

ALMOST OPTIMAL LOCAL WELL-POSEDNESS  
 FOR THE (3+1)-DIMENSIONAL  
 MAXWELL–KLEIN–GORDON EQUATIONS

MATEI MACHEDON AND JACOB STERBENZ

1. INTRODUCTION

This paper contains a detailed study of the local-in-time regularity properties of the Maxwell–Klein–Gordon (MKG) equations. The MKG equations represent a physical model for the interaction of a spin 0 particle, sometimes referred to as a Higgs or scalar field, with the classical equations of electrodynamics, i.e., Maxwell’s equations. In terms of the calculus of variations, the MKG system consists of the Euler–Lagrange equations associated to the density

$$L = -\frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - \frac{1}{2}D_{\alpha}\phi\overline{D^{\alpha}\phi},$$

where

$$F_{\alpha\beta} = \partial_{\alpha}A_{\beta} - \partial_{\beta}A_{\alpha}$$

is the electro-magnetic field associated to the set  $\{A_{\alpha}\}$  of real-valued potentials on the Minkowski space,

$$D_{\alpha} = \partial_{\alpha} + \sqrt{-1}A_{\alpha}$$

is the corresponding covariant derivative with respect to those potentials, and

$$\phi : \mathbb{R}^{3+1} \rightarrow \mathbb{C}$$

is a complex scalar field. In terms of the field  $\{F^{\alpha\beta}\}$ , the Euler–Lagrange equations read

$$(1a) \quad \partial_{\alpha}F^{\alpha\beta} = -\Im\left(\phi\overline{D^{\beta}\phi}\right),$$

$$(1b) \quad D_{\alpha}D^{\alpha}\phi = 0.$$

We note here that our Minkowski metric is  $\eta_{\alpha\beta} := \text{diag}(-1, 1, 1, 1)$ , and we are observing the usual summation convention with respect to repeated upper and lower indices (although repeated lower or upper indices will not be summed over). Also, Greek indices run over the set  $\{0, 1, 2, 3\}$ , while Latin indices have values  $\{1, 2, 3\}$ .

Expanding out the contraction  $\partial_{\alpha}F^{\alpha\beta}$  in terms of the field potentials and setting  $\square := \partial_{\alpha}\partial^{\alpha}$ , we can write the system (1) as the following set of nonlinear second

---

Received by the editors October 16, 2002.

2000 *Mathematics Subject Classification*. Primary 35Q60, 35L70.

Both authors were supported by NSF grant DMS-0100406.

order equations:

$$(2a) \quad \square A^\beta = -\Im \left( \overline{\phi} D^\beta \phi \right) + \partial^\beta \partial_\alpha A^\alpha ,$$

$$(2b) \quad \square \phi = -2\sqrt{-1} (A_\alpha \partial^\alpha \phi) + A_\alpha A^\alpha \phi - \sqrt{-1} (\partial_\alpha A^\alpha) \phi .$$

A key feature of the formulation (2) of the Maxwell–Klein–Gordon equations is that they exhibit an  $SU(1)$  *gauge symmetry*. This means that solutions to this system are invariant (still solutions) if one performs the *gauge transformations*

$$\begin{aligned} A_\alpha &\rightsquigarrow A_\alpha + \partial_\alpha \varphi , \\ \phi &\rightsquigarrow e^{-i\varphi} \phi , \end{aligned}$$

where  $\varphi$  is some real-valued function on Minkowski space-time. This facet of the equations will be central to our local analysis because it allows us to impose a choice of gauge that causes several striking and deep cancellations in the nonlinear terms on the right-hand side of (2b).

The main problem we will be concerned with here is the low-regularity local-in-time Cauchy (or initial value) problem for the system (2). That is, on the slice  $\{t = 0\} \times \mathbb{R}^3$ , one poses initial data at some given time, say  $t = 0$ ,

$$(3a) \quad \{A_\alpha(0)\} = \{a_\alpha\} , \quad \{\partial_t A_\alpha(0)\} = \{a'_\alpha\} ,$$

$$(3b) \quad \phi(0) = f , \quad \partial_t \phi(0) = g ,$$

which satisfy the constraint equation

$$(4) \quad \Delta a_0 = \frac{1}{2\sqrt{-1}} (\bar{f}g - f\bar{g}) + a_0 |f|^2 + \partial^i a'_i ,$$

and then solves the system (2) for these initial conditions. It is now of interest to analyze these solutions to understand what properties can be read off from the initial data. In particular, we will be interested here in discovering how much smoothness and decay the initial data need to possess in order that there be a unique (in an appropriate sense) solution to the system (2). This follows a long line of investigation initiated in a series of papers by S. Klainerman and M. Machedon; see [12], [14], and [15]. For an introduction to the general program for semilinear<sup>1</sup> equations see [13], and see the next subsection for a brief history of the equations we are considering here. The search for the optimal local well-posedness of equations like (2) begins with the following simple observation: Note that solutions to (2) are invariant under the scale transformation

$$(5) \quad \{A_\alpha(\cdot)\} \rightsquigarrow \{\lambda A_\alpha(\lambda \cdot)\} ,$$

$$(6) \quad \phi(\cdot) \rightsquigarrow \lambda \phi(\lambda \cdot) .$$

Also, note that at fixed time, the homogeneous Sobolev space  $\dot{H}^{\frac{1}{2}}$  is also invariant with respect to the same scale transformations. Thus, a local existence theorem for initial data with regularity  $\dot{H}^{\frac{1}{2}}$  would immediately imply global existence (at the same regularity even!) simply by scaling the initial data. Therefore, as one reaches the critical regularity, that is the scale invariant Sobolev space, the local regularity properties of the equations (2) begin to blend with the global regularity properties (at least for small data). At this point, the fine structure of the equations (2) comes

<sup>1</sup>There has also been a recent push toward understanding these questions for more general quasilinear wave equations, in particular for the Einstein equations; see for example [20], [21] and [36].

into focus, and one begins to see striking differences between these equations and more general nonlinear wave equations that have very different local and global regularity properties.<sup>2</sup>

The local and global existence theory for the nonlinear wave equations of mathematical physics has seen exciting and fast-paced progress recently for a particular set of equations, those of the wave-maps from Minkowski space into a complete Riemannian manifold. See for instance [32] and [34], [30]–[31], [19], [25], [10], [11], and finally [35]. While the small data regularity problem for these equations is now largely complete, the corresponding problem for gauge field equations seems to be much more difficult and is still far from understood.<sup>3</sup>

In order to understand (one reason) why progress for the Maxwell–Klein–Gordon equations has been much slower than for wave-maps, it is helpful to compare the difference in the connection structure inherent in these two sets of equations.<sup>4</sup> Let us suppose for the moment that we only need to control the equation

$$(7) \quad D_\alpha D^\alpha \phi = 0 .$$

Now, if the field potentials (connection)  $\{A_\alpha\}$  had vanishing curvature, the equation (7) would be easy to estimate because one could make a gauge change so that it would become  $\square\phi = 0$ . However, in general the potentials will have a nontrivial curvature, say  $F$ , and the most we can hope for is that we can get enough estimates on  $F$  to control (7). For wave-maps, this can be accomplished because the equations take (roughly) the form (7) where the curvature satisfies

$$(8) \quad F_{WM} \approx (\nabla\phi)^2 .$$

Unfortunately, for the Maxwell–Klein–Gordon equations, a short calculation shows that the curvature itself must be a solution to a nonlinear wave equation, i.e., one has

$$(9) \quad \square F_{MKG} \approx Q(\phi, \bar{\phi}) ,$$

where  $Q$  stands for a tensor whose components contain the “null forms”  $Q_{\alpha\beta}(\phi, \psi) := \partial_\alpha\phi\partial_\beta\psi - \partial_\beta\phi\partial_\alpha\psi$ .

Although the “null forms” on the right-hand side of (9) do exhibit extra cancellations with respect to the symbol of  $\square$ , it is not hard to see that from the point of view of space-time estimates, controlling an equation of the form (9) is much more difficult than controlling the expression (8). In fact, the curvature for the MKG equations is so ill behaved that there seems to be no space-time estimates for it once one gets close to the scaling (that is at regularities  $H^{\frac{1}{2}+\epsilon}$  for  $\epsilon$  small) other than  $L^\infty(L^2)$  with the appropriate number of derivatives. That is, as one gets down to the scaling, one starts to lose dispersion for the field potentials  $\{A_\alpha\}$ .

<sup>2</sup>As a simple example, take a generic system of equations of the form  $\square\Psi^I = \Psi^I\nabla\Psi^I$ . It can be shown that these will not be locally well posed for regularities at the level of  $H^1$  or below (see [24]), even though they have the same scale transformations as (5)–(6). Furthermore, smooth solutions to these generic equations will blow up in finite time even for small initial data. This is in stark contrast to the equations we consider here, which exhibit both improved local regularity, and global stability; see [26].

<sup>3</sup>After this paper had been submitted, there has been exciting progress by I. Rodnianski and T. Tao for the MKG equations in higher dimensions at the critical regularity; see [27].

<sup>4</sup>This discussion is really relevant only for wave-maps at the scaling, but it is still useful to keep in mind because the connection structure plays a major role for the MKG equations with regularities at the level of  $H^1$  and below.

See Remark 3.1 for a more detailed discussion. This is a huge blow to any usual type of iteration procedure, because dispersive type space-time estimates of the form<sup>5</sup>  $L^q(L^r)$  play a central role in obtaining inductive estimates.<sup>6</sup> Nevertheless, it turns out that the dangerous part of the curvature  $F_{MKG}$  which does not exhibit any extra space-time estimates miraculously cancels itself where it appears in the equation (7) for the scalar field  $\phi$ ! This allows one to recover enough space-time estimates for solutions to that equation so that one may prove local existence and uniqueness (again, in the appropriate sense) for the system (2) with initial data taken in the Sobolev spaces  $H^s$  for any  $\frac{1}{2} < s$ . This provides the first example of an equation coming from  $(3+1)$ -dimensional gauge field theory where (part of) the local well-posedness conjecture contained in [13] can be verified.

**1.1. Use of the Coulomb gauge and progress based on the model equations.** In order to exploit the extra cancellations occurring in the equation (9), as well as the algebraic structure on the left-hand side of (2b), we shall couple the system (2) with the equation

$$(10) \quad (\operatorname{div})\{A_i\} = 0 .$$

An important property of the field potentials  $\{A_\alpha\}$  which satisfy the condition (10) is that they can be recovered easily from their curvature via the equations

$$(11) \quad A_0 = -\frac{1}{\Delta}\partial_i F^{i0} , \quad A_j = \frac{1}{\Delta}\partial_i F^{ij} .$$

This means that estimates for the curvature,  $\{F_{\alpha\beta}\}$ , translate in a straightforward way into estimates for the  $\{A_\alpha\}$ .

The other important effect of the gauge condition (10) is the cancellations which it causes in the contraction  $A_\alpha\partial^\alpha\phi$ . To see these, note that by a direct application of (10), the system (2) can be written as

$$(12a) \quad \square A_i = -\Im(\phi\overline{D_i\phi}) + \partial_i\partial_0 A^0 ,$$

$$(12b) \quad \Delta A_0 = -\Im(\phi\overline{D_0\phi}) ,$$

$$(12c) \quad \square\phi = -2\sqrt{-1}(A_\alpha\partial^\alpha\phi) + A_\alpha A^\alpha\phi + \sqrt{-1}(\partial_0 A_0)\phi ,$$

$$(12d) \quad \partial_i A^i = 0 .$$

We now use the fact that the spatial field potentials in (12) are divergence free by applying the projection  $\mathcal{P} := -\frac{(\operatorname{curl})^2}{\Delta}$  to both sides of the equation (12a). Using the fact that  $(\operatorname{curl})\nabla_x \equiv 0$  and that the resulting equation for the  $\{A_i\}$  now forces them to be divergence free (so we can drop (12d)), we can rewrite the above system

<sup>5</sup>These are mixed Lebesgue spaces to be defined in a moment.

<sup>6</sup>Even worse, this failure of dispersion for the gauge potentials raises serious questions as to what happens in the case of nonabelian gauge fields. It would seem at first glance that this could cause the Yang–Mills equations to be ill-posed close to the scale invariant Sobolev space (also  $\dot{H}^{\frac{1}{2}}$ ). Perhaps incorporating the gauge structure into the iteration scheme in a more fundamental way could fix this problem.

of equations as

$$(13a) \quad \square A_i = -\Im \mathcal{P}(\phi \overline{D_i \phi}) ,$$

$$(13b) \quad \Delta A_0 = -\Im(\phi \overline{D_0 \phi}) ,$$

$$(13c) \quad \square \phi = -2\sqrt{-1}(A_\alpha \partial^\alpha \phi) + A_\alpha A^\alpha \phi + \sqrt{-1}(\partial_0 A_0)\phi ,$$

$$(13d) \quad \Delta \partial_t A_0 = -\partial^i \Im(\phi \overline{D_i \phi}) .$$

We remark that equations (13a)–(13c) uniquely determine a smooth solution for given smooth initial data satisfying the condition (12d), and that (13d) is automatically satisfied for this solution.

A direct calculation of the quadratic nonlinearity in the equation (13a) shows that it can be written as a sum of terms of the form (no summing over repeated lower indices!)

$$(14) \quad \frac{\partial_i}{\Delta} Q_{ij}(\phi, \overline{\phi}) .$$

Likewise, one can also write the quadratic part of (13c) which contains the spatial field potentials in terms of the  $Q_{ij}$ . To see this, we use (11) to write

$$(15) \quad A_i \partial^i \phi = \partial_i \phi \cdot \frac{1}{\Delta} \partial_j F^{ji} = - \sum_{i < j} Q_{ij}(\phi, \frac{1}{\Delta} F_{ij}) .$$

If we now pay attention to only those parts of the nonlinearity in (13) which contain the null structures (14)–(15), i.e., we throw away the cubic terms as well as any instances of the “elliptic” variable  $A_0$  (although it turns out that this term is *not* easier to treat than the “hyperbolic” terms), we arrive at the following schematic:

$$(16a) \quad \square \psi^I = |D_x|^{-1} Q(\psi^I, \psi^I) ,$$

$$(16b) \quad \square \psi^I = Q(|D_x|^{-1} \psi^I, \psi^I) .$$

We will refer to the system (16) as the *MKG–CG model equations*.

In recent years, the initial value problem for the system (13) has been extensively studied in terms of the model equations (16). The first major breakthrough, due to Klainerman–Machedon [14], was to prove that the system (13) is locally well posed in the energy space  $H^1$ , and hence globally well posed due to the approximate  $H^1$  conservation law for this system, improving the previous result of Eardley and Moncrief, [4]–[5], thus removing the artificial restriction on the smoothness of the initial data imposed in that earlier work. The approach of [14] was based on estimating the oscillatory integral

$$(17) \quad \left\| \int_0^t \sin((t-s)|\xi|) \widehat{G}(s, \xi) ds \right\|_{L_t^\infty(L_x^2)} \lesssim t^{\frac{1}{2}} \|G\|_{L^2([0,t] \times \mathbb{R}^3)} ,$$

via a simple Cauchy–Schwarz, reducing the task of proving well-posedness to show the space-time estimates of the form

$$\begin{aligned} \left\| |D_x|^{-1} Q_{ij}(\psi^I, \psi^I) \right\|_{L^2([0,t] \times \mathbb{R}^3)}^2 &\lesssim \left( \|\psi^I(0, \cdot)\|_{H^1(\mathbb{R}^3)} + \int_0^t \|\square \psi^I(s, \cdot)\|_{L^2(\mathbb{R}^3)} ds \right)^2 , \\ \left\| Q_{ij}(|D_x|^{-1} \psi^I, \psi^I) \right\|_{L^2([0,t] \times \mathbb{R}^3)}^2 &\lesssim \left( \|\psi^I(0, \cdot)\|_{H^1(\mathbb{R}^3)} + \int_0^t \|\square \psi^I(s, \cdot)\|_{L^2(\mathbb{R}^3)} ds \right)^2 , \end{aligned}$$

for arbitrary test functions  $\psi^I$ . While the above estimates are sharp (for solutions of the homogeneous wave equation), their application via Picard iteration to the system (16) falls  $\frac{1}{2}$  a derivative short of the scaling because the bound (17) destroys the oscillations under the integral sign, losing the extra decay for large  $\xi$ .

In order to push local regularity below the energy space  $H^1$ , it is necessary to take advantage of the extra smoothing properties of  $\square$  which come from the oscillations in the integral (17). To accomplish this, one first writes (16) as a system of integral equations (at least for a short time):

$$(18a) \quad \psi^I = \chi_T \left[ \psi^{I(0)} + \frac{1}{\square} \left( |D_x|^{-1} Q(\psi^I, \psi^I) \right) \right],$$

$$(18b) \quad \psi^I = \chi_T \left[ \psi^{I(0)} + \frac{1}{\square} \left( Q(|D_x|^{-1} \psi^I, \psi^I) \right) \right],$$

where  $\chi_T$  is a compactly supported smooth bump function identically equal to one on a fixed time interval containing  $t = 0$ , and  $\frac{1}{\square}$  is the standard wave parametrix which inverts the wave equation with zero initial data. Here  $\psi^{I(0)}$  denotes propagation of the initial data as free waves. This procedure allows one to replace the symbol of  $\square$  with inhomogeneous weights, so well-posedness can be recovered via an iteration of the system (18) in weighted spaces once one can prove bilinear estimates of the form (away from the light cone, of course)

$$(19a) \quad \left\| \frac{1}{\sqrt{\square}} \cdot \frac{1}{|D|^{\frac{3}{2}-s}} Q_{ij}(\psi^I, \psi^I) \right\|_{L^2(\mathbb{R}^{3+1})} \lesssim \left\| \sqrt{\square} \cdot |D|^{s-\frac{1}{2}} \psi^I \right\|_{L^2(\mathbb{R}^{3+1})}^2,$$

$$(19b) \quad \left\| \frac{1}{\sqrt{\square}} \cdot |D|^{s-\frac{1}{2}} Q_{ij}(|D|^{-1} \psi^I, \psi^I) \right\|_{L^2(\mathbb{R}^{3+1})} \lesssim \left\| \sqrt{\square} \cdot |D|^{s-\frac{1}{2}} \psi^I \right\|_{L^2(\mathbb{R}^{3+1})}^2,$$

where  $s < 1$  is the desired regularity. Using  $L^2$  bilinear estimates for the wave equation, Cuccagna [3] was able to show that (a slight variant of) the estimates (19) hold in the range where  $\frac{3}{4} < s < 1$ .

To push beyond the  $s = \frac{3}{4}$  regularity becomes quite a bit more technical. To start with, the estimate (19a) by itself is not true for  $s$  in the range  $s < \frac{2}{3}$ , even if one only considers functions  $\psi^I$  supported in an  $O(1)$  neighborhood of the light cone. To see this, notice that if  $\psi_\chi^{I(0)}$  are solutions to the homogeneous wave equation that have been cut off smoothly in time, then one has the heuristic identity<sup>7</sup>

$$Q_{ij}(\psi_\chi^{I(0)}, \psi_\chi^{I(0)}) \approx \sqrt{\square} \left( |D|^{\frac{1}{2}} \psi_\chi^{I(0)} \cdot |D|^{\frac{1}{2}} \psi_\chi^{I(0)} \right).$$

Thus, (19) is roughly equivalent to trying to estimate

$$\left\| \frac{1}{|D|^{\frac{3}{2}-s}} \left( |D|^{\frac{1}{2}} \psi_\chi^{I(0)} \cdot |D|^{\frac{1}{2}} \psi_\chi^{I(0)} \right) \right\|_{L^2(\mathbb{R}^{3+1})} \lesssim \left\| |D|^s \psi_\chi^{I(0)} \right\|_{L^2(\mathbb{R}^{3+1})}^2.$$

In other words, trying to show that

$$(20) \quad \left\| \frac{1}{|D|^{\frac{3}{2}-s}} \left( |D|^{\frac{1}{2}} \psi^{I(0)} \cdot |D|^{\frac{1}{2}} \psi^{I(0)} \right) \right\|_{L^2([0,1] \times \mathbb{R}^3)} \lesssim \left\| \psi^I(0, \cdot) \right\|_{H^s(\mathbb{R}^3)}^2,$$

where  $\psi^{I(0)}$  are solutions to the homogeneous wave equation.

Using certain standard eccentric initial data sets (see the next subsection and the work [6]), it is possible to show that (20) can only be true in the range  $\frac{2}{3} \leq s$ .

<sup>7</sup> See the null form bound (28).

This failure is a serious setback to setting up an iteration procedure for the model equations, because it means that the first iterate to the corresponding time cutoff system is not in any subspace of the wave-Sobolev spaces  $X^{s, \frac{1}{2}}$  for  $s$  close to the scaling. It is important to notice that this failure is not just a problem for the model; in fact, one sees immediately that the integral (20) would come up when trying to put the first iterate  $A_i^{(1)}$  of the true equations in the  $X^{s, \frac{1}{2}}$  spaces.

*Remark 1.1.* We remark that in dimension  $n = 4$ , the analog of (20) is a true estimate for  $s$  arbitrarily close to the scaling, which in this case is  $s_c = 1$ . The problem is that (19b) is not true for functions other than the  $\psi_\chi^{I(0)}$ , i.e., ones that are not localized in an  $O(1)$  neighborhood of the light cone in Fourier space. However, one can salvage the estimate (19b) for arbitrary test functions by placing some extra norms on the right-hand side. The task is then to be able to recover these additional norms via appropriate bilinear estimates, and thus prove well-posedness of the model (16) for  $s$  arbitrarily close to the scaling  $s_c$ . This has been carried out in [18] for the Maxwell–Klein–Gordon equations using the norms first introduced in [16], and in [23] for the related Yang–Mills. Also, the corresponding critical Besov space result for these equations can most likely be proved by using a combination of the null structures we discuss in this paper and the setup used in [28] to treat higher-dimensional generic equations. However, the problem of well-posedness for the Yang–Mills or Maxwell–Klein–Gordon equations in the critical Sobolev space,  $\dot{H}^1$ , would be much more difficult, and is sure to require the notion of covariant  $X^{s, \theta}$  spaces as well as bilinear estimates for covariant wave equations.

**1.2. Failure of the model and statement of the main result.** In dimension  $n = 3$ , the question still remains if one can somehow get around the problems inherent to (19), and prove well-posedness for the model equations. It turns out that this is not the case. In fact, one can show that any iteration procedure based solely on the model equations will not work. More specifically, we will give a heuristic proof of the following.

**Theorem 1.2.** *For any  $\frac{1}{2} \leq s < \frac{3}{4}$ , there exists a sequence of initial data such that*

$$\|\psi_N^{I(0)}(0)\|_{H^s} \leq 1, \quad \|\partial_t \psi_N^{I(0)}(0)\|_{H^{s-1}} \leq 1,$$

*and such that the second iterate  $\psi_N^{I(2)}$  to the system (18) has the property that*

$$\|\psi_N^{I(2)}(t)\|_{H^s} \rightarrow \infty,$$

*as  $N \rightarrow \infty$ , for any fixed time  $t > 0$ .*

The question of well-posedness now shifts back to the original equations, (13). It is a remarkable fact, and the main result of this paper, that one still has well-posedness of the true equations arbitrarily close to the scaling.

**Theorem 1.3.** *If  $s$  is any number such that  $\frac{1}{2} < s$ , then there exists an  $\epsilon(s) > 0$  with the property that if the initial data for the system (13) are specified such that*

$$\|\phi(0)\|_{H^s}, \quad \|\partial_t \phi(0)\|_{H^{s-1}}, \quad \|\{A_i(0)\}\|_{H^s}, \quad \|\{\partial_t A_i(0)\}\|_{H^{s-1}} \leq \epsilon(s),$$

*with the gauge condition*

$$(\operatorname{div})\{A_i(0)\} = 0, \quad (\operatorname{div})\{\partial_t A_i(0)\} = 0,$$

satisfied, then there exists a sequence of smooth functions  $\phi_n$  and  $\{A_n\}$ , each of which is a solution to the system (13) on the interval  $[-1, 1]$ , and such that

$$\begin{aligned}\phi_n(0) &\rightarrow \phi(0) , & \{A_n(0)\} &\rightarrow \{A(0)\} , \\ \partial_t \phi_n(0) &\rightarrow \partial_t \phi(0) , & \{\partial_t A_n(0)\} &\rightarrow \{\partial_t A(0)\} ,\end{aligned}$$

in  $H^s$ . Furthermore, the above sequence is a Cauchy sequence in the space

$$\| \cdot \|_{C[-1,1](H^s)} \cap \| \cdot \|_{C^{(1)}[-1,1](H^{s-1})} ,$$

and converges on the interval  $[-1, 1]$  to a solution  $\phi, \{A\}$  of (13) on that interval which depends continuously on the data in the sense that

$$\begin{aligned}\| \phi(t, \cdot) \|_{C[-1,1](H^s)} + \| \phi \|_{C^{(1)}[-1,1](H^{s-1})} &\lesssim \| \phi(0) \|_{H^s} + \| \partial_t \phi(0) \|_{H^{s-1}} , \\ \| \{A_i\} \|_{C[-1,1](H^s)} + \| \{A_i\} \|_{C^{(1)}[-1,1](H^{s-1})} &\lesssim \| \{A_i(0)\} \|_{H^s} + \| \{\partial_t A_i(0)\} \|_{H^{s-1}} .\end{aligned}$$

Finally, this is the only solution to the system (13) which can be obtained as a limit of solutions with regularized data.

In fact, we will prove more. We will show that there exists a Banach space  $B^s$  such that our solution to the system (13) on the interval  $[-1, 1]$  can be extended to a solution of an appropriate time cutoff problem, and that this extension is unique in the space  $B^s$ . We leave this out of the statement of the main theorem because  $B^s$  is not a fixed time space, and we prefer the statement of the main theorem to be purely local. Also, our philosophy here is that special function spaces play a secondary role, and that the real focus is on expanding formulas for iterates in terms of previous iterates and controlling the resulting expressions “by hand”.

The remainder of this paper is laid out in the following way. In the next subsection we will give a heuristic proof of Theorem 1.2, which relies on some simplifying assumptions as well as some approximate identities. We will also show there why the counterexample is well posed when plugged into the true equations. We encourage those readers who are familiar with the model equations (16), and the structure of these sort of problems in general, to read this next subsection, and then proceed directly to subsections 3.1 and 3.5, as the major part of this paper is centered around proving enough estimates (which are more or less standard) that things can be reduced to a sum of “counterexamples”, which are essentially the same as the one discussed in the next subsection.

In the second section, we discuss the two main technical tools we borrow from the recent literature. These are the Fourier space decomposition of the symbol of the wave equation due to Tataru (see [33]) and T. Tao’s dual scale machine for proving multilinear estimates for the wave equation, [29] and [22]. We should emphasize here that this paper contains no “new technology” with respect to multilinear estimates for the wave equation. Rather, the main difficulty of this work is in understanding the structure of the equations (13).

In the third section, we set up a time cutoff system for the equations (13) which takes advantage of the extra identity, (13d), for the elliptic term  $A_0$ . We will then break this time cutoff system down in a series of subsections organized around the various terms involved and the interaction in those terms between the “elliptic” and “hyperbolic” variables.

1.3. **Proof of Theorem 1.2.** To understand what can go wrong with the model equations (16), we will work out the first two Picard iterates for that system with certain eccentric initial data sets. Because the complete calculations would be unduly tedious, we will make the following two simplifying assumptions. Firstly, in a usual time-cutoff iteration procedure, the initial Picard iterate would be the time-cutoff solution to the homogeneous wave equation with the given initial data, say

$$\chi_T \phi^{(0)} \quad , \quad \{ \chi_T A^{(0)} \} \quad ,$$

with

$$\begin{aligned} \phi^{(0)}(0) &= \phi(0) \quad , & \{ A^{(0)}(0) \} &= \{ A(0) \} \quad , \\ \partial_t \phi^{(0)}(0) &= \partial_t \phi(0) \quad , & \{ \partial_t A^{(0)}(0) \} &= \{ \partial_t A(0) \} \quad , \\ \square \phi^{(0)} &= 0 \quad , & \square \{ A^{(0)} \} &= 0 \quad . \end{aligned}$$

In our example, we will take  $\{ A^{(0)}(0) \} = \{ \partial_t A^{(0)}(0) \} = 0$ , and we will replace  $\chi_T \phi^{(0)}$  with the following approximate initial iterate:

$$(21) \quad \widetilde{\phi_{approx}^{(0)}}(\tau, \xi) = \int \eta(\tau - \lambda) \widetilde{\phi^{(0)}}(\lambda, \xi) d\lambda,$$

where  $\eta$  is a smooth, even, and positive function such that  $\eta \equiv 1$  on  $[-1, 1]$  and  $\eta$  vanishes outside of  $[-2, 2]$ . It is apparent that (21) is essentially a  $\chi_T \phi^{(0)}$  cutoff in an  $O(1)$  neighborhood of the light cone, and that we have not lost anything essential due to the exponential decay of  $\chi_T \phi^{(0)}$  away from the light cone in Fourier space.

Our second simplifying assumption will be to replace the time-cutoff parametrix to the wave equation occurring in the integral formulation of the system (16), with the following operator which has a simpler behavior in a neighborhood of the light cone, but still has the same smoothing properties as the time-cutoff wave parametrix:

$$\widetilde{VF}(\tau, \xi) = \left[ \frac{1}{1 + |\xi|} \eta(|\tau| - |\xi|) + \frac{1}{\tau^2 - |\xi|^2} \left( 1 - \eta(|\tau| - |\xi|) \right) \right] \widetilde{F}(\tau, \xi) \quad .$$

With these choices in mind, we see that finding two iterates of the system (16) is roughly equivalent to solving the following system of equations:

$$(22) \quad A^{(1)} = V \left[ \frac{(curl)^2}{\Delta} \left( \phi_{approx}^{(0)} \nabla \overline{\phi_{approx}^{(0)}} \right) \right] \quad ,$$

$$(23) \quad \phi^{(2)} = -V \left[ A_i^{(1)} \partial^i \phi_{approx}^{(0)} \right] \quad .$$

We now let  $N$  be a large parameter, which we will allow to go to infinity, and let  $\frac{1}{2} < s < 1$ . We consider initial data such that

$$\widetilde{\phi_{approx}^{(0)}}(\tau, \xi) = \eta(\tau - |\xi|) \frac{1}{N^{\frac{3}{2} + 2s}} \chi_{B(N)}(\xi) \quad ,$$

where  $B(N)$  is the box of dimensions  $N \times N \times N$  oriented along the positive  $\xi_1$  axis at distance  $N^2$  from the origin, as shown in Figure 1.

We will now show that the  $H^s$  norm of  $\phi^{(2)}$  at fixed time is unbounded as a function of  $N$  if  $s < \frac{3}{4}$ .

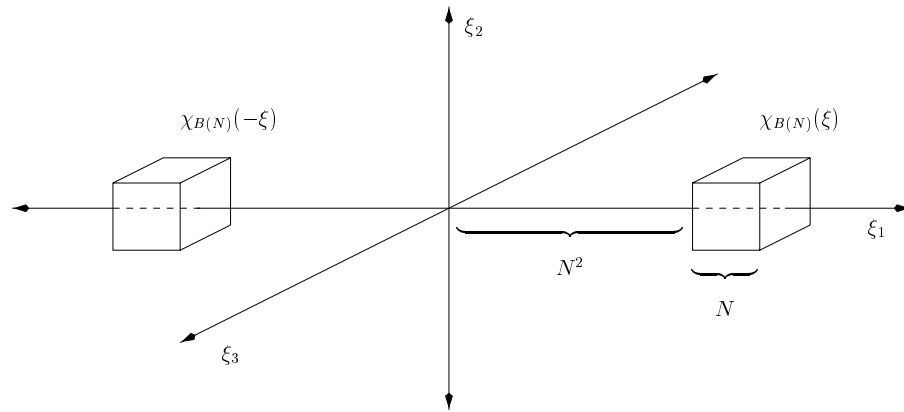


FIGURE 1. Initial data for  $\phi^{(0)}$ .

Our first order of business will be to isolate that portion of  $A$  which contributes the most to the right-hand side of (23). Because the order of  $\partial_2 \phi_{approx}^{(0)}$  and  $\partial_3 \phi_{approx}^{(0)}$  is a factor of  $N^{-1}$  smaller than that of  $\partial_1 \phi_{approx}^{(0)}$ , this is

$$\phi_{worst\ term}^{(2)} = -V \left[ A_1^{(1)} \partial_1 \phi_{approx}^{(0)} \right] .$$

In what follows, we will ignore the lower order terms. The reader can easily see that these are negligible.

A quick calculation involving the supports of  $\widetilde{\phi_{approx}^{(0)}}$  and  $\overline{\phi_{approx}^{(0)}}$  shows that the support of  $\widetilde{A_1^{(1)}}$  is a null plane of thickness  $O(1)$ , directed along the  $\xi_1$  axis. It is then apparent that the support of  $\widetilde{A_1^{(1)}} * \widetilde{\phi_{approx}^{(0)}}$  is nothing but a thickening of the original support of  $\widetilde{\phi_{approx}^{(0)}}$ . With these support considerations in mind, a simple counting of the weights involved shows that we have the approximate identity (ignoring signs in Fourier space for the moment):

$$\phi_{worst\ term}^{(2)} \approx A_1^{(1)} \cdot \phi_{approx}^{(0)} .$$

We now focus our attention on isolating the highest order term of  $A_1^{(1)}$ . A short calculation counting the weights involved shows that this is given by the expression

$$(24) \quad A_1^{(1)}{}_{worst\ term} = -V \left[ \frac{\partial_2^2 + \partial_3^2}{\Delta} \left( \phi_{approx}^{(0)} \partial_1 \overline{\phi_{approx}^{(0)}} \right) \right] .$$

Again, all the other terms in  $A_1^{(1)}$  are of the order  $N^{-1}$  smoother. Using the fact that  $\widetilde{A_1^{(1)}}$  is essentially a null plane directed along the positive  $\xi_1$  axis, we have

$$V(\partial_2^2 + \partial_3^2) \approx 1 ,$$

so we can write (again ignoring signs)

$$A_1^{(1)}{}_{worst\ term} \approx \phi_{approx}^{(0)} \cdot \overline{\phi_{approx}^{(0)}} .$$

With these calculations in mind, we see that the highest order term of  $\phi^{(2)}$  can be written (approximately) as the following trilinear expression of the initial data:

$$(25) \quad \phi_{highest\ term}^{(2)} \approx \phi_{approx}^{(0)} \cdot \overline{\phi_{approx}^{(0)}} \cdot \phi_{approx}^{(0)}.$$

Calculating the above expression in Fourier space reveals that

$$\widetilde{\phi_{approx}^{(0)}} * \overline{\widetilde{\phi_{approx}^{(0)}}} * \widetilde{\phi_{approx}^{(0)}} \approx N^3 \cdot \frac{1}{N^{3+4s}} \cdot N^3 \cdot \frac{1}{N^{\frac{3}{2}+2s}} \eta\left(\frac{1}{4}(\tau - |\xi|)\right) \chi_{B(N)}\left(\frac{1}{4}\xi\right).$$

Testing the above expression for  $H^s$  at a fixed time we see that

$$\|\phi_{highest\ term}^{(2)}(t)\|_{H^s} \approx N^{3-4s},$$

which is unbounded unless  $\frac{3}{4} \leq s$ . We arrive at the conclusion that it is not possible to iterate a time-cutoff problem associated with the model equations (16) without dropping out of the space  $H^s$  at fixed time.

If we plug the above counterexample into the true equations, we immediately come upon the following realization: the expression

$$A_0^{(1)} = -\frac{1}{\Delta} \left( \phi_{approx}^{(0)} \partial_t \overline{\phi_{approx}^{(0)}} \right),$$

is easily seen to be the same size (with the same sign!) as (24) on account of the relation

$$\partial_t \phi_{approx}^{(0)} = \partial_1 \phi_{approx}^{(0)} (1 + O(N^{-1})).$$

This means that for the true equations, the correct expression for the highest order term in  $\phi^{(2)}$  is

$$\phi_{true}^{(2)} = -V \left[ \frac{1}{\Delta} (I - V(\partial_2^2 + \partial_3^2)) (\phi_{approx}^{(0)} \partial_1 \overline{\phi_{approx}^{(0)}}) \partial_1 \phi_{approx}^{(0)} \right].$$

The cancellations between the elliptic and hyperbolic terms now takes the form of the following approximate identity:

$$(26) \quad (I - V(\partial_2^2 + \partial_3^2)) \left( \phi_{approx}^{(0)} \partial_1 \overline{\phi_{approx}^{(0)}} \right) \approx O(N^{-1}) \phi_{approx}^{(0)} \partial_1 \overline{\phi_{approx}^{(0)}}.$$

Strictly speaking, (26) is only true in the region of Fourier space where the support of  $\widetilde{\psi^{(0)}} * \overline{\widetilde{\psi^{(0)}}}$  is a distance  $\sim N$  from the light cone, but it turns out that the symbol of  $(I - V(\partial_2^2 + \partial_3^2))$  contributes enough at every distance to make things work out, so we will just assume that (26) holds everywhere.

This shows that for the true equations, the initial data we have taken for our counterexample yield a  $\phi_{true}^{(2)}$  that is an order  $N^{-1}$  smoother than the  $\phi^{(2)}$  previously constructed. Testing this for  $H^s$  at fixed time yields

$$\|\phi_{true}^{(2)}(t)\|_{H^s} \approx N^{2-4s},$$

which is bounded in the range  $\frac{1}{2} \leq s$ .

## 2. NOTATION AND PRELIMINARY ESTIMATES

For quantities  $A$  and  $B$ , we write  $A \lesssim B$  to mean that  $A \leq C \cdot B$  for some large constant  $C$ . The constant  $C$  may change from line to line, but will always remain fixed for any given instance where this notation appears. Likewise we use the notation  $A \sim B$  to mean that  $\frac{1}{C} \cdot B \leq A \leq C \cdot B$ . We also use the notation  $A \ll B$  to mean that  $A \leq \frac{1}{C} \cdot B$  for some large constant  $C$ . This is the notation we will use throughout the paper to break down quantities into the standard cases:

$A \sim B$ , or  $A \ll B$ , or  $B \ll A$ ; and  $A \lesssim B$ , or  $B \ll A$ , without ever discussing which constants we are using.

For any quantity  $A$ , we will employ the notation  $A+$ , (resp.  $A-$ ) to denote the expression  $A + \epsilon$  (resp.  $A - \epsilon$ ), for  $\epsilon > 0$  and  $\epsilon$  sufficiently small. It is crucial to keep in mind that we *do not* assume any uniformity in the  $A+$  notation. Specifically, the implicit constant in any estimate in which this notation appears may depend on the choice of  $\epsilon$ . We employ this notation to eliminate complicated formulas of small numbers which depend on various fixed parameters in our proof.

For a given function of two variables  $(t, x) \in \mathbf{R} \times \mathbf{R}^3$  we write the spatial and space-time Fourier transform as

$$\begin{aligned} \widehat{f}(t, \xi) &= \int e^{-2\pi i \xi \cdot x} f(t, x) \, dx, \\ \widetilde{f}(\tau, \xi) &= \int e^{-2\pi i(\tau t + \xi \cdot x)} f(t, x) \, dt dx, \end{aligned}$$

respectively. At times, we will also write  $\mathcal{F}[f] = \widetilde{f}$ . We also use the mixed Lebesgue space notation

$$\|f\|_{L^q(L^r)} = \left( \int \|f(t, \cdot)\|_{L^r}^q \, dt \right)^{\frac{1}{q}}.$$

At times, we will use the notation  $L^2$  as a substitute for  $L^2(L^2)$ . Because of this, we will distinguish the spatial version of  $L^2$  with the notation  $L_x^2$ . Also, in certain places in the sequel, it will be convenient for us to employ the same type of notation in Fourier space. Since this will cause us to switch the roles of time and space, we will use a subscript to emphasize the order of integration, for example we will use  $L_\xi^2(L_\tau^1)$  with the obvious meaning. Also, we will often replace  $L_\tau^2(L_\xi^2)$  with the shorter  $L^2$ .

Throughout this paper, we will employ the standard differential multipliers

$$\begin{aligned} \widehat{|D_x|^s} &= |\xi|^s, \\ \widehat{\langle D_x \rangle^s} &= (1 + |\xi|^2)^{\frac{s}{2}}, \\ \widetilde{\langle D_{t,x} \rangle^s} &= (1 + |(\tau, \xi)|^2)^{\frac{s}{2}}. \end{aligned}$$

We will also make use of the standard homogeneous and inhomogeneous Sobolev spaces of the spatial variable. These are defined via the formulas

$$\begin{aligned} \|f\|_{\dot{H}_x^s} &= \||D_x|^s f\|_{L_x^2}, \\ \|f\|_{H_x^s} &= \|\langle D_x \rangle^s f\|_{L_x^2}. \end{aligned}$$

Let  $\varphi$  be a smooth bump function (i.e., supported on the set  $|s| \leq 2$  such that  $\varphi = 1$  for  $|s| \leq 1$ ). For  $\lambda \in 2^{\mathbb{Z}}$ , we denote its dyadic scaling by  $\varphi_\lambda(s) = \varphi(\frac{s}{\lambda})$ . Following Tataru [33], we list the standard Littlewood–Paley type cutoff functions used to decompose the symbol of the wave and Laplace equations in Table 1.

For each of the cutoff functions listed in Table 1, we will use the same uppercase letter to denote the corresponding Fourier multiplier operator, e.g.  $\widehat{P_\lambda f} = p_\lambda \widehat{f}$ . We also use a multiindex notation to denote certain compositions of these operators:

$$\begin{aligned} P_{\lambda,1} &= P_\lambda C_1, & P_{\lambda,d} &= P_\lambda C_d, \\ S_{\lambda,1} &= S_\lambda C_1, & S_{\lambda,d} &= S_\lambda C_d. \end{aligned}$$

TABLE 1.

Cutoff	Description
$p_{\leq 1}(\xi) = \varphi( \xi )$	Inhomogeneous spatial cutoff
$p_\lambda(\xi) = \varphi_{2\lambda}( \xi ) - \varphi_{\frac{1}{2}\lambda}( \xi )$	Spatial cutoff at frequency $\lambda$
$s_{\leq 1}(\tau, \xi) = \varphi( (\tau, \xi) )$	Inhomogeneous space-time cutoff
$s_\lambda(\tau, \xi) = \varphi_{2\lambda}( (\tau, \xi) ) - \varphi_{\frac{1}{2}\lambda}( (\tau, \xi) )$	Space-time cutoff for $\lambda > 1$
$c_{\leq 1}(\tau, \xi) = \varphi( \tau  -  \xi )$	Inhomogeneous cone cutoff
$c_d(\tau, \xi) = \varphi_{2d}( \tau  -  \xi ) - \varphi_{\frac{1}{2}d}( \tau  -  \xi )$	Cutoff at distance $d > 1$ from the cone

At times we will also use the notation  $S_{\lambda,d}^+$  (resp.  $S_{\lambda,d}^-$ ) to denote that the multiplier in question is supported in the upper (resp. lower) half-space. We will also make use of certain other multipliers whose meaning will be clear from subscripts; e.g. we will write  $S_{|\xi| \ll |\tau|}$  to denote cutoff in the space-time region  $|\xi| \ll |\tau|$ . Also, in the sequel, we will write  $I - C_{\leq 1} = S_{d \geq 1}$ .

We now describe a device which will be crucial for the decompositions which arise in this paper. For each angle  $\alpha$  of size  $|\alpha| \sim \sqrt{\frac{d}{\lambda}}$  in the spatial variable in Fourier space, we can restrict the multiplier  $S_{|\tau| \lesssim |\xi|} S_{\lambda,d}$  so that its spatial support lies in this solid angle. We label the smooth, uniformly finitely overlapping partition of unity corresponding to this by  $\{S_{\lambda,d}^\alpha\}$ . One has the reconstruction formula

$$S_{|\tau| \lesssim |\xi|} S_{\lambda,d} = \sum_{\alpha \in \mathbb{S}^2} S_{\lambda,d}^\alpha.$$

Notice that each of the multipliers  $S_{\lambda,d}^\alpha$  has Fourier support in a parallelepiped of size  $\lambda \times \sqrt{\lambda d} \times \sqrt{\lambda d} \times d$  oriented in the radial direction along and at distance  $\sim d$  from the light cone. These will be building blocks for some of the norms we introduce in the sequel.

If  $X_i$  and  $Y$  are Banach spaces, and  $L_i$  are any linear maps on the  $X_i$ , then for any constant  $M_0$  we write

$$\prod_i L_i(X_i) \subseteq M_0 Y$$

to mean that for any set of functions  $\phi^i \in X_i$ , the following bound holds:

$$(27) \quad \left\| \prod_i L_i(\phi^i) \right\|_Y \lesssim M_0 \prod_i \|\phi^i\|_{X_i}.$$

For example, in this notation one can write the frequency-localized 3-dimensional Sobolev embeddings as

$$P_\lambda(L_x^2) \subseteq \lambda^{\frac{3}{2}-\frac{3}{p}}L_x^p, \quad 2 \leq p.$$

If  $\phi^i$  is a set of functions, and  $Y$  is a Banach space, then for linear operators  $L_i$ , we will employ the embedding notation

$$\prod_i L_i(\phi^i) \hookrightarrow M_0 Y,$$

to mean the same as (27), where the spaces  $X_i$  are defined by the context.

If  $F$  is a function on space-time, we will denote by  $W F(0)$  the solution to the homogeneous wave equation with initial data  $F(0)$  and  $\partial_t F(0)$ .

Let  $E$  denote any fundamental solution to the homogeneous wave equation, i.e., one has the formula  $\square E = \delta$ . We define the standard Cauchy parametrix for the wave equation by the formula

$$\frac{1}{\square} F = E * F - W E * F(0).$$

Explicitly, one has the identity

$$\widehat{\frac{1}{\square} F}(t, \xi) = - \int_0^t \frac{\sin(2\pi|\xi|(t-s))}{2\pi|\xi|} \widehat{F}(s, \xi) ds.$$

For any test function  $F$ , we also define division by the symbol of the wave equation away from the light cone by

$$\frac{1}{\Xi}(I - C_1)F = E * (I - C_1)F.$$

Of course, the definition of  $\frac{1}{\Xi}$  does not depend on  $E$  as long as we stay away from the light cone, and for us this will always be the case. Also, one has the formula

$$\mathcal{F} \left[ \frac{1}{\Xi}(I - C_1)F \right] (\tau, \xi) = \frac{1}{4\pi^2(\tau^2 - |\xi|^2)} (1 - c_1) \widetilde{F}(\tau, \xi).$$

We end this subsection with the following well-known bound on the symbol of the projection  $\mathcal{P}$ , adapted to the notation of this section (for a proof see e.g. [16]). This will be used many times in the sequel:

**Lemma 2.1** (Null Form Bound). *Given arbitrary test functions  $\phi^i$ , one has the bound*

$$(28) \quad \left| \mathcal{F} \left[ \mathcal{P} S_{\lambda,d} (S_{\mu_1,\delta_1} \phi^1 \nabla S_{\mu_2,\delta_2} \phi^2) \right] \right| \lesssim \frac{\max\{d, \delta_1, \delta_2\}^{\frac{1}{2}}}{\lambda^{\frac{1}{2}}} s_{\lambda,d} \left( \mu_1^{\frac{1}{2}} s_{\mu_1,\delta_1} |\widehat{\phi^1}| * \mu_2^{\frac{1}{2}} s_{\mu_2,\delta_2} |\widehat{\phi^2}| \right).$$

**2.1. Some multilinear estimates for the wave equation.** One of the main tools we will use in our proof of Theorem 1.3 will be the the standard device of foliating certain multilinear estimates for the homogeneous wave equation. Throughout this section, we employ the following convention: Given  $f$ , some function of space only, we denote the forward and backward wave propagators by the formula

$$\widetilde{W^\pm f}(\tau, \xi) = \delta(\tau \mp |\xi|) \widehat{f}(\xi).$$

By taking inverse Fourier transforms, we have the formulas:

$$(29) \quad \phi(t) = W^\pm f(t) = e^{\pm 2\pi i t \sqrt{-\Delta}} f .$$

All of the solutions to the wave equation we discuss in this section will be assumed to be of the form (29) above. This will streamline the discussion of some of the estimates used here because we will not have to worry about time derivatives when considering initial data. We now begin by stating the standard linear Strichartz estimates on (3 + 1)-dimensional Minkowski space; see [7].

**Lemma 2.2** (Strichartz Estimates). *Let  $2 < q$  and  $2 \leq r < \infty$  be such that  $\frac{1}{q} + \frac{1}{r} \leq \frac{1}{2}$ ; then if  $\square\phi = 0$ , the following estimate holds:*

$$(30) \quad \|\phi\|_{L^q(L^r)} \lesssim \|\phi(0)\|_{\dot{H}^s} ,$$

where the scaling relation  $\frac{1}{q} + \frac{3}{r} = \frac{3}{2} - s$  holds.

We now move on to multilinear estimates; these can be divided into two groups. The first will be the standard set of bilinear Strichartz estimates; see [22] and [23].

**Lemma 2.3** (Bilinear Strichartz Estimate). *Let  $2 < q$  and  $2 \leq r < \infty$  be such that  $\frac{1}{q} + \frac{1}{r} \leq \frac{1}{2}$ ; then if  $\square\phi^i = 0$ , the following estimate holds:*

$$(31) \quad \||D_x|^{-\sigma}(\phi^1\phi^2)\|_{L^{\frac{q}{2}}(L^{\frac{r}{2}})} \lesssim \prod_i \|\phi^i(0)\|_{\dot{H}^{s_0}} ,$$

where  $s_0 + \frac{\sigma}{2} + \frac{1}{q} + \frac{3}{r} = \frac{3}{2}$  and  $\sigma < 3 - \frac{4}{q} - \frac{6}{r}$ .

We will also have an occasion to use certain bilinear estimates which are stronger than (31) but rely on the additional assumption that the waves involved have disjoint directions in Fourier space. Specifically, we call a *+ wave* (resp. *- wave*) a solution to the homogeneous wave equation which is supported on the forward (resp. backward) light cone in Fourier space. We now state a precise  $L^2$  bilinear Strichartz estimate for waves whose initial data are separated by an angle  $\Theta$ . We mention that this idea appeared in Bourgain’s appendix to [17].

**Lemma 2.4** (Angular Strichartz Estimate). *Let  $d_1 \leq d_2$ . Let  $\phi_{B^\alpha}, \psi_{B^\beta}$  be waves with data  $f_{B^\alpha}$  and  $g_{B^\beta}$ , respectively. Assume  $\hat{f}_{B^\alpha}$  is supported in a parallelepiped  $B^\alpha$  of size  $\lambda \times \sqrt{\lambda d_1} \times \sqrt{\lambda d_1}$  and  $\hat{g}_{B^\beta}$  is supported in a parallelepiped  $B^\beta$  of size  $\lambda \times \sqrt{\lambda d_2} \times \sqrt{\lambda d_2}$ . Also assume that  $\phi_{B^\alpha}$  is a + wave and  $\psi_{B^\beta}$  is a  $\pm$  wave, and further that the angle between  $B^\alpha$  and  $\pm B^\beta$  is bounded below by  $\Theta_{min}$ . Then*

$$(32) \quad \|\phi_{B^\alpha} \psi_{B^\beta}\|_{L^2(L^2)} \lesssim \left(\frac{\lambda\sqrt{\lambda d_1}}{\Theta_{min}}\right)^{1/2} \|f_{B^\alpha}\|_{L^2} \|g_{B^\beta}\|_{L^2} .$$

*Proof of (32).* A direct equation shows that we have the formula

$$(33) \quad \widetilde{\phi_{B^\alpha} \psi_{B^\beta}}(\tau, \xi) = \int \delta(\tau \mp |\eta| - |\xi - \eta|) \hat{f}_{B^\alpha}(\xi - \eta) \hat{g}_{B^\beta}(\eta) d\eta .$$

Thus, using Plancherel and Cauchy–Schwarz with measures, it suffices to prove that

$$\begin{aligned} & \int \delta(\tau \mp |\eta| - |\xi - \eta|) \chi_{B^\alpha}(\xi - \eta) \chi_{B^\beta}(\eta) \, d\eta \\ &= \int_{R(\tau, \xi)} \chi_{B^\alpha}(\xi - \eta) \chi_{B^\beta}(\eta) \frac{dS_{R(\tau, \xi)}(\eta)}{|\nabla_\eta(|\xi - \eta| \pm |\eta|)|} \\ &\lesssim \frac{\lambda \sqrt{\lambda d_1}}{\Theta}, \end{aligned}$$

where  $R(\tau, \xi)$  is the ellipse or hyperboloid defined by the formula  $\tau = \pm|\eta| + |\xi - \eta|$ , and  $dS_{R(\tau, \xi)}$  is the corresponding Euclidean surface measure. The above inequality now follows directly from the bounds

$$\begin{aligned} \Theta_{min} &\lesssim |\nabla_\eta(|\xi - \eta| \pm |\eta|)|, \\ |B^\alpha \cap R(\tau, \xi)| &\lesssim \lambda \sqrt{\lambda d_1}. \end{aligned}$$

Notice that the orientation of the blocks  $B^\alpha$  and  $B^\beta$  does not matter here. The only things of relevance are the minimum angle between these two regions and the fact that any sufficiently regular hypersurface can only intersect these blocks in a cross section that contains the long direction  $\lambda$  at most once.  $\square$

As a first application of (32), we prove the following strengthening of (31) in the  $L^2$  case, based on the assumption that both waves in the product are supported on the same cone:

**Lemma 2.5** (Special Bilinear Strichartz Estimate). *If  $\square\phi^i = 0$ , and both  $\phi^1$  and  $\phi^2$  are + waves (or both are - waves), then for any  $0 < s_0 \leq \frac{1}{2}$  the following estimate holds:*

$$(34) \quad \||D_x|^{-1+2s_0}(\phi^1\phi^2)\|_{L^2(L^2)} \lesssim \prod_i \|\phi^i(0)\|_{\dot{H}^{s_0}}.$$

*Proof of (34).* Our first step is to decompose everything into dyadic frequencies, using the Plancherel theorem to trade derivatives and weights. Here and in the sequel, we use the following standard subscript notation for truncation in Fourier space:

$$\phi_{\mu_i}^i = P_{\mu_i} \phi^i.$$

We now compute that

$$\||D_x|^{-1+2s_0}(\phi^1\phi^2)\|_{L^2(L^2)} \lesssim \sum_{\lambda, \mu_i} \lambda^{-1+2s_0} \|P_\lambda(\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2)\|_{L^2(L^2)}.$$

We now break into cases depending on the size of the various frequencies involved.

**Case 1:**  $\mu_1 \ll \mu_2$  or  $\mu_2 \ll \mu_1$ . We will assume without loss of generality that  $\mu_1 \ll \mu_2$ , the other case being the same by symmetry. In this case, we must have  $\lambda \sim \mu_2$ , so by a direct application of Hölder’s inequality and classical  $L^4(L^4)$

Strichartz, we can bound

$$\begin{aligned}
& \sum_{\substack{\lambda, \mu_i \\ \mu_1 \ll \mu_2}} \lambda^{-1+2s_0} \| P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2) \|_{L^2(L^2)} \\
& \lesssim \sum_{\mu_1 \ll \mu_2} \mu_2^{-1+2s_0} \| \phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \|_{L^2(L^2)} \\
& \lesssim \sum_{\mu_1 \ll \mu_2} \left( \frac{\mu_1}{\mu_2} \right)^{\frac{1}{2}-s_0} (\mu_1 \mu_2)^{s_0} \| \phi_{\mu_1}^1(0) \|_{L^2} \| \phi_{\mu_2}^2(0) \|_{L^2} \\
& \lesssim \prod_{i=1,2} \left( \sum_{\mu_1 \ll \mu_2} \left( \frac{\mu_1}{\mu_2} \right)^{\frac{1}{2}-s_0} \| \phi_{\mu_i}^i(0) \|_{\dot{H}^{s_0}}^2 \right)^{\frac{1}{2}} \\
& \lesssim \| \phi^1(0) \|_{\dot{H}^{s_0}} \| \phi^2(0) \|_{\dot{H}^{s_0}} .
\end{aligned}$$

**Case 2:**  $\mu_1 \sim \mu_2$ . In this case we will assume without loss of generality that  $\mu_1 = \mu_2$ , and we begin by decomposing the Fourier support of each  $\phi_\mu^i$  into a sum of blocks of size  $\lambda \times \lambda \times \lambda$  with (uniformly) bounded overlap. We enumerate this collection of blocks by  $\alpha$  and write

$$\phi_\mu^i = \sum_\alpha \phi_\mu^{i\alpha} .$$

Then  $P_\lambda (\phi_\mu^1 \cdot \phi_\mu^2)$  becomes the essentially diagonal sum:

$$\sum_{|\alpha+\beta|<8} \phi_\mu^{1\alpha} \cdot \phi_\mu^{2\beta} .$$

Then, for a fixed  $\lambda$  and  $\mu$ , using the estimate (32) with  $d = \lambda$  we can bound

$$\begin{aligned}
& \| P_\lambda (\phi_\mu^1 \cdot \phi_\mu^2) \|_{L^2(L^2)} \\
& \lesssim \sum_{|\alpha+\beta|<8} \| P_\lambda (\phi_\mu^{1\alpha} \cdot \phi_\mu^{2\beta}) \|_{L^2(L^2)} \\
& \lesssim \lambda \sum_{|\alpha+\beta|<8} \| \phi_\mu^{1\alpha}(0) \|_{L^2} \| \phi_\mu^{2\beta}(0) \|_{L^2} \\
& \lesssim \lambda \prod_{i=1,2} \| \phi_\mu^i(0) \|_{L^2} .
\end{aligned}$$

Summing over  $\lambda$  and  $\mu$  we obtain

$$\begin{aligned}
& \sum_{\lambda, \mu} \lambda^{-1+2s_0} \| P_\lambda (\phi_\mu^1 \cdot \phi_\mu^2) \|_{L^2(L^2)} \\
& \lesssim \sum_{\lambda, \mu} \lambda^{2s_0} \| \phi_\mu^1(0) \|_{L^2} \| \phi_\mu^2(0) \|_{L^2} \\
& = \sum_{\lambda, \mu} \left( \frac{\lambda}{\mu} \right)^{2s_0} \| \phi_\mu^1(0) \|_{\dot{H}^{s_0}} \| \phi_\mu^2(0) \|_{\dot{H}^{s_0}} \\
& \lesssim \| \phi^1(0) \|_{\dot{H}^{s_0}} \| \phi^2(0) \|_{\dot{H}^{s_0}} .
\end{aligned}$$

□

The second group of estimates we will need is the trilinear analog of (31), which we will only state for the line  $\frac{1}{q} + \frac{1}{r} = \frac{1}{2}$ . Estimates of this type are originally due to Terence Tao [29].

**Lemma 2.6** (Trilinear Strichartz Estimate). *Let  $(q, r)$  be given with  $3 < q$  and  $\frac{1}{q} + \frac{1}{r} = \frac{1}{2}$ ; then if  $\square\phi^i = 0$ , the following estimate holds:*

$$(35) \quad \left\| |D_x|^{-\sigma} (\phi^1 \phi^2 \phi^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \lesssim \prod_i \|\phi^i(0)\|_{\dot{H}^{s_0}},$$

where  $s_0 + \frac{\sigma}{3} + \frac{1}{q} + \frac{3}{r} = \frac{3}{2}$  and  $\sigma < 1 - \frac{3}{r}$ ; hence  $s_0 > \frac{2}{3} - \frac{1}{r}$ .

*Proof of (35).* We first run a frequency decomposition on the left-hand side of (35) to obtain:

$$\text{L.H.S. of (35)} = \left\| \sum_{\lambda, \mu_i} |D_x|^{-\sigma} P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})}.$$

Next, we break into cases depending on the relative sizes of the various frequencies involved.

**Case 1:**  $\lambda \sim \max\{\mu_i\}$ . Without loss of generality, we will assume that  $\max\{\mu_i\} = \mu_3$ . In this case, we may use Hölder's inequality and the linear Strichartz estimate (30) to bound

$$\begin{aligned} & \left\| \sum_{\substack{\mu_i, \lambda \\ \mu_1, \mu_2 \leq \mu_3 \\ \lambda \sim \mu_3}} |D_x|^{-\sigma} P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \\ & \lesssim \sum_{\mu_1, \mu_2 \leq \mu_3} \mu_3^{-\sigma} \|\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \\ & \lesssim \sum_{\mu_1, \mu_2 \leq \mu_3} \mu_3^{-\sigma} \|\phi_{\mu_1}^1\|_{L^q(L^r)} \|\phi_{\mu_2}^2\|_{L^q(L^r)} \|\phi_{\mu_3}^3\|_{L^q(L^r)} \\ & \lesssim \sum_{\mu_1, \mu_2 \leq \mu_3} \left(\frac{\mu_1}{\mu_3}\right)^{\frac{\sigma}{3}} \left(\frac{\mu_2}{\mu_3}\right)^{\frac{\sigma}{3}} \|\phi_{\mu_1}^1(0)\|_{\dot{H}^{s_0}} \|\phi_{\mu_2}^2(0)\|_{\dot{H}^{s_0}} \|\phi_{\mu_3}^3(0)\|_{\dot{H}^{s_0}} \\ & \lesssim \sup_{\mu_3} \|\phi_{\mu_3}^3(0)\|_{\dot{H}^{s_0}} \cdot \prod_{i=1,2} \left( \sum_{\mu_1, \mu_2 \leq \mu_3} \left(\frac{\mu_1}{\mu_3}\right)^{\frac{\sigma}{3}} \left(\frac{\mu_2}{\mu_3}\right)^{\frac{\sigma}{3}} \|\phi_{\mu_i}^i(0)\|_{\dot{H}^{s_0}}^2 \right)^{\frac{1}{2}} \\ & \lesssim \|\phi^1(0)\|_{\dot{H}^{s_0}} \|\phi^2(0)\|_{\dot{H}^{s_0}} \|\phi^3(0)\|_{\dot{H}^{s_0}}. \end{aligned}$$

**Case 2:**  $\min\{\mu_i\} \leq \lambda$ , and  $\lambda \ll \max\{\mu_i\}$ . Without loss of generality, we will assume here that  $\mu_1 \leq \lambda$ . It then follows that  $\mu_2 \sim \mu_3$ , and we can bound

$$\begin{aligned} & \left\| \sum_{\substack{\mu_i, \lambda \\ \mu_1 \leq \lambda \\ \lambda \ll \mu_2 \sim \mu_3}} |D_x|^{-\sigma} P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \\ & \lesssim \sum_{\substack{\mu_i, \lambda \\ \mu_1 \leq \lambda \\ \lambda \ll \mu_2 \sim \mu_3}} \lambda^{-\sigma} \left\| P_\lambda (\phi_{\mu_1}^1 \cdot P_{\lesssim \lambda}(\phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3)) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \\ & \lesssim \sum_{\substack{\mu_i, \lambda \\ \mu_1 \leq \lambda \\ \lambda \ll \mu_2 \sim \mu_3}} \left\| \lambda^{-\frac{\sigma}{3}} \phi_{\mu_1}^1 \right\|_{L^q(L^r)} \cdot \left\| \lambda^{-\frac{2\sigma}{3}} P_{\lesssim \lambda} (\phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{2}}(L^{\frac{r}{2}})} \\ & \lesssim \left\| \phi^1 \right\|_{\dot{H}^{s_0}} \cdot \sum_{\lambda \ll \mu_2 \sim \mu_3} \left\| \lambda^{-\frac{2\sigma}{3}} P_{\lesssim \lambda} (\phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{2}}(L^{\frac{r}{2}})}. \end{aligned}$$

The condition  $\frac{2}{3}\sigma < \frac{2}{3} - \frac{2}{r}$  guarantees we are in the correct range to apply (31). However, we cannot simply do this directly because for fixed  $\mu_2 \sim \mu_3$  we need to sum over all of the  $\lambda$  involved. To get around this, we rely on an “improved” Strichartz estimate used in the proof of (31); see [23]. This states that there is an  $\epsilon > 0$  (depending of course on the various exponents involved) such that

$$\left\| \lambda^{-\frac{2\sigma}{3}} P_{\lesssim \lambda} (\phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{2}}(L^{\frac{r}{2}})} \lesssim \left( \frac{\lambda}{\mu_2} \right)^\epsilon \left\| \phi_{\mu_2}^2(0) \right\|_{\dot{H}^{s_0}} \left\| \phi_{\mu_3}^3(0) \right\|_{\dot{H}^{s_0}}.$$

The desired inequality now follows by summing through using a Cauchy–Schwarz.

**Case 3:**  $\lambda \leq \min\{\mu_i\}$ , and  $\lambda \ll \max\{\mu_i\}$ . In this case we are dealing with sums which are all of the form

$$\begin{aligned} (36) \quad & \left\| \sum_{\lambda \leq \mu_1 \leq \mu_2 \sim \mu_3} |D_x|^{-\sigma} P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \\ & \lesssim \sum_{\lambda \leq \mu_1 \leq \mu_2 \sim \mu_3} \lambda^{-\sigma} \left\| P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})}. \end{aligned}$$

In order to bound (36), we will show the following scale invariant estimate:

$$\begin{aligned} (37) \quad & \left\| P_\lambda (\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \\ & \lesssim \lambda^{1-\frac{3}{r}} \mu_1^{\frac{2}{3}-\frac{1}{r}} \mu_2^{\frac{2}{3}-\frac{1}{r}} \mu_3^{\frac{2}{3}-\frac{1}{r}} \left\| \phi_{\mu_1}^1(0) \right\|_{L^2} \left\| \phi_{\mu_2}^2(0) \right\|_{L^2} \left\| \phi_{\mu_3}^3(0) \right\|_{L^2}. \end{aligned}$$

This will be enough to add over the dyadic values of  $\lambda \leq \mu_1 \leq \mu_2 \sim \mu_3$  in the sum (36) because  $\sigma - (1 - \frac{3}{r}) > 0$  and

$$\begin{aligned} & \lambda^{1-\frac{3}{r}} \mu_1^{\frac{2}{3}-\frac{1}{r}} \mu_2^{\frac{2}{3}-\frac{1}{r}} \mu_3^{\frac{2}{3}-\frac{1}{r}} \\ & = \lambda^\sigma \mu_1^{s_0} \mu_2^{s_0} \mu_3^{s_0} \left( \frac{\lambda}{\mu_1} \right)^{-\frac{\sigma-(1-\frac{3}{r})}{3}} \left( \frac{\lambda}{\mu_2} \right)^{-\frac{\sigma-(1-\frac{3}{r})}{3}} \left( \frac{\lambda}{\mu_3} \right)^{-\frac{\sigma-(1-\frac{3}{r})}{3}}. \end{aligned}$$

We will now assume without loss of generality that  $\mu_2 = \mu_3 = 1$  in (37). What we are then trying to show is that

$$(38) \quad \left\| P_\lambda (\phi_\mu^1 \cdot \phi_\mu^2 \cdot \phi_\mu^3) \right\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \lesssim \lambda^{1-\frac{3}{r}} \mu^{\frac{2}{3}-\frac{1}{r}} \left\| \phi_\mu^1(0) \right\|_{L^2} \left\| \phi_\mu^2(0) \right\|_{L^2} \left\| \phi_\mu^3(0) \right\|_{L^2},$$

for  $\lambda \leq \mu \leq 1$ .

To prove (38), we use the two-scale physical space method which appeared in [29]. See also [22] for applications to bilinear Strichartz estimates. To implement this, we begin by tiling  $\mathbb{R}^3$  with cubes  $Q_{\frac{1}{\lambda}}$ , of side length  $\frac{1}{\lambda}$ . Using a direct computation involving the convolution kernel of  $P_\lambda$ , one can easily show that

$$\|P_\lambda(\phi_\mu^1 \cdot \phi_1^2 \cdot \phi_1^3)\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \lesssim \lambda^{3-\frac{q}{r}} \left\| \left( \sum_{Q_{\frac{1}{\lambda}}} \|\phi_\mu^1 \cdot \phi_1^2 \cdot \phi_1^3\|_{L_x^{\frac{r}{3}}(Q_{\frac{1}{\lambda}})} \right)^{\frac{3}{r}} \right\|_{L_t^{\frac{q}{3}}}.$$

After several rounds of Hölder’s inequality, the above estimate reduces the proof of (38) to showing that

$$\left\| \left( \sum_{Q_{\frac{1}{\lambda}}} \|\phi_\mu\|_{L_x^r(Q_{\frac{1}{\lambda}})}^r \right)^{\frac{1}{r}} \right\|_{L_t^q} \lesssim \lambda^{\frac{2}{r}-\frac{2}{3}} \mu^{\frac{2}{3}-\frac{1}{r}} \|\phi_\mu(0)\|_{L^2}.$$

This last estimate can be found by interpolating between the plain Strichartz estimate

$$\left\| \left( \sum_{Q_{\frac{1}{\lambda}}} \|\phi_\mu\|_{L_x^r(Q_{\frac{1}{\lambda}})}^r \right)^{\frac{1}{r}} \right\|_{L_t^q} \lesssim \mu^{1-\frac{2}{r}} \|\phi_\mu(0)\|_{L^2},$$

and the “improved” Strichartz estimate of T. Tao (taken from [22]),

$$\left\| \left( \sum_{Q_{\frac{1}{\lambda}}} \|\phi_\mu\|_{L_x^2(Q_{\frac{1}{\lambda}})}^r \right)^{\frac{1}{r}} \right\|_{L_t^q} \lesssim \lambda^{-\frac{2}{q}} \mu^{\frac{1}{q}} \|\phi_\mu(0)\|_{L^2}.$$

□

Finally, we state a rebalanced version of (35) that will be useful in applications:

**Corollary 2.7.** *Given  $s > \frac{1}{2}$ , then if  $\epsilon_0 > 0$  is sufficiently small, there exists a  $q > 3$  depending on both  $s$  and  $\epsilon_0$ , such that the following estimate holds:*

$$(39) \quad \|\langle D_x \rangle^{-s}(\phi^1 \phi^2 \phi^3)\|_{L^{\frac{q}{3}}(L^2)} \lesssim \|\phi^1(0)\|_{H^{1-s-\epsilon_0}} \prod_{i=2,3} \|\phi^i(0)\|_{H^s}.$$

*Proof of (39).* Notice that the multipliers and Sobolev spaces in (39) are all inhomogeneous. This allows us to reduce the problem to one on fixed dyadic frequencies. Furthermore, when discussing low frequencies, we will assume that the  $P_1$  multiplier in the  $\phi_i^i$  notation is replaced by  $P_{\leq 1}$ . We begin by choosing  $\epsilon_0$  so small that the  $r$  appearing in

$$\|\langle D_x \rangle^{\frac{1}{2}}(\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3)\|_{L^{\frac{q}{3}}(L^{\frac{r}{3}})} \lesssim \prod_{i=1,2,3} \|\phi_{\mu_i}^i(0)\|_{H^{\frac{1}{2}-2\epsilon_0}},$$

is close enough to 6 (notice that  $r < 6$ ) that one has

$$\langle D_x \rangle^{-\frac{s-1/2}{2}} L^{\frac{r}{3}} \subseteq L^2.$$

That this is possible follows from the fact that  $\sigma = s_0 = \frac{1}{2}$ ,  $q = 3$ , and  $r = 6$  is an end point of (35) (see the remark below), and if  $r$  is made smaller, the scaling makes  $s_0$  smaller. If  $\mu_1 \lesssim \max\{\mu_2, \mu_3\}$  in (40), then a straightforward rebalancing of the weights and Sobolev’s inequality show that

$$(40) \quad \begin{aligned} \|\langle D_x \rangle^{-s} P_\lambda(\phi_{\mu_1}^1 \cdot \phi_{\mu_2}^2 \cdot \phi_{\mu_3}^3)\|_{L^{\frac{q}{3}}(L^2)} \\ \lesssim \lambda^{-\frac{s-1/2}{2}} \|\phi_{\mu_1}^1(0)\|_{H^{1-s-2\epsilon_0}} \cdot \prod_{i=2,3} (\|\phi_{\mu_i}^i(0)\|_{H^{s-2\epsilon_0}}). \end{aligned}$$

Furthermore, if  $\max\{\mu_2, \mu_3\} \ll \mu_1$ , then  $\mu_1 \sim \lambda$  in (40), so by a direct application of Hölder’s inequality and three rounds of Strichartz, this estimate is still true. The extra room in all the dyadic variables now allows us to add this estimate directly over all dyadic variables to achieve the desired result which is (39).  $\square$

**2.2. Foliations and some function spaces.** We now discuss the the main norms that will be used in this paper. These are Tataru’s Besov space version,  $X_p^{s,\theta}$ , of the “classical”  $H^{s,\theta}$  spaces, the “outer block” norms  $Y^{s,\theta}$  which first appeared in [33], some modified mixed Lebesgue spaces, and various Sobolev type spaces based on space and space-time weights.

The basic building blocks for the  $X$  norms are the following family of seminorms:

$$\|u\|_{X_{\lambda,p}^{\theta}}^p = \sum_{1 \leq d} d^{\theta p} \|S_{\lambda,d} u\|_{L^2}^p, \quad \lambda \in 2^{\mathbb{Z}^+}.$$

Note that the summation in the first norm is taken over all (dyadic) values of  $1 \leq d$ , thus, these norms are inhomogeneous. The  $X$  and  $Y$  norms themselves are then formed as follows:

$$\begin{aligned} \|u\|_{X_p^{s,\theta}}^2 &= \sum_{\lambda \in 2^{\mathbb{Z}^+}} \lambda^{2s} \|u\|_{X_{\lambda,p}^{\theta}}^2, \\ \|u\|_{Y^{s,\theta}} &= \lambda^s d^{\theta} \sup_{d,\lambda,i} \|S_{|\tau|<2|\xi|} S_{\lambda,d}^{\alpha} u\|_{\bar{L}^{1+(\bar{L}^{\infty})}}. \end{aligned}$$

We note that the summation in the first norm above is taken over all dyadic values of  $1 \leq \lambda$ . We also mention briefly that our  $Y$  norms here are *not* a single norm but in fact a collection of norms. This means that some care needs to be taken when interpreting the notation  $u \hookrightarrow Y^{s,\theta}$ . We will always take this to mean that for any  $0 < \epsilon$ , and  $\epsilon$  sufficiently small, there is a  $C_{\epsilon}$  such that

$$\lambda^s d^{\theta} \sup_{d,\lambda,\alpha} \|S_{\lambda,d}^{\alpha} u\|_{\bar{L}^{1+\epsilon}(\bar{L}^{\infty})} \leq C_{\epsilon} \|u\|.$$

We note here that the need for this notation is due to the failure of the  $L^2(L^{\infty})$  estimate for the  $(3 + 1)$ -dimensional wave equation. In principle, in a normal time cutoff iteration procedure for a problem not at the scaling, one should be able to work with an  $L^2(L^{\infty})$  norm by adding a few derivatives where necessary. However, this option will not be open to us here because the specific structure of our iteration argument only cuts off in time around the light cone. Therefore, all our integrations will be effectively global in time, so we will need a genuine substitute to  $L^2(L^{\infty})$ .

The  $\bar{L}^{1+\epsilon}(\bar{L}^{\infty})$  notation above refers to the spaces

$$\|f\|_{\bar{L}^q(\bar{L}^r)} = \inf\{\|g\|_{L^q(L^r)} \mid |f| \leq \tilde{g}\}.$$

Notice that all Strichartz estimates and Sobolev embeddings carry through to the  $\bar{L}^q(\bar{L}^r)$  spaces with the obvious modification of the above notation for measures supported on the cone. Also, any  $L_{\tau,\xi}^{\infty}$  multiplier is trivially bounded on these spaces. In order to streamline our proof, we will always use these norms in this paper, wherever it would be natural to use  $L^q(L^r)$  even if it is not entirely necessary. However, a little care needs to be taken in order for this to work, because for instance, one does *not* have the inclusion  $\bar{L}^q(\bar{L}^r) \subseteq L^q(L^r)$ . Due to these considerations, we point out that all of the norms used in this paper, with the exception, of course, of the energy type estimates  $L^{\infty}(H^s)$ , depend only on the size of the Fourier

transform; thus all Fourier transforms will be tacitly assumed to be positive, except in those cases where we need to show  $L^\infty(H^s)$  bounds by hand.

We will also have occasion to use some multiweighted space-time Sobolev spaces when dealing with certain “elliptic” variables. We define these as

$$\|u\|_{H^{q,r,s}} = \|\langle D_{t,x} \rangle^q \langle D_x \rangle^r |D_x|^s u\|_{L^2(L^2)} .$$

We will also use the following notation to denote certain spatial energy type spaces based on space-time weights:

$$\|u\|_{\bar{L}^q(H_{t,x}^s)} = \|\langle D_{t,x} \rangle^s u\|_{\bar{L}^q(L^2)} .$$

The final “norm” we introduce here is not really a norm but a convenient notation for testing embeddings via duality. We will write

$$\|u\|_{\mathcal{L}^1} = \int u .$$

The advantage to the  $\mathcal{L}^1$  notation, as opposed to using the space  $L^1(L^1)$ , is that it allows us to trade weights via the Plancherel theorem while it also supports Hölder’s inequality for the  $\bar{L}^q(\bar{L}^r)$  spaces. For instance, it follows from the notations introduced above that one may write

$$\|u^1 \cdot |D_x|^{-1} u^2\|_{\mathcal{L}^1} \leq \| |D_x|^{-\gamma} u^1\|_{\bar{L}^q(\bar{L}^r)} \| |D_x|^{-1+\gamma} u^2\|_{\bar{L}^{q'}(\bar{L}^{r'})} .$$

Finally, we briefly recall the standard mechanism for transferring multilinear estimates for the wave equation to the spaces  $X_1^{s,\theta}$ .

If  $u$  is a function on space-time, then we denote by  $u_s^+$  the Fourier transform of  $\tilde{u}$  restricted to the  $s^{th}$  translate of the positive light cone, i.e.,

$$\widehat{u_s^+}(\xi) = \int \delta(\tau - s - |\xi|) \tilde{u}(\tau, \xi) \, d\tau .$$

Let  $u_s^-$  denote the same integral with  $\delta(\tau - s - |\xi|)$  replaced by  $\delta(\tau + s + |\xi|)$ . By refoliating things, we get the reproducing formula:

$$\tilde{u}(\tau, \xi) = \int \delta(s \mp \tau + |\xi|) \widehat{u_s^\mp}(\xi) \, ds .$$

This notation and the Fourier inversion formula allow us to write  $u$  as a sum over forward or backward waves:

$$(41) \quad u(t, x) = \int e^{\pm 2\pi i t s} W^\pm u_s^\pm(t, x) \, ds .$$

Using the formula (41) along with the fact that  $\|u\|_{L_{t,x}^2} = \|u_s^\pm\|_{L_{s,x}^2}$ , we extend the linear multilinear Strichartz estimates of the previous section as follows. If  $u^i$  are test functions whose Fourier transforms lie in the upper half-plane, and if we are in the range stated in Lemma 2.3, then we can estimate

$$\begin{aligned} & \| |D_x|^{-\sigma} (u^1 u^2) \|_{\bar{L}^{\frac{q}{2}}(\bar{L}^{\frac{r}{2}})} \\ & \lesssim \int \| |D_x|^{-\sigma} (W^+(u^1)_{s_1}^+ \cdot W^+(u^2)_{s_2}^+) \|_{\bar{L}^{\frac{q}{2}}(\bar{L}^{\frac{r}{2}})} \, ds_1 ds_2 \\ & \lesssim \int \| (u^1)_{s_1}^+ \|_{H^\gamma} \cdot \| (u^2)_{s_2}^+ \|_{H^\gamma} \, ds_1 ds_2 \\ & \lesssim \|u_1\|_{X_1^{\gamma, \frac{1}{2}}} \|u_2\|_{X_1^{\gamma, \frac{1}{2}}} . \end{aligned}$$

By decomposing a given set of test functions into sums of pieces supported on the upper and lower half-spaces in Fourier space, we have shown that:

**Lemma 2.8** (Foliated Bilinear Strichartz Estimate). *If  $u^i$  are arbitrary test functions, then if  $2 < q$  and  $\frac{1}{q} + \frac{1}{r} = \frac{1}{2}$ , one has the estimates*

$$(42) \quad \left\| |D_x|^{-\sigma} (u^1 u^2) \right\|_{\dot{L}^{\frac{q}{2}}(\dot{L}^{\frac{r}{2}})} \lesssim \|u_1\|_{X_1^{\gamma, \frac{1}{2}}} \|u_2\|_{X_1^{\gamma, \frac{1}{2}}}$$

for  $\frac{\sigma}{2} + \gamma + \frac{1}{q} + \frac{3}{r} = \frac{3}{2}$  and  $\sigma < 3 - \frac{4}{q} - \frac{6}{r}$ .

There are similar estimates for the  $X$  spaces based on the other Strichartz estimates proven in the previous subsection. We omit the details.

**2.3. Some frequency interactions and a fundamental decomposition.** We begin here by recording the following notation which will be used for the rest of the paper: For any product, or more generally, any quadratic form  $u \cdot v$ , we define

$$(43) \quad \langle u \cdot v \rangle_{H \times H \Rightarrow L}^{close\ cone} = \sum_{\substack{1 \leq \lambda \ll \sigma_1 \sim \sigma_2 \\ \max\{\eta_1, \eta_2\} \ll d}} S_{\lambda, d} [S_{\sigma_1, \eta_1}(u) \cdot S_{\sigma_2, \eta_2}(v)] ,$$

$$(44) \quad \langle u \cdot v \rangle_{H \times H \Rightarrow L}^{medium\ cone} = \sum_{\substack{1 \leq \lambda \ll \sigma_1 \sim \sigma_2 \\ d \lesssim \max\{\eta_1, \eta_2\} \lesssim \lambda}} S_{\lambda, d} [S_{\sigma_1, \eta_1}(u) \cdot S_{\sigma_2, \eta_2}(v)] ,$$

$$(45) \quad \langle u \cdot v \rangle_{H \times L \Rightarrow H}^{close\ cone} = \sum_{\substack{1 \leq \lambda \ll \sigma_1 \sim \sigma_2 \\ \max\{\eta_1, \eta_2\} \ll d}} S_{\sigma_2, \eta_2} [S_{\lambda, d}(u) \cdot S_{\sigma_1, \eta_1}(v)] ,$$

$$(46) \quad \langle u \cdot v \rangle_{H \times L \Rightarrow H}^{medium\ cone} = \sum_{\substack{1 \leq \lambda \ll \sigma_1 \sim \sigma_2 \\ d \lesssim \max\{\eta_1, \eta_2\} \lesssim \lambda}} S_{\sigma_2, \eta_2} [S_{\lambda, d}(u) \cdot S_{\sigma_1, \eta_1}(v)] .$$

It is helpful to notice that a *High*  $\times$  *High*  $\Rightarrow$  *Low* interaction is dual to *High*  $\times$  *Low*  $\Rightarrow$  *High* interaction. Also, the interactions (43) and (44) only happen when  $S_{\sigma_1, \eta_1}$  and  $S_{\sigma_2, \eta_2}$  are in a  $+$   $-$  interaction (that is, they correspond to a product of a  $+$  and a  $-$  wave), while (45) and (46) only happen when  $S_{\sigma_1, \eta_1}$  and  $S_{\sigma_2, \eta_2}$  are in a  $++$  or  $--$  interaction. Also, the decompositions  $\langle \cdot \rangle^{close\ cone}$  and  $\langle \cdot \rangle^{medium\ cone}$  are disjoint with respect to the Fourier supports of the  $S_{\mu_i, \delta_i}$ . This guarantees there is no overlapping in a sum of these decompositions.

We mention here that a quick calculation shows, because of the condition  $\max\{\delta_1, \delta_2\} \ll d$ , that one has the identity

$$S_{1 \leq |\tau| < 2|\xi|} \langle \cdot \rangle_{H \times H \Rightarrow L}^{close\ cone} = \langle \cdot \rangle_{H \times H \Rightarrow L}^{close\ cone} .$$

This will be useful in situations where the  $Y$  norms appear.

The frequency interactions (43)–(46) encapsulate the main difficulties we will encounter in this paper. Roughly speaking, the *High*  $\times$  *High*  $\Rightarrow$  *Low* interaction will be the worst part of the potentials  $\{A\}$ , while the *High*  $\times$  *Low*  $\Rightarrow$  *High* interactions will be the most difficult part of  $\phi$  to control.

In order to save space in the last section of the paper, we will also employ the notation

$$\begin{aligned} \langle u \cdot v \rangle_{H \times H \Rightarrow L} &= \sum_{\substack{\lambda \ll \sigma_1 \sim \sigma_2 \\ \max\{\eta_1, \eta_2\} \lesssim \lambda}} S_{\lambda, d} [S_{\sigma_1, \eta_1}(u) \cdot S_{\sigma_2, \eta_2}(v)] , \\ \langle u \cdot v \rangle_{H \times L \Rightarrow H} &= \sum_{\substack{\lambda \ll \sigma_1 \sim \sigma_2 \\ \max\{\eta_1, \eta_2\} \lesssim \lambda}} S_{\sigma_2, \eta_2} [S_{\lambda, d}(u) \cdot S_{\sigma_1, \eta_1}(v)] . \end{aligned}$$

Apparently, one has the identities

$$\begin{aligned} \langle u \cdot v \rangle_{H \times H \Rightarrow L} &= \langle u \cdot v \rangle_{H \times H \Rightarrow L}^{close\ cone} + \langle u \cdot v \rangle_{H \times H \Rightarrow L}^{medium\ cone} , \\ \langle u \cdot v \rangle_{H \times L \Rightarrow H} &= \langle u \cdot v \rangle_{H \times L \Rightarrow H}^{close\ cone} + \langle u \cdot v \rangle_{H \times L \Rightarrow H}^{medium\ cone} . \end{aligned}$$

Because of its importance in later sections, we will now give a precise decomposition of the terms in the sum  $\langle \cdot \rangle_{H \times H \Rightarrow L}^{close\ cone}$ . The idea will be to restrict the angular regions where the product in these terms may interact in Fourier space. This method has been used in various forms for some time now (see e.g. [33] and Bourgain’s appendix of [17]), and it is precisely the tool we need in order to prove the multilinear estimates which appear in the sequel.

In what follows, we will restrict our attention to a single term in the right-hand side of (43). In later sections we will also need this decomposition in the range (44); this will involve a straightforward modification of the procedure below. We will also, without loss of generality, assume that we are in the (+−) case. Our task is now to decompose the convolution

$$(47) \quad s_{\lambda, d} \left( s_{\mu_1, \delta_1}^+ * s_{\mu_2, \delta_2}^- \right) ,$$

where we are in the range  $\lambda \ll \mu_1 \sim \mu_2$  and  $\max\{\delta_1, \delta_2\} \ll \lambda$ .

As mentioned above, this will be accomplished by isolating the angles where the  $s_{\mu_1, \delta_1}^+$  and  $s_{\mu_2, \delta_2}^-$  can interact. For  $(\tau, \xi) \in \text{supp}(s_{\mu_1, \delta_1}^+)$  and  $(\tau', \xi') \in \text{supp}(s_{\mu_2, \delta_2}^-)$  we compute that

$$\begin{aligned} d &\sim \left| |\tau + \tau'| - |\xi + \xi'| \right| \\ &= \left| \left| |\xi| - |\xi'| + O(\delta_1) + O(\delta_2) \right| - |\xi + \xi'| \right| \\ &= \left| \left| |\xi| - |\xi'| + O(\delta_1) + O(\delta_2) \right| - \left| |\xi| - |\xi'| \right| - |\xi + \xi'| + \left| |\xi| - |\xi'| \right| \right| \\ &\gtrsim \left| O(\delta_1) + O(\delta_2) + \frac{\mu^2}{\lambda} O(\theta_{\xi, -\xi'}^2) \right| . \end{aligned}$$

Because of the restriction  $\max\{\delta_1, \delta_2\} \ll d$ , an immediate consequence of the above calculation in the following restriction on the angle between the spatial projections of the supports of  $s_{\mu_1, \delta_1}^+$  and  $s_{\mu_2, \delta_2}^-$ :

$$(48) \quad \Theta_{\xi, -\xi'} \lesssim \frac{\sqrt{\lambda d}}{\mu} .$$

To take advantage of (48), we decompose the spatial supports of  $s_{\mu_1, \delta_1}^+$  and  $s_{\mu_2, \delta_2}^-$  into (essentially) disjoint blocks of dimension  $\lambda \times \sqrt{\lambda d} \times \sqrt{\lambda d}$  which are oriented in the radial direction.

Next, label these blocks by  $B^\alpha$  and  $B^\beta$  and denote the cutoff of the multipliers  $s_{\mu_1, \delta_1}$  and  $s_{\mu_2, \delta_2}$  along these labeled blocks by  $s_{\mu_1; \lambda, d, \delta_1}^\alpha$  and  $s_{\mu_2; \lambda, d, \delta_2}^\beta$ , respectively.

Notice that the multipliers  $s_{\mu_1; \lambda, d, \delta_1}^\alpha$  and  $s_{\mu_2; \lambda, d, \delta_2}^\beta$  are supported on sets of size roughly  $\lambda \times \sqrt{\lambda d} \times \sqrt{\lambda d} \times \delta_i$ . We emphasize the fact that the supports here are *not* parallelepipeds, because  $\delta_i \ll d$ . Therefore, some care needs to be taken to distinguish these multipliers from the  $S_{\lambda, d}^\alpha$  multipliers introduced earlier. We mention here that the usefulness of the  $s_{\mu_i; \lambda, d, \delta_i}^\alpha$  comes from the fact that their supports have a specific size, which will always turn out to be small enough to get the estimates we need through fairly elementary means.

Using this angular decomposition, we can discard the  $s_{\lambda, d}$  on the left-hand side of (47) and write the convolution as an essentially diagonal sum. We record this fact as

**Lemma 2.9.** *Let  $\widetilde{\phi}^i$  be positive test functions; then if*

- (1)  $\lambda \ll \mu_1 \sim \mu_2$ , and
- (2)  $\max\{\delta_1, \delta_2\} \ll d$ ,

one can bound:

$$(49) \quad s_{\lambda, d} \left( s_{\mu_1, \delta_1}^+ \widetilde{\phi}^1 * s_{\mu_2, \delta_2}^- \widetilde{\phi}^2 \right) \lesssim \sum_{|\Theta_{B^\alpha, -B^\beta}| \lesssim \frac{\sqrt{\lambda d}}{\mu}} s_{\mu_1; \lambda, d, \delta_1}^\alpha \widetilde{\phi}^1 * s_{\mu_2; \lambda, d, \delta_2}^\beta \widetilde{\phi}^2 .$$

Due to the fact that the above sum is essentially diagonal, we often abusively write

$$\sum_{|\Theta_{B^\alpha, -B^\beta}| \lesssim \frac{\sqrt{\lambda d}}{\mu}} s_{\mu_1; \lambda, d, \delta_1}^\alpha \widetilde{\phi}^1 * s_{\mu_2; \lambda, d, \delta_2}^\beta \widetilde{\phi}^2 = \sum_{\alpha} s_{\mu_1; \lambda, d, \delta_1}^\alpha \widetilde{\phi}^1 * s_{\mu_2; \lambda, d, \delta_2}^{-\alpha} \widetilde{\phi}^2 .$$

### 3. THE ITERATION PROBLEM

Our solution to the system (13) will be obtained through Picard iteration on an auxiliary system whose solutions agree with those of (13) on the time interval  $[-1, 1]$ . However, the usual time-cutoff procedure based on the operator  $\chi_T \frac{1}{\square}$  will be too crude for our purposes. The reason for this is that we will be forced to expand our iterates at the  $k^{th}$  level in terms of the previous three iterates. In this type of procedure, taking commutators with the time cutoff function  $\chi_T$  becomes too much trouble because keeping track of the various Fourier supports is difficult. The type of operator we would like to use is  $\frac{1}{\Xi} S_{\lambda, d}$ , which does not change Fourier supports, as opposed to  $\chi_T \frac{1}{\square} S_{\lambda, d}$  which does. To realize this, for the part of the nonlinearity  $\mathcal{N}$  which are supported away from the light cone in Fourier space, we shall replace the expression

$$(50) \quad \chi_T \frac{1}{\square} S_{\lambda, d} \mathcal{N}$$

by the formula

$$\frac{1}{\Xi} S_{\lambda, d} \mathcal{N} - \chi_T W \frac{1}{\Xi} S_{\lambda, d} \mathcal{N}(0) .$$

However, we will still use (50) in an  $O(1)$  neighborhood of the light cone.

To save space, we write the nonlinear terms of (13) as:

$$\mathcal{N}_1(\{A\}, \phi) = -\mathcal{P}\mathfrak{S} \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} - \sqrt{-1} A^{(k-1)} |\phi^{(k-1)}|^2 \right),$$

$$\mathcal{N}_2(\{A\}, \phi) = -\mathfrak{S} \left( \phi^{(k-1)} \overline{\partial_t \phi^{(k-1)}} - \sqrt{-1} A_0^{(k-1)} |\phi^{(k-1)}|^2 \right),$$

$$\mathcal{N}_3(\{A\}, \phi) = -\partial^i \mathfrak{S} \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} - \sqrt{-1} A_i^{(k-1)} |\phi^{(k-1)}|^2 \right),$$

$$\begin{aligned} \mathcal{N}_4(\{A\}, \phi) &= -2\sqrt{-1} \left( -A_0^{(k-1)} \partial_t \phi^{(k-1)} + A \cdot \nabla \phi^{(k-1)} \right) \\ &\quad - |A_0^{(k-1)}|^2 \phi^{(k-1)} + |A^{(k-1)}|^2 \phi^{(k-1)} + \sqrt{-1} \mathcal{B}_0^{(k-1)} \phi^{(k-1)}. \end{aligned}$$

The system we will iterate is then

$$\begin{aligned} (51a) \quad A_i^{(k)} &= \chi_T \left( A_i^{(0)} + \frac{1}{\square} S_{\cdot,1} \mathcal{N}_1(\{A^{(k-1)}\}, \phi^{(k-1)}) \right. \\ &\quad \left. - W \frac{1}{\Xi} S_{\cdot,d>1} \mathcal{N}_1(\{A^{(k-1)}\}, \phi^{(k-1)})(0) \right) \\ &\quad + \frac{1}{\Xi} S_{\cdot,d>1} \mathcal{N}_1(\{A^{(k-1)}\}, \phi^{(k-1)}), \end{aligned}$$

$$(51b) \quad A_0^{(k)} = \frac{1}{\Delta} \mathcal{N}_2(\{A^{(k-1)}\}, \phi^{(k-1)}),$$

$$(51c) \quad \mathcal{B}_0^{(k)} = \frac{1}{\Delta} \mathcal{N}_3(\{A^{(k-1)}\}, \phi^{(k-1)}),$$

$$(51d) \quad \phi^{(k)} = \chi_T \left( \phi^{(0)} + \frac{1}{\square} \mathcal{N}_4(\{A^{(k-1)}\}, \phi^{(k-1)}) \right).$$

Notice that we have replaced the field  $\partial_t A_0$  in the nonlinearity  $\mathcal{N}_4$  with the field  $\mathcal{B}_0$ . Since we will be approximating our  $H^{\frac{1}{2}+\epsilon}$  solution by a sequence of smooth solutions, we need to know that for any sufficiently smooth solution to the system (51), i.e., a solution to those equations with  $k = \infty$ , one has that  $\partial_t A_0 = \mathcal{B}_0$ . This can be seen as follows: Suppose  $\{\{A\}, \mathcal{B}_0, \phi\}$  is a smooth, say at the level of  $H^2$  or better, solution to the above iteration problem. Then on the time interval  $[-1, 1]$ , a direct calculation shows that

$$\begin{aligned} \Delta \partial_t A_0 + \partial^i \mathfrak{S} \left( \phi \overline{\partial_i \phi} - \sqrt{-1} A_i |\phi|^2 \right) \\ = \partial_t A_0 |\phi|^2 + \frac{\partial_i}{\Delta} \mathfrak{S} \left( \phi \overline{\partial_i \phi} - \sqrt{-1} A_i |\phi|^2 \right) |\phi|^2, \end{aligned}$$

so on this time interval, the function

$$F = \partial_t A_0 + \frac{\partial_i}{\Delta} \Im \left( \phi \overline{\partial_i \phi} - \sqrt{-1} A_i |\phi|^2 \right)$$

solves the following equation at fixed time:

$$-\Delta F + F |\phi|^2 = 0 .$$

An integration by parts argument then shows that one must have  $F \equiv 0$  on the interval  $[-1, 1]$ , thus

$$\partial_t A_0 = \mathcal{B}_0$$

on this time interval. This shows that  $\{A, \phi\}$  is in fact a solution to (13) on this time interval.

The iteration procedure begins by setting

$$\square A_i^{(0)} = \square \phi^{(0)} = A_0^{(0)} = \mathcal{B}_0^{(0)} = 0 ,$$

with the initial conditions

$$\begin{aligned} A_i^{(0)}(0) &= A_i(0) \in H^s, & \partial_t A_i^{(0)}(0) &= \partial_t A_i(0) \in H^{s-1}, \\ \phi^{(0)}(0) &= \phi(0) \in H^s, & \partial_t \phi^{(0)}(0) &= \partial_t \phi(0) \in H^{s-1}, \\ \operatorname{div} A_i(0) &= 0, & \operatorname{div} \partial_t A_i(0) &= 0. \end{aligned}$$

All of the above  $H^s$  and  $H^{s-1}$  norms are assumed to be less than some small  $\epsilon(s)$  which will be implicitly determined in the proof. Specifically, due to the quadratic and higher nature of the nonlinearity, we choose  $\epsilon(s)$  so small that all of the embeddings listed below hold with the same uniform constants.

Now, for any  $N$ , we have the inclusions:

$$(52) \quad \chi_T \phi^{(0)}, \chi_T A^{(0)} \hookrightarrow X_1^{s,N},$$

which follow from the fact that both  $\chi_T \phi^{(0)}$  and  $\chi_T A^{(0)}$  have exponential decay away from the light cone in Fourier space.

The bootstrapping assumption we will use in our iteration is the following: for  $l = 1, 2, 3$ , we will assume that there are splittings

$$(53) \quad A_0^{(k-l)} = A_0^{(k-l)}_{H^{s-\frac{1}{2}, 1-s, s}} + A_0^{(k-l)}_{(\bar{L}^\infty \cap L^2)(H^s_{t,x})},$$

$$(54) \quad \mathcal{B}_0^{(k-l)} = \mathcal{B}_0^{(k-l)}_{H^{s-\frac{1}{2}, 0, 0}} + \mathcal{B}_0^{(k-l)}_{\langle D_{t,x} \rangle^{-s-\frac{1}{2}} (\bar{L}^\infty \cap L^2)(\bar{L}^1)},$$

and that the following uniform inclusions hold:

$$(55a) \quad A_i^{(k-l)} \hookrightarrow (\bar{L}^\infty \cap L^2)(H_{t,x}^s),$$

$$(55b) \quad \phi^{(k-l)} \hookrightarrow X_1^{s, \frac{1}{2}},$$

$$(55c) \quad A_0^{(k-l)} \underset{H^{s-\frac{1}{2}, 1-s, s}}{\hookrightarrow} H^{s-\frac{1}{2}, 1-s, s},$$

$$(55d) \quad A_0^{(k-l)} \underset{(\bar{L}^\infty \cap L^2)(H_{t,x}^s)}{\hookrightarrow} (\bar{L}^\infty \cap L^2)(H_{t,x}^s),$$

$$(55e) \quad \mathcal{B}_0^{(k-l)} \underset{H^{s-\frac{1}{2}, 0, 0}}{\hookrightarrow} H^{s-\frac{1}{2}, 0, 0},$$

$$(55f) \quad \mathcal{B}_0^{(k-l)} \underset{\langle D_{t,x} \rangle^{-s-\frac{1}{2}}(\bar{L}^\infty \cap L^2)(\bar{L}^1)}{\hookrightarrow} \langle D_{t,x} \rangle^{-s-\frac{1}{2}}(\bar{L}^\infty \cap L^2)(\bar{L}^1).$$

The main focus of this paper will be to show that if this inductive hypothesis is satisfied, then the above inclusions hold with the same uniform constants for  $l = 0$ . We mention briefly here that there is a similar splitting for the spatial components of the vector  $\{A^{(k-l)}\}$  which involves the space  $X_1^{s, \frac{1}{2}}$ . The reason we do not list the corresponding embeddings here is that their assumption is not needed in order to prove those same estimates at the level of the next iterate. Also, making the bootstrapping assumption (55), we will show that the following embeddings hold, with the same uniform constants as those mentioned above:

$$(56a) \quad A_i^{(k)} \hookrightarrow L^\infty(H_x^s), \quad \partial_t A_i^{(k)} \hookrightarrow L^\infty(H_x^{s-1}),$$

$$(56b) \quad \phi^{(k)} \hookrightarrow L^\infty(H_x^s), \quad \partial_t \phi^{(k)} \hookrightarrow L^\infty(H_x^{s-1}).$$

Finally, we note here that even though we will not spell things out explicitly, the reader will see that our method of proof can be altered in a straightforward way to accommodate differences. Therefore, to prove Theorem 1.3, it will suffice to prove the inductive estimates that we have listed above.

**3.1. Estimates for the spatial potentials  $A_i^{(k)}$ .** We begin this section by decomposing the  $A_i^{(k)}$ , collectively referred to as  $A^{(k)}$  for the remainder of this section, into a sum of pieces, each of which will be estimated separately.

In order to show our induction, all that is required of us is that we prove the inclusion (55a). This, however, is much too weak for our purposes. What we will

need to know is, roughly speaking, that  $A^{(k)}$  is in the space  $X_1^{s, \frac{1}{2}}$ . Of course, as we know from the range of the estimate (20), this cannot hold even at the level of the first iterate if  $s < \frac{2}{3}$ . As a substitute, what we can show is that *most* of  $A^{(k)}$  is in the space  $X_1^{s, \frac{1}{2}}$ . Furthermore, we will show that the piece of  $A^{(k)}$  not in  $X_1^{s, \frac{1}{2}}$  is given by a very specific formula, (58) below. This means that we are in a position to apply Lemma 2.9 to it. This will be our main tool used to control terms that show up later in the paper containing (58), as well as to show that this piece of  $A^{(k)}$  is in the space  $(\bar{L}^\infty \cap L^2)(H_{t,x}^s)$ .

Our first step is to isolate those terms of  $A^{(k)}$  which do not possess exponential decay away from the light cone in Fourier space:

$$(57) \quad A_{cubic}^{(k)} = \frac{1}{\Xi} S_{\cdot, d \geq 1} \mathcal{P} \left[ A^{(k-1)} |\phi^{(k-1)}|^2 \right] ,$$

$$(58) \quad A_{bad\ quadratic}^{(k)} = \left\langle \frac{1}{\Xi} S_{\cdot, d \geq 1} \mathcal{P} \left[ \Im \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right] \right\rangle_{H \times H \Rightarrow L}^{close\ cone} ,$$

$$(59) \quad A_{good\ quadratic}^{(k)} = \frac{1}{\Xi} S_{\cdot, d \geq 1} \mathcal{P} \left[ \Im \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right] - A_{bad\ quadratic}^{(k)} .$$

For the remainder of  $A^{(k)}$  we write:

$$(60) \quad A_{remainder}^{(k)} = A^{(k)} - A_{good\ quadratic}^{(k)} - A_{bad\ quadratic}^{(k)} - A_{cubic}^{(k)} .$$

Our main task is now, assuming the inductive hypothesis (55a)–(55b), to prove the following more specific estimates which easily imply (55a) with  $l = 0$ , and whose specific form will be used in the later sections:

$$(61) \quad A_{remainder}^{(k)} \hookrightarrow X_1^{s, N}, N > 0 ,$$

$$(62) \quad A_{cubic}^{(k)} \hookrightarrow X_1^{s, \frac{1}{2}} ,$$

$$(63) \quad A_{good\ quadratic}^{(k)} \hookrightarrow X_1^{s, \frac{1}{2}} ,$$

$$(64) \quad A_{good\ quadratic}^{(k)} - \left\langle A_{good\ quadratic}^{(k)} \right\rangle_{H \times H \Rightarrow L}^{medium\ cone} \hookrightarrow Y^{s-1, \frac{1}{2}} ,$$

$$(65) \quad A_{bad\ quadratic}^{(k)} \hookrightarrow (L^\infty \cap \bar{L}^\infty \cap L^2)(H_{t,x}^s) ,$$

$$(66) \quad \partial_t A^{(k)} \hookrightarrow L^\infty(H_x^{s-1}) .$$

*Proof of (61).* Using the same notation as in (51a) we write

$$A_{remainder}^{(k)} = \chi_T \left( A_i^{(0)} + \frac{1}{\square} S_{\cdot,1} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) - W \frac{1}{\Xi} S_{\cdot,d>1} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) (0) \right).$$

The inclusions

$$\chi_T A_i^{(0)} \hookrightarrow X_1^{s,N}$$

and

$$\chi_T \left( W \frac{1}{\Xi} S_{\cdot,d>1} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) (0) \right) \hookrightarrow X_1^{s,N}$$

follow from the same reasoning used to prove the inclusion (52), keeping in mind that the embeddings (62)–(66) will imply that

$$\begin{aligned} \frac{1}{\Xi} S_{\cdot,d>1} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) (0) &\hookrightarrow H_x^s, \\ \partial_t \frac{1}{\Xi} S_{\cdot,d>1} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) (0) &\hookrightarrow H_x^{s-1}. \end{aligned}$$

It remains to deal with the term

$$\chi_T \frac{1}{\square} S_{\cdot,1} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right).$$

Due to the well-known embedding

$$\chi_T \frac{1}{\square} : X_1^{s-1,N-1} \hookrightarrow X_1^{s,N},$$

it suffices to show that for any  $N$  the following uniform embeddings hold:

$$(67) \quad S_{\cdot,1} \mathcal{P} \left[ X_1^{s,\frac{1}{2}} \cdot \nabla X_1^{s,\frac{1}{2}} \right] \hookrightarrow X_1^{s-1,N-1},$$

$$(68) \quad S_{\cdot,1} \mathcal{P} \left[ (\bar{L}^\infty \cap L^2)(H^s) \cdot (X_1^{s,\frac{1}{2}})^2 \right] \hookrightarrow X_1^{s-1,N-1}.$$

Notice that, because the above terms are an  $O(1)$  distance from the light cone, all we are really requiring here is that they belong to the space  $X_1^{s,0}$ . This corresponds to the case  $d = 1$  of the embeddings (62)–(65), which we will also show below.  $\square$

*Proof of (62).* This is an immediate consequence of the following:

$$(69) \quad (\bar{L}^\infty \cap L^2)(H^s) \cdot (X_1^{s,\frac{1}{2}})^2 \subseteq X_1^{s-1,-\frac{1}{2}}.$$

To show (69), we proceed by duality. What we are trying to prove is that

$$(\bar{L}^\infty \cap L^2)(H^s) \cdot X_\infty^{1-s,\frac{1}{2}} \cdot (X_1^{s,\frac{1}{2}})^2 \subseteq \mathcal{L}^1.$$

Due to the weight trading properties of the  $\mathcal{L}^1$  “space”, we see that in order to prove the above estimate, it is enough to show the following inclusion:

$$(70) \quad (\bar{L}^\infty \cap L^2)(L^2) \cdot \langle D_{t,x} \rangle^{-s} \left( X_\infty^{1-s,\frac{1}{2}} \cdot (X_1^{s,\frac{1}{2}})^2 \right) \subseteq \mathcal{L}^1.$$

Via interpolation, one sees that

$$(\bar{L}^\infty \cap L^2)(L^2) \subseteq \bar{L}^{\infty-}(L^2).$$

Therefore, after using Hölder’s inequality for the  $\mathcal{L}^1$  “space”, (70) will follow once we can show that

$$\langle D_x \rangle^{-s} \left( X_\infty^{1-s,\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} \right) \subseteq \bar{L}^{1+}(L^2).$$

This follows immediately from the foliated version of (39) and the inclusion

$$X_\infty^{1-s, \frac{1}{2}} \subseteq X_1^{(1-s)-, \frac{1}{2}} .$$

□

Notice that an immediate consequence of (69) is that we have the embedding (68).

Our next task here will be to show the inclusions (63)–(65). We will start by trying to estimate as much of the term

$$\square A_{quadratic}^{(k)} = S_{\cdot, d \geq 1} \mathcal{P} \left[ \Im \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right]$$

as possible in the space  $X_1^{s-1, -\frac{1}{2}}$ . What we will see is that this works for all frequency interactions except for the piece  $\square A_{bad\ quadratic}^{(k)}$  which contains the bad *High* × *High* ⇒ *Low* interaction. For this last piece, we prove the inclusion (65).

In what follows, we will estimate the expression

$$\square A_{full\ quadratic}^{(k)} = \mathcal{P} \left[ \Im \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right] .$$

Having done this, the embedding (68) will follow from the estimates below. Also, we will prove the inclusions (63)–(64) together as the case breakdown is similar for both.

*Proof of (63)–(64).* We begin with a Littlewood–Paley type decomposition with respect to the space-time variable and the distance from the cone:

$$(71) \quad \square A_{full\ quadratic}^{(k)} = \sum_{\substack{\lambda, d \\ \mu_i, \delta_i}} \mathcal{P} S_{\lambda, d} \left[ \Im \left( S_{\mu_1, \delta_1}(\phi^{(k-1)}) \partial_i S_{\mu_2, \delta_2}(\overline{\phi^{(k-1)}}) \right) \right] .$$

Because all the norms we are considering here depend only on the size of the Fourier transform, we can, after using Lemma 2.1, replace the sum (71) by the following expression:

$$(72) \quad \sum_{\substack{\lambda, d \\ \mu, \delta}} \frac{\max\{d, \delta_1, \delta_2\}^{\frac{1}{2}}}{\lambda^{\frac{1}{2}}} S_{\lambda, d} \left( \mu_1^{\frac{1}{2}} S_{\mu_1, \delta_1}(\phi^{(k-1)}) \cdot \mu_2^{\frac{1}{2}} S_{\mu_2, \delta_2}(\overline{\phi^{(k-1)}}) \right) .$$

We remark briefly that all of the dyadic variables in the above sum can be taken to be > 1 or = 1. In the latter case we use the inhomogeneous cutoff  $S_{\leq 1}$ .

We now break down into cases depending on the various sizes of the frequencies involved.

**Case1:**  $\min\{\mu_1, \mu_2\} \lesssim \lambda$ , **and**  $\max\{d, \delta_1, \delta_2\} \neq d$ . Because of the restriction  $\min\{\mu_1, \mu_2\} \lesssim \lambda$ , we see that we are also in the region where  $\max\{\mu_1, \mu_2\} \sim \lambda$ . To show (63) in this range, after some weight trading and using duality, it suffices to show that

$$X_\infty^{1-s, \frac{1}{2}} \cdot X_1^{s-\frac{1}{2}, 0} \cdot X_1^{s, \frac{1}{2}} \subseteq \mathcal{L}^1 .$$

This is an immediate consequence of Hölder’s inequality and the following two inclusions:

$$\begin{aligned} X_1^{s-\frac{1}{2}, 0} &\subseteq \bar{L}^{2+}(\bar{L}^{2+}) && \text{(Sobolev),} \\ X_\infty^{1-s, \frac{1}{2}} \cdot X_1^{s, \frac{1}{2}} &\subseteq \bar{L}^{2-}(\bar{L}^{2-}) && \text{(Strichartz interpolated with } L^2(L^2)\text{).} \end{aligned}$$

We now show (64) in this range. Again, after some weight trading, we see that we are trying to show that

$$\frac{1}{d^{\frac{1}{2}\lambda}} S_{\lambda,d}^\alpha (X_1^{s-\frac{1}{2},0} \cdot X_1^{1,\frac{1}{2}}) \subseteq \bar{L}^{1+}(\bar{L}^\infty),$$

for any given  $S_{\lambda,d}^\alpha$ . This follows from the inclusions

$$\begin{aligned} X_1^{s-\frac{1}{2},0} &\subseteq \bar{L}^2(\bar{L}^{2+}) && \text{(Sobolev),} \\ X_1^{1,\frac{1}{2}} &\subseteq \bar{L}^{2+}(\bar{L}^{\infty-}) && \text{(Strichartz),} \\ \frac{1}{d^{\frac{1}{2}\lambda}} S_{\lambda,d}^\alpha (\bar{L}^{1+}(L^2)) &\subseteq \bar{L}^{1+}(\bar{L}^\infty) && \text{(Local Sobolev).} \end{aligned}$$

**Case2:**  $\min\{\mu_1, \mu_2\} \lesssim \lambda$ , **and**  $\max\{d, \delta_1, \delta_2\} = d$ . To show (63) in this case, we again use the observation that  $\max\{\mu_1, \mu_2\} \sim \lambda$  to trade some weights and see that by duality it is enough to show that

$$X_\infty^{0,s-\frac{1}{2}} \cdot X_1^{\frac{1}{2},\frac{1}{2}} \cdot X_1^{\frac{1}{2},\frac{1}{2}} \subseteq \mathcal{L}^1.$$

This in turn follows from the embeddings

$$\begin{aligned} (X_1^{\frac{1}{2},\frac{1}{2}})^2 &\subseteq L^2(L^2) && \text{(Strichartz),} \\ X_\infty^{0,s-\frac{1}{2}} &\subseteq X_1^{0,0} \subseteq L^2(L^2). \end{aligned}$$

The proof of (64) in this range follows from some weight trading and the embeddings

$$\begin{aligned} X_1^{\frac{s}{2}+\frac{3}{4},\frac{1}{2}} \cdot X_1^{\frac{s}{2}+\frac{3}{4},\frac{1}{2}} &\subseteq \bar{L}^{1+}(\bar{L}^\infty) && \text{(Strichartz \& Sobolev),} \\ S_{\lambda,d}^\alpha (\bar{L}^{1+}(\bar{L}^\infty)) &\subseteq \bar{L}^{1+}(\bar{L}^\infty). \end{aligned}$$

**Case3a:** (63) **with**  $\lambda \ll \mu_1 \sim \mu_2$ , **and**  $d \lesssim \max\{\delta_1, \delta_2\}$ . A short computation of the weights involved shows that by duality, it is enough to prove the inclusion

$$X_\infty^{1,\frac{1}{2}} \cdot X_1^{0,0} \cdot X_1^{s-\frac{1}{2},\frac{1}{2}} \subseteq \mathcal{L}^1.$$

The above can in turn be found by multiplying together the following embeddings:

$$\begin{aligned} X_\infty^{1,\frac{1}{2}} &\subseteq X_1^{1-,\frac{1}{2}} \subseteq \bar{L}^{2+}(\bar{L}^{\infty-}) && \text{(Strichartz),} \\ X_1^{s-\frac{1}{2},\frac{1}{2}} &\subseteq \bar{L}^{\infty-}(\bar{L}^{2+}) && \text{(Energy interp. w/ } L^2 \text{ \& Sobolev),} \\ X_1^{0,0} &\subseteq L^2(L^2). \end{aligned}$$

**Case3b:** (64) **with**  $\lambda \ll \mu_1 \sim \mu_2$ , **and**  $d \lesssim \min\{\delta_1, \delta_2\}$ . We note here that this case is more than enough to take care of the remaining terms in the estimate (64) because the frequency restrictions for this estimate are in the range where  $\lambda \ll \min\{\delta_1, \delta_2\} \sim \max\{\delta_1, \delta_2\}$ . After some redistribution of the weights, we are trying to show that

$$\frac{1}{\lambda^2 d} S_{\lambda,d}^\alpha (X_1^{s-\frac{1}{2},0} \cdot X_1^{0,0}) \subseteq \bar{L}^{1+}(\bar{L}^\infty).$$

This follows from multiplying together the inclusions

$$\begin{aligned} X_1^{s-\frac{1}{2},0} \cdot X_1^{0,0} &\subseteq \bar{L}^{1+}(\bar{L}^1) && (L^2 \text{ \& Sobolev}), \\ \frac{1}{\lambda^2 d} S_{\lambda,d}^\alpha(\bar{L}^{1+}(\bar{L}^1)) &\subseteq \bar{L}^{1+}(\bar{L}^\infty) && (\text{Local Sobolev}). \end{aligned}$$

□

We are finally left with the sum (72) in the region where  $\lambda \ll \mu_1 \sim \mu_2$  and  $\max\{\delta_1, \delta_2\} \ll d$ , in the presence of either a (+−) or a (−+) interaction.

This exhausts what can be done with the sum (71) in the context of the  $X_1^{s,\theta}$  spaces. The problem is that there is no hope of proving the inclusion

$$\sum_{\substack{\lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} \left(\frac{d}{\lambda}\right)^{\frac{1}{2}} S_{\lambda,d} \left(\mu_1^{\frac{1}{2}} S_{\mu_1, \delta_1}^+(\phi^{(k-1)}) \cdot \mu_2^{\frac{1}{2}} S_{\mu_2, \delta_2}^-(\overline{\phi^{(k-1)}})\right) \hookrightarrow X_1^{s-1, -\frac{1}{2}}.$$

We emphasize that this is not merely because we assume only that  $\phi^{(k-1)} \hookrightarrow X_1^{s, \frac{1}{2}}$ . The failure of the above embedding is complete in the sense that it is not even true for functions supported in an  $O(1)$  neighborhood of the light cone in the range  $s < \frac{2}{3}$ . This corresponds to the fact that the estimate (20) is false in this range. The most one can do is to prove (65).

*Proof of (65).* To prove this it suffices to show the (+−) case. What we are trying to show is that the following inclusion holds:

$$(73) \quad \sum_{\substack{\lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} \frac{1}{d^{\frac{1}{2}} \lambda^{\frac{3}{2}}} S_{\lambda,d} \left(\mu_1^{\frac{1}{2}} S_{\mu_1, \delta_1}^+(\phi^{(k-1)}) \cdot \mu_2^{\frac{1}{2}} S_{\mu_2, \delta_2}^-(\overline{\phi^{(k-1)}})\right) \hookrightarrow \bar{L}^\infty(H^s) \cap X_1^{s,0}.$$

We will prove first the  $(\bar{L}^\infty \cap L^\infty)(L^2)$  type estimate. To proceed, we will estimate everything in Fourier space using the dual bound

$$\|\varphi\|_{(\bar{L}^\infty \cap L^\infty)(L^2)} \leq \|\tilde{\varphi}\|_{L_\xi^2(L_\tau^1)}.$$

With this reduction, and after some weight trading, we see that it suffices to show the following convolution inequality in Fourier space:

$$(74) \quad \sum_{\substack{\lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} \frac{1}{d^{\frac{1}{2}} \lambda^{s+\frac{1}{2}}} s_{\lambda,d} \left(\mu_1^s s_{\mu_1, \delta_1}^+(\widetilde{\phi^{(k-1)}}) * \mu_2^s s_{\mu_2, \delta_2}^-(\widetilde{\overline{\phi^{(k-1)}}})\right) \hookrightarrow L_\xi^2(L_\tau^1).$$

The extra factor of  $\lambda^{\frac{1}{2}-s}$  allows us to sum over the various dyadic values of  $\lambda$  and  $d$ , while the fact that  $\mu_1 \sim \mu_2$  allows us to sum out these dyadic values by using a Cauchy–Schwarz. With this in mind, it suffices to look at a single dyadic piece and prove the estimate

$$s_{\lambda,d} \left(s_{\mu_1, \delta_1}^+(L^2) * s_{\mu_2, \delta_2}^-(L^2)\right) \subseteq \lambda(\delta_1 \cdot \delta_2 \cdot d)^{\frac{1}{2}} L_\xi^2(L_\tau^1).$$

We are now in a position to use the decomposition of Lemma 2.9. This allows us to write:

$$(75) \quad s_{\lambda,d} \left(s_{\mu_1, \delta_1}^+(L^2) * s_{\mu_2, \delta_2}^-(L^2)\right) \approx \sum_\alpha s_{\mu; \lambda, d, \delta_1}^\alpha(L^2) * s_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2).$$

Because this is an (essentially) diagonal, it suffices to bound a single term of the expression on the left-hand side of (75).

We are now trying to show that

$$(76) \quad s_{\mu;\lambda,d,\delta_1}^\alpha(L^2) * s_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \subseteq \lambda(\delta_1 \cdot \delta_2 \cdot d)^{\frac{1}{2}} L_\xi^2(L_\tau^1) .$$

To prove (76), we foliate the spaces  $s_{\mu;\lambda,d,\delta_1}^\alpha(L^2)$  and  $s_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2)$  by forward and backward facing light cones, respectively. This leaves us with trying to show that

$$s_{\mu;\lambda,d,\delta_1}^\alpha \widetilde{W^+}(L_\xi^2) * s_{\mu;\lambda,d,\delta_2}^{-\alpha} \widetilde{W^-}(L_\xi^2) \subseteq \lambda d^{\frac{1}{2}} L_\xi^2(L_\tau^1) .$$

The proof of this last estimate follows from a direct integration in  $\tau$  variable, followed by a Cauchy–Schwarz based on the fact that the size of the spatial projection of the set sum of the supports of the  $s_{\mu;\lambda,d,\delta_1}^\alpha$  and  $s_{\mu;\lambda,d,\delta_2}^{-\alpha}$  multipliers is a box of size  $\sim \lambda \times \sqrt{\lambda d} \times \sqrt{\lambda d}$ . This completes the proof of (74).

It remains to prove the  $X_2^{s,0}$  embedding. In this case we are trying to show that

$$\sum_{\substack{\lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} \frac{1}{d^{\frac{1}{2}} \lambda^{s+\frac{1}{2}}} S_{\lambda,d} \left( \mu_1^s S_{\mu_1,\delta_1}^+(\phi^{(k-1)}) \cdot \mu_2^s S_{\mu_2,\delta_2}^-(\overline{\phi^{(k-1)}}) \right) \hookrightarrow L^2(L^2) .$$

As mentioned above, there is enough extra room in the  $\lambda$  variable to sum over all the dyadic values of  $\lambda$  and  $d$ , so it suffices to look at a single dyadic piece and prove that

$$S_{\lambda,d} \left( S_{\mu_1,\delta_1}^+(L_{t,x}^2) \cdot S_{\mu_2,\delta_2}^-(L_{t,x}^2) \right) \subseteq \lambda(\min\{\delta_1, \delta_2\} \cdot d)^{\frac{1}{2}} L^2(L^2) .$$

This will in turn be a consequence of the angular decomposition of Lemma 2.9, followed by a Cauchy–Schwarz on the corresponding diagonal sum, and the following two embeddings:

$$\begin{aligned} S_{\lambda,d}^\alpha(L^2) &\subseteq \lambda d^{\frac{1}{2}} L^2(\bar{L}^\infty) && \text{(Local Sobolev),} \\ S_{\mu_1;\lambda,d,\delta_1}^\alpha(L^2) \cdot S_{\mu_2;\lambda,d,\delta_2}^{-\alpha}(L^2) &\subseteq (\min\{\delta_1, \delta_2\})^{\frac{1}{2}} L^2(\bar{L}^1) && \text{(Energy} \cdot L^2), \end{aligned}$$

where  $s_{\lambda,d}^\alpha$  denotes the block of size  $\sim \lambda \times \sqrt{\lambda d} \times \sqrt{\lambda d} \times d$ , which contains the support of the convolution  $s_{\mu_1;\lambda,d,\delta_1}^\alpha * s_{\mu_2;\lambda,d,\delta_2}^{-\alpha}$ . This completes the proof of (65).  $\square$

*Remark 3.1.* There are a few things to point out here. The first is that the failure of  $A_{bad}^{(k)}$  quadratic to be in the  $X_1^{s,\frac{1}{2}}$  space is really a failure of it to disperse. One can see this by trying to prove any other Strichartz estimate other than that of  $L^\infty(H^s)$  without a loss of derivatives for the expression (24) of the counterexample we introduced in Section 1.3.

We also point out here what seems to be one of the most obvious difficulties that would arise from an attempt to extend our result to the scale invariant Besov space  $\dot{B}^{\frac{1}{2},1}$ . This comes from considering the embedding (73) at this regularity, for the two large frequencies  $\mu_i$  essentially equal and fixed. With the decomposition of Lemma 2.9 in mind, this is morally no easier than trying to show the elliptic embedding

$$(77) \quad |D_x|^{-3} (P_\mu(L_x^2) \cdot P_\mu(L_x^2)) \subseteq B^{0,1} .$$

Of course (77) is false because there is no extra (dimensionless) savings in the higher frequencies which would allow one to sum over all of the lower frequencies in the product  $P_\mu(L_x^2) \cdot P_\mu(L_x^2)$ . It is important to note that this problem would come up when trying to estimate the curvature (9) itself at the scaling, so choosing another

gauge seems unlikely to eliminate this problem. The only hope we can think of is that when one takes into account the angular separation in the product (73), there is enough savings to make the sum over the lower frequencies. Unfortunately, angular separation does not seem to play a role for  $L^\infty(L^1)$  type estimates for products. In fact, using Duhamel’s principle, it would probably be easy to write down a counterexample to the  $\dot{B}^{\frac{1}{2},1}$  energy estimate for a solution to the equation

$$\square A^{(1)} = \frac{\sqrt{\square}}{|D_x|} \left( \phi^{(0)} \cdot \phi^{(0)} \right) ,$$

with  $\phi^{(0)}$  in  $L_x^2$  at fixed frequency.

*Proof of (66).* Due to the estimates (61)–(63) and (65), our task is reduced to showing that

$$\frac{\partial_t}{\Xi} S_{|\xi| \ll |\tau|} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) \hookrightarrow L^\infty(H_x^{s-1}) .$$

After a calculation of the weights involved, a foliation, and restriction to a single dyadic frequency, this is further reduced to showing that

$$(78) \quad \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \mathcal{N}_1 \left( \{A^{(k-1)}\}, \phi^{(k-1)} \right) \hookrightarrow L^2(L^2) .$$

We first establish (78) for the cubic term in  $\mathcal{N}_1$ . In this case we are reduced to the estimate

$$\langle D_x \rangle^{(s-\frac{3}{2})+} \left( \bar{L}^\infty(H_x^s) \cdot (X_1^{s,\frac{1}{2}})^2 \right) \subseteq L^2(L^2) .$$

This in turn follows from the embeddings

$$\begin{aligned} (X_1^{s,\frac{1}{2}})^2 &\subseteq L^2(\bar{L}^{2+}) && \text{(Strichartz \& Sobolev),} \\ \langle D_x \rangle^{(s-\frac{3}{2})+} (\bar{L}^\infty(H_x^s) \cdot L^2(\bar{L}^{2+})) &\subseteq L^2(L^2) && \text{(Sobolev).} \end{aligned}$$

To establish (78) for the quadratic term in  $\mathcal{N}_1$ , it is enough to show that

$$\sum_{\mu_i} \langle D_x \rangle^{s-1} \lambda^{-\frac{1}{2}+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(X_1^{s-\frac{1}{2},\frac{1}{2}}) \cdot S_{\mu_2}(X_1^{s-\frac{1}{2},\frac{1}{2}}) \right) \subseteq L^2(L^2) .$$

Splitting into cases depending on the relative values of the  $\mu_i$  and trading some weights, we are reduced to the inclusions

$$\langle D_x \rangle^{s-1} \left( X_1^{(s-\frac{1}{4})-, \frac{1}{2}} \cdot X_1^{(s-\frac{1}{4})-, \frac{1}{2}} \right) \subseteq L^2(L^2) \quad \text{(Bi-Strichartz),}$$

and

$$\sum_{\mu_1 \sim \mu_2} \langle D_x \rangle^{s-1} \lambda^{-\frac{1}{2}+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(X_1^{s-\frac{1}{2},\frac{1}{2}}) \cdot S_{\mu_2}(X_1^{s-\frac{1}{2},\frac{1}{2}}) \right) \subseteq L^2(L^2) ,$$

where in the last embedding we have  $\lambda \ll \mu_i$ . We may assume that  $\mu_1 = \mu_2 = \mu$  is fixed here, and we further decompose the resulting estimate as follows:

$$(79) \quad \langle D_x \rangle^{s-1} \lambda^{-\frac{1}{2}+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu,\delta_1}(X_1^{s-\frac{1}{2},\frac{1}{2}}) \cdot S_{\mu,\delta_2}(X_1^{s-\frac{1}{2},\frac{1}{2}}) \right) \subseteq L^2(L^2) .$$

To show (79) in the case where  $\lambda \lesssim \max\{\delta_1, \delta_2\}$ , after some weight trading, we are reduced to the embedding

$$\langle D_x \rangle^{(-s-1)+} \left( X_1^{0,\frac{1}{2}} \cdot X_1^{0,0} \right) \subseteq L^2(L^2) ,$$

which in turn follows from the estimates

$$\begin{aligned} X_1^{0, \frac{1}{2}} \cdot X_1^{0,0} &\subseteq L^2(\bar{L}^1) && \text{(Energy \& } L^2), \\ \langle D_x \rangle^{(-1-s)+} L^2(\bar{L}^1) &\subseteq L^2(L^2) && \text{(Sobolev)}. \end{aligned}$$

It remains to show (79) when  $\max\{\delta_1, \delta_2\} \ll \lambda$ . In this case, the presence of the  $S_{|\xi| \ll |\lambda|}$  multiplier assures us that we are in either a  $(++)$  or  $(--)$  interaction. The desired estimate then follows from

$$\langle D_x \rangle^{(s-\frac{3}{2})+} \left( S_{\mu, \delta_1}^\pm (X_1^{s-\frac{1}{2}, \frac{1}{2}}) \cdot S_{\mu, \delta_1}^\pm (X_1^{s-\frac{1}{2}, \frac{1}{2}}) \right) \subseteq L^2(L^2) \quad \text{(Special Strichartz).}$$

□

**3.2. Estimates for the elliptic potentials  $A_0^{(k)}$  and  $\mathcal{B}_0^{(k)}$ .** The goal of this section is parallel to that of Section 3.1 and is, assuming the inductive hypothesis (55), to isolate the term

$$A_0^{(k)} \text{ bad quadratic} = \left\langle \frac{1}{\Delta} \mathfrak{F} \left( \phi^{(k-1)} \overline{\partial_t \phi^{(k-1)}} \right) \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}},$$

as well as prove certain inclusions which are stronger than those of (55). To state these, we write the other terms in  $A_0^{(k)}$  as

$$\begin{aligned} A_0^{(k)} \text{ cubic} &= \frac{1}{\Delta} A_0^{(k-1)} |\phi^{(k-1)}|^2, \\ A_0^{(k)} \text{ good quadratic} &= \frac{1}{\Delta} \mathfrak{F} \left( \phi^{(k-1)} \overline{\partial_t \phi^{(k-1)}} \right) - A_0^{(k)} \text{ bad quadratic}. \end{aligned}$$

We will prove the following inclusions:

---

---


$$(80) \quad A_0^{(k)} \text{ cubic} \hookrightarrow H^{s-\frac{1}{2}, 1-s, s},$$

$$(81) \quad A_0^{(k)} \text{ good quadratic} \hookrightarrow H^{s-\frac{1}{2}, 1-s, s},$$

$$(82) \quad A_0^{(k)} \text{ good quadratic} - \left\langle A_0^{(k)} \text{ good quadratic} \right\rangle_{H \times H \Rightarrow L}^{\text{medium cone}} \hookrightarrow Y^{s-\frac{1}{2}, 0},$$

$$(83) \quad A_0^{(k)} \text{ bad quadratic} \hookrightarrow (\bar{L}^\infty \cap L^2)(H_{t,x}^s),$$

$$(84) \quad \mathcal{B}_0^{(k)} - \left\langle \frac{\partial^i}{\Delta} \mathfrak{F} \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}} \hookrightarrow H^{s-\frac{1}{2}, 0, 0},$$

$$(85) \quad \left\langle \frac{\partial^i}{\Delta} \mathfrak{F} \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}} \hookrightarrow \langle D_{t,x} \rangle^{-s-\frac{1}{2}} (\bar{L}^\infty \cap L^2)(\bar{L}^1).$$


---

---

*Proof of (80).* In our induction, we will assume that the splitting (53) takes the form

$$\begin{aligned} A_0^{(k-1)} \text{ bad quadratic} &\hookrightarrow \bar{L}^\infty(H_{t,x}^s), \\ A_0^{(k-1)} - A_0^{(k-1)} \text{ bad quadratic} &\hookrightarrow H^{s-\frac{1}{2},1-s,s}. \end{aligned}$$

Note that this is consistent with the estimates (81) and (83). We are now reduced to showing that

$$(86) \quad \langle D_{t,x} \rangle^{s-\frac{1}{2}} \langle D_x \rangle^{1-s} |D_x|^{s-2} \left( \bar{L}^\infty(H_{t,x}^s) \cdot (X_1^{s,\frac{1}{2}})^2 \right) \subseteq L^2(L^2),$$

$$(87) \quad \langle D_{t,x} \rangle^{s-\frac{1}{2}} \langle D_x \rangle^{1-s} |D_x|^{s-2} \left( H^{s-\frac{1}{2},1-s,s} \cdot (X_1^{s,\frac{1}{2}})^2 \right) \subseteq L^2(L^2).$$

To prove (86)–(87) for an  $O(1)$  neighborhood of the spatial origin in Fourier space, it is enough to show that

$$\begin{aligned} \langle D_{t,x} \rangle^{s-\frac{1}{2}} P_{\leq 1} \left( \bar{L}^\infty(H_{t,x}^s) \cdot (X_1^{s,\frac{1}{2}})^2 \right) &\subseteq L^2(\bar{L}^1), \\ \langle D_{t,x} \rangle^{s-\frac{1}{2}} P_{\leq 1} \left( H^{s-\frac{1}{2},1-s,s} \cdot (X_1^{s,\frac{1}{2}})^2 \right) &\subseteq L^2(\bar{L}^1). \end{aligned}$$

These in turn follow from the Leibniz rule used to distribute the  $\langle D_{t,x} \rangle^{s-\frac{1}{2}}$  weights, and the embeddings

$$\begin{aligned} \bar{L}^\infty(H_{t,x}^{\frac{1}{2}}) &\subseteq \bar{L}^\infty(L^2), \\ (X_1^{\frac{1}{2},\frac{1}{2}})^2 &\subseteq L^2(L^2) \quad (\text{Strichartz}), \end{aligned}$$

and

$$\begin{aligned} H^{0,1-s,s} &\subseteq L^2(\bar{L}^6) \quad (\text{Sobolev}), \\ X_1^{\frac{1}{2},\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} &\subseteq \bar{L}^\infty(\bar{L}^{\frac{6}{5}}) \quad (\text{Energy \& Sobolev}), \end{aligned}$$

respectively.

To prove (86) for  $P_{\lambda>1}$ , after some weight trading with respect to the spatial variable and using the Leibniz rule used to distribute  $\langle D_{t,x} \rangle^{s-\frac{1}{2}}$ , it is enough to show that

$$\langle D_x \rangle^{-1} \left( \bar{L}^\infty(H_{t,x}^{\frac{1}{2}}) \cdot X_1^{\frac{1}{2},\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} \right) \subseteq L^2(L^2),$$

which follows from the two estimates

$$\begin{aligned} X_1^{\frac{1}{2},\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} &\subseteq L^2(\bar{L}^{2+}) \quad (\text{Strichartz \& Sobolev}), \\ \langle D_x \rangle^{-1} \left( \bar{L}^\infty(H_{t,x}^{\frac{1}{2}}) \cdot L^2(\bar{L}^{2+}) \right) &\subseteq L^2(L^2) \quad (\text{Sobolev}). \end{aligned}$$

To prove (87) for  $P_{\lambda>1}$ , after using Leibniz, we are reduced to showing that

$$\langle D_x \rangle^{-1} \left( H^{0,1-s,s} \cdot X_1^{\frac{1}{2},\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} \right) \subseteq L^2(L^2),$$

which follows from Sobolev’s inequality and the estimates

$$\begin{aligned} H^{0,1-s,s} &\subseteq L^2(\bar{L}^6) \quad (\text{Sobolev}), \\ X_1^{\frac{1}{2},\frac{1}{2}} \cdot X_1^{s,\frac{1}{2}} &\subseteq \bar{L}^\infty(\bar{L}^{\frac{3}{2}}) \quad (\text{Energy \& Sobolev}). \end{aligned}$$

□

*Proof of (81).* Here we decompose

$$A_0^{(k)} \text{ quadratic} = A_0^{(k)} \text{ good quadratic} + A_0^{(k)} \text{ bad quadratic}$$

into a sum of its various dyadic frequencies, replacing derivatives by their numerical dyadic values as follows:

$$(88) \quad \sum_{\mu_i, \delta_i} \frac{1}{\Delta} P_{\leq 1} \left( S_{\mu_1, \delta_1} \phi^{(k-1)} \cdot \mu_2 S_{\mu_2, \delta_2} \overline{\phi^{(k-1)}} \right) + \sum_{\substack{1 \leq \lambda \\ \mu_i, \delta_i}} \frac{1}{\lambda^2} P_\lambda \left( S_{\mu_1, \delta_1} \phi^{(k-1)} \cdot \mu_2 S_{\mu_2, \delta_2} \overline{\phi^{(k-1)}} \right) .$$

We now follow the strategy of Section 3.1, and put as much of the sum (88) into the space  $H^{s-\frac{1}{2}, 1-s, s}$  as possible by trying to prove the estimates

$$(89) \quad \sum_{\mu_i, \delta_i} \max\{\mu_1, \mu_2\}^{s-\frac{1}{2}} P_{\leq 1} \left( S_{\mu_1, \delta_1} \phi^{(k-1)} \cdot \mu_2 S_{\mu_2, \delta_2} \overline{\phi^{(k-1)}} \right) \hookrightarrow L^2(\bar{L}^1) ,$$

$$(90) \quad \sum_{\substack{1 \leq \lambda \\ \mu_i, \delta_i}} \lambda^{-1} \max\{\mu_1, \mu_2\}^{s-\frac{1}{2}} P_\lambda \left( S_{\mu_1, \delta_1} \phi^{(k-1)} \cdot \mu_2 S_{\mu_2, \delta_2} \overline{\phi^{(k-1)}} \right) \hookrightarrow L^2(L^2) .$$

We will see that, as in Section 3.1, this will work for most regions of Fourier space, and that what remains will be precisely  $A_0^{(k)} \text{ bad quadratic}$ . We now break down into cases depending on the relative sizes of the various frequencies involved:

**Case 1:**  $\mu_1 \ll \mu_2$  or  $\mu_2 \ll \mu_1$ . We will assume without loss of generality here that  $\mu_1 \ll \mu_2$ , as the other case follows by symmetry.

If we also have that  $\lambda \ll \mu_2$ , then we see that we are also in the region where  $\delta_2 \sim \mu_2$ , so (89) follows from the inclusion

$$X_1^{s, \frac{1}{2}} \cdot X_1^{0, 0} \subseteq L^2(\bar{L}^1) \quad (\text{Energy} \cdot L^2),$$

while (90) follows from the embedding

$$\langle D_x \rangle^{-1} \left( X_1^{s, \frac{1}{2}} \cdot X_1^{0, 0} \right) \subseteq L^2(L^2) \quad (\text{Energy} \cdot L^2 \ \& \ \text{Sobolev}).$$

If we are in the region where  $\mu_2 \lesssim \lambda$ , then after some weight trading, we see that (89)–(90) follow from the embeddings

$$(91) \quad \begin{aligned} X_1^{s, \frac{1}{2}} \cdot X_1^{\frac{1}{2}, \frac{1}{2}} &\subseteq L^2(\bar{L}^1) && (L^4(L^2) \cdot L^4(L^2)), \\ X_1^{s, \frac{1}{2}} \cdot X_1^{\frac{1}{2}, \frac{1}{2}} &\subseteq L^2(L^2) && (\text{Strichartz}). \end{aligned}$$

**Case 2a:**  $\mu_1 \sim \mu_2$  with either a  $(++)$  or  $(--)$  interaction. In this case, the embedding (89) follows from the inclusion

$$(92) \quad X_1^{s-\frac{1}{2}, \frac{1}{2}} \cdot X_1^{0, \frac{1}{2}} \subseteq L^2(\bar{L}^1) \quad (\text{Energy} \cdot L^2).$$

while if  $\mu_i \lesssim \lambda$ , (90) follows from the inclusion (91). Thus we may assume that we are trying to prove (90) with  $\lambda \ll \mu_1 \sim \mu_2$ .

The desired result follows from

$$\langle D_x \rangle^{-1} \left( S^\pm(X_1^{\frac{s}{2}-\frac{1}{4}, \frac{1}{2}}) \cdot S^\pm(X_1^{\frac{s}{2}-\frac{1}{4}, \frac{1}{2}}) \right) \subseteq L^2(L^2) \quad (\text{Spec. Strichartz}),$$

which is a direct consequence of the foliated version of the special Strichartz estimate (34). Here, the multiplier  $S^\pm$  denotes a cutoff in the upper (resp. lower) half-plane in Fourier space.

**Case 2b:**  $\mu_1 \sim \mu_2$  with either a  $(+-)$  or  $(-+)$  interaction. In this case (89) follows again from the inclusion (92), while (90) in the case  $\mu_i \lesssim \lambda$  follows again from (91). Thus, we will assume that we are trying to prove (90) with  $\lambda \ll \mu_1 \sim \mu_2$ .

Our first step will be to decompose (90) further, taking into account the distance from the cone. What we would like to do is to be able to show that

$$(93) \quad \sum_{\substack{\lambda, d, \mu_i, \delta_i \\ \lambda \ll \mu_1 \sim \mu_2}} \lambda^{-1} P_{\lambda, d} \left( \mu_1^s S_{\mu_1, \delta_1}^\pm \phi^{(k-1)} \cdot \mu_2^{\frac{1}{2}} S_{\mu_2, \delta_2}^\mp \overline{\phi^{(k-1)}} \right) \hookrightarrow L^2(L^2).$$

If we further restrict the above sum to the region where  $\lambda \lesssim \max\{\delta_1, \delta_2\}$ , then by weight trading, we see that (90) follows from the estimate

$$\langle D_x \rangle^{-\frac{3}{2}} \left( X_1^{\frac{s}{2} - \frac{1}{4}, \frac{1}{2}} \cdot X_1^{0,0} \right) \subseteq L^2(L^2) \quad (\text{Energy \& Sobolev}).$$

Likewise, in the region  $d \lesssim \max\{\delta_1, \delta_2\}$ , weight trading and duality show that (90) follows from the estimate

$$\langle D_x \rangle^{-1} X_2^{\frac{s}{2} - \frac{1}{4}, \frac{1}{2}} \cdot X_1^{\frac{s}{2} - \frac{1}{4}, \frac{1}{2}} \cdot X_1^{0,0} \subseteq \mathcal{L}^1,$$

which in turn follows from multiplying through the embeddings

$$\begin{aligned} \langle D_x \rangle^{-1} X_2^{\frac{s}{2} - \frac{1}{4}, \frac{1}{2}} &\subseteq \langle D_x \rangle^{-1} X_1^{0, \frac{1}{2}} \subseteq \bar{L}^{2+}(\bar{L}^{\infty-}) && (\text{Strichartz}), \\ X_1^{\frac{s}{2} - \frac{1}{4}, \frac{1}{2}} &\subseteq \bar{L}^{\infty-}(\bar{L}^{2+}) && (\text{Energy interp. w/ } L^2 \text{ \& Sobolev}), \\ X_1^{0,0} &\subseteq L^2(L^2). \end{aligned}$$

If we consider the sum (93) in the region where  $\max\{\delta_1, \delta_2\} \ll \min\{\lambda, d\}$ , then by inspecting the Fourier support of the product, we see that we may replace the multiplier  $P_\lambda$  by  $S_\lambda$ . The sum then becomes

$$\sum_{\substack{\lambda, d, \mu_i, \delta_i \\ \lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} \lambda^{s-\frac{3}{2}} S_{\lambda, d} \left( \mu_1^{\frac{1}{2}} S_{\mu_1, \delta_1} \phi^{(k-1)} \cdot \mu_2^{\frac{1}{2}} S_{\mu_2, \delta_2} \overline{\phi^{(k-1)}} \right),$$

which is exactly the frequency interaction corresponding to  $A_0^{(k)}$  bad quadratic. This completes the proof of (81).  $\square$

*Proof of (82).* This follows directly from the proof of the embedding (64). The important thing to note here is that the  $Y$  norms only see the region of Fourier space where  $1 \leq |\tau| \leq 2|\xi|$ , so one may replace the symbol of  $\frac{1}{\Delta}$  with  $\frac{1}{\langle D_{t,x} \rangle^2}$ . This, of course, is better than the symbol of  $\frac{1}{\Xi}$ , so one may use the proof of (64) verbatim.  $\square$

*Proof of (83).* Again, because the Fourier support of the interaction  $\langle \cdot \rangle_{H \times H \Rightarrow L}^{close\ cone}$  is contained in the region where  $1 \leq |\tau| \leq 2|\xi|$ , we may replace the symbol of  $\frac{1}{\Delta}$  with that of  $\frac{1}{\langle D_{t,x} \rangle^2}$  and follow the proof of (65) above.  $\square$

*Proof of (84).* We will first concentrate on the cubic term and prove the inclusion

$$\frac{\partial_i}{\Delta} \mathfrak{S} \left( A_i^{(k-1)} |\phi^{(k-1)}|^2 \right) \hookrightarrow H^{s-\frac{1}{2},0,0} .$$

Using the energy estimate for  $A_i^{(k-1)}$ , this follows from the embedding

$$\frac{1}{|D_x|} \left( \bar{L}^\infty(H_{t,x}^s) \cdot (X_1^{s,\frac{1}{2}})^2 \right) \subseteq H^{s-\frac{1}{2},0,0} ,$$

which may be proved in exactly the same manner as (86) above.

It remains to deal with the quadratic term in  $\mathcal{B}_0^{(k)}$ . It is enough to show that

$$\begin{aligned} & |D_x|^{-1} \langle D_x \rangle^{-1} \mathfrak{S} \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \\ & - \left\langle |D_x|^{-1} \langle D_x \rangle^{-1} \mathfrak{S} \left( \phi^{(k-1)} \overline{\partial_i \phi^{(k-1)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \hookrightarrow H^{s-\frac{1}{2},1,0} . \end{aligned}$$

This follows from the exact same procedure used in the proof of (81) above. This completes the proof of (84).  $\square$

*Proof of (85).* Expanding out the  $\langle \cdot \rangle_{H \times H \Rightarrow L}^{close\ cone}$  notation and trading all derivatives for their dyadic values, it is enough to show that

$$X_1^{s-\frac{1}{2},\frac{1}{2}} \cdot X_1^{0,\frac{1}{2}} \subseteq (\bar{L}^\infty \cap L^2)(\bar{L}^1) ,$$

which follows directly from the energy and  $L^2$  estimates for the  $X_1^{0,\frac{1}{2}}$  spaces.  $\square$

**3.3. Estimates for  $\phi^{(k)}$  that do not involve the “elliptic cancellations”, part I: Terms involving  $A_i^{(k-1)}$ .** The goal of this and the final two subsections of the paper is to prove the inclusion (55b) for  $l = 0$ . To accomplish this we assume (55), as well as the more specific embeddings (61)–(65), proved in the previous sections. We will only be concerned here with estimating those terms in  $\phi^{(k)}$  which can be placed in the space  $X_1^{s,\frac{1}{2}}$  based solely on the analysis of the spatial potentials  $A_i^{(k-1)}$ . These will turn out to be all terms in  $\phi^{(k)}$  involving the  $A_i^{(k-1)}$  except for certain exceptional *High*  $\times$  *Low*  $\Rightarrow$  *High* interactions.

In order to list the embeddings proved in this section, we first use the following notation to isolate the part of  $\square \phi^{(k)}$  involving  $A_i^{(k-1)}$  *quadratic* which can be analyzed without adding in terms involving  $A_0^{(k-1)}$ . Recalling the notation

$$A_{quadratic}^{(k-1)} = \frac{1}{\Xi} S_{\cdot, d \geq 1} \mathcal{P} \left[ \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_i \phi^{(k-2)}} \right) \right] ,$$

as well as the decomposition

$$A_{quadratic}^{(k-1)} = A_{good\ quadratic}^{(k-1)} + A_{bad\ quadratic}^{(k-1)} ,$$

we write

$$\begin{aligned} \phi_{good\ quadratic\ A_i}^{(k)} &= A_i^{(k-1)}_{quadratic} \cdot \partial^i \phi^{(k-1)} \\ &- \left\langle \left\langle A_i^{(k-1)}_{good\ quadratic} \right\rangle_{H \times H \Rightarrow L}^{medium\ cone} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ &- \left\langle A_i^{(k-1)}_{bad\ quadratic} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ &- \left\langle A_i^{(k-1)}_{bad\ quadratic} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{medium\ cone} . \end{aligned}$$

Then we will show that

$$(94) \quad |A_i^{(k-1)}|^2 \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

$$(95) \quad A_i^{(k-1)}_{remainder} \cdot \partial^i \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

$$(96) \quad A_i^{(k-1)}_{cubic} \cdot \partial^i \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

$$(97) \quad \phi_{good\ quadratic\ A_i}^{(k)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

*Proof of (94).* Using the results of subsection 3.1, we can write

$$A_i^{(k-1)} = A_i^{(k-1)}_{X_1^{s, \frac{1}{2}}} + \left\langle A_i^{(k-1)}_{quadratic} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} ,$$

where

$$A_i^{(k-1)}_{X_1^{s, \frac{1}{2}}} \hookrightarrow X_1^{s, \frac{1}{2}} .$$

Substituting this into the left-hand side of (94), and expanding out the square, we first estimate the terms which do not involve  $(A_i^{(k-1)}_{quadratic})^2$  with the embeddings

$$X_\infty^{1-s, \frac{1}{2}} \cdot (X_1^{s, \frac{1}{2}})^3 \subseteq \mathcal{L}^1 \quad (\text{Strichartz}),$$

$$\bar{L}^\infty(H_{t,x}^s) \cdot (X_1^{s, \frac{1}{2}})^2 \cdot X_\infty^{1-s, \frac{1}{2}} \subseteq \mathcal{L}^1 \quad (\text{Tri-Strichartz}).$$

The embedding labeled (Tri-Strichartz) has been previously shown and is estimate (70).

We now deal with the term involving  $(A_i^{(k-1)}_{quadratic})^2$ . What we are trying to do is to show that

$$(98) \quad \left( \sum_{\substack{1 < \lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} \frac{1}{\Xi} S_{\lambda, d} \mathcal{P} \left[ S_{\mu_1, \delta_1} \phi^{(k-2)} \cdot \nabla S_{\mu_2, \delta_2} \overline{\phi^{(k-2)}} \right] \right)^2 \cdot \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

Replacing all derivatives by their numerical dyadic values, using the null form estimate (28), and using the fact that there is enough room in the  $\lambda$  variable to add over the various dyadic values of  $\lambda$  and  $d$ , we can reduce the above embedding to showing that

$$(99) \quad S_{\lambda_1, d_1} \left( S_{\mu_1, \delta_1} (X_1^{0, \frac{1}{2}}) \cdot S_{\mu_2, \delta_2} (X_1^{0, \frac{1}{2}}) \right) \cdot S_{\lambda_2, d_2} \left( S_{\sigma_1, \eta_1} (X_1^{0, \frac{1}{2}}) \cdot S_{\sigma_2, \eta_2} (X_1^{0, \frac{1}{2}}) \right) \cdot X_1^{s, \frac{1}{2}} \\ \subseteq \max\{\lambda_1, \lambda_2\}^{0+} \lambda_1^{\frac{3}{2}} \lambda_2^{\frac{3}{2}} d_1^{\frac{1}{2}} d_2^{\frac{1}{2}} X_1^{s-1, -\frac{1}{2}},$$

where we are in the range

- (1)  $\lambda_1 \ll \mu_1 \sim \mu_2$ ;
- (2)  $\max\{\delta_1, \delta_2\} \ll d_1$ ;
- (3)  $\lambda_2 \ll \sigma_1 \sim \sigma_2$ ;
- (4)  $\max\{\eta_2, \eta_2\} \ll d_1$ .

For the remainder of this section we will assume, without loss of generality, that  $\mu_1 = \mu_2 = \mu$  and  $\sigma_1 = \sigma_2 = \sigma$ . We now employ the angular decomposition of Lemma 2.9 to the first two terms in the product on the left-hand side of (99). Focusing on a single term in the resulting sum, we see that it is enough to prove the following weighted block type embedding:

$$(100) \quad S_{\mu; \lambda_1, d_1, \delta_1}^\alpha (L^2) \cdot S_{\mu; \lambda_1, d_1, \delta_2}^{-\alpha} (L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_1}^\beta (L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_2}^{-\beta} (L^2) \\ \cdot X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \subseteq \max\{\lambda_1, \lambda_2\}^{0+} \lambda_1^{\frac{3}{2}} \lambda_2^{\frac{3}{2}} (d_1 d_2 \delta_1 \delta_2 \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1,$$

for the range stated above.

We will now assume that  $\lambda_1 \leq \lambda_2$ . With this in mind, we see that we may replace the product  $X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}}$  appearing in the left-hand side of (100) with the expression

$$(101) \quad S \lesssim_{\lambda_2} \left[ X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \right].$$

We first estimate the term (101) by using the foliated bilinear Strichartz estimate (42), followed by a Sobolev embedding:

$$(102) \quad S \lesssim_{\lambda_2} \left[ X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \right] \subseteq \lambda_2^{1+} \bar{L}^{1+}(\bar{L}^\infty).$$

Thus, to complete the proof of (100), we need to be able to show the inclusion

$$(103) \quad S_{\mu; \lambda_1, d_1, \delta_1}^\alpha (L^2) \cdot S_{\mu; \lambda_1, d_1, \delta_2}^{-\alpha} (L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_1}^\beta (L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_2}^{-\beta} (L^2) \\ \subseteq \lambda_1^{\frac{3}{2}} \lambda_2^{\frac{1}{2}} (d_1 d_2 \delta_1 \delta_2 \eta_1 \eta_2)^{\frac{1}{2}} \bar{L}^{\infty-}(\bar{L}^1).$$

We now break down into two cases depending on the relative sizes of  $\sqrt{\lambda_1 d_1}$  and  $\sqrt{\lambda_2 d_2}$ .

**Case 1:**  $\sqrt{\lambda_1 d_1} \lesssim \sqrt{\lambda_2 d_2}$ . This case follows directly from multiplying the following embeddings and then trading  $\sqrt{\lambda_1 d_1}$  for  $\sqrt{\lambda_2 d_2}$ :

$$S_{\mu; \lambda_1, d_1, \delta_1}^\alpha (L^2) \cdot S_{\mu; \lambda_1, d_1, \delta_2}^{-\alpha} (L^2) \subseteq \lambda_1^2 d_1 (\delta_1 \delta_2)^{\frac{1}{2}} \bar{L}^\infty(\bar{L}^\infty) \quad (\text{Local Sobolev}), \\ S_{\sigma; \lambda_2, d_2, \eta_1}^\beta (L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_2}^{-\beta} (L^2) \subseteq (\eta_1 \eta_2)^{\frac{1}{2}} \bar{L}^{\infty-}(\bar{L}^1) \quad (\text{Energy} \cdot L^2).$$

**Case 2:**  $\sqrt{\lambda_2 d_2} \ll \sqrt{\lambda_1 d_1}$ . Here we will need to regroup the terms in the product (103) and compute the convolutions in Fourier space with a bit more care. What we will do is to multiply and interpolate the following embeddings:

$$(104) \quad S_{\mu; \lambda_1, d_1, \delta_1}^\alpha(L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_1}^\beta(L^2) \subseteq (\lambda_1 \lambda_2 d_2 \delta_1 \eta_1)^{\frac{1}{2}} \bar{L}^\infty(L^2) \cap L^2(L^2),$$

$$(105) \quad S_{\mu_1; \lambda_1, d_1, \delta_2}^{-\alpha}(L^2) \cdot S_{\sigma; \lambda_2, d_2, \eta_2}^{-\beta}(L^2) \subseteq (\lambda_1 \lambda_2 d_2 \delta_2 \eta_2)^{\frac{1}{2}} \bar{L}^\infty(L^2) \cap L^2(L^2).$$

The estimates (104)–(105) will follow from a direct calculation in Fourier space using the  $L_\xi^2(L_\tau^2)$  and  $L_\xi^2(L_\tau^1)$  dual bounds. Thus, for a given set of positive test functions  $\tilde{\phi}^i$ , we need to decompose the convolution

$$(106) \quad s_{\mu; \lambda_1, d_1, \delta_1}^\alpha \tilde{\phi}^1 * s_{\sigma; \lambda_2, d_2, \eta_1}^\beta \tilde{\phi}^2.$$

We begin by tiling the spatial domain with boxes,  $W^\gamma$ , of size  $\sim \lambda_1$  with the property that the sides of the  $W^\gamma$  are parallel to the sides of the spatial projection of the multiplier  $s_{\sigma; \lambda_2, d_2, \eta_1}^\beta$  and such that the spatial projection of  $s_{\mu; \lambda_1, d_1, \delta_1}^\alpha$  is completely contained in one of these boxes.

Next, we cut the box, say  $W^{\gamma_0}$ , containing  $s_{\mu; \lambda_1, d_1, \delta_1}^\alpha$  into disjoint rectangles of size  $\lambda_1 \times \sqrt{\lambda_2 d_2} \times \sqrt{\lambda_2 d_2}$  with sides parallel to  $W^{\gamma_0}$ . We will call these  $B^k$ . We also cut the spatial support of  $s_{\sigma; \lambda_2, d_2, \eta_1}^\beta$  into rectangles of the same size and orientation, which we denote by  $C^l$ . We label the (spatial) cutoff functions associated with the  $B^k$  and  $C^l$  by  $b^k$  and  $c^l$ , respectively.

The key fact about the geometry of these blocks which we will use is that if we fix  $(k_0, l_0)$ , then  $\text{supp}\{b^{k_0} * c^{l_0}\}$  intersects only a fixed number of the sets  $\text{supp}\{b^k * c^l\}$  for varying indices  $(k, l)$ , independently of the various length parameters appearing in the decomposition described above. What this means is that as far as norms are concerned, we can write the convolution (106) as a sum of functions that behave essentially as if their supports were disjoint. For instance, in the case of the  $L_\xi^2(L_\tau^1)$  embedding in (104)–(105) we can write:

$$\begin{aligned} \left\| \sum_{k,l} b^k s_{\mu; \lambda_1, d_1, \delta_1}^\alpha \tilde{\phi}^1 * c^l s_{\sigma; \lambda_2, d_2, \eta_1}^\beta \tilde{\phi}^2 \right\|_{L_\xi^2(L_\tau^1)}^2 \\ \lesssim \sum_{k,l} \| b^k s_{\mu; \lambda_1, d_1, \delta_1}^\alpha \tilde{\phi}^1 * c^l s_{\sigma; \lambda_2, d_2, \eta_1}^\beta \tilde{\phi}^2 \|_{L_\xi^2(L_\tau^1)}^2. \end{aligned}$$

After a foliation, this reduces the proof of (104)–(105) in the  $L_\xi^2(L_\tau^1)$  case to showing that

$$b^k s_{\mu; \lambda_1, d_1, \delta_1}^\alpha \widetilde{W^\pm(L_\xi^2)} * c^l s_{\sigma; \lambda_2, d_2, \eta_1}^\beta \widetilde{W^\pm(L_\xi^2)} \subseteq (\lambda_1 \lambda_2 d_2)^{\frac{1}{2}} L_\xi^2(L_\tau^1).$$

This last estimate follows from a direct integration in the  $\tau$  variable, followed by a Cauchy–Schwarz in the spatial variable based on the fact that the  $L^2$  norm of the spatial support of the above multipliers is  $(\lambda_1 \lambda_2 d_2)^{\frac{1}{2}}$ .

To obtain the pure  $L^2$  estimate, a direct Cauchy–Schwarz reveals that

$$b^k s_{\mu; \lambda_1, d_1, \delta_1}^\alpha(L^2) * c^l s_{\sigma; \lambda_2, d_2, \eta_1}^\beta(L^2) \subseteq (\lambda_1 \lambda_2 d_2 \cdot \min\{\delta_1, \eta_1\})^{\frac{1}{2}} L_\xi^2(L_\tau^2),$$

and the desired result follows. This concludes the proof of (94). □

*Proof of (95).* Recall the embedding

$$A_i^{(k-1)} \text{ remainder} \hookrightarrow X_1^{s,N}, \quad N > 0.$$

We will first prove the following more general estimate, which will be useful in the remainder of this section:

$$(107) \quad \mathcal{P}(X_1^{s, \frac{1}{2}}) \cdot \nabla X_1^{s, \frac{1}{2}} - \left\langle \mathcal{P}(X_1^{s, \frac{1}{2}}) \cdot \nabla X_1^{s, \frac{1}{2}} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \subseteq X_1^{s-1, -\frac{1}{2}} .$$

By duality and the adjointness of  $\mathcal{P}$ , to prove (107) it suffices to show the estimate

$$X_1^{s, \frac{1}{2}} \cdot \mathcal{P}(X_\infty^{s-1, \frac{1}{2}} \cdot \nabla X_1^{s, \frac{1}{2}}) - X_1^{s, \frac{1}{2}} \cdot \left\langle \mathcal{P}(X_\infty^{s-1, \frac{1}{2}} \cdot \nabla X_1^{s, \frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \mathcal{L}^1 .$$

But this is essentially identical to the estimates used to prove the inclusion (63). The only difference is the switch between a pair of the  $X_1^{s, \frac{1}{2}}$  and  $X_\infty^{1-s, \frac{1}{2}}$ . Some care needs to be taken to leave room for the sum away from the light cone in the spaces  $X_\infty^{1-s, \frac{1}{2}}$ . We leave this as a (straightforward) exercise for the reader.

To prove (95), it remains for us to show that

$$\left\langle A_i^{(k-1)}_{remainder} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

By duality, the above estimate will follow from the estimate

$$X_1^{s, N} \cdot \left\langle \mathcal{P}(X_\infty^{s-1, \frac{1}{2}} \cdot \nabla X_1^{s, \frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \mathcal{L}^1 .$$

By the properties of the  $\mathcal{L}^1$  notation, we may replace all derivatives in the above expression by their numerical dyadic value. Thus, we are reduced to showing that

$$X_1^{s+\frac{1}{2}, N-\frac{1}{2}} \cdot X_1^{0, \frac{1}{2}} \cdot X_\infty^{0, \frac{1}{2}} \subseteq \mathcal{L}^1 ,$$

which follows from the weight trading properties of the “space”  $\mathcal{L}^1$  and the inclusions

$$\begin{aligned} X_1^{1, N-\frac{1}{2}} &\subseteq \bar{L}^{2+}(\bar{L}^{\infty-}) && \text{(Strichartz),} \\ \langle D_{t,x} \rangle^{\frac{1}{2}-s} (X_\infty^{0, \frac{1}{2}} \cdot X_1^{0, \frac{1}{2}}) &\subseteq \bar{L}^{2-}(\bar{L}^{1+}) && (L^2 \cdot \text{Energy interp. w/ } L^2 \text{ \& Sobolev}). \end{aligned}$$

□

*Proof of (96).* Using the embedding

$$(108) \quad A_i^{(k-1)}_{cubic} \hookrightarrow X_1^{s, \frac{1}{2}} ,$$

and the inclusion (107), we are reduced to showing that

$$(109) \quad \left\langle A_i^{(k-1)}_{cubic} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

This will require more than just the embedding (108).

To prove (109), we expand  $A_i^{(k-1)}_{cubic}$  in terms of its previous iterates as follows:

$$(110) \quad A_i^{(k-1)}_{cubic} = \mathcal{P}S_{,d>1} \frac{1}{\Xi} \left[ A_i^{(k-2)}_{X_1^{s, \frac{1}{2}}} \cdot |\phi^{(k-2)}|^2 + \left\langle A_i^{(k-2)}_{quadratic} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot |\phi^{(k-2)}|^2 \right] .$$

To deal with the first term in (110) inserted into (109), we prove the following embedding:

$$(111) \quad \left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{,d>1} (X_1^{s,\frac{1}{2}})^3 \cdot \nabla X_1^{s,\frac{1}{2}} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \subseteq X_1^{s-1,-\frac{1}{2}}.$$

By duality, this is equivalent to trying to show that

$$(X_1^{s,\frac{1}{2}})^3 \cdot \left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{,d>1} (X_\infty^{1-s,\frac{1}{2}} \nabla X_1^{s,\frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \mathcal{L}^1.$$

An argument similar to the one used to prove (65) can be used to show that

$$\left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{,d>1} (X_\infty^{1-s,\frac{1}{2}} \nabla X_1^{s,\frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq (\bar{L}^\infty \cap L^2)(H_{t,x}^{\frac{1}{2}-}),$$

where the  $(s - \frac{1}{2})_+$  loss is due to the weights involved and the  $X_\infty$  Besov structure. Thus, in order to prove (111), it is enough to show that

$$\langle D_{t,x} \rangle^{(-\frac{1}{2})_+} (X_1^{s,\frac{1}{2}})^3 \subseteq \bar{L}^q(L^2),$$

for some  $1 < q < 2$ . This follows from the foliated version of the trilinear Strichartz estimate (35) after a little weight trading.

To deal with the second term in (110) inserted into (109), we need to show the following:

$$\begin{aligned} \left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{,d>1} \left( \left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{,d>1} (X_1^{s,\frac{1}{2}} \cdot \nabla X_1^{s,\frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot (X_1^{s,\frac{1}{2}})^2 \right) \cdot \nabla X_1^{s,\frac{1}{2}} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ \subseteq X_1^{s-1,-\frac{1}{2}}. \end{aligned}$$

By duality this is equivalent to showing the embedding

$$\begin{aligned} (X_1^{s,\frac{1}{2}})^2 \cdot \left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{\lambda,d>1} (X_1^{s,\frac{1}{2}} \cdot \nabla X_1^{s,\frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \\ \cdot \left\langle \mathcal{P}_{\frac{1}{\Xi}} S_{\lambda,d>1} (X_1^{s,\frac{1}{2}} \cdot \nabla X_\infty^{1-s,\frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \mathcal{L}^1. \end{aligned}$$

The proof of this last estimate is essentially identical to the proof of the embedding (98), following from a slight variant of the estimates which appeared there. This concludes the proof of (96).  $\square$

*Proof of (97).* We begin by observing that the inclusion

$$A_i^{(k-1)} \text{ good quadratic} \hookrightarrow X_1^{s,\frac{1}{2}},$$

together with the embedding (107) shows that one has the following estimate:

$$\begin{aligned} A_i^{(k-1)} \text{ good quadratic} \partial^i \phi^{(k-1)} \\ - \left\langle A_i^{(k-1)} \text{ good quadratic} \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \hookrightarrow X_1^{s-1,-\frac{1}{2}}. \end{aligned}$$

Therefore, our first order of business will be to prove the inclusion

$$(112) \quad \begin{aligned} \left\langle A_i^{(k-1)} \text{ good quadratic} \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ - \left\langle \left\langle A_i^{(k-1)} \text{ good quadratic} \right\rangle_{H \times L \Rightarrow H}^{medium\ cone} \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \hookrightarrow X_1^{s-1,-\frac{1}{2}}. \end{aligned}$$

To show (112), we will use the fact that

$$A_i^{(k-1)} \text{ good quadratic} - \left\langle A_i^{(k-1)} \text{ good quadratic} \right\rangle_{H \times L \Rightarrow H}^{\text{medium cone}} \hookrightarrow Y^{s-1, \frac{1}{2}}.$$

Then by duality, it is enough to show the following embedding holds:

$$Y^{s-1, \frac{1}{2}} \cdot \left\langle \mathcal{P} \left( X_\infty^{s-1, \frac{1}{2}} \nabla X_1^{s, \frac{1}{2}} \right) \right\rangle_{H \times L \Rightarrow H}^{\text{close cone}} \subseteq \mathcal{L}^1.$$

Expanding the notation out in the above product, we see that we are trying to show that

$$\sum_{\substack{\lambda \ll \mu_1 \sim \mu_2 \\ \max\{\delta_1, \delta_2\} \ll d}} S_{\lambda, d}(Y^{s-1, \frac{1}{2}}) \cdot \mathcal{P} \left( \nabla S_{\mu_1, \delta_1}(X_1^{s, \frac{1}{2}}) \cdot S_{\mu_2, \delta_2}(X_\infty^{s-1, \frac{1}{2}}) \right) \subseteq \mathcal{L}^1.$$

We note that because of the range of the various dyadic variables in the above expression, it is enough to prove it for a single dyadic piece at a loss of  $\lambda^{0+}$ . Thus, trading derivatives for their numerical dyadic values, and using the decomposition of Lemma 2.9 on the high frequency product, after a Cauchy–Schwarz we are reduced to showing that

$$\frac{d^{\frac{1}{2}}}{\lambda^{\frac{1}{2}-}} S_{\lambda, d}^\alpha(Y^{s-1, \frac{1}{2}}) \cdot \mu_1^{\frac{1}{2}} S_{\mu_1; \lambda, d, \delta_1}^\alpha(X_1^{s, \frac{1}{2}}) \cdot \mu_2^{\frac{1}{2}} S_{\mu_2; \lambda, d, \delta_2}^{-\alpha}(X_\infty^{s-1, \frac{1}{2}}) \subseteq \mathcal{L}^1,$$

where  $\text{supp}\{s_{\lambda, d}^\alpha\}$  contains  $\text{supp}\{s_{\mu_1; \lambda, d, \delta_1}^\alpha * s_{\mu_2; \lambda, d, \delta_2}^{-\alpha}\}$ . This last estimate follows from the fact that  $\delta_2 \ll \lambda$  and the following set of embeddings:

$$\begin{aligned} \frac{d^{\frac{1}{2}}}{\lambda^{\frac{1}{2}-}} S_{\lambda, d}^\alpha(Y^{s-1, \frac{1}{2}}) &\subseteq \bar{L}^{1-}(\bar{L}^\infty) && \text{(Definition of } Y^{s-1, \frac{1}{2}}), \\ X_1^{0, \frac{1}{2}} &\subseteq \bar{L}^{\infty-}(L^2) && \text{(Energy interp. w/ } L^2), \\ \delta_2^{0-} S_{\mu_2; \lambda, d, \delta_2}^{-\alpha}(X_\infty^{0, \frac{1}{2}}) &\subseteq X_1^{0, \frac{1}{2}} \subseteq \bar{L}^\infty(L^2) && \text{(Energy)}. \end{aligned}$$

It remains to prove the inclusion

$$\begin{aligned} A_i^{(k-1)} \text{ bad quadratic} \cdot \partial^i \phi^{(k-1)} - \left\langle A_i^{(k-1)} \text{ bad quadratic} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{close cone}} \\ - \left\langle A_i^{(k-1)} \text{ bad quadratic} \cdot \partial^i \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{medium cone}} \hookrightarrow X_1^{s-1, -\frac{1}{2}}. \end{aligned}$$

By duality, this is equivalent to showing that

$$\begin{aligned} (113) \quad A_i^{(k-1)} \text{ bad quadratic} \cdot \psi \partial^i \phi^{(k-1)} - A_i^{(k-1)} \text{ bad quadratic} \cdot \left\langle \psi \nabla \phi^{(k-1)} \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}} \\ - A_i^{(k-1)} \text{ bad quadratic} \cdot \left\langle \psi \nabla \phi^{(k-1)} \right\rangle_{H \times H \Rightarrow L}^{\text{medium cone}} \hookrightarrow \mathcal{L}^1, \end{aligned}$$

where  $\psi \in X_\infty^{1-s, \frac{1}{2}}$ .

The proof of (113) is of roughly the same form as (112). To see this, we begin by expanding out  $A_i^{(k-1)} \text{ bad quadratic}$  in terms of the previous iterates. Then under the integral sign, using duality and the fact that  $\mathcal{P}^2 = \mathcal{P}$ , we may write the first

term on the left-hand side of (113) as

$$\begin{aligned} & \int A_i^{(k-1)} \textit{bad quadratic} \cdot \psi \partial^i \phi^{(k-1)} \\ &= \int \frac{1}{\Xi} \mathcal{P} \left( \psi \nabla \phi^{(k-1)} \right) \cdot \left\langle \mathcal{P}(\phi^{(k-2)} \overline{\nabla \phi^{(k-2)}}) \right\rangle_{H \times H \Rightarrow L}^{\textit{close cone}}, \end{aligned}$$

and similarly for the second and third terms. Next, we note that a similar procedure as the one used in subsection 3.1 can be used to show that

$$\begin{aligned} & \frac{1}{\Xi} S_{.,d>1} \mathcal{P} \left( \psi \nabla \phi^{(k-1)} \right) - \left\langle \frac{1}{\Xi} \mathcal{P} \left( \psi \nabla \phi^{(k-1)} \right) \right\rangle_{H \times H \Rightarrow L}^{\textit{close cone}} \\ & \quad - \left\langle \frac{1}{\Xi} \mathcal{P} \left( \psi \nabla \phi^{(k-1)} \right) \right\rangle_{H \times H \Rightarrow L}^{\textit{medium cone}} \hookrightarrow Y^{-s, \frac{1}{2}}, \end{aligned}$$

where the loss of the extra  $2s - 1$  factor is due to the  $X_\infty$  Besov structure. This reduces the proof of (113) to showing the following embedding:

$$Y^{-s, \frac{1}{2}} \cdot \left\langle \mathcal{P}(X_1^{s, \frac{1}{2}} \nabla X_1^{s, \frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{\textit{close cone}} \subseteq \mathcal{L}^1.$$

The proof of this last embedding follows from essentially the same procedure and estimates used in the proof of (112) above. Notice that the extra  $2s - 1$  loss in the  $Y$  space can be made up for by the extra room in the product of the two  $X_1$  spaces due to the range of the indices involved. This completes the proof of (97).  $\square$

**3.4. Estimates for  $\phi^{(k)}$  that do not involve the “elliptic cancellations”, part II: Terms containing  $A_0^{(k-1)}$ .** In this section, we will work much as we did in the previous one. We first isolate the main terms we will be dealing with by writing

$$\begin{aligned} \phi_{\textit{good quadratic } A_0}^{(k)} &= A_0^{(k-1)} \textit{quadratic} \partial_t \phi^{(k-1)} \\ & - \left\langle \left\langle A_0^{(k-1)} \textit{good quadratic} \right\rangle_{H \times H \Rightarrow L}^{\textit{medium cone}} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\textit{close cone}} \\ & - \left\langle A_0^{(k-1)} \textit{bad quadratic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\textit{close cone}} \\ & \quad - \left\langle A_0^{(k-1)} \textit{bad quadratic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\textit{medium cone}}, \end{aligned}$$

where

$$A_0^{(k-1)} \textit{quadratic} = \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right),$$

and the decomposition

$$A_0^{(k-1)} \textit{quadratic} = A_0^{(k-1)} \textit{good quadratic} + A_0^{(k-1)} \textit{bad quadratic}$$

is the same as in subsection 3.2.

Using duality and the structure of the proofs in subsection 3.2, we will show the following embeddings:

$$(114) \quad \left( A_0^{(k-1)} \right)^2 \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

$$(115) \quad A_0^{(k-1)} \text{cubic} \partial_t \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

$$(116) \quad \phi_{\text{good quadratic } A_0}^{(k)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

$$(117) \quad \mathcal{B}_0^{(k-1)} \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

*Proof of (114).* Using (55), we may write

$$A_0^{(k-1)} = A_0^{(k-1)} \text{bad quadratic} + A_0^{(k-1)} \text{H}^{s-\frac{1}{2}, 1-s, s} ,$$

where

$$A_0^{(k-1)} \text{H}^{s-\frac{1}{2}, 1-s, s} \hookrightarrow \text{H}^{s-\frac{1}{2}, 1-s, s} .$$

Expanding the square on the left-hand side of (114) and using duality, our first order of business is to show that

$$(118) \quad \left( \text{H}^{s-\frac{1}{2}, 1-s, s} \right)^2 \cdot X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \subseteq \mathcal{L}^1 ,$$

$$(119) \quad (L^2 \cap \bar{L}^\infty)(\text{H}_{t,x}^s) \cdot \text{H}^{s-\frac{1}{2}, 1-s, s} \cdot X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \subseteq \mathcal{L}^1 .$$

The estimate (118) follows immediately from the two embeddings

$$\begin{aligned} \text{H}^{s-\frac{1}{2}, 1-s, s} &\subseteq L^2(\bar{L}^{6+}) && \text{(Sobolev)}, \\ X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} &\subseteq \bar{L}^\infty(\bar{L}^{\frac{3}{2}-}) && \text{(Energy \& Sobolev)}, \end{aligned}$$

while the estimate (119) follows from the estimates

$$\begin{aligned} (L^2 \cap \bar{L}^\infty)(\text{H}_{t,x}^s) &\subseteq \bar{L}^{\infty-}(\bar{L}^{3+}) && \text{(Interpolation \& Sobolev)}, \\ \text{H}^{s-\frac{1}{2}, 1-s, s} &\subseteq L^2(\bar{L}^6) && \text{(Sobolev)}, \\ X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} &\subseteq \bar{L}^{2+}(\bar{L}^{2-}) && \text{(Strichartz)}. \end{aligned}$$

It now remains to prove the embedding

$$\left( A_0^{(k-1)} \text{bad quadratic} \right)^2 \phi^{(k-1)} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

Since  $A_0^{(k-1)} \text{bad quadratic}$  is roughly of the same form as  $A_i^{(k-1)} \text{bad quadratic}$  (but smoother!), the above inclusion follows from the same reasoning used to show (98). This completes the proof of (114).  $\square$

*Proof of (115).* The first thing we will show is the embedding

$$A_0^{(k-1)} \text{cubic} \partial_t \phi^{(k-1)} - \left\langle A_0^{(k-1)} \text{cubic} \partial_t \phi^{(k-1)} \right\rangle_{\text{H} \times \text{L} \Rightarrow \text{H}}^{\text{close cone}} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

By duality, this follows from the more general estimate

$$(120) \quad H^{s-\frac{1}{2}, 1-s, s} \cdot \partial_t X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} - H^{s-\frac{1}{2}, 1-s, s} \cdot \left\langle \partial_t X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \mathcal{L}^1.$$

To prove (120), it suffices to show that

$$\partial_t X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} - \left\langle \partial_t X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq H^{\frac{1}{2}-s, 0, -1}.$$

The proof of this last estimate is virtually identical to the proof of the embedding (81), except that one must be careful when evaluating the  $X_\infty$  Besov structure: If the (space-time) frequency of  $X_\infty^{1-s, \frac{1}{2}}$  is much larger than that of  $X_1^{s, \frac{1}{2}}$  in the above product, the extra  $\langle D_{t,x} \rangle^{\frac{1}{2}-s}$  weight takes care of the summation away from the cone, and this is why we had to carry this extra weight around. If we are in a region where the above product has small frequency, the extra weights will not allow us to add away from the cone in the  $X_\infty$  space. However, in this case, the  $X_\infty \cdot X_1$  product breaks down into a diagonal sum, so the addition in the cone variable is absorbed by the  $L^\infty - L^1$  duality.

It now remains to show that

$$(121) \quad \left\langle A_0^{(k-1)} \text{cubic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \hookrightarrow X_1^{s-1, -\frac{1}{2}}.$$

Just like in the previous section, this requires us to write out the expansion

$$(122) \quad A_0^{(k-1)} \text{cubic} = \frac{1}{\Delta} \left[ A_0^{(k-2)} \text{H}^{s-\frac{1}{2}, 1-s, s} |\phi^{(k-2)}|^2 + A_0^{(k-2)} \text{bad quadratic} |\phi^{(k-2)}|^2 \right].$$

Expanding the first term in (122) into the left-hand side of (121), and proceeding by duality, we are trying to show that

$$H^{s-\frac{1}{2}, 1-s, s} \cdot (X_1^{s, \frac{1}{2}})^2 \cdot \langle D_{t,x} \rangle^{-2} \left\langle X_1^{0, \frac{1}{2}} \cdot X_\infty^{0, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \mathcal{L}^1.$$

Due to the inclusion

$$\langle D_{t,x} \rangle^{(0)-} \left\langle X_1^{0, \frac{1}{2}} \cdot X_\infty^{0, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \subseteq \bar{L}^\infty(\bar{L}^1),$$

it is enough to show that

$$\langle D_x \rangle^{(-2)+} \left( (X_1^{s, \frac{1}{2}})^2 \cdot H^{s-\frac{1}{2}, 1-s, s} \right) \subseteq \bar{L}^1(\bar{L}^\infty),$$

which is an immediate consequence of Sobolev's inequality and the estimates

$$H^{s-\frac{1}{2}, 1-s, s} \subseteq L^2(\bar{L}^6) \quad (\text{Sobolev}),$$

$$(X_1^{s, \frac{1}{2}})^2 \subseteq L^2(\bar{L}^{2+}) \quad (\text{Strichartz}).$$

Expanding the second term in (122) into the left-hand side of (121), we see that we are trying to show that

$$\left\langle \frac{1}{\Delta} (X_1^{s, \frac{1}{2}} \cdot X_1^{s, \frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot \left\langle \frac{1}{\Delta} (X_\infty^{1-s, \frac{1}{2}} \cdot X_1^{s, \frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot (X_1^{s, \frac{1}{2}})^2 \subseteq \mathcal{L}^1.$$

Again, this is of roughly the same form as the estimate (98). This completes the proof of (115).  $\square$

*Proof of (116).* We begin here by noting that the inclusion

$$A_0^{(k-1)} \text{ good quadratic} \hookrightarrow H^{s-\frac{1}{2}, 1-s, s} ,$$

together with the embedding (120), shows that one has the following inclusion:

$$A_0^{(k-1)} \text{ good quadratic} \partial_t \phi^{(k-1)} - \left\langle A_0^{(k-1)} \text{ good quadratic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{close cone}} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

Therefore, we will first show that

$$(123) \quad \left\langle A_0^{(k-1)} \text{ good quadratic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{close cone}} - \left\langle \left\langle A_0^{(k-1)} \text{ good quadratic} \right\rangle_{H \times H \Rightarrow L}^{\text{medium cone}} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{close cone}} \hookrightarrow X_1^{s-1, -\frac{1}{2}} .$$

Using the fact that

$$A_0^{(k-1)} \text{ good quadratic} - \left\langle A_0^{(k-1)} \text{ good quadratic} \right\rangle_{H \times H \Rightarrow L}^{\text{medium cone}} \hookrightarrow Y^{s-\frac{1}{2}, 0} ,$$

the proof of (123) follows by duality and the embedding

$$Y^{s-\frac{1}{2}, 0} \cdot \left\langle X_\infty^{1-s, \frac{1}{2}} \partial_t X_1^{s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}} \subseteq \mathcal{L}^1 .$$

This last estimate follows from exactly the same procedure and estimates as in the proof of (112).

It remains to prove the inclusion

$$A_0^{(k-1)} \text{ bad quadratic} \partial_t \phi^{(k-1)} - \left\langle A_0^{(k-1)} \text{ bad quadratic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{close cone}} - \left\langle A_0^{(k-1)} \text{ bad quadratic} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{\text{medium cone}} \hookrightarrow X_1^{s-1, -\frac{1}{2}} ,$$

which, by duality, is equivalent to

$$(124) \quad A_0^{(k-1)} \text{ bad quadratic} \cdot \psi \partial_t \phi^{(k-1)} - A_0^{(k-1)} \text{ bad quadratic} \cdot \left\langle \psi \partial_t \phi^{(k-1)} \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}} - A_0^{(k-1)} \text{ bad quadratic} \cdot \left\langle \psi \partial_t \phi^{(k-1)} \right\rangle_{H \times H \Rightarrow L}^{\text{medium cone}} \hookrightarrow \mathcal{L}^1 ,$$

where  $\psi \in X_\infty^{1-s, \frac{1}{2}}$ .

The proof of (124) follows essentially the same pattern as the proof of (123). We begin by expanding out  $A_0^{(k-1)} \text{ bad quadratic}$  in terms of the previous iterates and noting that we can prove

$$\psi \partial_t \phi^{(k-1)} - \left\langle \psi \partial_t \phi^{(k-1)} \right\rangle_{H \times H \Rightarrow L}^{\text{close cone}} - \left\langle \psi \partial_t \phi^{(k-1)} \right\rangle_{H \times H \Rightarrow L}^{\text{medium cone}} \subseteq Y^{-s-\frac{3}{2}, 0} ,$$

using essentially the same procedure as in subsection 3.2. This reduces the proof of (124) to showing the following embedding:

$$\left\langle \frac{1}{\Delta} (X_1^{s, \frac{1}{2}} \partial_t X_1^{s, \frac{1}{2}}) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot Y^{-s-\frac{3}{2}, 0} \subseteq \mathcal{L}^1 .$$

This last estimate follows from the same procedure and estimates as those used in the proof of (112) above. □

*Proof of (117).* Using the fact that we may write

$$\mathcal{B}_0^{(k-1)} = \mathcal{B}_0^{(k-1)}_{H^{s-\frac{1}{2}, 0, 0}} + \mathcal{B}_0^{(k-1)}_{\langle D_{t,x} \rangle^{-s-\frac{1}{2}} (\bar{L}^\infty \cap L^2) (\bar{L}^1)} ,$$

where each term in the above sum satisfies the appropriate estimate, to prove (117) by duality, it is enough to show the following two embeddings:

$$\begin{aligned} H^{s-\frac{1}{2}, 0, 0} \cdot X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} &\subseteq \mathcal{L}^1 , \\ \langle D_{t,x} \rangle^{-s-\frac{1}{2}} (\bar{L}^\infty \cap L^2) (\bar{L}^1) \cdot X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}} &\subseteq \mathcal{L}^1 . \end{aligned}$$

These in turn follow from the weight trading properties of the  $\mathcal{L}^1$  spaces and the inclusions

$$\begin{aligned} \langle D_{t,x} \rangle^{s-\frac{1}{2}} H^{s-\frac{1}{2}, 0, 0} &\subseteq L^2(L^2) , \\ \langle D_{t,x} \rangle^{\frac{1}{2}-s} (X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}}) &\subseteq L^2(L^2) \quad (\text{Strichartz interp. w/ } L^2 \text{ \& Sobolev}), \end{aligned}$$

and

$$\begin{aligned} (\bar{L}^\infty \cap L^2) (\bar{L}^1) &\subseteq \bar{L}^{\infty-} (\bar{L}^1) && (\text{Interpolation}), \\ \langle D_{t,x} \rangle^{-s-\frac{1}{2}} (X_1^{s, \frac{1}{2}} \cdot X_\infty^{1-s, \frac{1}{2}}) &\subseteq \bar{L}^{1+} (\bar{L}^\infty) && (\text{Bi-Strichartz \& Sobolev}). \end{aligned}$$

respectively. □

**3.5. Estimates for the quadratic part of  $\phi^{(k)}$  involving the “elliptic” cancellation.** We now focus our attention on estimating those portions of the term

$$A_\alpha^{(k-1)} \text{quadratic} \partial^\alpha \phi^{(k-1)}$$

which have not been previously shown to be in the space  $X_1^{s-1, -\frac{1}{2}}$ . Expanding out the formula for  $A_0^{(k-1)}$  *quadratic* in terms of the previous iterates, these are

$$\begin{aligned} & \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ & + \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{medium\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ & + \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{medium\ cone} \\ & = \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H} \\ & + \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} \\ & - \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone}, \end{aligned}$$

with a similar formula for the terms involving the  $A_i^{(k-1)}$  *quadratic*. Because we only care about the size of the Fourier transform, by the triangle inequality, it is enough to find estimates for the terms

$$(125) \quad \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H} - \left\langle \left\langle \mathcal{P} \frac{1}{\Xi} \left( \phi^{(k-2)} \overline{\nabla \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot \nabla \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}$$

and

$$(126) \quad \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} - \left\langle \left\langle \mathcal{P} \frac{1}{\Xi} \left( \phi^{(k-2)} \overline{\nabla \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L} \cdot \nabla \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone}$$

and

$$(127) \quad \left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\partial_t \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \partial_t \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone} - \left\langle \left\langle \mathcal{P} \frac{1}{\Xi} \left( \phi^{(k-2)} \overline{\nabla \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot \nabla \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H}^{close\ cone}$$

separately.

Notice that by duality and due to the range of the frequencies involved, the terms (125) and (126) are essentially dual to each other. Also, note that the range of frequencies in the term (127) is essentially a subset of the range of frequencies occurring in the first two expressions (125) and (126). Therefore, for the remainder of this section, we will focus our attention on the term (125).

Before we continue, let us mention again why one cannot treat each piece of (125) separately. For example, if one were to try and prove the embedding

$$\left\langle \left\langle \mathcal{P} \frac{1}{\Xi} \left( \phi^{(k-2)} \overline{\nabla \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot \nabla \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H} \hookrightarrow X_1^{s-1, -\frac{1}{2}}$$

by duality and a quick computation of the weights involved, we see that one would be trying to show that

$$(128) \quad \langle D_{t,x} \rangle^{-2+} (X_1^{s-\frac{1}{2}, \frac{1}{2}} \cdot X_1^{s-\frac{1}{2}, \frac{1}{2}}) \cdot (X_1^{0, \frac{1}{2}} \cdot X_1^{0, \frac{1}{2}}) \subseteq \mathcal{L}^1 .$$

If we are in the range  $s > \frac{3}{4}$ , then (128) follows from the embeddings

$$\begin{aligned} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} (X_1^{s-\frac{1}{2}, \frac{1}{2}} \cdot X_1^{s-\frac{1}{2}, \frac{1}{2}}) &\subseteq L^2(L^2) , \quad 3/4 < s && \text{(Bi-Strichartz),} \\ \langle D_{t,x} \rangle^{(-\frac{3}{2})-} (X_1^{0, \frac{1}{2}} \cdot X_1^{0, \frac{1}{2}}) &\subseteq L^2(L^2) && \text{(Energy} \cdot L^2 \text{ \& Sobolev).} \end{aligned}$$

If we are in the range  $s < \frac{3}{4}$ , then (128) is in fact false even if all the test functions in the product are supported in an  $O(1)$  neighborhood of the light cone in Fourier space. This is because (128) is false at the level of solutions to the homogeneous wave equation. In fact, this is exactly what our counterexample in subsection 1.3 showed.

We now resume the matter at hand. In order to take advantage of the cancellations between the two terms in (125), we add and subtract from it the following:

$$\left\langle \left\langle \frac{1}{\Delta} \mathfrak{S} \left( \phi^{(k-2)} \overline{\nabla \phi^{(k-2)}} \right) \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \nabla \phi^{(k-1)} \right\rangle_{H \times L \Rightarrow H} .$$

By duality then, we see that in order to show that (125) is in the space  $X_1^{s-1, -\frac{1}{2}}$ , it is enough to show the following two inclusions:

$$(129) \quad \frac{1}{\Delta} \left( I + \frac{(curl)^2}{\Xi} \right) \left\langle X_1^{s, \frac{1}{2}} \cdot \nabla X_1^{s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot \left\langle X_1^{s, \frac{1}{2}} \cdot \nabla X_\infty^{1-s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L} \subseteq \mathcal{L}^1 ,$$

$$(130) \quad \left\langle X_1^{s, \frac{1}{2}} \cdot \partial_\alpha X_1^{s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L}^{close\ cone} \cdot \left\langle X_1^{s, \frac{1}{2}} \cdot \partial^\alpha X_\infty^{1-s, \frac{1}{2}} \right\rangle_{H \times H \Rightarrow L} \subseteq \mathcal{L}^1 .$$

These are the main null form estimates which are at the heart of this work.

*Proof of (129).* Due to the range of the various dyadic frequencies involved, it suffices to prove (129) at fixed dyadic values at the scaling, and at a loss of  $\lambda^{0+}$ , i.e., in order to show (129), it is enough to prove the following:

$$(131) \quad \mathcal{M} S_{\lambda,d} \left( S_{\mu_1, \delta_1} (X_1^{\frac{1}{2}, \frac{1}{2}}) \cdot S_{\mu_2, \delta_2} (X_1^{\frac{1}{2}, \frac{1}{2}}) \right) \cdot \left( S_{\sigma_1, \eta_1} (X_1^{\frac{1}{2}, \frac{1}{2}}) \cdot S_{\sigma_2, \eta_2} (X_1^{\frac{1}{2}, \frac{1}{2}}) \right) \subseteq \lambda^{0+} \mathcal{L}^1 ,$$

where we are in the range

$$(1) \quad \lambda \ll \mu_1 \sim \mu_2;$$

- (2)  $\lambda \ll \sigma_1 \sim \sigma_2$ ;
- (3)  $\max\{\delta_1, \delta_2\} \ll d$ ;
- (4)  $\max\{\eta_1, \eta_2\} \lesssim \lambda$ .

Here,  $\mathcal{M}$  denotes the multilinear operator

$$\mathcal{M}(\bullet \bullet \bullet) \cdot (\bullet \bullet \bullet) = \frac{1}{\Delta} \left( I + \frac{(\text{curl})^2}{\Xi} \right) (\bullet \nabla \bullet) \cdot (\bullet \nabla \bullet) .$$

For the rest of the proof, we will assume without loss of generality that we are also in the range where  $\mu_1 = \mu_2 = \mu$ ,  $\sigma_1 = \sigma_2 = \sigma$ , and  $\delta_2 \leq \delta_1 \leq \eta_1$ .

We now use Lemma 2.9, to decompose the first term in the product on the left-hand side of (131). Due to the (essential) diagonality in this sum, we can reduce the proof of (131), to the following weighted  $L^2$  multilinear estimate:

$$(132) \quad \mathcal{M} \left( S_{\mu; \lambda, d, \delta_1}^\alpha(L^2) \cdot S_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2) \right) \cdot S_{\lambda, d}^\alpha(S_{\sigma, \eta_1}(L^2) \cdot S_{\sigma, \eta_2}(L^2)) \\ \subseteq \lambda^{0+} \mu \sigma (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1 ,$$

where  $\text{supp}\{s_{\mu; \lambda, d, \delta_1}^\alpha * s_{\mu; \lambda, d, \delta_2}^{-\alpha}\} \subseteq \text{supp}\{s_{\lambda, d}^\alpha\}$ , and where  $\text{supp}\{s_{\mu; \lambda, d, \delta_1}^\alpha\}$  and  $\text{supp}\{s_{\mu; \lambda, d, \delta_2}^{-\alpha}\}$  are separated by at most an  $O(\sqrt{\frac{d}{\lambda}})$  angle.

Due to the invariance of  $\mathcal{M}$ , we can, after a spatial rotation if necessary, assume that we are trying to prove (132) where the spatial projections of  $\text{supp}\{s_{\mu; \lambda, d, \delta_1}^\alpha\}$  and  $\text{supp}\{s_{\mu; \lambda, d, \delta_2}^{-\alpha}\}$  are oriented along the  $\xi_1$  axis in Fourier space, while their set sum (i.e., the set  $\text{supp}_\xi\{s_{\lambda, d}^\alpha\}$ ) lies along the positive  $\xi_1$  axis. Of course, this is well defined up to an  $O(\sqrt{\frac{d}{\lambda}})$  angle, due to the restriction mentioned above, which is a consequence of Lemma 2.9.

With these reductions, we now proceed to give a precise bound on the symbol of  $\mathcal{M}$ . A direct calculation shows that

$$\mathcal{M}(\bullet \bullet \bullet) \cdot (\bullet \bullet \bullet) = \frac{1}{\Delta \cdot \Xi} \begin{pmatrix} -\partial_0^2 + \partial_1^2 & \partial_1 \partial_2 & \partial_1 \partial_3 \\ \partial_2 \partial_1 & -\partial_0^2 + \partial_2^2 & \partial_2 \partial_3 \\ \partial_3 \partial_1 & \partial_3 \partial_2 & -\partial_0^2 + \partial_3^2 \end{pmatrix} (\bullet \nabla \bullet) \cdot (\bullet \nabla \bullet) .$$

Thus, under the angular restrictions mentioned above, we see that in Fourier space we may replace the symbol of  $\mathcal{M}$  appearing in (132) by the following weighted multilinear expression:

$$\frac{1}{\lambda^2 \cdot \lambda d} \begin{pmatrix} (-|\widetilde{\partial_t}| + |\partial_1|)\lambda & \lambda\sqrt{\lambda d} & \lambda\sqrt{\lambda d} \\ \lambda\sqrt{\lambda d} & \lambda^2 & \lambda\sqrt{\lambda d} \\ \lambda\sqrt{\lambda d} & \lambda\sqrt{\lambda d} & \lambda^2 \end{pmatrix} \begin{pmatrix} (\bullet) * (\mu \bullet) \\ (\bullet) * (\sqrt{\lambda d} \bullet) \\ (\bullet) * (\sqrt{\lambda d} \bullet) \end{pmatrix} \cdot \begin{pmatrix} (\bullet) * (\sigma \bullet) \\ (\bullet) * (\widetilde{\partial_2} \bullet) \\ (\bullet) * (\widetilde{\partial_3} \bullet) \end{pmatrix} .$$

We now proceed to prove (132), using the above bound on the symbol of  $\mathcal{M}$ , treating each term in the matrix separately.

**The  $\mathcal{M}_{11}$  term.** In this case, to prove (132), we can after a Plancherel, assume that we are working in Fourier space. Using the weight bounds for  $\mathcal{M}_{11}$ , we see that it is enough to show that

$$(133) \quad (-|\widetilde{\partial_0}| + |\partial_1|) \left( s_{\mu; \lambda, d, \delta_1}^\alpha(L^2) * s_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2) \right) \cdot (s_{\sigma, \eta_1}(L^2) * s_{\sigma, \eta_2}(L^2)) \\ \subseteq \lambda^{2+} d (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1 .$$

To prove this, we will make a careful calculation of the symbol of  $|\partial_0| - |\partial_1|$  applied to the convolution  $s_{\mu;\lambda,d,\delta_1}^\alpha * s_{\mu;\lambda,d,\delta_2}^{-\alpha}$ . For  $\xi \in \text{supp}_\xi(s_{\mu;\lambda,d,\delta_1}^j)$  and  $\xi' \in \text{supp}_\xi(s_{\mu;\lambda,d,\delta_2}^{-j})$  we compute

$$\begin{aligned} |\partial_t| - |\partial_1| &= \widetilde{|\xi| - |\xi'| + O(\delta_1) + O(\delta_2)} - |\xi_1 + \xi'_1| \\ &\sim \widetilde{|\xi| - |\xi'|} - |\xi_1 + \xi'_1| + O(\max\{\delta_1, \delta_2\}). \end{aligned}$$

We will now assume that  $|\xi'| \leq |\xi|$ , and using the notation  $\hat{\xi}' = \frac{\xi'}{|\xi'|}$  as well as the observation that  $0 < (\xi + \xi')_1$  we write

$$\begin{aligned} &|\xi| - |\xi'| - |\xi_1 + \xi'_1| \\ &= (|\xi| - |\xi'| + (\xi + \xi') \cdot \hat{\xi}') - ((\xi + \xi')_1 + (\xi + \xi') \cdot \hat{\xi}') \\ &\lesssim \mu \cdot \theta_{\xi, -\xi'}^2 + \lambda \cdot \theta_{(1,0,0), -\xi'}^2 \\ &\lesssim \frac{\lambda d}{\mu}. \end{aligned}$$

Thus, after a Plancherel, we see that in order to prove (133) it is enough to show the following estimates:

$$(134) \quad \frac{\max\{\delta_1, \delta_2\}}{\lambda^2 d} \left( S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) \cdot S_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \right) \cdot (S_{\sigma,\eta_1}(L^2) * S_{\sigma,\eta_2}(L^2)) \subseteq (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1,$$

and

$$(135) \quad \frac{1}{\lambda^{1+\mu}} \left( S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) \cdot S_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \right) \cdot (S_{\sigma,\eta_1}(L^2) \cdot S_{\sigma,\eta_2}(L^2)) \subseteq (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1.$$

To prove the embedding (134), we recall that we are using the assumption  $\max\{\delta_1, \delta_2\} = \delta_1 \leq \eta_1$ . Thus, we may reduce to proving this estimate with  $\max\{\delta_1, \delta_2\}$  replaced by  $\eta_1$ . It then follows from multiplying the following embeddings:

$$\begin{aligned} S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) &\subseteq \lambda d^{\frac{1}{2}} L^2(\bar{L}^\infty) && \text{(Local Sobolev),} \\ S_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) &\subseteq \lambda d^{\frac{1}{2}} \delta_2^{\frac{1}{2}} \bar{L}^\infty(\bar{L}^\infty) && \text{(Local Sobolev),} \\ S_{\sigma,\eta_1}(L^2) \cdot S_{\sigma,\eta_2}(L^2) &\subseteq \eta_2^{\frac{1}{2}} L^2(\bar{L}^1) && \text{(Energy} \cdot L^2\text{).} \end{aligned}$$

To prove (135), noting the range of the various multipliers and the value of the dyadic weights, we see that the estimate follows from multiplying the following local embeddings:

$$\begin{aligned} S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) \cdot S_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) &\subseteq \lambda^{1+\mu} \mu \delta_1^{\frac{1}{2}} \delta_2^{\frac{1}{2}} \bar{L}^{1+}(\bar{L}^\infty) && \text{(Bi-Strichartz \& Sobolev),} \\ S_{\sigma,\eta_1}(L^2) \cdot S_{\sigma,\eta_2}(L^2) &\subseteq \eta_1^{\frac{1}{2}} \eta_2^{\frac{1}{2}} \bar{L}^{\infty-}(\bar{L}^1) && \text{(Energy interp. w/ } L^2 \cdot \text{Energy).} \end{aligned}$$

**The  $\mathcal{M}_{12}$  and  $\mathcal{M}_{13}$  terms.** Counting up the weights involved, we see that this boils down to an estimate which is exactly of the same form as (135) above.

**The  $\mathcal{M}_{21}$  and  $\mathcal{M}_{31}$  terms.** We will only deal with the  $\mathcal{M}_{21}$  term here, as the  $\mathcal{M}_{31}$  can be treated in exactly the same manner. What we are trying to show, then, is that

$$(136) \quad \left( s_{\mu;\lambda,d,\delta_1}^\alpha(L^2) * s_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \right) \cdot \left( s_{\sigma,\eta_1}(L^2) * |\tilde{\partial}_2| s_{\sigma,\eta_2}(L^2) \right) \subseteq \lambda^{\frac{3}{2}+} d^{\frac{1}{2}} \sigma(\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1 .$$

In order to show this, we will break down into two cases depending on the relative size of  $\max\{\eta_1, \eta_2\}$  and  $d$ . However, both of these cases follow essentially the same pattern, and the split is more for notational purposes as opposed to being really essential.

**Case 1:**  $\max\{\eta_1, \eta_2\} \lesssim d$ . In this case, we can run a similar decomposition as the one given in Lemma 2.9 on the product  $S_{\sigma,\eta_1}(L^2) \cdot S_{\sigma,\eta_2}(L^2)$  to obtain

$$S_{\sigma,\eta_1}(L^2) \cdot S_{\sigma,\eta_2}(L^2) \approx \sum_{\beta} S_{\sigma;\lambda,d,\eta_1}^\beta(L^2) \cdot S_{\sigma;\lambda,d,\eta_2}^{-\beta}(L^2) ,$$

where the spatial supports of the multipliers  $s_{\sigma;\lambda,d,\eta_i}^{\pm\beta}$  are still parallelepipeds of size  $\sim \lambda \times \sqrt{\lambda d} \times \sqrt{\lambda d}$ , but may be larger than the spatial supports of the  $s_{\mu;\lambda,d,\delta_i}^{\pm\alpha}$ . This is because the condition  $\max\{\eta_1, \eta_2\} \lesssim d$  still implies that the angle between the spatial support of  $s_{\sigma,\eta_1}$  and  $s_{\sigma,\eta_2}$  in the convolution cannot be too much larger than  $\frac{\sqrt{\lambda d}}{\sigma}$ .

We are now trying to show that

$$(137) \quad \left( s_{\mu;\lambda,d,\delta_1}^\alpha(L^2) * s_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \right) \cdot \left( s_{\sigma;\lambda,d,\eta_1}^\beta(L^2) * |\tilde{\partial}_2| s_{\sigma;\lambda,d,\eta_2}^{-\beta}(L^2) \right) \subseteq \lambda^{\frac{3}{2}+} d^{\frac{1}{2}} \sigma(\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1 .$$

To proceed further, we need to consider the **minimum** angle between the set  $supp_{\xi}\{s_{\sigma;\lambda,d,\eta_1}^\beta\} \cup supp_{\xi}\{s_{\sigma;\lambda,d,\eta_2}^{-\beta}\}$  and the set  $supp_{\xi}\{s_{\mu;\lambda,d,\delta_1}^\alpha\} \cup supp_{\xi}\{s_{\mu;\lambda,d,\delta_2}^{-\alpha}\}$ . We will call this angle  $\Theta_{min}$ . Then, in terms of this angle, we can bound the  $\tilde{\partial}_2$  appearing in the above estimate by

$$|\tilde{\partial}_2| \lesssim \sigma \left( \Theta_{min} + \frac{\sqrt{\lambda d}}{\sigma} + \frac{\sqrt{\lambda d}}{\mu} \right) .$$

Upon substituting  $\tilde{\partial}_2$  with the sum  $\sigma \left( \frac{\sqrt{\lambda d}}{\sigma} + \frac{\sqrt{\lambda d}}{\mu} \right)$  in (137), we see that we are left with estimates of the form (135), so it is enough to try to prove (137) with  $\tilde{\partial}_2$  replaced by  $\sigma \cdot \Theta_{min}$ . We will, of course, also assume now that  $0 < \Theta_{min}$ . After regrouping the product, we see that (137) follows from the embeddings

$$\begin{aligned} S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) \cdot S_{\sigma;\lambda,d,\eta_2}^{-\alpha}(L^2) &\subseteq (\lambda\sqrt{\lambda d})^{\frac{1}{2}} \Theta_{min}^{-\frac{1}{2}} (\delta_1 \eta_2)^{\frac{1}{2}} L^2(L^2) \quad (\text{Spec. Strichartz}), \\ S_{\mu;\lambda,d,\delta_2}^\beta(L^2) \cdot S_{\sigma;\lambda,d,\eta_1}^{-\beta}(L^2) &\subseteq (\lambda\sqrt{\lambda d})^{\frac{1}{2}} \Theta_{min}^{-\frac{1}{2}} (\delta_2 \eta_1)^{\frac{1}{2}} L^2(L^2) \quad (\text{Spec. Strichartz}). \end{aligned}$$

**Case 2:**  $d \ll \max\{\eta_1, \eta_2\}$ . In order to prove (136) in this case, we will use an angular decomposition of the product  $S_{\sigma,\eta_1}(L^2) \cdot S_{\sigma,\eta_2}(L^2)$ , based on a calculation of the maximum angle between the spatial supports of the corresponding convolution

factors in Fourier space. Using the exact same calculation as that used to prove Lemma 2.9, we see that this angle satisfies the bounds

$$\Theta \lesssim \frac{(\lambda \cdot \max\{\eta_1, \eta_2\})^{\frac{1}{2}}}{\sigma}.$$

Thus, just as we did for Lemma 2.9, we may write

$$S_{\sigma, \eta_1}(L^2) \cdot S_{\sigma, \eta_2}(L^2) \approx \sum_{\beta} S_{\sigma; \lambda, \max\{\eta_1, \eta_2\}, \eta_1}^{\beta}(L^2) \cdot S_{\sigma; \lambda, \max\{\eta_1, \eta_2\}, \eta_2}^{-\beta}(L^2),$$

where the above multipliers have the appropriate dimensions.

We can now, without loss of generality, assume that we are in the case where  $\max\{\eta_1, \eta_2\} = \eta_1$ , the other case following from the same method. What we are trying to prove, then, is that

$$(138) \quad \left( s_{\mu; \lambda, d, \delta_1}^{\alpha}(L^2) * s_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2) \right) \cdot \left( s_{\sigma; \lambda, \eta_1, \eta_1}^{\beta}(L^2) * |\tilde{\partial}_2| s_{\sigma; \lambda, \eta_1, \eta_2}^{-\beta}(L^2) \right) \subseteq \lambda^{\frac{3}{2}+} d^{\frac{1}{2}} \sigma (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1.$$

We now mimic the proof of Case 1 above and denote the **minimum** angle between the sets  $supp_{\xi}\{s_{\sigma; \lambda, \eta_1, \eta_1}^{\beta}\} \cup supp_{\xi}\{s_{\sigma; \lambda, \eta_1, \eta_2}^{-\beta}\}$  and  $supp_{\xi}\{s_{\mu; \lambda, d, \delta_1}^{\alpha}\} \cup supp_{\xi}\{s_{\mu; \lambda, d, \delta_2}^{-\alpha}\}$  by  $\Theta_{min}$ . Again, using this angle, we can bound the derivative appearing in the above inclusion by

$$|\tilde{\partial}_2| \lesssim \sigma (\Theta_{min} + \frac{\sqrt{\lambda \eta_1}}{\sigma} + \frac{\sqrt{\lambda d}}{\mu}).$$

The proof of (138) now follows from the same set of estimates as used in Case 1 above, except for the  $\sqrt{\lambda \eta_1}$  term. To deal with this value substituted for  $|\tilde{\partial}_2|$  in (138), it is enough to note the following embeddings:

$$\begin{aligned} S_{\mu; \lambda, d, \delta_1}^{\alpha}(L^2) \cdot S_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2) &\subseteq \lambda^{1+} d^{\frac{1}{2}} \delta_1^{\frac{1}{2}} \delta_2^{\frac{1}{2}} \bar{L}^{\infty-}(\bar{L}^{2+}) \quad (\text{Energy \& Local Sobolev}), \\ S_{\sigma; \lambda, \eta_1, \eta_1}^{\beta}(L^2) &\subseteq L^2(L^2), \\ S_{\sigma; \lambda, \eta_1, \eta_2}^{-\beta}(L^2) &\subseteq \sigma \eta_2^{\frac{1}{2}} \bar{L}^{2+}(\bar{L}^{\infty-}) \quad (\text{Strichartz}). \end{aligned}$$

**The  $\mathcal{M}_{22}$ ,  $\mathcal{M}_{23}$ ,  $\mathcal{M}_{32}$ , and  $\mathcal{M}_{33}$  terms.** In this case, we can reduce to studying the  $\mathcal{M}_{22}$  term, as the  $\mathcal{M}_{33}$  term is similar by symmetry, and the  $\mathcal{M}_{23}$  and  $\mathcal{M}_{32}$  terms are of a lower order. What we are trying to show, then, is that

$$(139) \quad \left( s_{\mu; \lambda, d, \delta_1}^{\alpha}(L^2) * s_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2) \right) \cdot \left( s_{\sigma, \eta_1}(L^2) * |\tilde{\partial}_2| s_{\sigma, \eta_2}(L^2) \right) \subseteq \lambda^{\frac{1}{2}+} d^{\frac{1}{2}} \mu \sigma (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1.$$

But we just proved (136), which is the above estimate with  $\mu$  replaced by  $\lambda$ . Therefore, there is nothing new to show here. This completes the proof of (129).  $\square$

*Proof of (130).* We begin by noting that, as in the previous case, it suffices to prove this estimate at the scaling, and at a loss of  $\lambda^{0+}$ . After an application of Lemma 2.9, we see that this amounts to showing that

$$(140) \quad \frac{1}{\Delta} \left( S_{\mu; \lambda, d, \delta_1}^{\alpha}(L^2) \partial_{\gamma} S_{\mu; \lambda, d, \delta_2}^{-\alpha}(L^2) \right) \cdot S_{\lambda, d}^{\alpha}(S_{\sigma, \eta_1}(L^2) \partial^{\gamma} S_{\sigma, \eta_2}(L^2)) \subseteq \lambda^{0+} \mu \sigma (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1,$$

where we are in the range

- (1)  $\lambda \ll \mu_1 \sim \mu_2$ ;
- (2)  $\lambda \ll \sigma_1 \sim \sigma_2$ ;
- (3)  $\max\{\delta_1, \delta_2\} \ll d$ ;
- (4)  $\max\{\eta_1, \eta_2\} \lesssim \lambda$ .

The proof now retraces the steps of the previous one, using the same decompositions and notation. The first thing we will do is to break down into cases depending on the relative sizes of  $\max\{\eta_1, \eta_2\}$  and  $d$ .

**Case 1:**  $\max\{\eta_1, \eta_2\} \leq d$ . After an angular decomposition in the second factor of (140), we see that we are trying to prove that

$$(141) \quad \frac{1}{\Delta} \left( S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) \partial_\gamma S_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \right) \cdot \left( S_{\sigma;\lambda,d,\eta_1}^\beta(L^2) \partial^\gamma S_{\sigma;\lambda,d,\eta_2}^{-\beta}(L^2) \right) \subseteq \lambda^{0+} \mu \sigma (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1 .$$

We now compute the symbol of  $\partial_\gamma \partial^\gamma$  in terms of  $\Theta_{max}$ , the **maximum** angle between the sets  $supp_\xi \{s_{\mu;\lambda,d,\delta_2}^{-\alpha}\}$  and  $-supp_\xi \{s_{\sigma;\lambda,d,\eta_2}^{-\beta}\}$ , as follows. For  $(\tau, \xi) \in supp \{s_{\mu;\lambda,d,\delta_2}^{-\alpha}\}$  and  $(\tau', \xi') \in supp \{s_{\sigma;\lambda,d,\eta_2}^{-\beta}\}$  we can assume without loss of generality that  $0 \leq \tau$  and  $\tau' \leq 0$ . Then we compute that

$$\begin{aligned} |\tau \tau' - \xi \cdot \xi'| &= |(\tau - |\xi|)\tau' + |\xi|(\tau' + |\xi'|) - |\xi||\xi'| - \xi \cdot \xi'| , \\ &\lesssim (\mu + \sigma)d + \mu \sigma \Theta_{max}^2 . \end{aligned}$$

We note that this bound can be made on the symbol of  $\partial_\gamma \partial^\gamma$ , while at the same time replacing  $\frac{1}{\Delta}$  by its dyadic bound  $\frac{1}{\lambda^2}$ .

As in the previous section, we can bound the maximum angle  $\Theta_{max}$  by some dyadic expressions plus the **minimum** angle between the sets  $supp_\xi \{s_{\sigma;\lambda,d,\eta_1}^\beta\} \cup supp_\xi \{s_{\sigma;\lambda,d,\eta_2}^{-\beta}\}$  and  $supp_\xi \{s_{\mu;\lambda,d,\delta_1}^\alpha\} \cup supp_\xi \{s_{\mu;\lambda,d,\delta_2}^{-\alpha}\}$ , as follows:

$$\Theta_{max} \lesssim \Theta_{min} + \frac{\sqrt{\lambda d}}{\sigma} + \frac{\sqrt{\lambda d}}{\mu} .$$

The estimate (141) now follows from the same sort of embeddings used in the proof of (129) above.

**Case 2:**  $d \leq \max\{\eta_1, \eta_2\}$ . We can assume without loss of generality that in fact  $\max\{\eta_1, \eta_2\} = \eta_1$ . In this case, we make the same angular decomposition we made in the second case of the previous proof to reduce to showing that

$$(142) \quad \frac{1}{\Delta} \left( S_{\mu;\lambda,d,\delta_1}^\alpha(L^2) \partial_\gamma S_{\mu;\lambda,d,\delta_2}^{-\alpha}(L^2) \right) \cdot \left( S_{\sigma;\lambda,\eta_1,\eta_1}^\beta(L^2) \partial^\gamma S_{\sigma;\lambda,\eta_1,\eta_2}^{-\beta}(L^2) \right) \subseteq \lambda^{0+} \mu \sigma (\delta_1 \delta_2 \cdot \eta_1 \eta_2)^{\frac{1}{2}} \mathcal{L}^1 .$$

Also, if we denote the **minimum** angle between the sets  $supp_\xi \{s_{\sigma;\lambda,d,\eta_1}^\beta\} \cup supp_\xi \{s_{\sigma;\lambda,d,\eta_2}^{-\beta}\}$  and  $supp_\xi \{s_{\mu;\lambda,d,\delta_1}^\alpha\} \cup supp_\xi \{s_{\mu;\lambda,d,\delta_2}^{-\alpha}\}$  by  $\Theta_{min}$ , we get the symbol bound

$$|\xi_\gamma \xi'^\gamma| \lesssim \mu d + \sigma \eta_1 + \sigma \mu (\Theta_{min} + \frac{\sqrt{\lambda \eta_1}}{\sigma} + \frac{\sqrt{\lambda d}}{\mu}) .$$

Substituting this last expression into the estimate (142), while replacing  $\frac{1}{\Delta}$  by  $\frac{1}{\lambda^2}$ , we are left with proving embeddings which are similar to those shown in the proof of (129); we leave the details to the reader.  $\square$

**3.6. Proof of the inclusion (56b).** This is the final section of the paper. Keeping in mind the bootstrapping estimate (55b) for  $l = 0$  (which we now know to be true), after a foliation, we are reduced to proving the estimate

$$\langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda, d} \mathcal{N}_4(\{A^{(k-1)}\}, \phi^{(k-1)}) \hookrightarrow L^2 .$$

We shall treat each term in  $\mathcal{N}_4$  in the above embedding separately as they require slightly different arguments. For the remainder of this section we will refer to spatial field potentials  $\{A_i\}$  collectively as  $A$ .

*Proof of (143) for the cubic terms in  $\mathcal{N}_4$ .* In this case, it suffices to drop multipliers and show the inclusions

$$(143) \quad \langle D_x \rangle^{(s-\frac{3}{2})+} \left( |A^{(k-1)}|^2 \phi^{(k-1)} \right) \hookrightarrow L^2 ,$$

$$(144) \quad \langle D_x \rangle^{s-1} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} \left( (A_0^{(k-1)})^2 \phi^{(k-1)} \right) \hookrightarrow L^2 .$$

We first prove (143) by decomposing it into two estimates based on a Littlewood–Paley decomposition with respect to the spatial variable as follows:

$$(145) \quad \sum_{\mu_1 \ll \mu_2} \langle D_x \rangle^{(s-\frac{3}{2})+} \left( P_{\mu_1}(|A^{(k-1)}|^2) \cdot P_{\mu_2}(\phi^{(k-1)}) \right) \hookrightarrow L^2(L^2) ,$$

$$(146) \quad \sum_{\mu_2 \lesssim \mu_1} \langle D_x \rangle^{(s-\frac{3}{2})+} \left( P_{\mu_1}(|A^{(k-1)}|^2) \cdot P_{\mu_2}(\phi^{(k-1)}) \right) \hookrightarrow L^2(L^2) .$$

After some weight trading, the estimate (145) follows from the embeddings

$$(147) \quad \begin{aligned} \langle D_x \rangle^{(s-1)-} (X_1^{s, \frac{1}{2}}) &\subseteq \bar{L}^{2+}(\bar{L}^\infty) && \text{(Strichartz \& Sobolev),} \\ (\bar{L}^\infty \cap L^2)(H_{t,x}^s) &\subseteq \bar{L}^{\infty-}(H_{t,x}^s) && \text{(Interpolation),} \\ \langle D_x \rangle^{(-\frac{1}{2})+} \left( (\bar{L}^{\infty-}(H_{t,x}^s))^2 \cdot \bar{L}^{2+}(\bar{L}^\infty) \right) &\subseteq L^2(L^2) && \text{(Sobolev).} \end{aligned}$$

To show (146), after some weight trading and an application of the estimate (147), we can reduce things to the following estimate, which can be proved with several rounds of the Sobolev embedding theorem:

$$\langle D_x \rangle^{(s-\frac{3}{2})+} \left( \bar{L}^{\infty-}(H_{t,x}^s) \cdot \bar{L}^{\infty-}(H_{t,x}^{(2s-1)-}) \cdot \bar{L}^{2+}(\bar{L}^\infty) \right) \subseteq L^2(L^2) .$$

To prove (144), we rely on the decomposition

$$(148) \quad A_0^{(k-1)} = A_0^{(k-1)} \bar{L}^\infty(H_{t,x}^s) + A_0^{(k-1)}_{H^{s-\frac{1}{2}, 1-s, s}} ,$$

to reduce things to proving estimates identical to the ones used to show (143) and the following:

$$\langle D_x \rangle^{s-1} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} \left( \left( H^{s-\frac{1}{2}, 1-s, s} \right)^2 \cdot X_1^{s, \frac{1}{2}} \right) \subseteq L^2(L^2) ,$$

which in turn follows from the embeddings

$$\begin{aligned} \left(H^{s-\frac{1}{2}, 1-s, s}\right)^2 &\subseteq \bar{L}^{1+}(\bar{L}^3) && \text{(Sobolev),} \\ \langle D_x \rangle^{s-1} \left(\bar{L}^{1+}(\bar{L}^3) \cdot X_1^{s, \frac{1}{2}}\right) &\subseteq \bar{L}^{1+}(L^2) && \text{(Energy \& Sobolev),} \\ \langle D_{t,x} \rangle^{(-\frac{1}{2})+} (\bar{L}^{1+}(L^2)) &\subseteq L^2(L^2) && \text{(Sobolev).} \end{aligned}$$

□

*Proof of (143) for the quadratic terms in  $\mathcal{N}_4$  containing  $\nabla_{t,x}\phi^{(k-1)}$ .* Here we need to show the estimates

$$(149) \quad \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( A^{(k-1)} \cdot \nabla \phi^{(k-1)} \right) \hookrightarrow L^2,$$

$$(150) \quad \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( A_0^{(k-1)} \partial_t \phi^{(k-1)} \right) \hookrightarrow L^2.$$

We first prove (149). We begin with proving the embedding

$$(151) \quad \sum_{\mu_1 \ll \mu_2} \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(\bar{L}^\infty(H_{t,x}^s)) \cdot \nabla S_{\mu_2}(X_1^{s, \frac{1}{2}}) \right) \subseteq L^2.$$

The multiplier  $S_{|\xi| \ll |\tau|}$  and the condition  $\mu_1 \ll \mu_2$  together imply that one may replace  $S_{\mu_2}(X^{s, \frac{1}{2}})$  in the above sum by  $S_{\mu_2, \mu_2}(X^{s, \frac{1}{2}})$ , i.e., one has that

$$\begin{aligned} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(\bar{L}^\infty(H_{t,x}^s)) \cdot \nabla S_{\mu_2}(X_1^{s, \frac{1}{2}}) \right) \\ \approx S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(\bar{L}^\infty(H_{t,x}^s)) \cdot S_{\mu_2}(X_1^{s-\frac{1}{2}, 0}) \right). \end{aligned}$$

Thus, to show (151), it suffices to note the estimate

$$(152) \quad \langle D_x \rangle^{(s-\frac{3}{2})+} \left( \bar{L}^\infty(H_{t,x}^s) \cdot L^2(H_{t,x}^{s-\frac{1}{2}}) \right) \subseteq L^2 \quad \text{(Sobolev).}$$

Next, we need to prove the estimate (151) where the sum is taken over the region where  $\mu_2 \lesssim \mu_1$ . Here we use the fact that  $A$  is divergence free to write  $A \cdot \nabla \phi = (div)(\phi A)$ . With this reduction and after some weight manipulation, we see that we are trying to show that

$$(\bar{L}^\infty \cap L^2)(L^2) \cdot X_1^{(s+\frac{1}{2})-, \frac{1}{2}} \subseteq L^2,$$

which in turn follows from the embeddings

$$\begin{aligned} (\bar{L}^\infty \cap L^2)(L^2) &\subseteq \bar{L}^{\infty-}(L^2) && \text{(Interpolation),} \\ X_1^{(s+\frac{1}{2})-, \frac{1}{2}} &\subseteq \bar{L}^{2+}(\bar{L}^\infty) && \text{(Strichartz \& Sobolev).} \end{aligned}$$

We now prove (150). Following the remarks at the end of (151), we may split this into the following two embeddings:

$$(153) \quad \sum_{\substack{\mu_i \\ \delta \sim \mu_2}} \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(A_0^{(k-1)}) \partial_t S_{\mu_2, \delta}(\phi^{(k-1)}) \right) \hookrightarrow L^2,$$

$$(154) \quad \sum_{\substack{\mu_2 \lesssim \mu_1 \\ \delta \ll \mu_2}} \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda,d} \left( S_{\mu_1}(A_0^{(k-1)}) \partial_t S_{\mu_2, \delta}(\phi^{(k-1)}) \right) \hookrightarrow L^2.$$

To deal with the first of the above embeddings, we note that the condition  $\delta \sim \mu_2$  allows us to replace  $\partial_t S_{\mu_2, \delta}(\phi^{(k-1)})$  with  $\mu_2^{\frac{1}{2}} \delta^{\frac{1}{2}} S_{\mu_2, \delta}(\phi^{(k-1)})$ , and we use the splitting (148) to reduce (153) to the estimate (152), and the following:

$$\langle D_x \rangle^{s-1} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} \left( H^{s-\frac{1}{2}, 1-s, s} \cdot X^{s-\frac{1}{2}, 0} \right) \subseteq L^2,$$

which in turn follows from the embeddings

$$\langle D_x \rangle^{s-1} \left( H^{s-\frac{1}{2}, 1-s, s} \cdot X^{s-\frac{1}{2}, 0} \right) \subseteq \bar{L}^{1+}(L^2) \quad (\text{Several Sobolev's in space}),$$

$$\langle D_{t,x} \rangle^{(-\frac{1}{2})+} (\bar{L}^{1+}(L^2)) \subseteq L^2(L^2) \quad (\text{Sobolev in time}).$$

To deal with the embedding (154), we first note that the condition  $\delta \ll \mu_2$  assures us that we may replace the term  $\partial_t S_{\mu_2, \delta}(\phi^{(k-1)})$  with  $\langle D_x \rangle S_{\mu_2, \delta}(\phi^{(k-1)})$ . We now use the splitting

$$A_0^{(k-1)} = A_0^{(k-1)}_{H^{s-\frac{1}{2}, 1-s, s}} + A_0^{(k-1)}_{\langle D_{t,x} \rangle^{-\frac{3}{2}-s} (\bar{L}^\infty \cap L^2)(\bar{L}^1)},$$

and some weight trading to reduce (154) to proving the following set of estimates:

$$(155) \quad \langle D_x \rangle^{(-\frac{3}{2})-} \left( L^2(L^2) \cdot X_1^{0, \frac{1}{2}} \right) \subseteq L^2(L^2),$$

$$(156) \quad H^{s-\frac{1}{2}, 1-s, s} \cdot X^{\frac{1}{2}, -\frac{1}{2}} \subseteq L^2(L^2),$$

$$(157) \quad \langle D_x \rangle^{(-\frac{3}{2})-} \left( (\bar{L}^\infty \cap L^2)(\bar{L}^1) \cdot X_1^{1+, \frac{1}{2}} \right) \subseteq L^2(L^2).$$

The estimates (155)–(156) follow easily from a combination of energy estimates for the  $X_1$  spaces and spatial Sobolev embeddings. The last estimate, (157), follows from the embeddings

$$X_1^{1+, \frac{1}{2}} \subseteq \bar{L}^{2+}(\bar{L}^\infty) \quad (\text{Strichartz \& Sobolev}),$$

$$(\bar{L}^\infty \cap L^2)(\bar{L}^1) \subseteq \bar{L}^{\infty-}(\bar{L}^1) \quad (\text{Interpolation}),$$

$$\langle D_x \rangle^{(-\frac{3}{2})-} (L^2(\bar{L}^1)) \subseteq L^2(L^2) \quad (\text{Sobolev}).$$

□

*Proof of (143) for the term  $\mathcal{B}_0^{(k-1)} \phi^{(k-1)}$ .* We are trying to show that

$$(158) \quad \langle D_x \rangle^{s-1} \lambda^{(-\frac{1}{2})+} S_{|\xi| \ll |\tau|} S_{\lambda, d} \left( \mathcal{B}_0^{(k-1)} \phi^{(k-1)} \right) \hookrightarrow L^2(L^2).$$

Using the splitting

$$\mathcal{B}_0^{(k-1)} = \mathcal{B}_0^{(k-1)}_{H^{s-\frac{1}{2}, 0, 0}} + \mathcal{B}_0^{(k-1)}_{\langle D_{t,x} \rangle^{-s-\frac{1}{2}} (\bar{L}^\infty \cap L^2)(\bar{L}^1)},$$

and getting rid of multipliers, (158) reduces to proving the following two embeddings:

$$(159) \quad \langle D_x \rangle^{s-1} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} \left( H^{s-\frac{1}{2}, 0, 0} \cdot X_1^{s, \frac{1}{2}} \right) \subseteq L^2(L^2),$$

$$(160) \quad \langle D_x \rangle^{s-1} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} \left( \langle D_{t,x} \rangle^{-s-\frac{1}{2}} (\bar{L}^\infty \cap L^2)(\bar{L}^1) \cdot X_1^{s, \frac{1}{2}} \right) \subseteq L^2(L^2).$$

The first of the above embeddings, (159), follows directly from the estimates

$$X_1^{s, \frac{1}{2}} \subseteq \bar{L}^\infty(H_{t,x}^s) \quad (\text{Energy}),$$

$$\langle D_x \rangle^{s-1} \langle D_{t,x} \rangle^{(-\frac{1}{2})+} \left( L^2(H_{t,x}^{s-\frac{1}{2}}) \cdot \bar{L}^\infty(H_{t,x}^s) \right) \subseteq L^2(L^2) \quad (\text{Sobolev}).$$

To prove (160), after a case splitting with respect to relative frequency sizes, it suffices to prove the following embeddings:

$$\begin{aligned} \langle D_x \rangle^{(-s-1)+} \left( \bar{L}^\infty(\bar{L}^1) \cdot X_1^{1+, \frac{1}{2}} \right) &\subseteq L^2(L^2), \\ \langle D_x \rangle^{(-\frac{1}{2})+} \left( \langle D_{t,x} \rangle^{-s-\frac{1}{2}} \bar{L}^\infty(\bar{L}^1) \cdot X_1^{1+, \frac{1}{2}} \right) &\subseteq L^2(L^2), \end{aligned}$$

both of which follow from the estimate

$$X_1^{1+, \frac{1}{2}} \subseteq \bar{L}^2(\bar{L}^\infty) \quad (\text{Strichartz}),$$

and several rounds of the Sobolev embedding theorem.  $\square$

This completes the proof of the inclusion (56b).

#### ACKNOWLEDGMENTS

Both authors would like to thank Sergiu Klainerman for continuing support and encouragement. The second author would especially like to thank Hans Lindblad and the mathematics department at UCSD for their hospitality while a portion of this work was being completed.

#### REFERENCES

- [1] Jean Bourgain, *Fourier transform restriction phenomena for certain lattice subsets and applications to non-linear evolution equations I, II*. *Geom. Funct. Analysis* **3** (1993), 107–156, 202–262. MR **95d**:35160a; MR **95d**:35160b
- [2] D. Christodoulou, S. Klainerman, *Asymptotic properties of linear field equations in Minkowski space*. *Comm. Pure Appl. Math.* **43** (1990), no. 2, 137–199. MR **91a**:58202
- [3] Scipio Cuccagna, *On the local existence for the Maxwell–Klein–Gordon system in  $\mathbb{R}^{3+1}$* . *Comm. P.D.E.* **24** (1999), no. 5-6, 851–867. MR **2000c**:35203
- [4] Douglas Eardley, Vincent Moncrief, *The global existence of Yang–Mills–Higgs fields in 4-dimensional Minkowski space I. Local existence and smoothness properties*. *Comm. Math. Phys.* **83** (1982), no. 2, 171–191. MR **83e**:35106a
- [5] Douglas Eardley, Vincent Moncrief, *The global existence of Yang–Mills–Higgs fields in 4-dimensional Minkowski space II. Completion of proof*. *Comm. Math. Phys.* **83** (1982), no. 2, 193–212. MR **83e**:35106b
- [6] D. Foschi, S. Klainerman, *Bilinear space-time estimates for homogeneous wave equations*. *Ann. Sci. École Norm. Sup. (4)* **33** (2000), no. 2, 211–274. MR **2001g**:35145
- [7] Markus Keel, Terence Tao, *Endpoint Strichartz estimates*. *Amer. J. Math.* **120** (1998), no. 5, 955–980. MR **2000d**:35018
- [8] Markus Keel, Terence Tao, *Global well-posedness of the Maxwell–Klein–Gordon Equation below the energy norm*. Preprint.
- [9] C. Kenig, G. Ponce, L. Vega, *The Cauchy problem for the K-dV equation in Sobolev spaces with negative indices*. *Duke Math Journal* **71** (1994), no. 1, 1–21. MR **94g**:35196
- [10] Joachim Krieger, *Global regularity of wave maps from  $R^{3+1}$  to surfaces*. *Comm. Math. Phys.* **238** (2003), no. 1-2, 333–366.
- [11] Joachim Krieger, *Global regularity of wave maps from  $R^{2+1}$  to the hyperbolic plane*. Preprint.
- [12] Sergiu Klainerman, Matei Machedon, *Space-time estimates for null forms and the local existence theorem*, *Comm Pure Appl. Math.* **46** (1993), 1221–1268.
- [13] Sergiu Klainerman, *Geometric and Fourier Methods in Nonlinear Wave Equations*, available at <http://www.math.princeton.edu/~seri/homepage/seri.htm>.
- [14] Sergiu Klainerman, Matei Machedon, *On the Maxwell–Klein–Gordon Equations with finite energy*. *Duke Math Journal* **74** (1994), no. 1, 19–44. MR **95f**:35210
- [15] Sergiu Klainerman, Matei Machedon, *Smoothing estimates for null forms and applications*. *Duke Math Journal* **81** (1995), no. 1, 99–103. MR **97h**:35022
- [16] Sergiu Klainerman, Matei Machedon, *Estimates for null forms and the spaces  $H^{s,\delta}$* . *Internat. Math. Res. Notices* **1996**, no. 17, 853–866. MR **98j**:46028

- [17] Sergiu Klainerman, Matei Machedon, with appendices by Jean Bourgain and Daniel Tataru, *Remark on Strichartz type inequalities*. Internat. Math. Res. Notices **1996**, no. 5, 201–220. MR **97g**:46037
- [18] Sergiu Klainerman, Matei Machedon, *On the optimal local regularity for gauge field theories*. Differential and Integral Equations **10** (1997), no. 6, 1019–1030. MR **2000d**:58043
- [19] Sergiu Klainerman, Igor Rodnianski, *On the global regularity of wave maps in the critical Sobolev norm*. Internat. Math. Res. Notices **2001**, no. 13, 655–677. MR **2002h**:58051
- [20] Sergiu Klainerman, Igor Rodnianski, *Improved local well-posedness for quasilinear wave equations in dimension three*. Duke Math. J. **117** (2003), no. 1, 1–124.
- [21] Sergiu Klainerman, Igor Rodnianski, *Rough Solutions to the Einstein Vacuum Equations*. Preprint.
- [22] Sergiu Klainerman, Igor Rodnianski, Terence Tao, *A physical space approach to wave equation bilinear estimates. Dedicated to the memory of Thomas H. Wolff*. J. Anal. Math. **87** (2002), 299–336.
- [23] Sergiu Klainerman, Daniel Tataru, *On the optimal local regularity for the Yang-Mills equations in  $\mathbb{R}^{4+1}$* . Journal of the AMS **12** (1999), no. 1, 93–116. MR **2000c**:58052
- [24] Hans Lindblad, *A sharp counterexample to the local existence of low-regularity solutions to nonlinear wave equations*. Duke Math Journal **72** (1993), no. 2, 503–539. MR **94h**:35165
- [25] Jalal Shatah, Michael Struwe, *The Cauchy problem for wave maps*. Internat. Math. Res. Notices **2002**, no. 11, 555–571.
- [26] Wei-Tong Shu, *Global existence of Maxwell-Higgs fields*. Nonlinear hyperbolic equations and field theory (Lake Como, 1990), 214–227, Pitman Res. Notes Math. Ser., **253**, Longman Sci. Tech., Harlow, 1992. MR **93h**:81092
- [27] Igor Rodnianski, Terence Tao, *Global regularity for the Maxwell–Klein–Gordon equations with small critical Sobolev norm in high dimensions*. Preprint.
- [28] Jacob Sterbenz, *Global regularity for generic non-linear wave equations in high dimensions I. Critical Besov space in  $(6 + 1)$ -D*. University of Maryland thesis.
- [29] Terence Tao, *Low regularity semi-linear wave equations*. Comm. P.D.E. **24** (1999), no. 3–4, 599–629. MR **2000a**:35175
- [30] Terence Tao, *Global regularity of wave maps. I. Small critical Sobolev norm in high dimension*. Internat. Math. Res. Notices **2001**, no. 6, 299–328. MR **2001m**:35200
- [31] Terence Tao, *Global regularity of wave maps II. Small energy in two dimensions*. Comm. Math. Phys. **224** (2001), no. 2, 443–544. MR **2002h**:58052
- [32] Daniel Tataru, *Local and global results for wave maps I*. Comm. Partial Differential Equations **23** (1998), no. 9–10, 1781–1793. MR **99j**:58209
- [33] Daniel Tataru, *On the equation  $\square u = |\nabla u|^2$  in  $5 + 1$  dimensions*. Math. Res. Lett. **6** (1999), no. 5–6, 469–485. MR **2001k**:35219
- [34] Daniel Tataru, *On global existence and scattering for the wave maps equation*. Amer. J. Math. **123** (2001), no. 1, 37–77. MR **2002c**:58045
- [35] Daniel Tataru, *Rough solutions for the Wave-Maps equation*. Preprint.
- [36] Hart Smith, Daniel Tataru, *Sharp local well-posedness results for the nonlinear wave equation*. Preprint.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742  
*E-mail address:* `mxm@math.umd.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742  
*E-mail address:* `jks@math.umd.edu`

*Current address:* Department of Mathematics, Princeton University, Princeton, New Jersey 08544

*E-mail address:* `sterbenz@math.princeton.edu`