rangle\,dx = -\int_{\mathbb{R}^2} |\partial_t U(t,x)|^2\,dx + O(r(R))$
- $\frac{d}{dt}\sum_{\alpha=0}^2 \int_{\mathbb{R}^2} x_i\phi(x/R)\frac{1}{2}|\partial_\alpha U(t,x)|^2\,dx = -\int_{\mathbb{R}^2}\langle \partial_i U, \partial_t U\rangle\,dx + O(r(R))$

where ϕ is a fixed bump function which is equal to one on $|x| \le 1$ and

$$r(R) := \int_{[|x|\ge R]} \sum_{\alpha=0}^2 |\partial_\alpha U(t,x)|^2\,dx$$

Proof. These are standard calculations for smooth wave maps. By Proposition 10.1 one can then pass to the limit. □

Note that one could alternatively express these in terms of Ψ_α^∞. We will now closely follow the arguments in [13].

10.0.1. *Some preliminary properties of the limiting profiles.* We begin with the following consequence of finite propagation speed. Let $I^+ := I \cap [0,\infty)$ where I is the life span of Ψ_α^∞.

Lemma 10.3. *Let $M > 0$ have the property that*

$$\int_{|x|>\frac{M}{2}} \sum_{\alpha=0}^2 |\Psi_\alpha^\infty(0,x)|^2\,dx < \varepsilon \tag{10.2}$$

Then

$$\int_{|x|>2M+t} \sum_{\alpha=0}^2 |\Psi_\alpha^\infty(t,x)|^2\,dx < C\varepsilon \tag{10.3}$$

for all $t \in I^+$. Here C is an absolute constant.

Proof. By definition, there exist $\mathbf{u}^n = (\mathbf{x}^n, \mathbf{y}^n) : I^+ \to \mathbb{H}^2$ which are admissible wave maps such that (9.90) holds. Now define

$$(\mathbf{x}_2^n, \mathbf{y}_2^n)(0,\cdot) := \Big(\chi_{[|x|>M]}\frac{\mathbf{x}^n(0,\cdot) - \mathbf{x}_0^n}{\mathbf{y}_0^n},\ e^{\chi_{[|x|>\frac{M}{2}]}\log[\frac{\mathbf{y}^n}{\mathbf{y}_0^n}(0,\cdot)]}\Big)$$

where $\chi_{[|x|>M]}$ is a smooth cutoff to the set $\{|x| > M\}$ which equals one on $\{|x| > \frac{5}{4}M\}$, say, and

$$\mathbf{x}_0^n := \fint_{[M<|x|<\frac{5}{4}M]} \mathbf{x}^n(x)\,dx_1dx_2, \qquad \mathbf{y}_0^n := \exp\Big(\fint_{[\frac{M}{2}<|x|<\frac{5}{8}M]} \log \mathbf{y}^n(x)\,dx_1dx_2\Big)$$

The construction here is such that $\mathbf{y}_2^n = \frac{\mathbf{y}^n}{\mathbf{y}_0^n}$ on the set $\{\nabla\chi_{[|x|>M]} \neq 0\}$. Let $\tilde{\mathbf{u}}^n$ be the wave map evolution of the data

$$\Big((\mathbf{x}_2^n, \mathbf{y}_2^n)(0,\cdot), \big(\frac{\partial_t \mathbf{x}^n(0,\cdot)}{\mathbf{y}_0^n}, \frac{\partial_t \mathbf{y}^n(0,\cdot)}{\mathbf{y}_0^n}\big)\Big)$$

By construction, the energy of $\tilde{\mathbf{u}}^n$ does not exceed $C\varepsilon$. This requires the use of Poincaré's inequality as in the proof of Lemma 7.22. One now concludes by means of finite propagation speed for classical wave maps, and by passing to the limit $n \to \infty$. □

Next, one has the following lower bound on $\lambda(t)$ in Corollary 9.36.

Lemma 10.4. *Assume I^+ is finite. After rescaling, we may assume that $I^+ = [0, 1)$. There exists a constant $C_0(K)$ depending on the compact set K in Corollary 9.36, such that*

$$0 < \frac{C_0(K)}{1-t} \le \lambda(t) \tag{10.4}$$

for all $0 \le t < 1$.

Proof. Take any sequence $t_j \to 1$. Consider the limiting profile $\{\tilde{\Psi}^\infty_{\alpha,j}\}_{\alpha=0}^2$ with data $\lambda(t_j)^{-1}\Psi^\infty_\alpha(t_j, (\cdot - \bar{x}(t_j))\lambda(t_j)^{-1})\}_{\alpha=0}^2$. By the well-posedness theory of the limiting profiles in Section 7.2, one infers that the $\{\tilde{\Psi}^\infty_{\alpha,j}\}_{\alpha=0}^2$ have a fixed life span independent of j which depends only on the compact set K. By the uniqueness property of the solutions and rescaling, $(1 - t_j)\lambda(t_j) \ge C_0(K)$ as claimed. □

Next, combining this with Lemma 10.3 one concludes the following support property of the Ψ^∞_α with finite life span.

Lemma 10.5. *Let Ψ^∞_α be as in the previous lemma. Then there exists $x_0 \in \mathbb{R}^2$ such that*

$$\operatorname{supp}(\Psi^\infty_\alpha(t, \cdot)) \subset B(x_0, 1 - t)$$

for all $0 \le t < 1$, $\alpha = 0, 1, 2$.

Proof. This follows the exact same reasoning as in Lemma 4.8 of [13]. One uses Lemma 10.3 instead of their Lemma 2.17 and Lemma 10.4 instead of their Lemma 4.7. □

Next, we turn to the vanishing moment condition of Propositions 4.10 and 4.11 in [13].

Proposition 10.6. *Let Ψ^∞_α be as above and assume that I^+ is finite. Then for $i = 1, 2$,*

$$\int_{\mathbb{R}^2} \langle \partial_i U, \partial_t U\rangle \, dx = \operatorname{Re} \int_{\mathbb{R}^2} \Psi^\infty_i \bar{\Psi}^\infty_0 \, dx = 0$$

for all times in I^+.

Proof. Assume that

$$\operatorname{Re} \int_{\mathbb{R}^2} \Psi^\infty_1 \bar{\Psi}^\infty_0 \, dx > \gamma > 0$$

This implies that the approximating sequence $\mathbf{u}^n$ satisfies

$$\int_{\mathbb{R}^2} \langle \partial_1 \mathbf{u}^n, \partial_t \mathbf{u}^n\rangle \, dx > \gamma > 0$$

for large n. Following [13] we apply a Lorentz transformation

$$L_d(t, x) := \Big(\frac{t - dx_1}{\sqrt{1-d^2}}, \frac{x_1 - dt}{\sqrt{1-d^2}}, x_2\Big)$$

to the u^n. Note that for any $\varepsilon > 0$ one has from Lemma 10.5 that

$$\sum_{\alpha=0}^{2} \int_{|x| \ge 1-t} |\partial_\alpha \mathbf{u}^n(t, x)|^2 \, dx < \varepsilon$$

for all $t \in I^+ = [0, 1)$ and sufficiently large n. Then the argument in [13] implies that there exists d small with the property that

$$\limsup_{n\to\infty} E(\mathbf{u}^n \circ L_d) < E_{crit}$$

By our induction hypothesis, $\|\psi^{n,d}\|_{S(I^+\times\mathbb{R}^2)} < M < \infty$ for all sufficiently large n. Here $\psi^{n,d}$ are the Coulomb components of the admissible wave maps $\mathbf{u}^n \circ L_d$. Note that the Coulomb components $\psi^{n,d}$ do not obey a simple transformation law relative to the Coulomb components ψ^n of $\mathbf{u}^n$. Nonetheless, it is possible to conclude from this that

$$\limsup_{n\to\infty} \|\psi^n\|_{S(I^+\times\mathbb{R}^2)} < M_1 < \infty$$

via Remark 7.8 which gives us the desired contradiction. Thus, we need to prove that for each $k_1 > k_2$

$$\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2} P_{k_1,\kappa_1}\psi^n P_{k_2,\kappa_2}\psi^n = f_{k_1,k_2} + g_{k_1,k_2} \tag{10.5}$$

where m_0 is a large depending on E_C, where we have the bounds (7.20) for f_{k_1,k_2} and g_{k_1,k_2}. Furthermore, we need to show that

$$P_k Q_{>k}\psi^n = h_k + i_k$$

with the bounds stated in Remark 7.8. We establish this for the bilinear expression, the corresponding computations for $P_k Q_{>k}\psi^n$ being similar. First, we claim the following bound for $\psi^{n,d}$:

$$\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\psi^{n,d}P_{k_2,\kappa_2}\psi^{n,d}\|_{L^2_{t,x}}^2 < \Lambda' \tag{10.6}$$

This, however, is immediate from the angular separation and (2.30) with a constant Λ' which depends on M and E_{crit}. In fact, we need something slightly stronger due to the usual tail issues:

$$\sup_y \sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\psi^{n,d}\tau_y P_{k_2,\kappa_2}\psi^{n,d}\|_{L^2_{t,x}}^2 < \Lambda' \tag{10.7}$$

where τ_y is a translation by $y\in\mathbb{R}^2$. Next, we claim the following estimate:

$$\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\phi^{n,d}P_{k_2,\kappa_2}\phi^{n,d}\|_{L^2_{t,x}}^2 < \Lambda' \tag{10.8}$$

where $\phi^{n,d}$ are the derivative components of the $\mathbf{u}^n\circ L_d$. This is the same as

$$\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2 < \Lambda'$$

where we wrote the phase $-i\partial^{-1}\phi^{n,d} = -i\mathrm{Re}\sum_{j=1}^2(-\Delta)^{-1}\partial_j\phi^{n,d}$ schematically. This follows from (10.6) and the Strichartz estimate

$$\Big(\sum_{k\in\mathbb{Z}} 2^{-\frac{3}{2}k}\sup_{j\geq 10}\sum_{c\in\mathcal{D}_{k,j}} 2^{-(1-2\varepsilon)j}\|P_c\phi^{n,d}\|_{L^4_tL^\infty_x}^2\Big)^{\frac12} \lesssim M \tag{10.9}$$

To prove (10.9), one uses the corresponding bound on $\psi^{n,d}$ (which is part of the S-norm), energy conservation, and a simple Littlewood-Paley trichotomy. To prove (10.8), one argues as follows. Split

$$\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2$$

$$\lesssim \sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}P_{<k_1-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}P_{<k_2-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2 \tag{10.10}$$

$$+\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}P_{<k_1-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}P_{>k_2-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2 \tag{10.11}$$

$$+\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}P_{>k_1-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}P_{<k_2-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2 \tag{10.12}$$

$$+\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}P_{>k_1-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}P_{>k_2-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2 \tag{10.13}$$

In (10.10) one reduces matters to (10.7) by placing the exponential in $L^\infty_t L^\infty_x$. Next, to bound (10.11) one notes that

$$\big\|P_{>k_2-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big\|_{L^4_t L^\infty_x}\lesssim 2^{-\frac{k_2}{4}}M$$

where the implicit constant depends on E_{crit}. Therefore,

$$\begin{aligned}
&\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}\big(\psi^{n,d}P_{<k_1-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\cdot P_{k_2,\kappa_2}\big(\psi^{n,d}P_{>k_2-m_0}e^{-i\partial^{-1}\phi^{n,d}}\big)\|_{L^2_{t,x}}^2\\
&\lesssim \sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1}\psi^{n,d}\|_{L^\infty_t L^2_x}^2\Big(\sum_{\ell>k_2}\sum_{\substack{c,c'\in\mathcal{D}_{\ell,k_2-\ell}\\ \mathrm{dist}(c,c')\lesssim 2^{k_2}}}\|P_cP_\ell\psi^{n,d}\|_{L^4_tL^\infty_x}\|P_{c'}P_{\ell+O(m_0)}(e^{-i\partial^{-1}\phi^{n,d}}-1)\|_{L^4_tL^\infty_x}\Big)^2\\
&\lesssim M^6
\end{aligned}$$

using the Strichartz estimate from above. The remaining terms are the same. This concludes the proof of (10.8). By the same logic, one also obtains

$$\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0}}} 2^{-k_2}\|P_{k_1,\kappa_1}Q_{\le k_1+C_2}\phi^{n,d}P_{k_2,\kappa_2}Q_{\le k_2+C_2}\phi^{n,d}\|_{L^2_{t,x}}^2<\Lambda'$$

where C_2 is a large constant depending only on the energy which will be determined later. This then implies the following version *without* the Lorentz transforms

$$\sum_{k_1>k_2}\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0'}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0'}}} 2^{-k_2}\|P_{k_1,\kappa_1}Q_{\le k_1+C_2}\phi^{n}P_{k_2,\kappa_2}Q_{\le k_2+C_2}\phi^{n}\|_{L^2_{t,x}}^2<\Lambda'$$

provided d is chosen small enough, but depending only on E_{crit} (so that m_0' is close to m_0). Finally we claim that

$$P_{k_1}Q_{>k_1+C_2}\phi^nP_{k_2}\phi^n \tag{10.14}$$

$$P_{k_1}\phi^nP_{k_2}Q_{>k_2+C_2}\phi^n \tag{10.15}$$

can both be included in g_{k_1,k_2}. To see this, one first expands

$$P_{k_1}Q_{>k_1+C_2}\phi^n = P_{k_1}Q_{>k_1+C_2}\left[\psi^n e^{-i\partial^{-1}\phi^n}\right]$$

$$= \sum_{\ell>k_1+C_2-10} P_{k_1}Q_{>k_1+C_2}\left[P_\ell\psi^n P_\ell e^{-i\partial^{-1}\phi^n}\right] \tag{10.16}$$

$$+ \sum_{k_1<\ell\leq k_1+C_2-10} P_{k_1}Q_{>k_1+C_2}\left[P_\ell\psi^n P_\ell e^{-i\partial^{-1}\phi^n}\right] \tag{10.17}$$

$$+ P_{k_1}Q_{>k_1+C_2}\left[P_{<k_1-5}\psi^n P_{k_1} e^{-i\partial^{-1}\phi^n}\right] \tag{10.18}$$

$$+ P_{k_1}Q_{>k_1+C_2}\left[P_{k_1}\psi^n P_{<k_1-5} e^{-i\partial^{-1}\phi^n}\right] \tag{10.19}$$

and then inserts these decompositions into (10.14). For (10.16) one places $P_{k_2}\phi^n$ into $L_t^4L_x^\infty$, and its contribution to

$$P_{k_1}Q_{>k_1+C_2}\phi^n$$

into $L_t^4L_x^2$ followed by an application of (10.9) with caps of size 2^{k_1}; more precisely, $P_\ell e^{-i\partial^{-1}\phi^n}$ goes into $L_t^4L_x^\infty$ as before, and $P_\ell\psi^n$ gets placed into $L_t^\infty L_x^2$ (see Lemma 2.18 for the issue of square-summing the $L_t^\infty L_x^2$-norm of ψ^n over caps of size 2^{k_1}). Note that one gains a smallness factor of the form $2^{-\frac{C_2}{10}}$ due to the improved Strichartz bounds. Next, we consider (10.19) and the remaining terms (10.17) and (10.18) will follow similar arguments. Now we decompose further:

$$P_{k_1}Q_{>k_1+C_2}\left[P_{k_1}\psi^n P_{<k_1-5} e^{-i\partial^{-1}\phi^n}\right]$$

$$= P_{k_1}Q_{>k_1+C_2}\left[Q_{>k_1+C_2-10}P_{k_1}\psi^n P_{<k_1-5} e^{-i\partial^{-1}\phi^n}\right] \tag{10.20}$$

$$+ P_{k_1}Q_{>k_1+C_2}\left[Q_{\leq k_1+C_2-10}P_{k_1}\psi^n P_{<k_1-5}Q_{>k_1+C_2-10} e^{-i\partial^{-1}\phi^n}\right] \tag{10.21}$$

For the contribution of (10.20) to (10.14) one estimates

$$\|P_{k_1}Q_{>k_1+C_2}\left[Q_{>k_1+C_2-10}P_{k_1}\psi^n P_{<k_1-5} e^{-i\partial^{-1}\phi^n}\right]\, P_{k_2}\phi^n\|_{L^2_{t,x}}$$

$$\lesssim 2^{k_2}\|Q_{>k_1+C_2-10}P_{k_1}\psi^n\|_{L^2_{t,x}}\|P_{k_2}\phi^n\|_{L_t^\infty L_x^2} \lesssim 2^{k_2}2^{-\frac{k_1+C_2}{2}}\|P_{k_1}\psi^n\|_{S[k_1]}\|P_{k_2}\phi^n\|_{L_t^\infty L_x^2}$$

which is sufficient since it gains the smallness $2^{-\frac{C_2}{2}}$. Finally, we use (1.6) for the case when we substitute (10.21) for $P_{k_1}Q_{>k_1+C_2}\phi^n$; one can then write

$$P_{k_1}Q_{>k_1+C_2}\phi^n P_{k_2}\phi^n$$

$$= P_{k_1}Q_{>k_1+C_2}\left[Q_{\leq k_1+C_2-10}P_{k_1}\psi^n\partial_t^{-1}P_{<k_1-5}Q_{>k_1+C_2-10}((\phi^n+\nabla^{-1}(\phi^n\phi^n))e^{-i\partial^{-1}\phi^n})\right]P_{k_2}\phi^n$$

where we have written (1.6) schematically in the form

$$\partial_t\partial^{-1}\phi^n = \phi^n + \nabla^{-1}(\phi^n\phi^n)$$

The contribution of ϕ^n is easy, it is placed again in $L_t^4L_x^\infty$ (of course after applying the usual trichotomy to $\phi^n e^{-i\partial^{-1}\phi^n}$). On the other hand, due to the determinant structure of $\nabla^{-1}(\phi^n\phi^n)$ we have

$$\nabla^{-1}(\phi^n\phi^n) = \nabla^{-1}(\psi^n\psi^n)$$

By using a further Hodge decomposition of the inputs on the right, we have for each $k \in \mathbb{Z}$

$$\|P_k\nabla^{-1}(\psi^n\psi^n)\|_{L_t^2\dot{H}^{\frac{1}{2}}} \lesssim \|\psi^n\|_S^2,$$

and from here we get

$$\|\partial_t^{-1}P_{<k_1-5}Q_{>k_1+C_2-10}(\nabla^{-1}(\psi^n\psi^n)e^{-i\partial^{-1}\phi^n})\|_{L_t^2L_x^\infty} \ll 2^{-\frac{k_1}{2}}\|\psi\|_S^2 \tag{10.22}$$

and from here we get

$$\|P_{k_1}Q_{>k_1+C_2}\left[Q_{\leq k_1+C_2-10}P_{k_1}\psi^n\partial_t^{-1}P_{<k_1-5}Q_{>k_1+C_2-10}(\nabla^{-1}(\phi^n\phi^n)e^{-i\partial^{-1}\phi^n})\right]P_{k_2}\phi^n\|_{L^2_{t,x}}$$

$$\ll 2^{k_2-\frac{k_1}{2}}\|P_{k_2}\phi^n\|_{L_t^\infty L_x^2}\|\psi\|_S^2$$

This concludes the proof that (10.14) may be included into g_{k_1,k_2}. For (10.15) one argues similarly. By following the same Littlewood-Paley trichotomies, one is eventually lead to the most difficult case

$$\begin{aligned}&P_{k_1}\phi^n P_{k_2}Q_{>k_2+C_2}\phi^n\\&=P_{k_1}\phi^n P_{k_2}Q_{>k_2+C_2}\big[Q_{\leq k_2+C_2-10}P_{k_2}\psi^n\partial_t^{-1}P_{<k_2-5}Q_{>k_2+C_2-10}((\phi^n+\nabla^{-1}(\phi^n\phi^n))e^{-i\partial^{-1}\phi^n})\big]\end{aligned}$$

where we again used the curl equation (1.6). The ϕ^n term is again easier, whereas for the nonlinear term we again use

$$\nabla^{-1}(\phi^n\phi^n)=\nabla^{-1}(\psi^n\psi^n)$$

Then as before we use (10.22), in order to infer that

$$\begin{aligned}&\|P_{k_1}\phi^n P_{k_2}Q_{>k_2+C_2}\big[Q_{\leq k_2+C_2-10}P_{k_2}\psi^n\partial_t^{-1}P_{<k_2-5}Q_{>k_2+C_2-10}(\nabla^{-1}(\phi^n\phi^n))e^{-i\partial^{-1}\phi^n})\big]\|_{L^2_{t,x}}\\&\lesssim\|P_{k_1}\phi^n\|_{L^\infty_t L^2_x}\|Q_{\leq k_2+C_2-10}P_{k_2}\psi^n\|_{L^\infty_{t,x}}\|\partial_t^{-1}P_{<k_2-5}Q_{>k_2+C_2-10}(\nabla^{-1}(\phi^n\phi^n)e^{-i\partial^{-1}\phi^n})\|_{L^2_tL^\infty_x}\\&\ll 2^{\frac{k_2}{2}}\|\psi\|_S^2\|P_{k_1}\phi^n\|_{L^\infty_t L^2_x},\end{aligned}$$

which justifies us in including it into g_{k_1,k_2}. In conclusion, we have now shown that we can write

$$\sum_{\substack{\kappa_{1,2}\in\mathcal{C}_{m_0'}\\ \mathrm{dist}(\kappa_1,\kappa_2)\gtrsim 2^{-m_0'}}}P_{k_1,\kappa_1}\phi^n P_{k_2,\kappa_2}\phi^n=\tilde f_{k_1,k_2}+\tilde g_{k_1,k_2}\tag{10.23}$$

with bounds as in (7.20). The goal is now to deduce (10.5) from this estimate. For this purpose, fix $k_1>k_2+C_1$ and caps $\kappa_1,\kappa_2\in\mathcal{C}_{m_0'}$ as above. We now describe how to break up

$$P_{k_1,\kappa_1}\psi^n\cdot P_{k_2,\kappa_2}\psi^n=P_{k_1,\kappa_1}(\phi^n e^{-i\partial^{-1}\phi^n})\cdot P_{k_2,\kappa_2}(\phi^n e^{-i\partial^{-1}\phi^n})$$

into various pieces which then constitute f_{k_1,k_2} and g_{k_1,k_2}, respectively when summed over the caps. First, write

$$\begin{aligned}&P_{k_1,\kappa_1}(\phi^n e^{-i\partial^{-1}\phi^n})\cdot P_{k_2,\kappa_2}(\phi^n e^{-i\partial^{-1}\phi^n})\\&=\sum_{i_1,i_2=1}^{3}P_{k_1,\kappa_1}(\phi^n A_{i_1}e^{-i\partial^{-1}\phi^n})\cdot P_{k_2,\kappa_2}(\phi^n B_{i_2}e^{-i\partial^{-1}\phi^n})\end{aligned}\tag{10.24}$$

where

$$A_1=P_{<k_1-m_0'-10},\quad A_2=P_{k_1-m_0'-10\leq\cdot<k_1+C_2},\quad A_3=P_{\geq k_1+C_2}$$

and similarly for B_i. Here C_2 is large depending on E_{crit}. If $i_2=3$, then one estimates

$$\begin{aligned}&\|P_{k_1,\kappa_1}(\phi^n A_{i_1}e^{-i\partial^{-1}\phi^n})\cdot P_{k_2,\kappa_2}(\phi^n B_{i_2}e^{-i\partial^{-1}\phi^n})\|_{L^2_{t,x}}\\&\lesssim\sum_m 2^{-\sigma|k_1-m|}\|P_m\phi^n\|_{L^\infty_tL^2_x}\sum_{\ell\geq k_2+C_2}\sum_{\substack{c_1,c_2\in\mathcal{D}_{\ell,k_2-\ell}\\ \mathrm{dist}(c_1,c_2)\lesssim 2^{k_2}}}2^{-\ell}\|P_{c_1}\phi^n\|_{L^4_tL^\infty_x}\|P_{c_2}[\phi^n e^{-i\partial^{-1}\phi^n}]\|_{L^4_tL^\infty_x}\\&\lesssim 2^{-\frac{C_2}{10}}2^{\frac{k_2}{2}}\sum_m 2^{-\sigma|k_1-m|}\|P_m\phi^n\|_{L^\infty_tL^2_x}\Big(\sum_{\ell>k_2}2^{-\sigma(\ell-k_2)}\|P_\ell\psi^n\|_{S[\ell]}\Big)^2\end{aligned}$$

with an implicit constant which is allowed to depend on the energy. Therefore, this is placed in g_{k_1,k_2}. The case where $i_1=3$ is similar. Next, suppose that $i_1=1$ *and* $i_2=1$. Then the cap localization passes on to the ϕ^n and due to (10.23) one places the resulting expression into $f_{k_1,k_2}+g_{k_1,k_2}$. We are left with three cases: $i_1=1,i_2=2$, and $i_1=2,i_2=1$, and $i_1=i_2=2$. Next, observe that we may assume that

$$P_{k_1,\kappa_1}(\phi^n A_{i_1}e^{-i\partial^{-1}\phi^n})=P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n A_{i_1}e^{-i\partial^{-1}\phi^n})$$

and

$$P_{k_2,\kappa_2}(\phi^n B_{i_2}e^{-i\partial^{-1}\phi^n})=P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n B_{i_2}e^{-i\partial^{-1}\phi^n})$$

for otherwise one obtains smallness from Bernstein's inequality. For example, consider now $i_1 = 1$, $i_2 = 2$ which is

$$P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n}) \cdot P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}e^{-i\partial^{-1}\phi^n})$$
$$= P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n}) \cdot P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}])$$

Now we distinguish two more cases: either the exponential in the second factor has frequency $< 2^{k_2-m_0'-20}$ or not. In the former case, one obtains

$$P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n}) \cdot P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}])$$
$$= P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n}) \cdot P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n P_{<k_2-m_0'-20}e^{-i\partial^{-1}\phi^n}])$$

Now perform a cap decomposition of the first and second ϕ^n factors inside the P_{k_2,κ_2} term. Observe that due to the fact that the frequencies of these factors are approximately 2^{k_2} at least one of them has to have angular separation with the cap κ_1 from the first factor by an amount comparable to $2^{-m_0'}$. We may therefore place this expression into $f_{k_1,k_2} + g_{k_1,k_2}$ in view of (10.23). If, on the other hand, the exponential in the second factor has frequency $> 2^{k_2-m_0'-20}$, then one writes

$$= P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n}) \cdot P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n P_{\geq k_2-m_0'-20}e^{-i\partial^{-1}\phi^n}])$$
$$= P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n})\cdot$$
$$\cdot P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n P_{\geq k_2-m_0'-20}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}]])$$

The idea here is to place the entire expression into $L^2_{t,x}$ by putting the first factor into $L^\infty_t L^2_x$, i.e., estimating

$$\|P_{k_1,\kappa_1}(P_{>k_1-C_2}\phi^n P_{<k_1-m_0'-10}e^{-i\partial^{-1}\phi^n})\|_{L^\infty_t L^2_x} \lesssim \|P_{k_1}\phi^n\|_{L^\infty_t L^2_x}$$

followed by the estimate

$$\|P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n P_{\geq k_2-m_0'-20}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}]])\|_{L^2_t L^\infty_x}$$
$$\lesssim 2^{k_2}\|P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[\phi^n P_{\geq k_2-m_0'-20}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}]])\|_{L^2_t L^2_x}$$
$$\lesssim 2^{k_2}\|P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[P_{<k_2-m_0'-C_4}\phi^n P_{\geq k_2-m_0'-20}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}]])\|_{L^2_t L^2_x} \tag{10.25}$$
$$+ 2^{k_2}\sum_{\substack{|k-k'|\leq m_0'+C_4\\ k\geq k_2-m_0'-C_4}}\|P_{k_2,\kappa_2}(P_{>k_2-C_2}\phi^n P_{k_2-m_0'-10\leq\cdot<k_2+C_2}\partial^{-1}[P_k\phi^n P_{k'}\partial^{-1}[\phi^n e^{-i\partial^{-1}\phi^n}]])\|_{L^2_t L^2_x} \tag{10.26}$$

Note that we may reduce (10.26) to (with possibly very large $O(1)$ but only depending on the energy)

$$2^{k_2}\|P_{k_2,\kappa_2}(P_{k_2+O(1)}\phi^n\partial^{-1}P_{k_2+O(1)}[P_{k_2+O(1)}\phi^n P_{k_2+O(1)}\partial^{-1}[P_{k_2+O(1)}\phi^n\, e^{-i\partial^{-1}\phi^n}]])\|_{L^2_t L^2_x} \tag{10.27}$$

since the extremely large frequencies give a gain of a smallness factor whence that case can be place entirely into the bootstrap term g_{k_1,k_2}. We chose C_4 here so large that the entire expression (10.25) is placed in the bootstrap term g_{k_1,k_2}. To see this, one estimates

$$\begin{aligned}
(10.25) &\lesssim 2^{-k_2}\|P_{k_2+O(1)}\phi^n\|_{L^6_{t,x}}\|P_{<k_2-m_0'-C_4}\phi^n\|_{L^6_t L^6_x}\|P_{k_2+O(1)}[\phi^n e^{-i\partial^{-1}\phi^n}]\|_{L^6_t L^6_x}\\
&\lesssim 2^{-k_2}\|P_{k_2+O(1)}\phi^n\|_{L^6_{t,x}}\sum_{\ell<k_2-m_0'-C_4}\|P_\ell\phi^n\|_{L^6_t L^6_x}\|P_{k_2+O(1)}[\phi^n e^{-i\partial^{-1}\phi^n}]\|_{L^6_t L^6_x}\\
&\lesssim 2^{-k_2}\|P_{k_2+O(1)}\phi^n\|_{L^6_{t,x}}\sum_{\ell<k_2-m_0'-C_4}2^{\frac{\ell}{2}}\|P_\ell\psi^n\|_{S[\ell]}\|P_{k_2+O(1)}[\phi^n e^{-i\partial^{-1}\phi^n}]\|_{L^6_t L^6_x}\\
&\lesssim 2^{-\frac{C_4}{2}}2^{\frac{k_2}{2}}\Big(\sum_\ell 2^{-\frac14|\ell-k_2|}\|P_\ell\psi^n\|_{S[\ell]}\Big)^3
\end{aligned}$$

Second, with each $S_0 := \sum_j P_j Q_{\leq j+C_3}$ and $S_1 := \sum_j P_j Q_{>j+C_3}$ where C_3 is a large constant depending only on the energy,

$$(10.27) \lesssim \sum_{i_1,i_2,i_3=0,1} 2^{-k_2} \|P_{k_2+O(1)} S_{i_1}\phi^n\|_{L^6_{t,x}} \|P_{k_2+O(1)} S_{i_2}\phi^n\|_{L^6_t L^6_x} \|P_{k_2+O(1)} S_{i_3}\phi^n\|_{L^6_t L^6_x}$$

Now note the following:

$$\begin{aligned}\sum_{k\in\mathbb{Z}} 2^{-k}\|P_k Q_{\leq k+C_3}\phi^n\|^2_{L^6_{t,x}} &\lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\|P_k Q_{\leq k+C_3'}\phi^{n,d}\|^2_{L^6_{t,x}} \lesssim \sum_{k\in\mathbb{Z}} 2^{-k}\|P_k\psi^{n,d}\|^2_{L^6_{t,x}}\\ &\lesssim \sum_{k\in\mathbb{Z}} \|P_k\psi^{n,d}\|^2_{S[k]} = \|\psi^{n,d}\|^2_S \lesssim M^2\end{aligned}$$

see above. On the other hand, the elliptic piece satisfies

$$\|P_k Q_{>k+C_3}\phi^n\|_{L^6_{t,x}} \lesssim 2^{-\frac{C_3}{10}} 2^{\frac{k}{2}} \sum_{\ell\in\mathbb{Z}} 2^{-\frac{1}{4}|\ell-k|}\|P_\ell\psi^n\|_{S[\ell]}$$

via the same arguments we used in the elliptic case earlier in this proof. The remaining cases $i_1 = 2, i_2 = 1$, $i_{1,2} = 2$, are treated similarly. This now concludes the proof of (10.5), and therefore of the proposition. □

Next, we formulate the analogue of Proposition 4.11 in our context.

Proposition 10.7. *Let $I^+ = [0,\infty)$ and assume that $\lambda(t) > \lambda_0 > 0$ for all $t \geq 0$. Then for $i = 1, 2$,*

$$\int_{\mathbb{R}^2} \langle \partial_i U, \partial_t U\rangle\, dx = \mathrm{Re}\int_{\mathbb{R}^2} \Psi_i^\infty \bar{\Psi}_0^\infty\, dx = 0$$

for all times in I^+.

Proof. In view of Proposition 10.6 we may also assume that $I^- = (-\infty, 0]$. For a contradiction, assume that

$$\mathrm{Re}\int_{\mathbb{R}^2} \Psi_1^\infty \bar{\Psi}_0^\infty\, dx = \gamma > 0$$

As in [13] one now obtains the following statements, cf. (4.10) and (4.11) in [13]:

- Given $\varepsilon > 0$ there exists $R_0(\varepsilon) > 0$ so that for all $t \geq 0$ one has

$$\int_{\left|x+\frac{\bar{x}(t)}{\lambda(t)}\right|\geq R_0(\varepsilon)} |\Psi_\alpha^\infty(t,x)|^2\, dx \leq \varepsilon \tag{10.28}$$

- There exists $M > 0$ so that for all $t \geq 0$, one has $\left|\frac{\bar{x}(t)}{\lambda(t)}\right| \leq t + M$

These are a consequence of the compactness in Corollary 9.36 and Lemma 10.3. Recall from the proof of Proposition 10.1 that upon passing to a suitable subsequence of the approximating maps $\mathbf{u}^n$, we may extract an L^2-limit for the standard derivative components ϕ_α^n; denote this by Φ_α^∞ (which, in contrast to Ψ_α^∞, we do not claim to be canonical). Now define for each $d > 0$, $R > 0$,

$$Z_\alpha^{d,R}(t,x) := \Phi_\alpha^{\infty,R}\Big(\frac{t-dx_1}{\sqrt{1-d^2}}, \frac{x_1-dt}{\sqrt{1-d^2}}, x_2\Big)$$

where

$$\Phi_\alpha^{\infty,R}(s,y) := R\Phi_\alpha^\infty(Rs, Ry)$$

These rescaled limiting profiles again have energy E_{crit}. Now define θ to be a smooth cutoff function supported on $|x| \leq 2$ and $\theta = 1$ on $|x| \leq 1$. The main calculation in the proof of Proposition 4.11 of [13] now reveals that, see (4.20) there, uniformly in $t_0 \in [1,2]$,

$$\sum_{\alpha=0}^{2}\int_{\mathbb{R}^2} \theta^2(x)|Z_\alpha^{d,R}(t_0,x)|^2\, dx = E_{crit} - \gamma d + d\eta(R,d) + \tilde{\eta}(R,d) + O(d^2) \tag{10.29}$$

with $\eta(R,d)$ and $\tilde\eta(R,d)\to 0$ as $R\to\infty$, uniformly in $0<d<d_0$ and with $O(d^2)$ uniform in R. Furthermore, the argument in [13] yields that for fixed $\varepsilon>0$, $R>0$, $d>0$ as above, one may find $t_0\in[1,2]$ such that

$$\int_{\frac12\le|x|\le 2}|Z_\alpha^{d,R}(t_0,x)|^2\,dx\le\varepsilon$$

We shall later pick ϵ, R depending on γ, d and d depending on γ, E_{crit}. Now for fixed choices of these parameters, pick n large enough such that for $\mathbf{u}^n=(\mathbf{x}^n,\mathbf{y}^n)$ an element of the approximating sequence of wave maps from $\mathbb{R}^{2+1}\to\mathbb{H}^2$, denoting by $\psi_\alpha^{n,d,R}$ the Coulomb components of $\mathbf{u}^n\circ L_d$ dilated by factor R as above, and similarly by $\phi_\alpha^{n,d,R}$ the standard derivative components, an averaging argument over different time-like foliations yields that we may also assume

$$\int_{\mathbb{R}^2}|\phi_\alpha^{n,d,R}(t_0,x)-Z_\alpha^{d,R}(t_0,x)|^2\,dx<\varepsilon.$$

Note that now t_0 may depend on n, but this does not affect the argument. The idea now is to truncate the data

$$(\mathbf{u}^n\circ L_d(Rt_0,Rx),\quad R\partial_t\mathbf{u}^n\circ L_d(Rt_0,Rx))$$

solve the Cauchy problem backwards, and undo the Lorentz transform. We thereby obtain a good approximation to the original essentially singular sequence ψ_α^n, but which satisfies good S-estimates, which gives us the desired contradiction. Thus, write $\mathbf{u}^n\circ L_d(Rt,Rx)=(\mathbf{x}^{n,d,R},\mathbf{y}^{n,d,R})$. To do this, we consider data

$$h^{n,d,R}(t_0,\cdot):=\Big(\chi_{[|x|<\frac12]}\frac{\mathbf{x}^{n,d,R}(t_0,\cdot)-\mathbf{x}_0^{n,d,R}}{\mathbf{y}_0^{n,d,R}},\ e^{\chi_{[|x|<1]}\log[\frac{\mathbf{y}^{n,d,R}}{\mathbf{y}_0^{n,d,R}}(t_0,\cdot)]}\Big),$$

where $\chi_{[|x|>M]}$ is a smooth cutoff to the set $\{|x|>M\}$ which equals one on $\{|x|>\frac54 M\}$, say, and $\chi_{[|x|<M]}:=1-\chi_{[|x|>M]}$. Moreover,

$$\mathbf{x}_0^{n,d,R}:=\fint_{[\frac14<|x|<\frac12]}\mathbf{x}^{n,d,R}(x)\,dx_1dx_2,\qquad \mathbf{y}_0^{n,d,R}:=\exp\Big(\fint_{[\frac12<|x|<1]}\log\mathbf{y}^{n,d,R}(x)\,dx_1dx_2\Big)$$

Also, denote by $h^{n,d,R}(t,\cdot)$ the above expressions with t_0 replaced by t. As in the proof of Lemma 10.3, one then checks that for these data we have

$$\int e(h^{n,d,R})(t_0,\cdot)\,dx<E_{crit}-\frac{\gamma d}{2}$$

where e is the energy density, provided we choose R large enough, ε and d small enough, and then n large enough. Now consider the wave maps evolution of the data

$$H^{n,d,R}(t_0,\cdot):=\big(h^{n,d,R}(t_0,\cdot),\quad\partial_t h^{n,d,R}(t_0,\cdot)\big)$$

Our energy induction hypothesis implies that this evolution is defined globally in time, and upon denoting the corresponding Coulomb derivative components by

$$\psi_{\chi,\alpha}^{n,d,R},$$

we obtain a global bound

$$\|\psi_{\chi,\alpha}^{n,d,R}\|_{S(\mathbb{R}^{2+1})}\le\Lambda(E_{crit},\,d,\,\gamma)<\infty$$

Denote the time evolution of the data $H^{n,d,R}(t_0,\cdot)$ by $H^{n,d,R}(t,\cdot)$, and the corresponding derivative components (not in the Coulomb Gauge) by

$$\phi_{\chi,\alpha}^{n,d,R}$$

We now undo the Lorentz transformation L_d, i.e., consider

$$h^{n,d,-d,R}(t,\cdot):=h^{n,d,R}(t,\cdot)\circ L_{-d}$$

The argument in the proof of the preceding proposition then yields that we also can conclude that the Coulomb derivative components of $h^{n,d,-d,R}(t,\cdot)$, which we denote by $\psi_{\chi,\alpha}^{n,d,-d,R}$, also satisfy a bound of the form

$$\|\psi_{\chi,\alpha}^{n,d,-d,R}\|_{S(\mathbb{R}^{2+1})}\le\Lambda'(E_{crit},\,d,\,\gamma)<\infty$$

Furthermore, denoting the standard derivative components of $h^{n,d,-d,R}(t,\cdot)$ by $\phi_{\chi,\alpha}^{n,d,-d,R}$, by finite propagation speed we have

$$\phi_{\chi,\alpha}^{n,d,-d,R}(0,x) = \phi_\alpha^{n,R}(0,x), \quad \alpha = 0,\,1,\,2,$$

provided $|x| < \frac{1}{10}$, say, where $\phi_\alpha^{n,R}(0,\cdot)$ are the standard derivative components of $\mathbf{u}^n(Rt,\,Rx)$ at time $t=0$. To conclude the proof of the proposition, we note that by the convergence of the ψ_α^n at time $t=0$ in the L^2-sense, picking R large enough and then also n large enough, we may arrange that (for suitable constants $\gamma_{nm} \in \mathbb{R}$)

$$\|\psi_{\chi,\alpha}^{n,d,-d,R}(0,\cdot) - e^{i\gamma_{nm}}\psi_\alpha^m(0,\cdot)\|_{L_x^2} \le \varepsilon_1, \quad m \ge n$$

where ε_1 is as in Proposition 7.11, with $A = \Lambda'(E_{crit}\,,\,d,\,\gamma)$. But this then yields the contradiction

$$\limsup_{m\to\infty} \|\psi_\alpha^m\|_{S(\mathbb{R}^{2+1})} < \infty$$

and we are done. □

10.0.2. *Rigidity I: harmonic maps and reduction to the self-similar case.* As in [13] one now has the following rigidity theorem.

Proposition 10.8. *With $\{\Psi_\alpha^\infty\}_{\alpha=0}^2$ as above, and with life span $(-T_0,T_1)$ one cannot have T_1 or T_0 finite. Moreover, if $\lambda(t) \ge \lambda_0 > 0$ for all $t\in\mathbb{R}$, one necessarily has $\Psi_\alpha^\infty = 0$ for $\alpha = 0,1,2$.*

The proof of it will follow from a sequence of lemmas, and only be completed after Proposition 10.17. We begin with the case where $T_1 = \infty$ and $\lambda(t) \ge \lambda_0 > 0$ on $[0,\infty)$. Assuming that Ψ_α^∞ do not all vanish, the logic then is to extract a nonconstant harmonic map of finite energy into the compact Riemann surface $\mathcal{S}$, leading to a contradiction. The following lemma is the analogue of Lemma 5.4 in [13]. While the statement is identical with that in [13], its proof is slightly different and invokes in a crucial way the geometry of the target. In the statement, we use a function $\epsilon \to R_0(\epsilon)$, defined as follows: by compactness, for every $\varepsilon > 0$ there exists $R_0(\varepsilon) > 0$ such that for all $t \ge 0$ one has

$$\int_{\left|x+\frac{\bar{x}(t)}{\lambda(t)}\right| \ge R_0(\varepsilon)} |\partial_\alpha U(t,x)|^2\,dx \le \varepsilon$$

since $\lambda(t) \ge \lambda_0 > 0$ for all $t \ge 0$.

Lemma 10.9. *There exists $\varepsilon_1 > 0$, $C > 0$ such that if $\varepsilon \in (0,\varepsilon_1)$ there exists $R_0(\varepsilon)$ so that if $R > 2R_0(\varepsilon)$ then there exists $t_0 = t_0(R,\varepsilon)$, $0 \le t_0 \le CR$ with the property that for all $0 < t < t_0$ one has*

$$\Big|\frac{\bar{x}(t)}{\lambda(t)}\Big| < R - R_0(\varepsilon), \qquad \Big|\frac{\bar{x}(t_0)}{\lambda(t_0)}\Big| = R - R_0(\varepsilon)$$

Proof. As a preliminary argument, we show that there exists $\alpha \in \mathbb{R}$ with

$$\int_I \int_{\mathbb{R}^2} |\Psi_0^\infty|^2(t,x)\,dxdt \ge \alpha > 0 \tag{10.30}$$

for all intervals I of length one. If not, there exists a sequence of intervals $J_n := [t_n, t_n+1]$ with the property that $t_n \to \infty$ and

$$\int_{J_n} \int_{\mathbb{R}^2} |\Psi_0^\infty|^2(t,x)\,dxdt \le \frac{1}{n} \tag{10.31}$$

Then there exist times $s_n \in J_n$ with the property that $\|\Psi_0^\infty(s_n,\cdot)\|_2 \to 0$ as $n\to\infty$. By Corollary 9.36 one has that

$$\Big\{\lambda(s_n)^{-1}\Psi_\alpha^\infty\big(s_n,(\cdot-\bar{x}(s_n))\lambda(s_n)^{-1}\big)\Big\}_{n=0}^\infty$$

forms a compact set for $\alpha = 0,1,2$. Passing to a subsequence, we may assume that strongly in L^2

$$\lambda(s_n)^{-1}\Psi_\alpha^\infty\big(s_n,(\cdot-\bar{x}(s_n))\lambda(s_n)^{-1}\big) \to \Psi_\alpha^*(\cdot)$$

By Lemma 7.10 there exists some nonempty time interval I^* around zero such that

$$\lambda(s_n)^{-1}\Psi_\alpha^\infty\big(s_n + t\lambda(s_n)^{-1},(\cdot-\bar{x}(s_n))\lambda(s_n)^{-1}\big) \to \Psi_\alpha^*(t,\cdot)$$

in $L^\infty_{\mathrm{loc}}(I^*; L^2(\mathbb{R}^2))$. Distinguish two cases: $\{\lambda(s_n)\}$ is bounded or not. In the former case, note that $\lambda(t) \geq \lambda_0 > 0$ implies that there exists a nonempty $I^\dagger \subset I^*$ such that $s_n + \lambda(s_n)^{-1} I^\dagger \subset J_n$ for each n. Therefore, (10.31) implies that

$$\int_{I^\dagger} \int_{\mathbb{R}^2} |\Psi_0^*|^2(t,x)\, dxdt = 0$$

This implies that $\Psi_0^*(t,\cdot) = 0$ for all $t \in I^\dagger$. On the other hand, if $\{\lambda(s_n)\}$ is unbounded for every sequence $\{s_n\}$ with $s_n \in J_n$, we invoke the covering argument from [48]. Thus write for each n

$$J_n = \bigcup_{s \in J_n} [s - \lambda^{-1}(s), s + \lambda^{-1}(s)]$$

By the Vitali covering lemma, we may pick a disjoint subcollection of intervals $\{I_s\}_{s\in A^n}$, $I_s := [s - \lambda^{-1}(s), s + \lambda^{-1}(s)]$ for some subset $A^n \subset J_n$ with the property that

$$\bigcup_{s \in A^n} |I_s| \geq \frac{1}{5}$$

But then the defining property of the J_n implies that for each J_n, we may pick times $s_n \in J_n$ with the property that

$$\int_{I_{s_n} \cap J_n} \|\Psi_0^\infty(t,\cdot)\|_{L^2_x}^2\, dt = o(\lambda^{-1}(s_n))$$

Alternatively, this implies that as $n \to \infty$

$$\int_{-1}^{1} \|\big(\chi_{J_n}\Psi_0^\infty\big)(s_n + t\lambda^{-1}(s_n),\, \cdot)\|_{L^2_x}^2\, dt = o(1)$$

Now pick a converging subsequence of

$$\lambda(s_n)^{-1}\Psi_0^\infty(s_n + t\lambda^{-1}(s_n), (\cdot - \bar{x}(s_n))\lambda(s_n)^{-1})$$

to again obtain a limiting object Ψ_α^* with the property that

$$\Psi_0^*(t,\cdot) = 0$$

provided $t \in I^*$, the latter its lifespan interval.
We now deduce the desired contradiction from this situation: as in Proposition 10.1, we can associate a weak wave map U^* from $\mathbb{R}^{2+1} \to \mathcal{S}$ with the limiting object Ψ_α^*, and this wave map has the property that

$$\partial_t U^* = 0, \quad t \in I^*$$

Moreover, we have

$$\sum_{\alpha=1}^{2} \|\partial_\alpha U^*\|_{L^2_x}^2 = \sum_{\alpha=1,2} \|\Psi_\alpha^*\|_{L^2_x}^2 \neq 0$$

We have thus obtained a nonvanishing finite energy harmonic map $U^* : \mathbb{R}^2 \to \mathcal{S}$, which is impossible, see [37].

We therefore conclude that (10.30) holds. The remainder of the argument is essentially the same as that in Lemma 5.4 of [13]: by Corollary 10.2,

$$\frac{d}{dt}\sum_{i=1}^{2}\int_{\mathbb{R}^2} x_i \phi(x/R) \langle \partial_t U(t,x), \partial_i U(t,x)\rangle\, dx = -\int_{\mathbb{R}^2} |\partial_t U(t,x)|^2\, dx + O(r(R)) \tag{10.32}$$

where

$$r(R) := \int_{[|x|\geq R]} \sum_{\alpha=0}^{2} |\partial_\alpha U(t,x)|^2\, dx$$

Furthermore, by definition of $R_0(\varepsilon) > 0$, for all $t \geq 0$ one has

$$\int_{\left|x + \frac{\bar{x}(t)}{\lambda(t)}\right| \geq R_0(\varepsilon)} |\partial_\alpha U(t,x)|^2\, dx \leq \varepsilon$$

Therefore, if the lemma were to fail, then (assuming $\bar{x}(0) = 0$ as we may) one would have

$$\big|\frac{\bar{x}(t)}{\lambda(t)}\big| \leq R - R_0(\varepsilon)$$

for all $0 \leq t < CR$. In view of the preceding, one concludes that $r(R) \leq C_5\varepsilon$ for some absolute constant C_5. Now choose $\varepsilon > 0$ so small that

$$\int_I \Big(- \int_{\mathbb{R}^2} |\partial_t U(t,x)|^2 \, dx + O(r(R)) \Big) \, dt \leq -\frac{\alpha}{2}$$

for all I of unit length. In view of the a priori bound

$$\sup_t \Big| \int_{\mathbb{R}^2} x_i \phi(x/R) \langle \partial_t U(t,x), \partial_i U(t,x) \rangle \, dx \Big| \leq C_6 \, R E_{crit}$$

one obtains a contradiction by integrating (10.32) over a sufficiently large time interval. □

Next, we obtain a contradiction to Lemma 10.9 by means of Proposition 10.7. This is completely analogous to Lemma 5.5 in [13].

Lemma 10.10. *There exists $\varepsilon_2 > 0$, $R_1(\varepsilon) > 0$, $C_0 > 0$ such that if $R > R_1(\varepsilon)$, $t_0 = t_0(R, \varepsilon)$ are as in Lemma 10.9, then for $0 < \varepsilon < \varepsilon_2$ one has*

$$t_0(R, \varepsilon) > \frac{C_0 R}{\varepsilon}$$

Proof. This follows from Proposition 10.7 by the same argument as in [13]. □

Proof of Proposition 10.8 for $T_1 = \infty$. Choosing ε small in Lemma 10.9 and Lemma 10.10 leads to a contradiction. □

It remains to prove Proposition 10.8 in case $T_1 < \infty$. This will be lead to a contradiction as in [13], by a reduction to the case of a self-similar blow-up scenario. More precisely, recall from Lemma 10.4 above that

$$\lambda(t) \geq \frac{C_0(K)}{1-t}, \qquad 0 < t < 1$$

where we assumed that $T_1 = 1$ as we may. Recall also that in this case

$$\mathrm{supp}(\Psi_\alpha^\infty(t, \cdot)) \subset B(0, 1-t), \qquad 0 < t < 1$$

see Lemma 10.5. Next, we prove an upper bound on $\lambda(t)$ which places us in the self-similar context.

Lemma 10.11. *Assuming that $T_1 = 1$ there exists a constant $C_1(K)$ such that*

$$\frac{C_1(K)}{1-t} \geq \lambda(t), \qquad 0 < t < 1$$

Proof. Suppose this fails. Let

$$z(t) := \sum_{j=1}^{2} \int x_j \Psi_j^\infty(t,x) \bar{\Psi}_0^\infty(t,x) \, dx, \qquad 0 < t < 1$$

Note that $z(t) \to 0$ as $t \to 1$. Moreover, by Corollary 9.36 one has

$$z'(t) = -\int |\Psi_0^\infty(t,x)|^2 \, dx$$

Hence,

$$z(t) = \int_t^1 \int |\Psi_0^\infty(s,x)|^2 \, dx ds$$

We now distinguish two cases: either there exists $\alpha > 0$ such that

$$\int_t^1 \int |\Psi_0^\infty(s,x)|^2 \, dx ds \geq \alpha(1-t), \qquad 0 < t < 1 \tag{10.33}$$

or not, i.e., there exists a sequence $J_n = (t_n, 1)$ with $t_n \to 1$ such that

$$|J_n|^{-1} \int_{J_n} \int |\Psi_0^\infty(s,x)|^2 \, dxds \to 0 \quad \text{as } n \to \infty \tag{10.34}$$

If the first alternative (10.33) holds, then one is lead to a contradiction as in [13]. On the other hand, we will now reduce the second alternative (10.34) to the existence of a nontrivial harmonic map into $\mathcal{S}$ by a similar argument as in the proof of Lemma 10.9, see also Struwe [48]. By the Vitali argument from above, one selects intervals $J_n' := (s_n - \lambda(s_n)^{-1}, s_n + \lambda(s_n)^{-1})$ with $s_n \in J_n$ such that

$$|J_n'|^{-1} \int_{J_n'} \int |\Psi_0^\infty(s,x)|^2 \, dxds \to 0 \quad \text{as } n \to \infty$$

Now one uses compactness as in the proof of Lemma 10.9 to conclude that there exists a limiting wave map Ψ_α^* on some nonempty interval I^* with $\Psi_0^* = 0$ on I^*. Therefore, Ψ^* leads to a a harmonic map U^* of energy E_{crit} into $\mathcal{S}$, which gives the desired contradiction. □

This now allows us to reduce to the exactly self-similar case.

Corollary 10.12. *If $T_1 = 1$, then the set*

$$\Big\{(1-t)\Psi_\alpha^\infty(t,(1-t)x) \;:\; 0 < t < 1, \quad \alpha = 0,1,2\Big\}$$

is compact in L^2.

Proof. This is as in Proposition 5.7 of [13]. □

10.0.3. *Rigidity II: the self-similar case.* We now turn to the last step in the Kenig-Merle program (modulo the issue of removing the assumption $\lambda(t) > \lambda_0$ for infinite times) which consists of excluding the possibility of self-similar blow-up. As in [29], [30] we set

$$y = \frac{x}{1-t}, \quad s = -\log(1-t), \quad 0 < t < 1$$

and

$$W(y,s,0) := U(x,t) = U(e^{-s}y, 1 - e^{-s}), \quad 0 \le s < \infty$$

where U is a weak wave map as constructed in Proposition 10.1. By construction, $\nabla_{s,y} W$ is supported in $\{|y| \le 1\}$. Next, for $\delta > 0$, introduce

$$y = \frac{x}{1-t+\delta}, \quad s = -\log(1-t+\delta), \quad 0 < t < 1$$

$$W(y,s,\delta) := U(e^{-s}y, 1 + \delta - e^{-s}) \tag{10.35}$$

Then we have that $W(y,s,\delta)$ is defined for $0 \le s < -\log\delta$ and

$$\mathrm{supp}(\partial_\alpha W(\cdot,\delta)) \subset \{|y| \le 1-\delta\}$$

The W solve the equation in the distributional sense

$$\partial_s^2 W = \frac{1}{\rho}\mathrm{div}(\rho\nabla W - \rho(y\cdot\nabla W)y) - 2y\cdot\nabla\partial_s W - \partial_s W - A(W)((\partial_s + y\cdot\nabla_y)W, \nabla_y W) \tag{10.36}$$

where the nonlinearity stands for the second fundamental form on the Riemann surface $\mathcal{S}$ relative to its embedding into $\mathbb{R}^N$.

We now state the following properties of W. Henceforth, $|\cdot|$ when applied to derivatives of W will denote the metric on $\mathcal{S}$ and $W = W(\cdot,\delta)$.

Lemma 10.13. *For $\delta > 0$ fixed,*

- $\mathrm{supp}(\partial_\alpha W(\cdot,\delta)) \subset \{|y| \le 1-\delta\} \quad \alpha = 0,1,2$
- $\int (|\nabla_y W|^2 + |\partial_s W|^2)\, dy \le C$
- $\sum_{\alpha=0}^2 \int |\partial_\alpha W(s,y)|^2 \,|\log(1-|y|^2)|\, dy \le C|\log\delta|$
- $\sum_{\alpha=0}^2 \int |\partial_\alpha W(s,y)|^2 \,(1-|y|^2)^{-\frac{1}{2}}\, dy \le C\delta^{-\frac{1}{2}}$

Proof. By direct calculation. □

As in [13] one now introduces a Lyapunov functional

$$\tilde{E}(W(s)) := \frac{1}{2}\int_{\mathbb{D}} [|\partial_s W|^2 + |\nabla_y W|^2 - |y\cdot\nabla_y W|^2]\,(1-|y|^2)^{-\frac{1}{2}}\,dy$$

This quantity satisfies

Proposition 10.14. *For $0 < s_1 < s_2 < \log(\frac{1}{\delta})$, the following identities hold:*

(1) $\tilde{E}(W(s_2)) - \tilde{E}(W(s_1)) = \int_{s_1}^{s_2}\int_{\mathbb{D}} \frac{|\partial_s W|^2}{(1-|y|^2)^{3/2}}\,dy\,ds$
(2) $\lim_{s\to\log(\frac{1}{\delta})} \tilde{E}(W(s)) \le E_{crit}$.

Proof. This is proved as in [29], see Lemma 2.1 there. The difference is of course that we have a different equation, namely (10.36). However, the point is that the second fundamental form is perpendicular to $\partial_s W$ and $\nabla_y W$ whence it drops out of the calculation needed for the first identity.

The second property is verified as in [13]. □

As a corollary, one now has the following:

Lemma 10.15. *For each $\delta > 0$ there exists $\bar{s}_\delta \in (\frac{|\log\delta|}{2}, |\log\delta|)$ such that*

$$\int_{\bar{s}_\delta}^{\bar{s}_\delta + |\log\delta|^{\frac{1}{2}}}\int_{\mathbb{D}} \frac{|\partial_s W|^2}{(1-|y|^2)^{\frac{3}{2}}}\,dyds \le \frac{E_{\text{crit}}}{|\log\delta|^{\frac{1}{2}}}$$

Proof. By Proposition 10.14,

$$\int_0^{|\log\delta|}\int_{\mathbb{D}} \frac{|\partial_s W|^2}{(1-|y|^2)^{\frac{3}{2}}}\,dyds \le E_{crit}$$

whence the claim. □

The goal is now to obtain a limit W^* as $\delta\to 0$ and to show that W^* is a stationary solution of (10.36). To this end, select $\delta_j \to 0$ such that for each $\alpha = 0,1,2,$

$$(1-\bar{t}_{\delta_j})\Psi_\alpha^\infty(\bar{t}_{\delta_j}, (1-\bar{t}_{\delta_j})x) \to \Psi_\alpha^*(x)$$

strongly in L^2, see Corollary 10.12. In fact, we may arrange also that

$$(1+\delta_j-\bar{t}_{\delta_j})\Psi_\alpha^\infty(\bar{t}_{\delta_j}, (1+\delta_j-\bar{t}_{\delta_j})x) \to \Psi_\alpha^*(x) \tag{10.37}$$

in L^2. Now consider the evolution on the level of the Ψ with data given by the left-hand side of (10.37), see Section 7.2. By our perturbation theory of Section 7.2 we conclude from (10.37) that these evolutions $\Psi_\alpha^{j*}(t,x)$ exist on some fixed lifespan, and moreover,

$$\Psi_\alpha^{j*}(t,x) = (1+\delta_j-\bar{t}_{\delta_j})\Psi_\alpha^\infty(\bar{t}_{\delta_j} + (1+\delta_j-\bar{t}_{\delta_j})t, (1+\delta_j-\bar{t}_{\delta_j})x)$$

on that lifespan $[0,T^*)$ where we may assume that $T^* < 1$. Note that on account of this identity,

$$\text{supp}(\Psi_\alpha^{j*}(t,\cdot)) \subset \Big\{|y| \le \frac{1-\bar{t}_{\delta_j}}{1+\delta_j-\bar{t}_{\delta_j}} - t < 1-t\Big\}$$

for each $\alpha = 0,1,2$ and $0 < t < T^*$. Now note that by the construction in the proof of Proposition 10.1 we may arrange that the weak wave maps U^{j*} associated with Ψ_α^{j*} and U associated with Ψ_α^∞ satisfy

$$U^{j*}(t,x) = U(\bar{t}_{\delta_j} + (1+\delta_j-\bar{t}_{\delta_j})t, (1+\delta_j-\bar{t}_{\delta_j})x)$$

Note that for fixed times $t\in(0,T^*)$ one has that $\{\partial_\alpha U^{j*}(t,\cdot)\}$ form a compact set in L^2 whence the argument in the proof of Proposition 10.1 implies that up to passing to a subsequence

$$\partial_\alpha U^{j*}(t,\cdot) \to \partial_\alpha U^*(t,\cdot)$$

strongly in L^2 uniformly on compact subintervals of time. Moreover, U^* is a weak wave map and satisfies the conservation laws. Next, we switch to the (s,y) variables. Define

$$W_j^*(y,s) := U(\bar{t}_{\delta_j} + (1+\delta_j-\bar{t}_{\delta_j})t, (1+\delta_j-\bar{t}_{\delta_j})x)$$

with the same relation between (s, y) and (t, x) as above. Similarly, define

$$W^*(y, s) = U^*(t, x)$$

Then by the preceding, uniformly in $0 \leq s \leq -\log(1 - T^*/2) =: \tilde{T}$ and for $\alpha = 0, 1, 2$,

$$\partial_\alpha W_j^*(\cdot, s) \to \partial_\alpha W^*(\cdot, s)$$

in the strong L^2 sense. Moreover, with W as in (10.35), one has with $\bar{s}_{\delta_j} = -\log(1 + \delta_j - \bar{t}_{\delta_j})$,

$$W_j^*(y, s) = W(y, \bar{s}_{\delta_j} + s, \delta_j)$$

and therefore also

$$\partial_\alpha W(y, \bar{s}_{\delta_j} + s, \delta_j) \to \partial_\alpha W^*(\cdot, s) \tag{10.38}$$

strongly in L^2 uniformly in $0 \leq s \leq \tilde{T}$. Moreover, W^* is a solution of (10.36) and

$$\mathrm{supp}(\partial_\alpha W^*(s, \cdot)) \subset \{|y| \leq 1\}$$

as well as

$$\mathrm{trace}(W^*(s, \cdot)) = const$$

where trace is the L^2-trace.

Lemma 10.16. *Let W^* be as above. Then,*

$$W^*(y, s) = W^*(y) \text{ and } W^* \not\equiv const.$$

Proof. With $S = -\log(1 - \tilde{T})$ and j large one has

$$\int_0^S \int_{\mathbb{D}} \frac{|\partial_s W^*(y, s)|^2}{(1 - |y|^2)^{3/2}} \, dyds \leq \varliminf_{j \to \infty} \int_0^S \int_{\mathbb{D}} \frac{|\partial_s W(y, \bar{s}_{\delta_j} + s, \delta_j)|^2}{(1 - |y|^2)^{3/2}} \, dyds$$

by (10.38). The right-hand side is bounded by

$$\varliminf_{j \to \infty} \int_{\bar{s}_{\delta_j}}^{S + \bar{s}_{\delta_j}} \int_{\mathbb{D}} \frac{|\partial_s W(y, s, \delta_j)|^2}{(1 - |y|^2)^{3/2}} \, dyds \lesssim \lim_{j \to \infty} |\log \delta_j|^{-1/2} = 0,$$

by Lemma 10.15. This shows that $W^*(y, s) = W^*(y)$ as claimed. The fact that $W^* \not\equiv const$ follows as in [13]. ☐

In other words, we have now obtained a stationary, nonconstant, distributional solution to (10.36) with finite energy (relative to the y variable) (as well as finite $\tilde{E}(W^*)$). The following proposition now leads to the desired contradiction.

Proposition 10.17. *Let W^* be a distributional stationary solution to* (10.36) *of finite energy*

$$\int_{\mathbb{D}} |\nabla W^*(y)|^2 \, dy < \infty$$

Then $W^* = const$. This thus contradicts the preceding construction of W^* and completes the proof of Proposition 10.8.

Proof. We follow the argument of Shatah-Struwe, see [40]: first, W^* is a weakly harmonic map from $\mathbb{D} \to \mathcal{S}$ where $\mathbb{D}$ is equipped with the hyperbolic metric

$$\frac{d\rho^2}{(1 - \rho^2)^2} + \frac{\rho^2}{1 - \rho^2} \, d\omega^2$$

where (ρ, ω) are polar coordinates on $\mathbb{D}$. This means that

$$-(\rho \sqrt{1 - \rho^2} \, W_\rho^*)_\rho + \frac{\Delta_\omega W^*}{\rho \sqrt{1 - \rho^2}} \perp T_{W^*} \mathcal{S}$$

Note that by Helein's theorem, this holds in the classical sense in the interior. Integrating by parts against $\rho\sqrt{1-\rho^2}\, W^*_\rho$ implies that

$$\frac{d}{d\rho}\Big(\int_{S^1}\rho^2(1-\rho^2)|W^*_\rho|^2\,d\omega-\int_{S^1}|W^*_\omega|^2\,d\omega\Big)=0$$

and thus

$$\int_{S^1}\rho^2(1-\rho^2)|W^*_\rho|^2\,d\omega-\int_{S^1}|W^*_\omega|^2\,d\omega=C_0$$

Setting $\rho=0$ one concludes that $C_0=0$ and sending $\rho\to 1$ along a suitable subsequence ρ_j implies that

$$\lim_{\rho_j\to 1}\int_{S^1}|W^*_\omega(\rho_j\omega)|^2\,d\omega\to 0$$

On the other hand, by the trace theorem, $\sup_{\frac12<\rho<1}\|W^*(\rho\omega)\|_{\dot H^{\frac12}(S^1)}\le C\|W^*\|_{H^1(\mathbb{D})}$. Since clearly also $\sup_{\frac12<\rho<1}\|W^*(\rho\omega)\|_{L^2(S^1)}<\infty$, one concludes via interpolation that trace$(W^*)=const$ as the L^2 trace on S^1. The change of variables

$$\sigma(\rho)=\exp\Big(-\int_\rho^1\frac{du}{u\sqrt{1-u^2}}\Big)$$

provides a conformal equivalence between the hyperbolic disk and the disk $\mathbb{D}$ with the Euclidean metric. In fact,

$$d\sigma^2+\sigma^2d\omega^2=\Big(\frac{\sigma}{\rho}\Big)^2(1-\rho^2)\Big(\frac{d\rho^2}{(1-\rho^2)^2}+\frac{\rho^2}{1-\rho^2}\,d\omega^2\Big)$$

By the conformal invariance of the Dirichlet energy in two dimensions, it follows that $v(\sigma,\omega):=W^*(\rho,\omega)$ is a weakly harmonic map $\mathbb{D}\to\mathcal{S}$ with the Euclidean disk $\mathbb{D}$. Moreover, one checks that v has finite $\dot H^1$ energy relative to the (σ,ω)-coordinates and that trace$(v)=const$ in this setting as well. By a result of Qing [35], it follows that v is C^∞ on $\bar{\mathbb{D}}$. And then the result of Lemaire [26] gives the desired conclusion that $W^*=const$. □

The only remaining case is to show that $\lambda(t)$ does not approach zero along some subsequence. This case is handled as in [13] or [28]. We follow the argument [13] essential verbatim.

Lemma 10.18. *Let Ψ^∞_α be the limiting object as above and suppose that $T_1=\infty$. Then $\lambda(t)>\lambda_0>0$ for all $t\ge 0$.*

Proof. Suppose this fails. Then there exist $t_n\to\infty$ so that $\lambda(t_n)\to 0$; in fact, one may assume even that

$$\lambda(t_n)\le\inf_{t\in[0,t_n]}\lambda(t).$$

¿From Corollary 9.36 one has

$$\Psi^n_\alpha:=\lambda(t_n)^{-1}\Psi^\infty_\alpha\big(t_n,(\cdot-\bar x(t_n))\lambda(t_n)^{-1}\big)\to\Psi^\dagger_\alpha$$

strongly in L^2. Then $E(\Psi^\dagger)=E_{crit}$ and we may assume that the lifespan $(-T^\dagger_0,T^\dagger_1)$ of $\Psi^\dagger_\alpha$ has the property that $T^\dagger_0<\infty$. Otherwise one obtains a contradiction from Proposition 10.8. Now define $\Psi^n_\alpha(\tau,x)$ and $\Psi^\dagger_\alpha(\tau,x)$ to be the evolutions of Ψ^n_α and $\Psi^\dagger_\alpha$. By the perturbation theory of Section 7.2 we conclude that $\liminf_{n\to\infty}T_0(\Psi^n_\alpha)=\infty$ and

$$\Psi^n_\alpha(\tau,x)\to\Psi^\dagger_\alpha(\tau,x)$$

in $L^\infty_{\rm loc}((-\infty,0]\times L^2)$. By uniqueness of the Ψ-evolutions

$$\Psi^n_\alpha(\tau,x)=\lambda(t_n)^{-1}\Psi^\infty_\alpha(t_n+\tau\lambda(t_n)^{-1},(x-\bar x(t_n))\lambda(t_n)^{-1})$$

for all $0\le t_n+\frac{\tau}{\lambda(t_n)}$. We claim that $\tau_n:=-t_n\lambda(t_n)$ satisfies

$$\underline{\lim}_n(-\tau_n)=\infty$$

so that for all $\tau\in(-\infty,0]$, for n large, $0\le t_n+\frac{\tau}{\lambda(t_n)}\le t_n$. In fact, if $-\tau_n\to-\tau_0<\infty$, then

$$\Psi^n_\alpha(x,-\tau_n)=\lambda(\tau_n)^{-1}\Psi^\infty_\alpha\big(\frac{x-x(t_n)}{\lambda(t_n)},0\big)$$

would converge to $\Psi^\dagger_\alpha(x,-\tau_0)$ in L^2, with $\lambda(t_n)\to 0$, which contradicts $\Psi^\dagger_\alpha \neq 0$.

We now make the further claim that $\|\Psi^\dagger_\alpha\|_{S(-\infty,0)} = +\infty$. Otherwise, by the perturbation theory of Section 7.2 for n large, $T_0(\Psi^n_\alpha)=\infty$ and $\|\Psi^n_\alpha\|_{S(-\infty,0)} \le M$, uniformly in n, which contradicts our assumption that $\|\Psi^\infty_\alpha\|_{S(0,+\infty)} = +\infty$. This is on account of Corollary 7.14, since for every interval $[0,\tilde\tau]$, one may find $[-\tilde\tau_1, 0]$ with the property that the map $\tau \to t_n + \frac{\tau}{\lambda(t_n)}$ takes the latter interval into the former.

Now fix $\tau\in(-\infty,0]$. Then for n sufficiently large, $t_n + \frac{\tau}{\lambda(t_n)} \ge 0$ and $\lambda(t_n+\frac{\tau}{\lambda(t_n)})$ is defined. Let

$$\lambda(t_n+\frac{\tau}{\lambda(t_n)})^{-1}\Psi^\infty_\alpha\Big(\frac{x-x(t_n+\frac{\tau}{\lambda(t_n)})}{\lambda(t_n+\frac{\tau}{\lambda(t_n)})}, t_n+\frac{\tau}{\lambda(t_n)}\Big) = \widetilde\lambda_n(\tau)^{-1}\Psi^n_\alpha\left(\frac{x-\widetilde x_n(\tau)}{\widetilde\lambda_n(\tau)},\tau\right)\in K,$$

with

$$\widetilde\lambda_n(\tau) = \frac{\lambda(t_n+\frac{\tau}{\lambda(t_n)})}{\lambda(t_n)} \ge 1,\ \widetilde x_n(\tau) = x(t_n+\frac{\tau}{\lambda(t_n)}) - \frac{x(t_n)}{\widetilde\lambda_n(\tau)}. \tag{10.39}$$

Now, since $\lambda_n^{-1} f\left(\frac{x-x_n}{\lambda_n}\right) \xrightarrow[n\to\infty]{} f$ strongly in L^2 with either $\lambda_n\to 0$ or $+\infty$, or $|x_n|\to\infty$ implies that $f\equiv 0$, we see that we can assume, after passing to a subsequence, that $\widetilde\lambda_n(\tau)\to\widetilde\lambda(\tau)$, $1\le \widetilde\lambda(\tau)<\infty$ and $\widetilde x_n(\tau)\to\widetilde x(\tau)\in\mathbb{R}^2$. This implies that

$$\widetilde\lambda(\tau)^{-1}\Psi^\dagger_\alpha\left(\frac{x-\widetilde x(\tau)}{\widetilde\lambda(\tau)},\tau\right)\in\overline{K}.$$

Hence, by Proposition 10.7 and 10.8, $\Psi^\dagger_\alpha = 0$, which is a contradiction. □

Proof of Theorem 1.1. We first address global existence and regularity and the global control of the S-norms. In fact, instead of (1.2) we of course require the stronger

$$\|\Psi_\alpha\|_S \le K(E_{crit})$$

from which (1.2) then follows by standard Littlewood-Paley calculus and the Strichartz component of the S-norm. Assume that this strengthened assertion of the theorem fails. Recall that E_{crit} was defined as the smallest energy with the property that there exists an essentially singular sequence of admissible maps at energy E_{crit}. In other words, there exists a sequence $\{\mathbf{u}^n\}_{n=1}^\infty$ of admissible wave maps $(-T_0^n, T_1^n)\times\mathbb{R}^2\to\mathbb{H}^2$ with associated gauged derivative components $\{\psi^n_\alpha\}_{n=1}^\infty$ and such that

- $E(\mathbf{u}^n)\to E_{crit}$
- $\max_{\alpha=0,1,2}\|\psi^n_\alpha\|_{S((-T_0^n,T_1^n)\times\mathbb{R}^2)}\to\infty$

as $n\to\infty$. The Bahouri-Gerard decomposition of Section 9 together with the Kenig-Merle argument of this section now lead to a contradiction whence such an essentially singular sequence cannot exist. This now gives the result, at least up to the scattering statement. As for the latter, we argue as follows. It suffices to carry this out for $\mathbb{H}^2$. Then by applying Lemma 7.6 we may represent the gauged derivative components ψ for any $\delta>0$ in the form

$$\psi = \psi_L^{(\delta)} + \psi_{NL}^{(\delta)}$$

on a time interval of the form (T_0,∞) where $\|\psi^{(\delta)}_{NL}\|_{L^\infty_t L^2_x}<\delta$ and $\psi^{(\delta)}_L$ is a free wave. The scattering for the free wave is automatic, and the $\psi^{(\delta)}_{NL}$ error can be iterated away. □

11. Appendix

11.1. **Completing the proof of Lemma 7.6.** We need to show, see (7.13), that there exist time intervals I_j, $j=1,2,\ldots,M_1$, with M_1 only depending on $\|\psi\|_S,\varepsilon_0$, with the property that

$$\max_{1\le j\le M_1}\sum_{\ell\in\mathbb{Z}}\|P_\ell F_\alpha(\psi)\|^2_{N[\ell](I_j\times\mathbb{R}^2)} < \varepsilon_0 C_0^6 \tag{11.1}$$

Here we need to verify this for F_α of at least quintic degree. In fact, the verification of this is more or less the same for all the higher order terms, and we explain it in detail for a quintic term of first type. From the discussion at the end of Section 6 we see that we may assume the expression to be reduced. Thus consider for example the expression

$$\nabla_{x,t}[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]$$

¿From Lemma 6.1 we infer that

$$\|\nabla_{x,t}P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi_3, P_{k_4}\psi_4)))]\|_{N[k]} \lesssim 2^{-\delta k_0}\|P_{k_0}\psi_0\|_{S[k_0]}$$

It then follows upon square summing over all $k \in \mathbb{Z}$ that the contribution from those expressions with $k_0 \gg k$ in the sense that $k_0 - k > C(\|\psi\|_S, \varepsilon_0)$ may be bounded by $\ll \varepsilon_0\|\psi_0\|_S$. In fact, similar reasoning allows us to reduce to the case when $r_1 < k_0 + O(1)$, $k = k_0 + O(1)$, where the implied constant $O(1)$ may of course be quite large depending on $\|\psi\|_S$ and ε_0, and furthermore we may assume that $k_i = r_j + O(1)$, $i = 1, 2, 3, 4,\ j = 1, 2, 3$. The proof of Lemma 6.1 also implies that we may assume all inputs other than the ones of the null-form $Q_{\nu k}(P_{k_3}\psi_3, P_{k_4}\psi_4)$ to be essentially in the hyperbolic regime, i.e., we may replace $P_{k_j}\psi_j$ by $P_{k_j}Q_{<k_j+O(1)}\psi_j$, $j = 1, 2, 3$, with $O(1)$ as before. Now assume at least one of the inputs of $Q_{\nu k}(P_{k_3}\psi_3, P_{k_4}\psi_4)$ is of elliptic type, in the sense that the difference between its modulation and frequency is large enough. W. l. o. g. write this as

$$\nabla_{x,t}P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{>k_3+C}\psi_3, P_{k_4}\psi_4)))]$$

where the implied constant C is large enough, depending on $\|\psi\|_S, \varepsilon_0$. Then if we write

$$P_{k_3}Q_{>k_3+C}\psi_3 = P_{k_3}Q_{[k_3+C,k_0+10]}\psi_3 + P_{k_3}Q_{>k_0+10}\psi_3,$$

the contribution of the first term on the right is seen to be very small, by placing the output into either $\dot{X}_{k_0}^{-1,-\frac{1}{2},1}$ or $L_t^1\dot{H}^{-1}$. On the other hand, consider now the contribution of the second term on the right. Here one places the output into $\dot{X}_{k_0}^{-\frac{1}{2}+\varepsilon,-1-\varepsilon,2}$ provided the output is in the elliptic regime, or else into $L_1^1\dot{H}^{-1}$. In either case, one verifies that provided $r_1 < -C$ is sufficiently negative, the contribution is small in the above sense. Hence assume now that $r_1 = O(1)$ (which again means an interval depending on $\|\psi\|_S$ as well as ε_0), and as before $P_{k_3}\psi_3 = P_{k_3}Q_{>k_0+10}\psi_3$. Then we may replace $P_{k_4}\psi_4$ by $P_{k_4}Q_{<k_4+\frac{C}{2}}\psi_4$, as otherwise it is again straightforward to see that we gain smallness. Hence we have now reduced to estimating

$$\nabla_{x,t}P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{>k_3+C}\psi_3, P_{k_4}Q_{<k_4+\frac{C}{2}}\psi_4)))],$$

but where now $k_j = r_i + O(1)$ for all i, j, and the output inherits the modulation from the large modulation term $P_{k_3}Q_{>k_3+C}\psi_3$, provided we dyadically localize the latter. But then a straightforward argument using the "divisibility" of $L^2_{t,x}$ reveals that we may pick intervals $\{I_j\}_{j=1}^{M_1}$ with $M_1 = M_1(\|\psi\|_S, \varepsilon_0)$ such that

$$\sum_{k_0\in\mathbb{Z}}\|\nabla_{x,t}\chi_{I_j}\big(P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{>k_3+C}\psi_3, P_{k_4}Q_{<k_4+\frac{C}{2}}\psi_4)))]\big)\|^2_{N[k_0]} < \varepsilon_0$$

Hence we have now reduced to establishing "divisibility" for the space-time frequency reduced expression (with $k_j = r_i + O(1)$ for all i, j)

$$\nabla_{x,t}P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{<k_3+C}\psi_3, P_{k_4}Q_{<k_4+C}\psi_4)))]$$

But since we may estimate this by

$$\begin{aligned}&\|\nabla_{x,t}P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{<k_3+C}\psi_3, P_{k_4}Q_{<k_4+C}\psi_4)))]\|_{N[k_0]}\\ &\lesssim \|P_{k_0}\psi_0\|_{S[k_0]}\|P_{k_1}\psi_1\|_{L_t^4L_x^\infty}\|P_{k_2}\psi_2\|_{L_t^4L_x^\infty}\|\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{<k_3+C}\psi_3, P_{k_4}Q_{<k_4+C}\psi_4)\|_{L_t^2L_x^\infty}\end{aligned}$$

Then use the bound

$$\begin{aligned}&\|\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{<k_3+C}\psi_3, P_{k_4}Q_{<k_4+C}\psi_4)\|_{L_t^2L_x^\infty}\\ &\lesssim \|P_{r_3}Q_{\nu k}(P_{k_3}Q_{<k_3+C}\psi_3, P_{k_4}Q_{<k_4+C}\psi_4)\|_{L_t^2L_x^2}\\ &\lesssim 2^{\frac{r_3}{2}}\prod_{j=3,4}\|P_{k_j}\psi_j\|_{S[k_j]}\end{aligned}$$

which follows from Lemma 4. 16, as well as Bernstein's inequality and our assumptions on the frequencies/modulations. But then again using the "divisibility" of the space $L^2_{t,x}$, we may pick time intervals $\{I_j\}$ as before such that

$$\sum_{k_0\in\mathbb{Z}}\|\nabla_{x,t}\chi_{I_j}P_k[P_{k_0}\psi_0\nabla^{-1}P_{r_1}(P_{k_1}\psi_1\nabla^{-1}P_{r_2}(P_{k_2}\psi_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}Q_{<k_3+C}\psi_3, P_{k_4}Q_{<k_4+C}\psi_4)))]\|^2_{N[k_0]} < \varepsilon_0$$

for all I_j. This furnishes the proof of claim (11.1) for the first type of quintilinear null-form. The remaining short error terms of either first or second type are treated similarly. For the higher order errors of long type (see the discussion at the end of Section 6 for the terminology), the claim follows from Proposition 6.5 as well as the divisibility of $L^8_{t,x}$.

11.2. **Completing the proof of Lemma 7.9.** Recall the setup in the proof of Lemma 7.9: we have a frequency envelope c_k controlling the data ψ at time t_j. We then make the bootstrapping assumption

$$\|P_k\psi\|_{S[k](I_j\times\mathbb{R}^2)} \leq A(C_0)c_k$$

The time intervals I_j have been chosen such that we have a clean separation

$$\psi|_{I_j} = \psi_L^{(j)} + \psi_{NL}^{(j)}$$

where we

$$\sum_{k\in\mathbb{Z}}\|P_k\psi_{NL}^{(j)}\|^2_{S[k](I_j\times\mathbb{R}^2)} < \varepsilon_0$$

$$\|\nabla_{x,t}\psi_L^{(j)}\|_{L_t^\infty\dot{H}^{-1}} \lesssim \|\psi\|_S^3\varepsilon_0^{-\frac{1}{M}}$$

for large M, say $M = 100$. We need to check that by refining each I_j if necessary into finitely many subintervals J_{ji} such that we have

$$\|P_kF_\alpha^{2l+1}(\psi)\|_{N[k](J_{ji}\times\mathbb{R}^2)} \ll c_k$$

where now $l = 2, 3, 4, 5$. We outline the argument for the quintic errors of first type, the remaining ones following a similar pattern. Thus consider the expression

$$\sum_{k_j,r_i}\nabla_{x,t}P_k[P_{k_0}\psi\nabla^{-1}P_{r_1}(P_{k_1}\psi\nabla^{-1}P_{r_2}(P_{k_2}\psi\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi, P_{k_4}\psi)))]$$

By picking M large enough, it is clear that the only contribution that matters is when we replace each factor $P_{k_j}\psi$, $j = 1, 2, 3, 4$, by $P_{k_j}\psi_L$. However, we note here in passing that one can also handle interactions of ψ_L and $\psi_N L$ terms with at least factors ψ_L present by means of the type of "divisibility" argument to follow. Hence consider now

$$\sum_{k_j,r_i}\nabla_{x,t}P_k[P_{k_0}\psi\nabla^{-1}P_{r_1}(P_{k_1}\psi_L\nabla^{-1}P_{r_2}(P_{k_2}\psi_L\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi_L, P_{k_4}\psi_L)))]$$

Due to Proposition 6.1, it is clear that we obtain the desired bound

$$\|\sum_{k_j,r_i}\nabla_{x,t}P_k[P_{k_0}\psi\nabla^{-1}P_{r_1}(P_{k_1}\psi_L\nabla^{-1}P_{r_2}(P_{k_2}\psi_L\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi_L, P_{k_4}\psi_L)))]\|_{N[k]} \ll c_k$$

provided either $|k_0 - k| \gg 1$, and similarly we may assume that $k_j = r_i + O(1)$ for $j = 1, 2, 3, 4$, $i = 1, 2, 3$. Thus we now reduce to estimating the expression where the summation is reduced to $k_0 = k + O(1), k_j = r_i + O(1)$ for $j = 1, 2, 3, 4$, $i = 1, 2, 3$. But in this case, the same type of divisibility argument used in the immediately preceding proof reveals that we may pick intervals J_{ji} whose number depends only on E_{crit} and which are independent of k such that

$$\sum_{r_3=k_3+O(1)=k_4+O(1)}\|P_{r_3}Q_{\nu k}(P_{k_3}\psi_L, P_{k_4}\psi_L)\|^2_{L_t^2\dot{H}^{-\frac{1}{2}}} \ll 1$$

and then the same estimates as in the preceding proof reveal that

$$\|\sum_{k_j,r_i}\nabla_{x,t}P_k[P_{k_0}\psi\nabla^{-1}P_{r_1}(P_{k_1}\psi_L\nabla^{-1}P_{r_2}(P_{k_2}\psi_L\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi_L, P_{k_4}\psi_L)))]\|_{N[k]} \ll c_k,$$

as desired. The argument for the remaining error terms is similar.

11.3. **Completion of the proof of Lemma 7.26.** To complete the proof, we need to show that the contributions of the χ-factors when implementing the Hodge decomposition for the factors of $|\nabla|^{-1}(\psi^2)$ in

$$\sum_{k\in\mathbb{Z}}\|P_k(\psi|\nabla|^{-1}(\psi^2))\|^2_{L^2_t\dot{H}^{-\frac{1}{2}}}$$

is also controllable in terms of $\|\psi\|_S$. Using the schematic relation

$$\chi=|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)],$$

we need to bound

$$\sum_{k\in\mathbb{Z}}\|P_k(\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\|^2_{L^2_t\dot{H}^{-\frac{1}{2}}}$$

$$\sum_{k\in\mathbb{Z}}\|P_k(\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]))\|^2_{L^2_t\dot{H}^{-\frac{1}{2}}}$$

We deal with the first expression, the second being treated along similar lines. Thus consider

$$\begin{aligned}P_k(\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi)) =&P_k(P_{[k-10,k+10]}\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\\&+P_k(P_{>k+10}\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\\&+P_k(P_{<k-10}\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\end{aligned}$$

Start with the first term on the right, the high-low interactions, which we further express as

$$\begin{aligned}&P_k(P_{[k-10,k+10]}\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\\&=\sum_{r<k+15}P_k(P_{[k-10,k+10]}\psi|\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\end{aligned}$$

Now assume the most delicate case, in which we have a high-high-low scenario inside the expression

$$|\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi)$$

with respect to the factors $|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]$, ψ. Thus in this case we can write

$$\begin{aligned}&|\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi)\\&=\sum_{r_1=r_2+O(1)>r+O(1)}|\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}(\psi^2)]P_{r_2}\psi)\\&=\sum_{r_1=r_2+O(1)>r+O(1)}|\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}P_{<r}(\psi^2)]P_{r_2}\psi)\\&+\sum_{r_1=r_2+O(1)>r+O(1)}|\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}P_{\geq r}(\psi^2)]P_{r_2}\psi)\end{aligned}$$

Now observe that for the first factor on the right we have the estimate

$$\begin{aligned}&\|\sum_{r_1=r_2+O(1)>r+O(1)}|\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}P_{<r}(\psi^2)]P_{r_2}\psi)\|_{L^2_tL^\infty_x}\\&=\|\sum_{r_1=r_2+O(1)>r+O(1)}\sum_{c_{1,2}\in\mathcal{D}_{r_1,r-r_1}\,\mathrm{dist}(c_1,-c_2)\lesssim 2^r}\nabla|^{-1}P_r(|\nabla|^{-1}P_{c_1}[\psi|\nabla|^{-1}P_{<r}(\psi^2)]P_{c_2}\psi)\|_{L^2_tL^\infty_x}\\&\lesssim 2^{-r}2^{(1-\varepsilon)(r-r_1)}2^{\frac{r_1}{2}}\|P_{r_1}\psi\|_{S[r_1]}\|P_{r_2}\psi\|_{S[r_2]}\|P_{<r}(\psi^2)]\|_{L^\infty_{t,x}}\\&\lesssim 2^{(1-\varepsilon)(r-r_1)}2^{\frac{r_1}{2}}\|P_{r_1}\psi\|_{S[r_1]}\|P_{r_2}\psi\|_{S[r_2]}\|\psi\|^2_E\end{aligned}$$

Hence we obtain the bound

$$\|P_k(P_{[k-10,k+10]}\psi \sum_{r_1=r_2+O(1)>r+O(1)} |\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}P_{<r}(\psi^2)]P_{r_2}\psi))\|_{L_t^2\dot{H}^{-\frac{1}{2}}}$$
$$\lesssim \sum_{r_1=r_2+O(1)>r+O(1)} 2^{\frac{r_1-k}{2}}2^{(1-\varepsilon)(r-r_1)}2^{\frac{r_1}{2}}\|P_{[k-10,k+10]}\psi\|_{L_t^\infty L_x^2}\|P_{r_1}\psi\|_{S[r_1]}\|P_{r_2}\psi\|_{S[r_2]}\|\psi\|_E^2$$

If we now square this expression and sum over $k \in \mathbb{Z}$, it is straightforward to check that we get the upper bound

$$\lesssim \|\psi\|_S^6\|\psi\|_E^4$$

Next, consider the contribution of the expression

$$\sum_{r_1=r_2+O(1)>r+O(1)} |\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}P_{\geq r}(\psi^2)]P_{r_2}\psi)$$
$$= \sum_{r_1=r_2+O(1)>r+O(1)} \sum_{\tilde{r}\geq r} |\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}P_{\tilde{r}}(\psi^2)]P_{r_2}\psi)$$
$$= \sum_{r_1=r_2+O(1)>r+O(1)} \sum_{\tilde{r}\geq r} \sum_{c_{1,2}\in\mathcal{D}_{r_1,\tilde{r}-r_1}\operatorname{dist}(c_1,-c_2)\lesssim 2^r} |\nabla|^{-1}P_r(|\nabla|^{-1}P_{c_1}[\psi|\nabla|^{-1}P_{\tilde{r}}(\psi^2)]P_{c_2}\psi)$$

Now for fixed $r, \tilde{r}, r_{1,2}$, we can estimate, using as before the improved Strichartz estimates as well as Bernstein's inequality

$$\|\sum_{c_{1,2}\in\mathcal{D}_{r_1,\tilde{r}-r_1}\operatorname{dist}(c_1,-c_2)\lesssim 2^r} |\nabla|^{-1}P_r(|\nabla|^{-1}P_{c_1}[\psi|\nabla|^{-1}P_{\tilde{r}}(\psi^2)]P_{c_2}\psi)\|_{L_t^2L_x^\infty}$$
$$\lesssim 2^{(1-\varepsilon)r}\|\sum_{c_{1,2}\in\mathcal{D}_{r_1,\tilde{r}-r_1}\operatorname{dist}(c_1,-c_2)\lesssim 2^r} P_r(|\nabla|^{-1}P_{c_1}[\psi|\nabla|^{-1}P_{\tilde{r}}(\psi^2)]P_{c_2}\psi)\|_{L_t^2L_x^{1+}}$$
$$\lesssim 2^{\frac{r_1}{2}}2^{(1-\epsilon)(\tilde{r}-r_1)}2^{-(1-\epsilon)\tilde{r}}\prod_{j=1,2}\|P_{r_j}\psi\|_{S[r_j]}\|\psi\|_E^2 \lesssim 2^{\frac{r}{2}}2^{(\frac{1}{2}-\varepsilon)(r-\tilde{r})}\prod_{j=1,2}\|P_{r_j}\psi\|_{S[r_j]}\|\psi\|_E^2,$$

and from here the estimate continues as before. The remaining frequency interactions inside

$$|\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi)$$

are handled similarly and omitted.

Next, consider the case of high-high interactions, i.e.,

$$P_k(P_{>k+10}\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))$$
$$= \sum_{k_1=r+O(1)>k+10} P_k(P_{k_1}\psi|\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))$$

We shall again consider the most delicate case when there are high-high interactions within

$$\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi)$$
$$= \sum_{r_1=r_2+O(1)>r+O(1)} \nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}(\psi^2)]P_{r_2}\psi)$$

But then arguing just as above one obtains the bound

$$\|\nabla|^{-1}P_r(|\nabla|^{-1}P_{r_1}[\psi|\nabla|^{-1}(\psi^2)]P_{r_2}\psi)\|_{L_t^2L_x^{2+}} \lesssim 2^{-(\frac{1}{2}-\varepsilon)r}2^{(\frac{1}{2}-\varepsilon)(r-r_1)}\prod_{j=1,2}\|P_{r_j}\psi\|_{S[r_j]}\|\psi\|_E^2,$$

and from here one obtains

$$\begin{aligned}&\Big\|\sum_{k_1=r+O(1)>k+10} P_k(P_{k_1}\psi|\nabla|^{-1}P_r(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi))\Big\|_{L_t^2\dot{H}^{-\frac12}}\\ &\lesssim \sum_{k_1=r+O(1)>k+10} 2^{(\frac12-\varepsilon)(k-r)}2^{(\frac12-\varepsilon)(r-r_1)}\|P_{k_1}\psi\|_{S[k_1]}\prod_{j=1,2}\|P_{r_j}\psi\|_{S[r_j]}\|\psi\|_E^2\end{aligned}$$

Squaring and summing over k again results in the same bound as before.
The case of low-high interactions, i.e.,

$$P_k(P_{<k-10}\psi|\nabla|^{-1}(|\nabla|^{-1}[\psi|\nabla|^{-1}(\psi^2)]\psi)),$$

is more of the same and omitted.

11.4. **Completion of the proof of Proposition 9.12, part I.** Here we show how to deal with the higher order terms encountered in the decomposition (9.37), i.e., the fifth term there. We shall again explain the method for the quintilinear terms of first type, the remaining higher order terms being treated similarly. Thus consider the expression

$$\nabla_{x,t}P_k[P_{k_0}\rho_0\nabla^{-1}P_{r_1}(P_{k_1}\rho_1\nabla^{-1}P_{r_2}(P_{k_2}\rho_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\rho_3,P_{k_4}\rho_4)))]$$

We use the letter ρ here to imply either a ψ-factor or one of $\epsilon_{1,2}$, the the setup in the proof of Proposition 9.12. Now we distinguish between a number of cases:

(1) At least one factor of both ϵ_1 and ϵ_2 is present. In this case, the entire expression contributes to ϵ_2, as follows from Proposition 6.1. Indeed, we can sum over all k_j, r_j and then square sum over $k\in Z$ and bound the entire expression by

$$\lesssim \|\epsilon_2\|_S\|\epsilon_1\|_S$$

where the implied constant only depends on E_{crit}. By choosing ε_0, which controls $\|\epsilon_1\|_S$, small enough, we can bootstrap.

(2) Only ϵ_1 factors in addition to ψ-factors. First, assume that there are at least two ϵ_1 factors. If one of them is ρ_0, then the output inherits the frequency envelope of ϵ_1 from Proposition 6.1, and the smallness follows from the presence of the extra factor ϵ_1. If the first factor ρ_0 is a ψ, then we need to show that the expression contributes to ϵ_2. But this again follows from Proposition 6.1, essentially as in Case (1) (d) of the proof of Proposition 9.12.
Next, assume that there is only one ϵ_1 factor present. If this factor is not ρ_0, then the expression contributes to ϵ_2, following the same reasoning as in Case (1), (b). Thus assume now that we have $\rho_0=\epsilon_1$, which is the expression

$$\nabla_{x,t}P_k[P_{k_0}\epsilon_1\nabla^{-1}P_{r_1}(P_{k_1}\psi\nabla^{-1}P_{r_2}(P_{k_2}\psi\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi,P_{k_4}\psi))]$$

Recall from the proof of Proposition 6.1 that here ψ really stands for ψ_L or $\psi_N L$, but we suppress this here. What matters is that $\|\psi\|_S$ depends on E_{crit} in a universal way independent of the stage of the iteration in the proof. As usual we may reduce to $k_j=r_i+O(1)$, $j=1,2,3,4$, $i=1,2,3$, and $k_0=k+O(1)>r_1+O(1)$. Furthermore, all inputs may be assumed to be in the hyperbolic regime (up to large constants only depending on E_{crit}). But then the smallness can be forced by shrinking I_j suitably and forcing that

$$\sum_{r\in\mathbb{Z}}\|\chi_{I_j}Q_{\nu k}P_r(P_{r+O(1)}\psi,P_{r+O(1)}\psi)\|^2_{L_t^2\dot{H}^{-\frac12}}\ll 1,$$

see the proof of Proposition 6.1. For the higher order errors of long type (recall the discussion in Section 6), the smallness is achieved by exploiting the "divisibility" of the norms $L^8_{t,x}$.

(3) Only ϵ_2 factors present in addition to factors ψ. All of these terms contribute to ϵ_2. If at least two factors ϵ_2 are present, we clearly obtain the desired smallness from Proposition 6.1. Hence now assume that only one such factor is present. If this factor is in the position of ρ_0, then we obtain smallness via "divisibility" or $L^2_{t,x}$ as in Case (2). If this factor is in the position of some ρ_j with $j=1,2,3,4$, one obtains smallness via a slightly different divisibility argument: first, reduce to the case when ρ_0 and one of

the ρ_j which represents a ψ have angular separation between their Fourier supports: to do this, consider for example

$$\nabla_{x,t}P_k[P_{k_0}\psi\nabla^{-1}P_{r_1}(P_{k_1}\psi\nabla^{-1}P_{r_2}(P_{k_2}\epsilon_2\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\psi,P_{k_4}\psi))]$$

Again we may assume that $k_j = r_i + O(1)$ for $j = 1,2,3,4$, $i = 1,2,3$, and $k_0 = k + O(1) > r_1 + O(1)$. Here we can use the divisibility of $L^2_{t,x}$ by placing $P_{r_3}Q_{\nu k}(P_{k_3}\psi,P_{k_4}\psi)$ into $L^2_t\dot{H}^{-\frac{1}{2}}$, see the proof of Proposition 6.1. On the other hand, for the expression

$$\nabla_{x,t}P_k[P_{k_0}\psi\nabla^{-1}P_{r_1}(P_{k_1}\psi\nabla^{-1}P_{r_2}(P_{k_2}\psi\nabla^{-1}P_{r_3}Q_{\nu k}(P_{k_3}\epsilon_2,P_{k_4}\psi))],$$

one obtains smallness from the divisibility of $L^4_tL^\infty_x$, more precisely, that of

$$\sum_{k\in\mathbb{Z}}\|P_k\psi\|^4_{L^4_tL^\infty_x}$$

11.5. **Completion of the proof of Proposition 9.12, part II.** . Here we show how to obtain the bootstrap for the elliptic part of ϵ, i.e., $Q_{\geq D}\epsilon$. Recall that we solve for $Q_{\geq D}\epsilon$ via the equation

$$\Box Q_{\geq D}\epsilon = Q_{\geq D}[\sum_{i=1}^5 F^{2i+1}_\alpha(\psi+\epsilon)] - Q_{\geq D}[\sum_{i=1}^5 F^{2i+1}_\alpha(\psi)]$$

where the F^{2i+1}_α are obtained as described in Section 3. In particular, $F^3_\alpha(\psi)$ constitutes the trilinear null-forms. Of course the proper interpretation of the right-hand side is that we substitute suitable Schwartz extensions for ψ and ϵ but which agree with the actual dynamic variables on the time interval that we work on. We start by considering the trilinear null-forms, which with the appropriate localizations we schematically write as

$$\nabla_{x,t}P_0Q_{\geq D}[(\psi+\epsilon)\nabla^{-1}\mathcal{Q}_{\nu j}(\psi+\epsilon,\psi+\epsilon)] - \nabla_{x,t}P_0Q_{\geq D}[\psi\nabla^{-1}\mathcal{Q}_{\nu j}(\psi,\psi)]$$

We need to show that we can write the above expression as the sum of two terms, which, when evaluated with respect to $\|\cdot\|_{N[0]}$, improve the bootstrap assumption (9.19). Now we distinguish between various cases:

(1) Here we consider the trilinear terms which are schematically of the form

$$\nabla_{x,t}P_0Q_{\geq D}[\epsilon\nabla^{-1}\mathcal{Q}_{\nu j}(\psi,\psi)]$$

We decompose this into two further terms according to the type of ϵ:

(1a): This is the expression $\nabla_{x,t}P_0Q_{\geq D}[\epsilon_1\nabla^{-1}\mathcal{Q}_{\nu j}(\psi,\psi)]$. Recalling the fine structure of the trilinear terms described in Section 3, we see that this can be decomposed into two types of terms

$$\begin{aligned}&\nabla_{x,t}P_0Q_{\geq D}[\epsilon_1\nabla^{-1}\mathcal{Q}_{\nu j}(\psi,\psi)]\\ &= \nabla_{x,t}P_0Q_{\geq D}[\epsilon_1\nabla^{-1}\mathcal{Q}_{\nu j}I^c(\psi,\psi)] \qquad (11.2)\\ &+ \nabla_{x,t}P_0Q_{\geq D}[(R_\mu)\epsilon_1\nabla^{-1}\mathcal{Q}_{\nu j}I(\psi,\psi)] \qquad (11.3)\end{aligned}$$

where in the last term an operator R_μ may be present or not. Start with the first term on the right, which we write as

$$\begin{aligned}&\nabla_{x,t}P_0Q_{\geq D}[\epsilon_1\nabla^{-1}\mathcal{Q}_{\nu j}I^c(\psi,\psi)]\\ &= \sum_{k_{1,2,3},r}\nabla_{x,t}P_0Q_{\geq D}[P_{k_1}\epsilon_1\nabla^{-1}P_r\mathcal{Q}_{\nu j}I^c(P_{k_2}\psi,P_{k_3}\psi)]\end{aligned}$$

Now the fundamental trilinear estimates in Section 5, see in particular (5.41), imply that under the bootstrap assumption

$$\|P_k\epsilon_1\|_{S[k]} \leq C_4 d_k$$

with some $C_4 = C_4(E_{crit})$, we have

$$\|\sum_{|k_1|\gg 1, k_{2,3}, r} \nabla_{x,t} P_0 Q_{\geq D}[P_{k_1}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2}\psi, P_{k_3}\psi)]\|_{N[0]} \ll C_4 d_0,$$

which is as desired. In fact, the proof of (5.41) cited above implies that one also obtains

$$\|\sum_{k_{1,2,3}, |r|\gg 1} \nabla_{x,t} P_0 Q_{\geq D}[P_{k_1}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2}\psi, P_{k_3}\psi)]\|_{N[0]} \ll C_4 d_0,$$

and finally, again the trilinear estimates from Section 5 imply that we may also assume $k_{2,3} = O(1)$ (implied constant depending on E_{crit}). Hence we may assume for the present term that all frequencies are $O(1)$. Thus we may now reduce to considering

$$\sum_{k_{1,2,3}+O(1)=r=O(1)} \nabla_{x,t} P_0 Q_{\geq D}[P_{k_1}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2}\psi, P_{k_3}\psi)]$$

Now if one of the inputs of the null-form $\mathcal{Q}_{\nu j} I^c(P_{k_2}\psi, P_{k_3}\psi)$ is of elliptic type, either at least one of ϵ_1 and the other input has at least comparable modulation, or else the output inherits the modulation from the large modulation input. In the former case, it is straightforward to obtain smallness: indeed, consider for example

$$\sum_{k_{1,2,3}+O(1)=r=O(1)} \sum_{l\gg 1} \nabla_{x,t} P_0 Q_{\geq D}[P_{k_1} Q_{<l-10}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2} Q_l \psi, P_{k_3} Q_{l+O(1)}\psi)]$$

We can estimate this by (using Bernstein's inequality)

$$\begin{aligned}
&\|\nabla_{x,t} P_0 Q_{\geq D}[P_{k_1} Q_{<l-10}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2} Q_l \psi, P_{k_3} Q_{l+O(1)}\psi)]\|_{N[0]} \\
&= \|\nabla_{x,t} P_0 Q_{[D,l+O(1)]}[P_{k_1} Q_{<l-10}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2} Q_l \psi, P_{k_3} Q_{l+O(1)}\psi)]\|_{N[0]} \\
&\lesssim \sum_{D\leq j\leq l+O(1)} 2^{-\varepsilon j} 2^{\frac{j}{2}} \|R_\nu P_{k_2} Q_l \psi\|_{L^2_{t,x}} \|P_{k_3} Q_{l+O(1)}\psi\|_{L^2_{t,x}} \|P_{k_1} Q_{<l-10}\epsilon_1\|_{L^\infty_t L^2_x} \\
&\ll C_4 d_0
\end{aligned}$$

The case when ϵ_1 has comparable modulation is of course similar. Hence we may assume that if one of the inputs $P_{k_{2,3}}\psi$ is of elliptic type, the output inherits its modulation. In order to obtain smallness in this case, we can form example use divisibility of $L^2_{t,x}$ by applying suitable cutoffs χ_{I_j} for which

$$\sum_{k_2\in\mathbb{Z}} \|\chi_{I_j} R_\nu P_{k_2} Q_{\gg k_2}\psi\|^2_{L^2_{t,x}} \ll 1$$

Next, assume that both inputs $P_{k_{2,3}}\psi$ of the null-form are of hyperbolic type. Then using the bilinear estimates of Section 4, we can estimate

$$\begin{aligned}
&\|\sum_{k_{1,2,3}+O(1)=r=O(1)} \nabla_{x,t} P_0 Q_{\geq D}[P_{k_1}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2} Q_{<k_2+O(1)}\psi, P_{k_3} Q_{<k_3+O(1)}\psi)]\|_{N[0]} \\
&\leq \|\sum_{k_{1,2,3}+O(1)=r=O(1)} \nabla_{x,t} P_0 Q_{\geq D}[P_{k_1}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I^c(P_{k_2} Q_{<k_2+O(1)}\psi, P_{k_3} Q_{<k_3+O(1)}\psi)]\|_{\dot{X}_0^{-\frac{1}{2}+\varepsilon,-1-\varepsilon,2}} \\
&\lesssim \|P_{k_1}\epsilon_1\|_{L^\infty_t L^2_x} \|\mathcal{Q}_{\nu j} I^c(P_{k_2} Q_{<k_2+O(1)}\psi, P_{k_3} Q_{<k_3+O(1)}\psi)\|_{L^2_{t,x}}
\end{aligned}$$

In this case, smallness is again forced by subdividing into suitable time intervals I_j with the property that

$$\|\chi_{I_j} \mathcal{Q}_{\nu j} I^c(P_{k_2} Q_{<k_2+O(1)}\psi, P_{k_3} Q_{<k_3+O(1)}\psi)\|_{L^2_{t,x}} \ll 1$$

This completes treatment of (11.2). Next we turn to (11.3). The same reasoning as for (11.2) shows that we may assume all frequencies $k_{1,2,3}, r$ (which we introduce in the same fashion as before) to be of size $O(1)$. Now a technical issue arises when the operator $R_\nu = R_0$. Indeed, in this case, it may happen that the output inherits the modulation of the first input $P_{k_1} R_\mu \epsilon_1$, and the remaining inputs necessarily need to be placed into the energy space which is "not divisible". However, this problem is somewhat artificial, since of course the Hodge decomposition for the temporal components becomes counterproductive for in the large modulation (elliptic) case. Thus for the expression

$$\sum_{k_{1,2,3}+O(1)=r=O(1)} \nabla_{x,t} P_0 Q_{\geq D}[R_0 P_{k_1} Q_{\gg 1}\epsilon_1 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I(P_{k_2}\psi, P_{k_3}\psi)],$$

it is best to re-combine it with the term

$$\sum_{k_{1,2,3}+O(1)=r=O(1)} \nabla_{x,t} P_0 Q_{\geq D}[R_0 P_{k_1} Q_{\gg 1}\epsilon_2 \nabla^{-1} P_r \mathcal{Q}_{\nu j} I(P_{k_2}\psi, P_{k_3}\psi)],$$

as well as the "elliptic" error χ_0 coming from

$$\epsilon_0 = R_0\epsilon + \chi_0$$

and replace it by $P_{k_1} Q_{\gg 1}\epsilon = P_{k_1} Q_{\gg 1}\epsilon_1 + P_{k_1} Q_{\gg 1}\epsilon_2$. Unfortunately, we encounter here the technical issue that the inputs $\epsilon_{1,2}$, ψ on the right-hand side are really Schwartz extensions of the actual components beyond the time interval I we work on, and hence do not exactly satisfy the div-curl system. The way around this is to work on a slightly smaller time interval $\tilde{I}$ obtained by removing small intervals $I_{1,2}$ from the endpoints of I with $I_{1,2}$ of length $\sim T_1$ with T_1 as in case 1 of the roof of Proposition 9.12. When we restrict the source terms to I, we may invoke the div-curl system for extremely elliptic (i.e., difference of modulation and frequency very large) terms up to negligible errors. This allows us to obtain bootstrapped bounds for $\epsilon_{1,2}$ on $\tilde{I}$, and at the endpoints, we can re-iterate the argument of Case 1. Then the $\epsilon_{1,2}$ on the full interval I can be re-assembled from these pieces via partition of unity with respect to time.
The preceding discussion reveals that we may as well suppress the operator R_μ. But once this is done, the divisibility argument used for (11.2) may be repeated to give the desired smallness upon suitably restricting the time intervals.

(1b): The argument for $\nabla_{x,t} P_0 Q_{\geq D}[\epsilon_1 \nabla^{-1} \mathcal{Q}_{\nu j}(\psi,\psi)]$ is exactly the same, one square sums over the output frequencies instead.

(2): Next we consider the schematically written terms of type $\nabla_{x,t} P_0 Q_{\geq D}[\psi \nabla^{-1} \mathcal{Q}_{\nu j}(\epsilon,\psi)]$. Again these split into two sub-types:

(2a): Terms of type $\nabla_{x,t} P_0 Q_{\geq D}[\psi \nabla^{-1} \mathcal{Q}_{\nu j}(\epsilon_1,\psi)]$. These contribute to ϵ_2, and indeed apart from the fact that one uses trilinear estimates from Section 5 for elliptic outputs, the smallness follows formally just as in Case 1 (b) (of the proof of Proposition 9.12).

(2b): Terms of type $\nabla_{x,t} P_0 Q_{\geq D}[\psi \nabla^{-1} \mathcal{Q}_{\nu j}(\epsilon_2,\psi)]$. Here one encounters again the issue with the terms containing $R_0\epsilon_2$ and of extremely large modulation. As in (1a) above this is handled by undoing the Hodge decomposition for these terms by restricting to a smaller time interval, up to negligible errors. This, as well as arguments as in (1) above, allow one to reduce to an expression of the form

$$\nabla_{x,t} P_0 Q_{\geq D}[P_{k_1}\psi \nabla^{-1} \mathcal{Q}_{\nu j}(P_{k_2}\epsilon_2, P_{k_3}\psi)]$$

where all inputs are of hyperbolic type (up to large constants depending on the energy alone). But then the smallness can be forced by reducing to frequency-separated inputs $P_{k_{1,3}}\psi$ (via the estimates of Section 5.3. But then divisibility is obtained by grouping the inputs $P_{k_{1,3}}\psi$ together and placing their product into $L^2_{t,x}$.

(3) The remaining trilinear null-forms with elliptic output are easier to handle, since they contain at least two factors of type $\epsilon_{1,2}$, and hence the smallness follows simply by the smallness assumptions on these factors (bootstrap assumptions), as well as the trilinear estimates of Section 5. We omit the details.

The higher order contributions from the

$$Q_{\geq D}[\sum_{i=2}^{5} F_{\alpha}^{2i+1}(\psi+\epsilon)] - Q_{\geq D}[\sum_{i=2}^{5} F_{\alpha}^{2i+1}(\psi)]$$

are estimated in a similar vein and omitted.

References

[1] Bahouri, H., Gérard, P. *High frequency approximation of solutions to critical nonlinear wave equations.* Amer. J. Math. 121 (1999), no. 1, 131–175.

[2] Bourgain, J. *Estimates for cone multipliers.* Geometric aspects of functional analysis (Israel, 1992–1994), 41–60, Oper. Theory Adv. Appl., 77, Birkhäuser, Basel, 1995.

[3] Cazenave, T., Shatah, J., Tahvildar-Zadeh, A. *Shadi Harmonic maps of the hyperbolic space and development of singularities in wave maps and Yang-Mills fields.* Ann. Inst. H. Poincaré Phys. Théor. 68 (1998), no. 3, 315–349.

[4] Christodoulou, D., Tahvildar-Zadeh, A. *On the asymptotic behavior of spherically symmetric wave maps.* Duke Math. J. 71 (1993), no. 1, 31–69.

[5] Christodoulou, D., Tahvildar-Zadeh, A. *On the regularity of spherically symmetric wave maps.* Comm. Pure Appl. Math. 46 (1993), no. 7, 1041–1091.

[6] Côte, R., Kenig, C., Merle, F. *Scattering below critical energy for the radial 4D Yang-Mills equation and for the 2D corotational wave map system.* Comm. Math. Phys. 284 (2008), no. 1, 203–225.

[7] Duoandikoetxea, J. *Fourier analysis.* Graduate Studies in Mathematics, 29. American Mathematical Society, Providence, RI, 2001.

[8] Fabes, E., Kenig, C., Serapioni, R. *The local regularity of solutions of degenerate elliptic equations.* Comm. Partial Differential Equations 7 (1982), no. 1, 77–116.

[9] Freire, A., Müller, S., Struwe, M. *Weak convergence of wave maps from $(1+2)$-dimensional Minkowski space to Riemannian manifolds.* Invent. Math. 130 (1997), no. 3, 589–617.

[10] Foschi, D., Klainerman, S. *Bilinear space-time estimates for homogeneous wave equations.*, Ann. Sci. École Norm. Sup. (4) 33 (2000), no. 2, 211–274.

[11] Gérard, P. *Oscillations and concentration effects in semilinear dispersive wave equations.* J. Funct. Anal. 141 (1996), no. 1, 60–98.

[12] Grotowski, J., Shatah, J. *Geometric evolution equations in critical dimensions.* Calc. Var. Partial Differential Equations 30 (2007), no. 4, 499–512.

[13] Kenig, C., Merle, F. *Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation.* Acta Math. 201 (2008), no. 2, 147–212.

[14] Kenig, C., Merle, F. *Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case.* Invent. Math. 166 (2006), no. 3, 645–675.

[15] Klainerman, S., Machedon, M. *On the optimal local regularity for gauge field theories.* Differential Integral Equations 10 (1997), no. 6, 1019–1030.

[16] Klainerman, S., Machedon, M. *On the regularity properties of a model problem related to wave maps.* Duke Math. J. 87 (1997), no. 3, 553–589.

[17] Klainerman, S., Machedon, M. *Smoothing estimates for null forms and applications. A celebration of John F. Nash, Jr.* Duke Math. J. 81 (1995), no. 1, 99–133.

[18] Klainerman, S., Rodnianski, I. *On the global regularity of wave maps in the critical Sobolev norm.* Internat. Math. Res. Notices 2001, no. 13, 655–677.

[19] Klainerman, S., Selberg, S. *Remark on the optimal regularity for equations of the wave maps type.* Comm. Partial Differential Equations 22 (1997), no. 5-6, 901–918.

[20] Klainerman, S., Selberg, S. *Bilinear estimates and applications to nonlinear wave equations.* Commun. Contemp. Math. 4 (2002), no. 2, 223–295.

[21] Klainerman, S., Tataru, D. *On the optimal local regularity for Yang-Mills equations in R^{4+1}.* J. Amer. Math. Soc. 12 (1999), no. 1, 93–116.

[22] Krieger, J. *Global regularity of wave maps from $\mathbf{R}^{2+1}$ to H^2. Small energy.* Comm. Math. Phys. 250 (2004), no. 3, 507–580.

[23] Krieger, J. *Null-form estimates and nonlinear waves.* Adv. Differential Equations 8 (2003), no. 10, 1193–1236.

[24] Krieger, J. *Global regularity of wave maps from R^{3+1} to surfaces.* Comm. Math. Phys. 238 (2003), no. 1-2, 333–366.

[25] Krieger, J., Schlag, W., Tataru, D. *Renormalization and blow up for charge one equivariant critical wave maps.* Invent. Math. 171 (2008), no. 3, 543–615.

[26] Lemaire, L. *Applications harmoniques de surfaces riemanniennes.* J. Differential Geom. 13 (1978), no. 1, 51–78.

[27] Lions, P. L. *The compensated compactness principle in the calculus of variations. The locally compact case I.* Ann. Inst. H. Poincaré Anal. Non Lineaire 1 (1984), 109–145.

[28] Merle, F. *Existence of blow-up solutions in the energy space for the critical generalized KdV equation*, J. Amer. Math. Soc. **14** (2001), 555–578.

[29] Merle, F., Zaag, H. *Determination of the blow-up rate for the semilinear wave equation.* Amer. J. Math. 125 (2003), no. 5, 1147–1164.
[30] Merle, F., Zaag, H. *A Liouville theorem for vector-valued nonlinear heat equations and applications.* Math. Ann. 316 (2000), no. 1, 103–137.
[31] Métivier, G., Schochet, S. *Trilinear resonant interactions of semilinear hyperbolic waves.* Duke Math. J. 95 (1998), no. 2, 241–304.
[32] Mockenhaupt, G. *A note on the cone multiplier.* Proc. Amer. Math. Soc. 117 (1993), no. 1, 145–152.
[33] Mockenhaupt, G., Seeger, A., Sogge, C. D. *Wave front sets, local smoothing and Bourgain's circular maximal theorem.* Ann. of Math. (2) 136 (1992), no. 1, 207–218.
[34] Nahmod, A., Stefanov, A., Uhlenbeck, K. *On the well-posedness of the wave map problem in high dimensions.* Comm. Anal. Geom. 11 (2003), no. 1, 49–83.
[35] Qing, J. *Boundary regularity of weakly harmonic maps from surfaces.* J. Funct. Anal. 114 (1993), no. 2, 458–466.
[36] Rodnianski, I., Sterbenz, J. *On the Formation of Singularities in the Critical $O(3)$ Sigma-Model.* To appear in Annals of Math.
[37] Schoen, R., Yau, S. T. *Lectures on harmonic maps.* Conference Proceedings and Lecture Notes in Geometry and Topology, II. International Press, Cambridge, MA, 1997.
[38] Segovia, C., Torrea, J. *Weighted inequalities for commutators of fractional and singular integrals.* Conference on Mathematical Analysis (El Escorial, 1989). Publ. Mat. 35 (1991), no. 1, 209–235.
[39] Shatah, J. *Weak solutions and development of singularities of the* SU(2) *σ-model.* Comm. Pure Appl. Math. 41 (1988), no. 4, 459–469.
[40] Shatah, J., Struwe, M. *Geometric wave equations.* Courant Lecture Notes in Mathematics, 2. New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1998.
[41] Shatah, J., Struwe, M. *The Cauchy problem for wave maps.* Int. Math. Res. Not. 2002, no. 11, 555–571.
[42] Shatah, J., Tahvildar-Zadeh, A. *On the Cauchy problem for equivariant wave maps.* Comm. Pure Appl. Math. 47 (1994), no. 5, 719–754.
[43] Shatah, J., Tahvildar-Zadeh, A. *Regularity of harmonic maps from the Minkowski space into rotationally symmetric manifolds.* Comm. Pure Appl. Math. 45 (1992), no. 8, 947–971.
[44] Stein, E. *Harmonic Analysis.* Princeton, 1994.
[45] Sterbenz, J., Tataru, D. *Regularity of Wave-Maps in dimension $2+1$.* Preprint 2009.
[46] Sterbenz, J., Tataru, D. *Energy dispersed large data wave maps in $2+1$ dimensions.* Preprint 2009.
[47] Struwe, M. *Variational methods, Applications to Nonlinear PDEs and Hamiltonian Systems.* Second edition, Springer Verlag, New York 1996.
[48] Struwe, M. *Equivariant wave maps in two space dimensions.* Comm. Pure Appl. Math. 56 (2003), no. 7, 815–823.
[49] Tao, T. *Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates.* Math. Z. 238 (2001), no. 2, 215–268.
[50] Tao, T. *An inverse theorem for the bilinear L^2 Strichartz estimate for the wave equation.* Preprint 2009.
[51] Tao, T. *Global regularity of wave maps VII. Control of delocalised or dispersed solutions.* Preprint 2009.
[52] Tao, T. *Global regularity of wave maps VI. Abstract theory of minimal-energy blowup solutions.* Preprint 2009.
[53] Tao, T. *Global regularity of wave maps V. Large data local wellposedness and perturbation theory in the energy class.* Preprint 2008.
[54] Tao, T. *Global regularity of wave maps IV. Absence of stationary or self-similar solutions in the energy class.* Preprint 2008.
[55] Tao, T. *Global regularity of wave maps III. Large energy from R^{1+2} to hyperbolic spaces.* Preprint 2008.
[56] Tao, T. *Global regularity of wave maps II. Small energy in two dimensions.* Comm. Math. Phys. 224 (2001), no. 2, 443–544.
[57] Tao, T. *Global regularity of wave maps. I. Small critical Sobolev norm in high dimension.* Internat. Math. Res. Notices 2001, no. 6, 299–328.
[58] Tataru, D. *Rough solutions for the wave maps equation.* Amer. J. Math. 127 (2005), no. 2, 293–377.
[59] Tataru, D. *The wave maps equation.* Bull. Amer. Math. Soc. 41 (2004), no. 2, 185–204.
[60] Tataru, D. *Null form estimates for second order hyperbolic operators with rough coefficients.* Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001), 383–409, Contemp. Math., 320, Amer. Math. Soc., Providence, RI, 2003.
[61] Tataru, D. *On global existence and scattering for the wave maps equation.* Amer. J. Math. 123 (2001), no. 1, 37–77.
[62] Tataru, D. *Local and global results for wave maps. I.* Comm. Partial Differential Equations 23 (1998), no. 9-10, 1781–1793.
[63] Wolff, T. *A sharp bilinear cone restriction estimate.* Ann. of Math. (2) 153 (2001), no. 3, 661–698.

JK: Department of Mathematics, The University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104, U.S.A.

WS: Department of Mathematics, The University of Chicago, 5734 South University Avenue, Chicago, IL 60615, U.S.A.

TABLE 2. Table of notations

Notation	Meaning	Instance
$\mathbf{u}$	wave map from $\mathbb{R}^{2+1}$ into $\mathbb{H}^2$	(1.1)
$\mathbf{x}, \mathbf{y}$	components of $\mathbf{u}$ in standard coordinates	(1.4)
$\mathbf{e_1}, \mathbf{e_2}$	standard orthonormal frame for $T\mathbb{H}^2$	before (1.4)
ϕ_α^j	derivative components of $\mathbf{u}$ with respect to $\mathbf{e_{1,2}}$	(1.5), (1.6)-(1.9)
ψ_α, ψ_α^j	Coulomb derivative components	(1.10)
ψ_α^n, ϕ_α^n, $\mathbf{x}^n, \mathbf{y}^n$	Components of an essentially singular sequence of wave maps	After Proposition 9.1
χ_β	elliptic term in the Hodge decomposition of ψ_β	(1.16)
E_{crit}	minimal blow up energy	(1.23)
t_ω, x_ω	null-frame coordinates	(2.5)
$\\|\cdot\\|_{S[k,\kappa]}$	null-frame component of the frequency localized norm $\\|\cdot\\|_{S[k]}$	(2.12)
$\\|\cdot\\|_{S[k]}$	Norm used to control the frequency localized Coulomb components $P_k\psi$	(2.17)
$\\|\cdot\\|_{\dot{X}_k^{p,q,r}}$	homogeneous Besov $X^{s,b}$-type norm	(2.1)
$\\|\cdot\\|_{N[\kappa]}$	null-frame component of the norms $\\|\cdot\\|_{N[k]}$ used to control the nonlinear source terms	(2.11)
$\\|\cdot\\|_{N[k]}$	Norm used to control the nonlinear source terms	Definition 2.9
$P_k, Q_j, Q_j^\pm$	Frequency and modulation cutoffs	Before (2.1)
$R_{k,\kappa}$	Rectangular slabs in Fourier space	(2.1)
ϕ_α^{na}	Frequency atoms of the Bahouri-Gerard frequency decomposition of the ϕ_α^n	Lemma 9.5
w_α^{nA}	weakly small error in Bahouri-Gerard frequency decomposition	Lemma 9.5
$\Phi_\alpha^{nA_0^{(0)}}$	Lowest frequency nonatomic derivative components	Lemma 9.7
$\Psi_\alpha^{nA_0^{(0)}}$	Lowest frequency nonatomic derivative Coulomb components	Prop. 9.9
$c_k^{(l)}, d_k$	various frequency envelopes	Prop. 9.11, Prop. 9.12
A_β	various versions of the Coulomb potential	After Prop. 9.14
$\Box_{A^n}$	Covariant wave operator	Def. 9.18
$S_{A^n}\big(u[0]\big)$	Covariant wave propagation associated with $\Box_{A^n}$ applied to data $u[0]$	Def. 9.22
V_ζ^{ab}	Concentration profiles obtained in second stage of covariant Bahouri-Gerard	Lemma 9.23
W_ζ^{naB}	Weakly small error of profile decomposition in second stage of covariant Bahouri-Gerard	Lemma 9.23