ENERGY DISPERSED SOLUTIONS FOR THE $(4+1)$-DIMENSIONAL MAXWELL-KLEIN-GORDON EQUATION

SUNG-JIN OH AND DANIEL TATARU

ABSTRACT. This article is devoted to the mass-less energy critical Maxwell-Klein-Gordon system in $4+1$ dimensions. In earlier work of the second author, joint with Krieger and Sterbenz, we have proved that this problem has global well-posedness and scattering in the Coulomb gauge for small initial data. This article is the second of a sequence of three papers of the authors, whose goal is to show that the same result holds for data with arbitrarily large energy. Our aim here is to show that large data solutions persist for as long as one has small energy dispersion; hence failure of global well-posedness must be accompanied with a non-trivial energy dispersion.

CONTENTS

1. INTRODUCTION

This article is concerned with the mass-less energy critical Maxwell-Klein-Gordon system (MKG) in the $4+1$ dimensional Minkowski space $\mathbb{R}^{1+4}$ equipped with the standard Lorentzian metric $\mathbf{m} = \mathrm{diag}(-1,1,1,1,1)$ in the standard rectilinear coordinates $(x^0,\dots,x^4)$. This system is generated by adding a scalar field component to the standard Maxwell Lagrangian,

$$\mathcal{S}_{\mathrm{M}}[A_\alpha] := \int_{\mathbb{R}^{1+4}} \frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} \, dxdt;$$

to obtain

$$\mathcal{S}[A_\alpha, \phi] := \int_{\mathbb{R}^{1+4}} \frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} + \frac{1}{2} \mathbf{D}_\alpha \phi \overline{\mathbf{D}^\alpha \phi} \, dxdt;$$

Here $\phi : \mathbb{R}^{1+4} \to \mathbb{C}$ is a scalar function, and $A_\alpha : \mathbb{R}^{1+4} \to \mathbb{R}$ is a real-valued connection 1-form, with curvature

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

The connection 1-form A_α is then used to define the covariant derivative

$$\mathbf{D}_\alpha \phi := (\partial_\alpha + iA_\alpha)\phi.$$

Introducing the covariant wave operator

$$\Box_A := \mathbf{D}^\alpha \mathbf{D}_\alpha$$

with the standard convention for raising/lowering and summing indices, we can write the *Maxwell-Klein-Gordon system* in the form

$$\begin{cases} \partial^\beta F_{\alpha\beta} = -J_\alpha, \\ \Box_A \phi = 0 \end{cases} \tag{1.1}$$

where the currents J_α are defined as

$$J_\alpha := -\mathrm{Im}(\phi \overline{\mathbf{D}_\alpha \phi}) . \tag{1.2}$$

The MKG system admits a positive definite formally conserved energy functional,

$$\mathcal{E}[A,\phi](t) = \mathcal{E}_{\{t\}\times\mathbb{R}^4}[A,\phi] := \int_{\{t\}\times\mathbb{R}^4} \Big(\frac{1}{4} \sum_{\alpha,\beta} F_{\alpha\beta}^2 + \frac{1}{2} \sum_\alpha |\mathbf{D}_\alpha \phi|^2 \Big) \, dx \tag{1.3}$$

and is also invariant under the scaling

$$\phi(t,x) \to \lambda^{-1}\phi(\lambda^{-1}t, \lambda^{-1}x), \ A_\alpha(t,x) \to \lambda^{-1}A(\lambda^{-1}t, \lambda^{-1}x).$$

Moreover, the energy functional is also invariant under the same scaling. Thus the $4+1$-dimensional MKG system is energy critical.

In order to state this system as a formally well-posed initial value problem, we need to take into account its gauge invariance. If (A_α, ϕ) is a solution, then so is $(A_\alpha - \partial_\alpha\chi, e^{i\chi}\phi)$ for any real-valued scalar function χ. In the gauge covariant setting, it is natural to define an *initial data set* for MKG to consist of a pair of 1-forms (a_j, e_j) and complex-valued functions f, g on $\mathbb{R}^4$. We say that (a,e,f,g) is the initial data set for a solution (A, ϕ) if

$$(A_j, F_{0j}, \phi, \mathbf{D}_t\phi)\restriction_{\{t=0\}} = (a_j, e_j, f, g),$$

where the latin indices only run over the spatial variables $x^1, \ldots, x^4$. The energy of the set (a,e,f,g), denoted by $\mathcal{E}[a,e,f,g]$, is defined in the obvious way from (1.3). The $\alpha = 0$ component of the MKG system imposes the *Gauss* (or *constraint*) *equation* for initial data sets, namely

$$\partial^\ell e_\ell = -J_0 = \mathrm{Im}(f\bar{g}). \tag{1.4}$$

To eliminate the gauge ambiguity, we add to the above system a single scalar gauge condition. Here we follow the approach in [13] and work with the *global Coulomb gauge*

$$\sum_{j=1}^4 \partial_j A_j = 0 \tag{1.5}$$

2

where latin summation indices are used for summations which are only with respect to spatial variables. Using this gauge, the MKG system can be written explicitly in the following form

$$\begin{cases} \Box A_i = \mathcal{P}_i J_x \\ \Box_A \phi = 0 \end{cases} \tag{1.6}$$

for the dynamic variables (A_i, ϕ). The operator $\mathcal{P}$ is the Leray projection onto divergence free vector fields,

$$\mathcal{P}_j v = I - \partial_j \Delta^{-1} \partial^\ell v_\ell$$

The second equation in (1.6) requires also the temporal component A_0, which is determined in an elliptic fashion, together with its time derivative, by the equations

$$\Delta A_0 = J_0, \qquad \Delta \partial_t A_0 = \partial^i J_i, \tag{1.7}$$

with standard decay conditions at infinity. Note that the first equation is precisely the Gauss equation. These equations uniquely determine both A_0 and $\partial_t A_0$ at fixed time.

Well-posedness theory of MKG at (scaling) sub-critical regularity has been studied extensively in various gauges. In dimensions $2+1$ and $3+1$, this system is energy sub-critical, and hence global well-posedness follows from an appropriate local well-posedness result; see [2, 3, 4, 5, 8, 15, 16, 21] and references therein. In $\mathbb{R}^{1+4}$, almost optimal local well-posedness of a model problem closely related to MKG and the Yang-Mills system was proved in [9]; this result was then further refined in [20, 22]. For a more detailed survey of earlier works on MKG, see [18, Section 1.3].

The subject of this article, as well as its companions [17, 18], is the energy critical MKG-CG problem in 4+1 dimensions. Given an arbitrary finite energy data set for the MKG problem, there exists a unique gauge equivalent data set of related size which satisfies the Coulomb gauge condition; see [17, Section 3]. Hence the main question now is to decide whether each finite energy MKG-CG initial data set can be extended to a global-in-time solution for the MKG-CG system. This is analogous to the celebrated *threshold conjecture* for energy critical wave maps, which has been recently answered in the affirmative [11, 23, 24, 26, 27, 28, 29, 30] (see also [14]).

The small data global well-posedness result was first obtained in high dimension $n \geq 6$ by Rodnianski-Tao [19]. The low dimensional result $n \geq 4$ was obtained more recently by Krieger-Sterbenz-Tataru [13]. The theorem in [13] asserts the following:

Theorem 1.1 ([13]). *There exists a universal constant $\epsilon_* > 0$ such that the following hold.*

(1) *(Existence and uniqueness) Let (a, e, f, g) be a C^∞ Coulomb data set (i.e., $\partial^\ell a_\ell = 0$) satisfying*

$$\mathcal{E}[a, e, f, g] < \epsilon_*^2 . \tag{1.8}$$

Then the MKG-CG system (1.5)-(1.6) admits a unique global smooth solution (A, ϕ) on $\mathbb{R}^{1+4}$ with these data.

(2) *(Continuous dependence) In addition, for every compact time interval J containing 0, the data-to-solution operator extends continuously on the set (1.8) to a map*

$$\mathcal{H}^1(\mathbb{R}^4) \ni (a, e, f, g) \to (A, \phi) \in C(J; \dot{H}^1(\mathbb{R}^4)) \cap \dot{C}^1(J; L^2(\mathbb{R}^4))$$

where the space $\mathcal{H}^1 = \mathcal{H}^1(\mathbb{R}^4)$ of finite energy initial data sets is defined by the norm

$$\|(a, e, f, g)\|_{\mathcal{H}^1}^2 := \|a\|_{\dot{H}^1}^2 + \|e\|_{L^2}^2 + \|f\|_{\dot{H}^1}^2 + \|g\|_{L^2}^2. \tag{1.9}$$

3

The last statement allows us to define the following notion of finite energy solutions:

Definition 1.2. Let I be a time interval. We define the space $C_t\mathcal{H}^1(I \times \mathbb{R}^4)$ by the norm

$$\|(A, \phi)\|_{C_t\mathcal{H}^1(I \times \mathbb{R}^4)} = \sup_{t \in I} \left(\sup_\mu \|A_\mu[t]\|_{\dot{H}^1 \times L^2} + \|\phi[t]\|_{\dot{H}^1 \times L^2} \right).$$

We say that a pair $(A, \phi) \in C_t\mathcal{H}^1(I \times \mathbb{R}^4)$ is an *admissible $C_t\mathcal{H}^1$ solution* to MKG on I if there exists a sequence $(A^{(n)}, \phi^{(n)})$ of classical solutions to MKG on $I \times \mathbb{R}^4$ such that

$$\|(A, \phi) - (A^{(n)}, \phi^{(n)})\|_{C_t\mathcal{H}^1(J \times \mathbb{R}^4)} \quad \text{as } n \to \infty$$

for every compact subinterval $J \subseteq I$. Here by *classical* solutions, we mean that

$$A, \phi \in \bigcap_{n,m=0}^\infty C_t^m(I; H^n).$$

In the process of proving the above result in [13], stronger spaces $S^1, Y^1 \subset C(\dot{H}^1) \cap \dot{C}^1(L^2)$ are introduced, and it is shown that the above solutions obeys the bound

$$\|A_0\|_{Y^1} + \|(A_x, \phi)\|_{S^1} \lesssim \|(a, e, f, g)\|_{\mathcal{H}^1} \tag{1.10}$$

with a continuous (but not uniformly continuous) data-to-solution map on each compact time interval. We provide the definition of the spaces S^1 and Y^1 in Section 3. We remark that the precise definition of S^1 differs in [13], [17] and in the present paper. The difference is however minor, and all the theorems stated here hold with respect to any of these three definitions. See also Remark 3.1.

Our goal, in a sequence of three papers, is to prove that a similar result holds for all finite energy data. The three steps in our proof are as follows:

Global Coulomb gauge [17]: Here we use the above small data result to show that the large data problem is locally well-posed in the Coulomb gauge, and that the solution can be extended for as long as energy concentration does not occur.

Energy dispersed solutions (present paper): Here we prove a more refined continuation criterion, namely that the solution can be extended for as long as it remains energy dispersed. Moreover, if the solution already exists up to $t = \infty$, then we prove that small energy dispersion implies scattering.

Blow-up analysis [18]: Here we complete the proof of the large data well-posedness result, showing that no blow-up is allowed at the tip of a light cone. We also prove the corresponding scattering result.

At least in a broad outline, the second and third steps above follow the scheme successfully developed in [23] and [24] in the context of wave maps. The first step in [17] is specific to the MKG problem, and is due to the long range effect of the Gauss equation as well as the inherent gauge ambiguity of MKG. Precisely, in order to truncate a large energy initial data into small energy data sets, the Gauss equation $\partial^\ell e_\ell = \text{Im}(f\bar{g})$ must be taken into account. Furthermore, the local gauges given by the small data result applied to these truncated data differ in their common domains, and need to be aggregated into a single global Coulomb gauge. An overview of the whole sequence is provided in [18, Sections 2 and 3].

Remark 1.3. To understand the issue of gauge invariance clearly, it is advantageous to take a more geometric point of view and consider ϕ as a section of a complex line bundle L with structure group $U(1) = \{e^{i\chi} : \chi \in \mathbb{R}\}$ over $\mathbb{R}^{1+4}$, and A as a connection on L. Since the

base manifold $\mathbb{R}^{1+4}$ is contractible, L is always topologically trivial; hence ϕ can be identified with a $\mathbb{C}$-valued scalar function, and A with a real-valued 1-form on $\mathbb{R}^4$ by using the trivial connection d as a reference. The choice of a gauge then corresponds to a particular choice of bases on the fibers to describe (A, ϕ). This viewpoint is taken to some extent in the other papers of the series [17, 18] to facilitate the usage of local gauges. In the present paper, however, we need not worry about such issues, as we work exclusively in the global Coulomb gauge.

Roughly speaking, the main result in [17] is local well-posedness of MKG-CG for data with any finite energy E, with a lower bound on the lifespan in terms of the *energy concentration scale*

$$r_c = r_c[a, e, f, g] := \sup\{r > 0 : \forall x \in \mathbb{R}^4, \, \mathcal{E}_{B_r(x)}[a, e, f, g] < \delta_0(E, \epsilon_*^2)\} \tag{1.11}$$

where $\delta_0(E, \epsilon_*^2) > 0$ is some fixed function (e.g. in [17] we use $\delta_0(E, \epsilon_*^2) \approx \epsilon_*^6 E^{-2}$ for $E > \epsilon_*^2$), ϵ_*^2 is the threshold in Theorem 1.1 and $\mathcal{E}_{B_r(x)}$ is the energy measured on the ball $B_r(x)$ of radius r centered at x. Observe that $r_c[a, e, f, g] > 0$ for any $(a, e, f, g) \in \mathcal{H}^1$.

The result in [17] also admits a formulation in terms of the S^1, Y^1 norms; for that we need a generalization of these norms to bounded time intervals, which we denote by $S^1[t_0, t_1]$, $Y^1[t_0, t_1]$ (see Section 3.2 for the definition). The precise statement is as follows.

Theorem 1.4 (Large energy local well-posedness theorem in global Coulomb gauge [17]). *Let (a, e, f, g) be an $\mathcal{H}^1$ initial data set satisfying the global Coulomb gauge condition $\partial^\ell a_\ell = 0$ with energy $\mathcal{E}[a, e, f, g] \le E$. Let $r_c = r_c[a, e, f, g]$ be defined as in (1.11). Then the following statements hold:*

(1) (Existence and uniqueness) There exists a unique admissible $C_t \mathcal{H}^1$ solution (A, ϕ) to MKG-CG on $[-r_c, r_c] \times \mathbb{R}^4$ with (a, e, f, g) as its initial data.

(2) (A-priori S^1 regularity) We have the additional regularity properties

$$A_0 \in Y^1[-r_c, r_c], \quad A_x, \phi \in S^1[-r_c, r_c].$$

(3) (Persistence of regularity) The solution (A, ϕ) is classical if (a, e, f, g) is classical. Here, by classical we mean $a, e, f, g \in \cap_{n=0}^\infty H^n$.

(4) (Continuous dependence) Consider a sequence $(a^{(n)}, e^{(n)}, f^{(n)}, g^{(n)})$ of $\mathcal{H}^1$ Coulomb initial data sets such that

$$\|(a^{(n)} - a, e^{(n)} - e, f^{(n)} - f, g^{(n)} - g)\|_{\mathcal{H}^1} \to 0 \quad \text{as } n \to \infty.$$

Then the lifespan of $(A^{(n)}, \phi^{(n)})$ eventually contains $[-r_c, r_c]$, and we have

$$\|A_0 - A_0^{(n)}\|_{Y^1[-r_c, r_c]} + \|(A_x - A_x^{(n)}, \phi - \phi^{(n)})\|_{S^1[-r_c, r_c]} \to 0 \quad \text{as } n \to \infty.$$

In other words, this result says that even if the initial data is large, we can continue the solution as a global Coulomb solution with good S^1 bounds for as long as energy does not concentrate to arbitrarily small balls.

Our main result here is based on the notion of energy dispersion introduced in [23]. Adapted to our context, the energy dispersed norm we use is

$$\|\phi\|_{ED(t_1, t_2)} = \sup_k 2^{-k} \|(P_k \phi, 2^{-k} P_k \phi_t)\|_{L^\infty[(t_1, t_2) \times \mathbb{R}^4]} \tag{1.12}$$

We measure the energy dispersion only for ϕ, and not for A.

The main theorem is as follows:

Theorem 1.5 (Energy Dispersed Regularity Theorem)**.** *There exist two functions* $1 \ll F(E)$ *and* $0 < \epsilon(E) \ll 1$ *of the energy* (1.3) *such that the following statement is true:*

If (A, ϕ) *is an admissible* $C_t \mathcal{H}^1$ *solution to MKG-CG on the open interval* (t_1, t_2) *with energy* $\leq E$ *and energy dispersion at most* $\epsilon(E)$, *i.e.,*

$$\|\phi\|_{ED(t_1, t_2)} \leq \epsilon(E),$$

then the following a-priori bound holds:

$$\|(A_x, \phi)\|_{S(t_1, t_2)} \leq F(E). \tag{1.13}$$

We remark that (1.13) implies the bound (see Theorem 5.1)

$$\|A_0\|_{Y^1(t_1, t_2)} + \|(A_x, \phi)\|_{S^1(t_1, t_2)} \lesssim_{F(E)} 1.$$

We also prove a continuation and scattering result, which may be applied in conjunction with Theorem 1.5.

Theorem 1.6 (Continuation and scattering of solutions with finite S^1 norm)**.** *Let* (A, ϕ) *be an admissible* $C_t \mathcal{H}^1$ *solution to MKG-CG on* $[0, T_+) \times \mathbb{R}^4$, *with* $0 < T_+ \leq \infty$, *obeying the bound*

$$\|A_0\|_{Y^1[0, T_+)} + \|(A_x, \phi)\|_{S^1[0, T_+)} < \infty.$$

Then the following statements hold.

(1) If $T_+ < \infty$, *then* (A, ϕ) *extends to an admissible* $C_t \mathcal{H}^1$ *solution with finite* S^1 *norm past* T_+.

(2) If $T_+ = \infty$, *then* (A_x, ϕ) *scatters as* $t \to \infty$ *in the following sense: There exists a solution* $(A_x^{(\infty)}, \phi^{(\infty)})$ *to the linear system*

$$\Box A_j^{(\infty)} = 0, \quad (\Box + 2iA_\ell^{free}\partial^\ell)\phi^{(\infty)} = 0,$$

with initial data $A_x^{(\infty)}[0], \phi^{(\infty)}[0] \in \dot{H}^1 \times L^2$ *such that*

$$\|A_x[t] - A_x^{(\infty)}[t]\|_{\dot{H}^1 \times L^2} + \|\phi[t] - \phi^{(\infty)}[t]\|_{\dot{H}^1 \times L^2} \to 0 \quad \text{as } T \to \infty.$$

Here A_x^{free} *is a homogeneous wave with* $A_x^{free}[0] = A_x[0]$.

Analogous statements hold in the past time direction as well.

We remark that the choice of A_x^{free} in the scattering statement above is somewhat robust, in that one can freely perturb $A_x^{free}[0]$ by any function in $\ell^1(H^1 \times L^2)$ where ℓ^1 stands for dyadic summation in frequency. In particular one can take $A_x^{free} = A_x^{(\infty)}$.

Our strategy for proving Theorem 1.5 is to use an induction on energy argument; this is imposed by the requirement to renormalize paradifferential interactions of the solution with itself. This is somewhat similar to the proof of the corresponding result for wave maps in [23]. See also [10] for an exposition of this argument in the context of wave maps, and Section 2 in the main paper of the sequence [18] for a brief summary of our strategy.

Remark 1.7. We remark that the same results hold in all higher dimensions for data in the scale invariant space $\dot{H}^{\frac{d}{2}-1} \times \dot{H}^{\frac{d}{2}-2}$. We have chosen to restrict our exposition to the more difficult case $d = 4$ in order to keep the notations simple, but our analysis easily carries over to higher dimension $d \geq 5$. The main difference in higher dimension is that we no longer have a conserved energy which is equivalent to the critical Sobolev norm. However, the small energy dispersion guarantees that the critical energy is almost conserved.

Remark 1.8. We note that an independent proof of global well-posedness and scattering of MKG-CG has been recently announced by Krieger-Lührmann, following a version of the Bahouri-Gérard nonlinear profile decomposition [1] and Kenig-Merle concentration compactness/rigidity scheme [6, 7] developed by Krieger-Schlag [11] for the energy critical wave maps.

1.1. **Notation and Conventions.** We use the asymptotic notation $A \lesssim B$ and $A = O(B)$ to mean $A \leq CB$ for some $C > 0$. We write $A \ll B$ if the implicit constant should be regarded as small. The dependence of the constant is specified by a subscript.

Our convention regarding indices is as follows. The greek indices $\alpha, \beta, \dots$ run over $0, \dots, 4$, whereas the latin indices $i, j, \dots$ only run over the spatial indices $1, \dots, 4$. We raise and lower indices using the Minkowski metric, and sum over repeated upper and lower indices.

We refer to each directional derivative by ∂_μ, and the full space-time gradient by ∇. We denote the (gauge) covariant derivative by $\mathbf{D}_\mu = \partial_\mu + iA_\mu$. For (Fourier) multipliers and pseudodifferential operators, it is convenient to use $D_\mu = \frac{1}{i}\partial_\mu$, whose symbol is ξ_μ.

Global small constants. We introduce a string of globally defined small constants, which are used in our main argument contained in Sections 2-8:

$$0 < \delta_{**} \ll \delta_* \ll \delta_0 \ll \delta_1 \ll c \ll \delta \ll 1.$$

Logically, each constant is chosen to be small enough depending on the one to the immediate right. For the convenience of the reader, we summarize the role of each constant as follows: δ is the exponent for dyadic gains in bilinear and multilinear estimates, most which come from [13]; c enters in the gain in large frequency gaps m; δ_1 is used for the gain in small energy dispersion; δ_0 is reserved for the definition of admissible frequency envelopes; δ_* and δ_{**} are the small constants used in the induction on energy argument in Section 6.

Littlewood-Paley projections. Let $m_{\leq 0}(r)$ be a smooth cutoff that equals 1 on $\{r \leq 1\}$ and vanishes on $\{r \geq 2\}$. For $k \in \mathbb{Z}$, let $m_{\leq k}(r) := m_{\leq 0}(r/2^k)$ and $m_k(r) := m_{\leq k}(r) - m_{\leq k-1}(r)$; then $\operatorname{supp} m_k \subseteq \{2^{k-1} \leq r \leq 2^{k+1}\}$ and forms a locally finite partition of unity, i.e., $\sum_k m_k = 1$. Using the space-time Fourier transform $\mathcal{F}$, we define various dyadic (or *Littlewood-Paley*) projections as follows:

$$P_k\varphi = \mathcal{F}^{-1}[m_k(|\xi|)\mathcal{F}[\varphi]], \quad Q_j\varphi = \mathcal{F}^{-1}[m_j(||\tau| - |\xi||)\mathcal{F}[\varphi]], \quad S_\ell\varphi = \mathcal{F}^{-1}[m_k(|(\tau, \xi)|)\mathcal{F}[\varphi]].$$

We also define $Q_j^\pm := Q^\pm Q_j$, where $Q^\pm := \mathcal{F}^{-1}[1_{[0,\infty)}(\pm\tau)\mathcal{F}[\varphi]]$ restricts to the $\pm$ frequency half-space. For an interval $I \subseteq \mathbb{Z}$, we define $P_I = \sum_{k \in I} P_k$, etc. At one place, we allow P_k to depend continuously on $k \in \mathbb{R}$; see the definition of $(\tilde{A}[0], \tilde{\phi}[0])$ in Section 6.

Frequency envelopes. For some more accurate bounds at various places we need to keep better track of the frequency distribution of norms. This is done using the language of frequency envelopes. An *admissible frequency envelope* will be any sequence $\{c_k\}_{k \in \mathbb{Z}}$ of positive numbers which is slowly varying,

$$c_j/c_k \leq 2^{\delta_0|j-k|}$$

with a small universal constant δ_0. Given such a sequence and a norm X, we define the norm

$$\|\phi\|_{X_c} = \sup_k c_k^{-1}\|P_k\phi\|_X.$$

We say that c is a frequency envelope for the data $(A_x[0], \phi[0])$ if for every $k \in \mathbb{Z}$, we have

$$\|(P_k A_x[0], P_k\phi[0])\|_{\dot{H}^1 \times L^2} \leq c_k.$$

Given any $A_x[0], \phi[0] \in \dot{H}^1 \times L^2$, we may construct such a c by convolving with $2^{-\delta_0|\cdot|}$, i.e.,

$$c_k := \sum_{k'} 2^{-\delta_0|k-k'|} \|(P_{k'} A_x[0], P_{k'} \phi[0])\|_{\dot{H}^1 \times L^2} \ .$$

By Young's inequality, we have $\|c\|_{\ell^2} \lesssim \|(A_x[0], \phi[0])\|_{\dot{H}^1 \times L^2}$.

1.2. **Structure of the paper.** In Section 2, we begin with some elliptic gauge related fixed time estimates. In particular these will help us relate the full nonlinear gauge independent energy with the linear energy associated to the MKG-CG system.

In the following section we switch to space-time analysis, and define the function spaces S^1 and N; with minor changes this follows [13]. We also recall some useful estimates from [13], and add to that some additional properties related to the interval decomposition of the S^1 and N spaces.

In Section 4 we describe the decomposition of the nonlinearity, and state the main bilinear and multilinear bounds which enter into the proof of our main result. To overcome difficulties related to large data, here we consider two additional classes of estimates, namely energy dispersed bounds and time divisible estimates.

In Section 5 we consider MKG waves of finite S^1 norm, and we establish further regularity properties for such waves. Based on these properties, we establish Theorem 1.6. We also consider the special case of MKG waves with small energy dispersion, and show that some other norms of such waves must also be small.

Section 6 contains the proof of our main result in Theorem 1.5. This is achieved using an induction of energy argument, following the principles introduced in [23].

The following two sections contain the proof of the bilinear and the trilinear estimates, where, in addition to results from [13], we bring in the energy dispersion and divisible norms. Heuristically, we will see that the role played by the small energy dispersion is to improve all the balanced frequency interactions in the bilinear estimates in Section 7. In the trilinear estimates in Section 8, there are possibly large unbalanced frequency interactions for which the small energy dispersion does not seem effective. Nevertheless, we show that the bulk can be bounded by a time divisible norm. This property allows us to carry out an induction on energy scheme as in Section 6.

Finally, the last section contains our paradifferential parametrix construction, based on those in [13, 19]. While very different technically, at the conceptual level this is similar to the argument in [23]. The main idea there is that a large frequency gap, rather than the small energy dispersion, is used to control the large paradifferential term.

2. Fixed time elliptic bounds and the energy

While the energy $\mathcal{E}[A, \phi]$ of the MKG system defined in (1.3) is gauge independent, when considering the system in the Coulomb gauge it is convenient to view (A_x, ϕ) as the main

dynamic variable, while A_0 and $\partial_t A_0$ are derived quantities obtained via the equations (1.7). Correspondingly, we view

$$(A_x[0], \phi[0]) = (A_x, \partial_t A_x, \phi, \partial_t \phi)(0)$$

as the initial data for the MKG-CG system, and determine the gauge covariant initial data set (a, e, f, g) via (1.7). We remark that $(A_x[0], \phi[0])$ can be freely prescribed up to the Coulomb condition $\partial^\ell A_\ell(0) = 0$. In this context, it is convenient to work with the linear energy

$$E_{lin}[A_x, \phi](t) = E_{lin}(A_x[t], \phi[t]) := \frac{1}{2} \int \sum_{\substack{\mu=0,\dots,4 \\ j=1,\dots,4}} |\partial_\mu A_j(t)|^2 + \sum_{\mu=0,\dots,4} |\partial_\mu \phi(t)|^2 \, dx. \qquad (2.1)$$

In order to justify this, we need to show that A_0 is indeed uniquely determined by (A_x, ϕ) at each time, and that the two energies are in some sense comparable. This is the goal of the main result here. In the process, we will also obtain some further solvability estimates for the equations (1.7) for A_0 that will also come in handy in the context of space-time bounds. We have:

Proposition 2.1. *The following statements hold.*
(1) *Let $(A_x, F_{0x}, \phi, \mathbf{D}_t \phi)(0)$ be a finite energy initial data set for the MKG-CG system. Then $(A_x[0], \phi[0]) \in \dot{H}^1 \times L^2$ and we have the estimate*

$$E_{lin}(A_x[0], \phi[0]) \lesssim \mathcal{E}[A, \phi] + \mathcal{E}[A, \phi]^2, \qquad (2.2)$$

where $\mathcal{E}[A, \phi]$ denotes the energy of the initial data set $(A_x, F_{0x}, \phi, \mathbf{D}_t \phi)(0)$.
(2) *Conversely, suppose that $(A_x[0], \phi[0]) \in \dot{H}^1 \times L^2$. Then there exist unique solutions $(A_0, \partial_t A_0) \in \dot{H}^1 \times L^2$ for the equations (1.7), depending smoothly on $(A_x[0], \phi[0])$ in the above topologies. Further, $\mathcal{E}(A, \phi)$ depends smoothly on $(A_x[0], \phi[0])$, and we have the energy relation*

$$\mathcal{E}[A, \phi] \lesssim E_{lin}(A_x[0], \phi[0]) + E_{lin}(A_x[0], \phi[0])^2. \qquad (2.3)$$

(3) *Assume in addition that $\phi[0]$ obeys the fixed time energy dispersion bound*

$$\|\phi[0]\|_{ED} := \sup_k 2^{-k} \|(P_k \phi, 2^{-k} P_k \partial_t \phi)(0)\|_{L^\infty(\mathbb{R}^4)} \leq \epsilon \qquad (2.4)$$

with $\epsilon \ll_{\mathcal{E}[A,\phi]} 1$. Then we have

$$\mathcal{E}[A, \phi] = E_{lin}(A_x[0], \phi[0]) + O_{\mathcal{E}[A,\phi]}(\epsilon^{\frac{1}{4}}). \qquad (2.5)$$

Proof. All estimates here are at fixed time, so we dispense with the time variable from the notations. We denote the two energies $\mathcal{E}[A, \phi]$ and $E_{lin}(A_x[0], \phi[0])$ simply by E and E_{lin}, respectively.

(1)). We begin with the spatial components of the energy, where we have

$$\frac{1}{2} \sum_{1 \leq j < k \leq 4} \|\partial_j A_k - \partial_k A_j\|_{L^2}^2 \leq E.$$

Combined with the gauge condition $\partial^j A_j = 0$, this gives the linear elliptic bound

$$\frac{1}{2} \|A_x\|_{\dot{H}^1}^2 = \frac{1}{2} \sum_{1 \leq j,k \leq 4} \|\partial_j A_k\|_{L^2}^2 \leq E,$$

9

and Sobolev embeddings further yield

$$\|A_x\|_{L^4}^2 \lesssim E.$$

On the other hand we also have

$$\frac{1}{2}\|\mathbf{D}_x\phi\|_{L^2}^2 \leq E.$$

By the diamagnetic inequality and Sobolev embeddings we obtain

$$\|\phi\|_{L^4}^2 \lesssim \|\nabla|\phi|\|_{L^2}^2 \leq \|\mathbf{D}_x\phi\|_{L^2}^2 \lesssim E.$$

Then we can further estimate

$$\|\phi\|_{\dot{H}^1}^2 = \|\nabla_x\phi\|_{L^2}^2 \leq \|\mathbf{D}_x\phi\|_{L^2}^2 + \|A_x\phi\|_{L^2}^2 \lesssim E + E^2$$

Next we turn our attention to the temporal components. We first have

$$\frac{1}{2}\|\partial_t A_x - \nabla_x A_0\|_{L^2}^2 \leq E.$$

Applying the divergence and using the Coulomb gauge condition we obtain

$$\|A_0\|_{\dot{H}^1}^2 \lesssim \|\Delta A_0\|_{\dot{H}^{-1}}^2 \lesssim E.$$

As the energy E also controls $\|\mathbf{D}_t\phi\|_{L^2}^2$, arguing as above we also obtain

$$\|\partial_t\phi\|_{L^2}^2 \lesssim E + E^2,$$

which concludes the proof of (2.2).

(2)). We begin with the analysis of the first equation in (1.7), which is rewritten as

$$(-\Delta + |\phi|^2)A_0 = -\text{Im}(\phi\overline{\partial_t\phi}).$$

We first need to know that this equation is solvable. More generally, we consider the inhomogeneous problem

$$(-\Delta + |\phi|^2)u = f \tag{2.6}$$

The solvability of this equation is dealt with via the following fixed time lemma:

Lemma 2.2. *Consider the equation (2.6) with $\phi \in \dot{H}^1$. Set $E_0 = \|\phi\|_{\dot{H}^1}$. Then*

(a) If $f \in \dot{H}^{-1}$ then there exists a unique solution $u \in \dot{H}^1$, satisfying

$$\|u\|_{\dot{H}^1} \lesssim \|f\|_{\dot{H}^{-1}} \tag{2.7}$$

Further, the map $(\phi, f) \to u$ is smooth in the $\dot{H}^1 \times \dot{H}^{-1} \to \dot{H}^1$ topology.
(b) If $f \in H^{-\frac{1}{2}}$ then there exists a unique solution $u \in \dot{H}^{\frac{3}{2}}$, satisfying

$$\|u\|_{\dot{H}^{\frac{3}{2}}} \lesssim_{E_0} \|f\|_{\dot{H}^{-\frac{1}{2}}} \tag{2.8}$$

Further, the map $(\phi, f) \to A_0$ is smooth in the $\dot{H}^1 \times \dot{H}^{-\frac{1}{2}} \to \dot{H}^{\frac{3}{2}}$ topology.
(c) In addition, for any frequency envelope $c \in \ell^2$ we have the bounds

$$\|u\|_{\dot{H}_c^1} \lesssim_{E_0} \|f\|_{\dot{H}_c^{-1}}, \qquad \|u\|_{\dot{H}_c^{\frac{3}{2}}} \lesssim_{E_0} \|f\|_{\dot{H}_c^{-\frac{1}{2}}} \tag{2.9}$$

Proof. (a) By Sobolev embeddings we have

$$\||\phi|^2\|_{\dot{H}^1 \to \dot{H}^{-1}} \lesssim \|\phi\|_{L^4}^2 \lesssim \|\phi\|_{\dot{H}^1}^2$$

Hence the operator $-\Delta + |\phi|^2$ is bounded from $\dot{H}^1 \to \dot{H}^{-1}$. It is also self-adjoint and coercive, so the bound

$$\|u\|_{\dot{H}^1} \le \|f\|_{\dot{H}^{-1}}$$

immediately follows. The regularity of the map $(\phi, f) \to u$ is obtained in a similar manner, by looking at the linearized equation.

(b) More generally, we will take $f \in \dot{H}^\sigma$ and prove that we can solve for $u \in \dot{H}^{\sigma+2}$ for any $-2 < \sigma < 0$. This in particular easily implies the frequency envelope bounds in part (c). By duality it suffices to consider the case $-1 \le \sigma < 0$.

To solve the problem perturbatively in $\dot{H}^{\sigma+2}$ it suffices to construct a multiplier Λ so that $\Lambda(\xi) \approx_{E_0} |\xi|^{\sigma+1}$ and

$$\|\Lambda|\phi|^2\Lambda^{-1} - |\phi|^2\|_{\dot{H}^1 \to \dot{H}^{-1}} \ll 1 \tag{2.10}$$

Then we can rewrite the equation as

$$(-\Delta + |\phi|^2)\Lambda u = \Lambda f - (\Lambda|\phi|^2\Lambda^{-1} - |\phi|^2)\Lambda u,$$

and the above estimate allows us to solve the above equation perturbatively based on the $\dot{H}^1$ solvability in part (a).

By duality and a Littlewood-Paley decomposition, (2.10) would follow if we had the stronger bound

$$I = \sum_{k_i} \left| \int \Lambda u_{k_1}\Lambda^{-1} u_{k_2}\phi_{k_3}\phi_{k_4} - u_{k_1}u_{k_2}\phi_{k_3}\phi_{k_4} dx \right| \ll \|u\|_{\dot{H}^1}^2 \tag{2.11}$$

We will denote each summand on the left hand side by $I(k_1, k_2, k_3, k_4)$. To achieve (2.11) we will choose Λ radial, with the property that $\Lambda(r)$ is non-decreasing and

$$\Lambda(s) \le \Lambda(r)\left(\frac{s}{r}\right)^{\sigma+1}, \qquad s > r.$$

Estimating each dyadic contribution using Sobolev embeddings we have

$$\left| \int u_{k_1}u_{k_2}\phi_{k_3}\phi_{k_4} dx \right| \lesssim 2^{-(k_{max}-k_{min})} \|u_{k_1}\|_{\dot{H}^1} \|u_{k_2}\|_{\dot{H}^1} \|\phi_{k_3}\|_{\dot{H}^1} \|\phi_{k_4}\|_{\dot{H}^1}$$

and similarly

$$\left| \int \Lambda u_{k_1}\Lambda^{-1} u_{k_2}\phi_{k_3}\phi_{k_4} dx \right| \lesssim 2^{\sigma(k_{max}-k_{min})} \|u_{k_1}\|_{\dot{H}^1} \|u_{k_2}\|_{\dot{H}^1} \|\phi_{k_3}\|_{\dot{H}^1} \|\phi_{k_4}\|_{\dot{H}^1}$$

where $k_{max} = \max\{k_1, \ldots, k_4\}$, $k_{min} = \min\{k_1, \ldots, k_4\}$. Hence contributions from widely separated frequencies are small. To measure that, we fix a frequency gap parameter m (which will be chosen depending only on E_0) and split

$$I = I_{close} + I_{far} := \sum_{k_{max}-k_{min}<m} I(k_1, \ldots, k_4) + \sum_{k_{max}-k_{min}\ge m} I(k_1, \ldots, k_4).$$

For I_{far} we have

$$I_{far} \lesssim 2^{\sigma\frac{m}{2}} \|u\|_{\dot{H}^1}^2 \|\phi\|_{\dot{H}^1}^2$$

11

which can be made sufficiently small by choosing m large enough compared to E_0. For I_{close} we use the off-diagonal decay to obtain

$$I_{close} \lesssim_m \|u\|_{\dot{H}^1}^2 \|\phi\|_{\dot{B}^{1,2}_\infty}^2$$

Hence only the large dyadic parts of ϕ have nontrivial contributions. To account for those, we choose a finite set of dyadic indices $K \subset \mathbb{Z}$ outside of which we have

$$\|\phi\|_{B^{1,2}_\infty(K^c)} := \sup_{k \in K^c} 2^k \|\phi_k\|_{L^2} \ll_{E_0} 1. \tag{2.12}$$

Note that the number of indices in K can be bounded by a constant depending only on E_0. Since I_{close} only allows interactions of frequencies at most m apart, it is natural to expand K by m to $K^m := \{k + k' : k \in K, |k'| \le m\}$. Then all unfavorable (i.e., large) interactions occur only for frequencies within K^m, i.e.,

$$I_{close} \le \sum_{k_i : [k_{min}, k_{max}] \cap K = \emptyset} I(k_1, \dots, I_{k_4}) + \sum_{k_i : [k_{min}, k_{max}] \cap K \ne \emptyset} I(k_1, \dots, I_{k_4})$$

$$\lesssim_m \|u\|_{\dot{H}^1}^2 \|\phi\|_{B^{1,2}(K^c)}^2 + \sum_{k_i : [k_{min}, k_{max}] \subseteq K^m} I(k_1, \dots, I_{k_4}).$$

The first term on the last line is small enough thanks to (2.12). The second term can be eliminated altogether by refining the choice of Λ. Precisely, we set $\Lambda(r)$ to be a piecewise smooth function which is constant for $\log_2 r \in K^{2m}$ and equals an appropriate constant multiple of $r^{\sigma+1}$ outside. Then it is easy to check that $I(k_1, \dots, k_4) = 0$ if $k_1, k_2 \in K^m$; hence (2.11) follows. Furthermore, since m and the number of indices in K are bounded by E_0, it follows that $\Lambda(r) \approx_{E_0} r^{\sigma+1}$ as required.

(c) For $\sigma_0 = -1, -\frac{1}{2}$ and $f = f_k$ we claim that

$$\|\phi_j\|_{\dot{H}^{\sigma_0+2}} \lesssim_{E_0} 2^{-\delta|j-k|} \|f_k\|_{\dot{H}^{\sigma_0}}. \tag{2.13}$$

In fact, a similar bound holds for any $-2 < \sigma_0 < 0$ with $\delta > 0$ depending on σ. By linearity, we may fix k, and by scaling (which leaves E_0 invariant), we may assume that $k = 0$. Then the bound (2.13) follows by applying (b) with $-2 < \sigma < \sigma_0$ and $\sigma_0 < \sigma < 0$ to control the solution ϕ in upper and lower Sobolev spaces, which implies that ϕ_j decays in L^2 away from $j = 0$. $\qquad \square$

We now continue the proof of part (2) of Proposition 2.1. From part (a) of the above lemma we obtain the estimate

$$\|A_0\|_{\dot{H}^1} \lesssim \|\phi\|_{\dot{H}^1} \|\partial_t \phi\|_{L^2} \lesssim E_{lin}.$$

Then, using the embedding $\dot{H}^1 \subset L^4$, we directly obtain the estimate (2.3).

(3)). Comparing E with E_{lin} we have

$$E = E_{lin} + E_{lin}^{\frac{1}{2}} O(\|A\phi\|_{L^2} + \|\nabla A_0\|_{L^2}) + O(\|A\phi\|_{L^2}^2 + \|\nabla A_0\|_{L^2}^2)$$

therefore it suffices to establish the bounds

$$\|A\phi\|_{L^2} \lesssim_{E_{lin}} \epsilon^{\frac{1}{4}}, \qquad \|\nabla A_0\|_{L^2} \lesssim_{E_{lin}} \epsilon^{\frac{1}{4}}. \tag{2.14}$$

The first is easily obtained using the standard Littlewood-Paley trichotomy. For high-low interactions we have

$$\|A_j\phi_k\|_{L^2} \lesssim 2^{k-j}\|A_j\|_{L^2}\|\phi_k\|_{ED}, \qquad j > k$$

For low-high interactions we have

$$\|A_j\phi_k\|_{L^2} \lesssim \|A_j\|_{L^8}\|\phi_k\|_{L^{\frac{8}{3}}} \lesssim 2^{\frac{1}{8}(j-k)}\|A_j\|_{\dot{H}^1}\|\phi\|_{\dot{H}^1}^{\frac{3}{4}}\|\phi\|_{ED}^{\frac{1}{4}}, \qquad j < k$$

Finally for high-high interactions we have

$$\|P_j(A_k\phi_k)\|_{L^2} \lesssim 2^{\frac{1}{2}(j-k)}\|A_j\|_{\dot{H}^1}\|\phi\|_{\dot{H}^1}^{\frac{1}{2}}\|\phi\|_{ED}^{\frac{1}{2}}, \qquad j \leq k.$$

In all cases we have favorable off-diagonal decay, so the l^2 dyadic summation for the output is inherited from A. Hence the first bound in (2.14) follows.

For the second bound in (2.14) we use the lemma to reduce it to

$$\|\phi\partial_t\phi\|_{\dot{H}^{-1}} \lesssim_{E_{lin}} \epsilon^{\frac{1}{4}}.$$

The argument for this is similar to the one above, and is left for the reader. $\qquad\square$

3. Space-time function spaces

3.1. **The S^1, N, Z and Y^1 spaces.** We begin our discussion with the function spaces introduced in [13], namely S^1 for the MKG waves (A, ϕ) and N for the inhomogeneous terms in both the $\Box$ and the $\Box_A$ equation. These are spaces of functions defined over all of $\mathbb{R}^{n+1}$, together with the related spaces S and N^*. They are all defined via their dyadic subspaces, with norms

$$\|\phi\|_X^2 = \sum_k \|\phi_k\|_{X_k}^2, \qquad X \in \{S, S^1, N\}$$

We recall the definition of their norms. With minor modifications at high modulations, we follow [13]. For N_k we set

$$N_k = L^1 L^2 + X_1^{0,-\frac{1}{2}}, \tag{3.1}$$

where

$$\|\phi\|_{X_r^{s,b}} := \left(\sum_k \left(\sum_j (2^{sk}2^{bj}\|P_kQ_j\phi\|_{L^2L^2})^r\right)^{\frac{2}{r}}\right)^{\frac{1}{2}}.$$

The N_k norm is the same as in [13].

The S_k space is a strengthened version of N_k^*,

$$X_1^{0,\frac{1}{2}} \subseteq S_k \subseteq L^\infty L^2 \cap X_\infty^{0,\frac{1}{2}} = N_k^*, \tag{3.2}$$

while S_k^1 is defined as

$$\|\phi\|_{S_k^1} = \|\nabla\phi\|_{S_k} + 2^{-\frac{k}{2}}\|\Box\phi\|_{L^2L^2} + 2^{-\frac{4k}{9}}\|\Box\phi\|_{L^{\frac{9}{5}}L^2}. \tag{3.3}$$

Compared to [13] we have loosened the ℓ^1 summability of the $\Box^{-1}L^2L^2$ norm and added the $\Box^{-1}L^{\frac{9}{5}}L^2$ norm above. Both of these modifications are of interest only at high modulations. The exact exponent 9/5 is not really important, for our purposes it only matters that it is less than two and greater than 5/3.

Remark 3.1. In [17], yet another definition of the S^1 norm is employed, namely

$$\|\phi\|_{S_k^1} = \|\nabla\phi\|_{S_k} + 2^{-\frac{k}{2}}\|\Box\phi\|_{L^2 L^2}.$$

Our justification for keeping the same notation S^1 (besides notational simplicity) is that the difference among these three definitions is minor. For a solution to MKG-CG, one can easily pass from one definition to another using the high modulation bounds in Propositions 4.4 and 4.10. In particular, in every theorem stated in the introduction, statements with respect to one of these definitions of S^1 easily implies those with respect to others.

We now recall the definition of the space S_k from [13]. The space S_k scales like free waves with $L^2 \times \dot{H}^{-1}$ initial data, and is defined by

$$\|\phi\|_{S_k}^2 = \|\phi\|_{S_k^{str}}^2 + \|\phi\|_{S_k^{ang}}^2 + \|\phi\|_{X_\infty^{0,\frac{1}{2}}}^2 ,$$

where:

$$\|\phi\|_{S_k^{str}} = \sup_{2\leq q,r\leq\infty,\ \frac{1}{q}+\frac{3/2}{r}\leq\frac{3}{4}} 2^{(\frac{1}{q}+\frac{4}{r}-2)k}\|(\phi, 2^{-k}\partial_t\phi)\|_{L^q L^r} , \qquad \|\phi\|_{S_k^{ang}} = \sup_{l<0}\|\phi\|_{S_{k,k+2l}^{ang}} ,$$

$$\|\phi\|_{S_{k,j}^{ang}}^2 = \sum_\omega \|P_l^\omega Q_{<k+2l}\phi\|_{S_k^\omega(l)}^2 \qquad \text{with } l = \lceil\frac{j-k}{2}\rceil.$$

(3.4)

The S_k^{str} norm controls all admissible Strichartz norms on $\mathbb{R}^{1+4}$. The ω-sum in the definition of $S_{k,j}^{ang}$ is over a covering of $\mathbb{S}^3$ by caps ω of diameter 2^l with uniformly finite overlaps, and the symbols of P_l^ω form a smooth partition of unity associated to this covering. The angular sector norm $S_k^\omega(l)$ combines the null frame space as in wave maps [25, 31] with additional square-summed norms over smaller radially directed blocks $\mathcal{C}_{k'}(l')$ of dimensions $2^{k'} \times (2^{k'+l'})^3$. We first define

$$\|\phi\|_{PW_\omega^\pm(l)} = \inf_{\phi=\int\phi^{\omega'}} \int_{|\omega-\omega'|\leq 2^l} \|\phi^{\omega'}\|_{L^2_{\pm\omega'}(L^\infty_{(\pm\omega')^\perp})} d\omega' ,$$

$$\|\phi\|_{NE} = \sup_\omega \|\nabla_\omega\phi\|_{L^\infty_\omega(L^2_{\omega^\perp})} ,$$

where the norms are with respect to $\ell_\omega^\pm = t \pm \omega\cdot x$ and the transverse variable in the $(\ell_\omega^\pm)^\perp$ hyperplane (i.e., constant $\ell_\omega^\pm$ hyperplanes). Moreover, ∇_ω denotes tangential derivatives on the $(\ell_\omega^+)^\perp$ hyperplane. As in [13], we set:

$$\|\phi\|_{S_k^\omega(l)}^2 = \|\phi\|_{S_k^{str}}^2 + 2^{-2k}\|\phi\|_{NE}^2 + 2^{-3k}\sum_\pm \|Q^\pm\phi\|_{PW_\omega^\mp(l)}^2$$

$$+ \sup_{\substack{k'\leq k,l'\leq 0 \\ k+2l\leq k'+l'\leq k+l}} \sum_{\mathcal{C}_{k'}(l')} \Big(\|P_{\mathcal{C}_{k'}(l')}\phi\|_{S_k^{str}}^2 + 2^{-2k}\|P_{\mathcal{C}_{k'}(l')}\phi\|_{NE}^2$$

$$+ 2^{-2k'-k}\|P_{\mathcal{C}_{k'}(l')}\phi\|_{L^2(L^\infty)}^2 + 2^{-3(k'+l')}\sum_\pm \|Q^\pm P_{\mathcal{C}_{k'}(l')}\phi\|_{PW_\omega^\mp(l)}^2 \Big) , \quad (3.5)$$

where the $\mathcal{C}_{k'}(l')$ sum runs over a covering of $\mathbb{R}^4$ by the blocks $\mathcal{C}_{k'}(l')$ with uniformly finite overlaps, and the symbols of $P_{\mathcal{C}_{k'}(l')}$ form an associated partition of unity. We also define the

14

smaller space $S_k^\sharp \subset S_k$ (see the bound (3.7) below) by

$$\|u\|_{S_k^\sharp} = \|\Box u\|_{N_k} + \|\nabla u\|_{L^\infty L^2}.$$

On occasion we need to separate the two characteristic cones $\{\tau = \pm|\xi|\}$. Thus we define the spaces $N_{k,\pm}$, $S_{k,\pm}^\sharp$ and $N_{k,\pm}^*$ in an obvious fashion, so that

$$N_k = N_{k,+} \cap N_{k,-}, \quad S_k^\sharp = S_{k,+}^\sharp + S_{k,-}^\sharp, \quad N_k^* = N_{k,+}^* + N_{k,-}^* .$$

Next we describe an auxiliary space of the type $L^1(L^\infty)$ which will be useful for decomposing the nonlinearity:

$$\|\phi\|_Z = \sum_k \|P_k\phi\|_{Z_k}, \quad \|\phi\|_{Z_k}^2 = \sup_{l<C} \sum_\omega 2^l \|P_l^\omega Q_{k+2l}\phi\|_{L^1(L^\infty)}^2 .$$

Note that as defined this space already scales like $\dot{H}^1$ free waves. In addition, note the following useful embedding which is a direct consequence of Bernstein's inequality:

$$\Box^{-1}L^1(L^2) \subseteq Z . \tag{3.6}$$

Finally, the function space for A_0 is simple to describe, since the A_0 equation is elliptic:

$$\|A_0\|_{Y^1}^2 = \|\nabla A_0\|_{L^\infty L^2}^2 + \|\nabla A_0\|_{L^2 \dot{H}^{\frac{1}{2}}}^2 ,$$

where we recall that ∇ denotes the full space-time gradient.

Let E denote the linear energy space, i.e.,

$$E = \dot{H}^1 \times L^2.$$

One of the results in [13] asserts that we have linear solvability for the d'Alembertian in our setting.

Proposition 3.2. *We have the linear estimates*

$$\|\nabla\phi\|_S \lesssim \|\phi[0]\|_E + \|\Box\phi\|_N , \tag{3.7}$$

$$\|\phi\|_{S^1} \lesssim \|\phi[0]\|_E + \|\Box\phi\|_{N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}}} . \tag{3.8}$$

Here (3.7) is the embedding $S^\sharp \subset S$, whereas (3.8) follows immediately from (3.7).

3.2. **Interval localization.** So far, we have described the global setting in [13]. However, in this article we work on compact time intervals, therefore we also need suitable interval localized function spaces. This is not straightforward, since our function spaces are defined using modulation localizations, which are nonlocal in time. To start with, we take the easy way out and define

$$\|\phi\|_{S^1[I]} = \inf_{\phi=\tilde{\phi}_{|I}} \|\tilde{\phi}\|_{S^1}, \qquad \|f\|_{N[I]} = \inf_{f=\tilde{f}_{|I}} \|\tilde{f}\|_N \tag{3.9}$$

However, the next result allows us to simplify somewhat these definitions:

Proposition 3.3. *(1) Consider a time interval I, and its characteristic function χ_I. Then we have the bounds*

$$\|\chi_I\phi\|_S \lesssim \|\phi\|_S, \qquad \|\chi_I f\|_N \lesssim \|f\|_N, \tag{3.10}$$

The latter norm is also continuous as a function of I. We also have the linear estimates

$$\|\nabla\phi\|_{S[I]} \lesssim \|\phi[0]\|_E + \|\Box\phi\|_{N[I]}, \tag{3.11}$$

$$\|\phi\|_{S^1[I]} \lesssim \|\phi[0]\|_E + \|\Box\phi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})[I]}. \tag{3.12}$$

(2) Consider any partition $I = \bigcup I_k$. Then the N norm is interval divisible, i.e.

$$\sum_k \|f\|^2_{N[I_k]} \lesssim \|f\|^2_{N[I]} \tag{3.13}$$

and the S and S^1 norms are interval square summable, i.e.

$$\|\phi\|^2_{S[I]} \lesssim \sum_k \|\phi\|^2_{S[I_k]}, \qquad \|\phi\|^2_{S^1[I]} \lesssim \sum_k \|\phi\|^2_{S^1[I_k]} \tag{3.14}$$

We remark that a consequence of part (1) is that, up to equivalent norms, we can replace the arbitrary extensions in (3.9) by the zero extension in the N case, respectively by homogeneous waves with $(\phi, \partial_t\phi)$ as the data at each endpoint outside I in the S^1 case.

Proof. (1)). It suffices to prove the desired bounds for frequency dyadic pieces of ϕ and f. In the N case it also suffices to work with the space $L^1L^2 + X_1^{0,-\frac{1}{2}}$. But in this case this is exactly the proof of (158) in [23], where just steps 1 and 2 are needed.

By duality, we have the same bound for $L^\infty L^2 \cap X_\infty^{0,\frac{1}{2}}$, which is a part of the S norm. We now consider the remaining parts of the S norm. The only difficulty is with the S_k^{ang} norm, due to the modulation localization. Fixing a modulation scale $j = k + 2l$, we consider either $Q_{<j}^+(\chi_I\phi)$ or $Q_{<j}^-(\chi_I\phi)$. There are two cases to consider:

(i) Short intervals, $|I| < 2^{-j}$. Then

$$\|Q_{<j}^+(\chi_I\phi)\|_{S_{k,j}^{ang}} \lesssim \|Q_{<j}^+(\chi_I\phi)\|_{X_1^{0,\frac{1}{2}}} \lesssim 2^{\frac{j}{2}}\|\chi_I\phi\|_{L^2L^2} \lesssim \|\phi\|_{L^\infty L^2}$$

(ii) Long intervals, $|I| > 2^{-j}$. Then we write

$$Q_{<j}^+(\chi_I\phi) = Q_{<j-30}\chi_I Q_{<j-2}^+\phi + Q_{<j}^+(Q_{>j-30}\chi_I Q_{<j-2}^+\phi) + Q_{<j}^+(\chi_I Q_{>j-2}^+\phi) \tag{3.15}$$

For the first term we use the bound

$$\|Q_{<j-30}\chi_I Q_{<j-2}^+\phi\|_{S_{k,j}^{ang}} \lesssim \|Q_{<j-30}\chi_I\|_{L^\infty}\|\phi\|_{S_k} \lesssim \|\phi\|_{S_k}$$

which was proved in [17, Lemma 7.1] (technically speaking, [17, Lemma 7.1] is stated for χ_I which decays in space, but we may simply approximate χ_I by smooth compactly supported functions). The other two terms in (3.15) are estimated in L^2L^2 as in (i): For the second term, we use

$$\|Q_{<j}^+(Q_{>j-30}\chi_I Q_{<j-2}^+\phi)\|_{S_{k,j}^{ang}} \lesssim 2^{\frac{j}{2}}\|Q_{>j-30}\chi_I Q_{<j-2}^+\phi\|_{L^2L^2} \lesssim 2^{\frac{j}{2}}\|Q_{>j-30}\chi_I\|_{L^2}\|Q_{<j-2}^+\phi\|_{L^\infty L^2}$$
$$\lesssim \|\phi\|_{L^\infty L^2}.$$

In the last inequality, we used the bound

$$\|Q_j\chi_I\|_{L^2} \lesssim 2^{-\frac{j}{2}},$$

which follows from Plancherel in t and the fact that the Fourier transform of the χ_I is a suitable rescaling and modulation of $\sin \tau / \tau$. Finally, the third term in (3.15) is treated as follows:

$$\|Q_{<j}^+(\chi_I Q_{>j-2}^+ \phi)\|_{S_{k,j}^{ang}} \lesssim 2^{\frac{j}{2}}\|Q_{>j-2}^+\phi\|_{L^2 L^2} \lesssim \|\phi\|_{S_k}.$$

(2)). The N bound (3.13) is exactly as in Proposition 5.4 in (159) in [23]. The S^1 bound (3.14) reduces easily to the corresponding S bound. The bound (3.14) for the N^* part of the S norm follows by duality from (3.13). Of the remaining components of the S part we have the same difficulty as in part (1), namely with the modulation localizations occurring in the S_k^{ang} norms. The solution is also the same as in part (1); precisely that for each modulation scale j we split the intervals into short and long, and estimate the two contributions as above:

(i) Short intervals, $|I_m| < 2^{-j}$. Then the modulation localization operator $Q_{<j}^+$ can cause significant overlapping of outputs coming from inputs in different intervals I_m. Hence our strategy is to harmlessly discard $Q_{<j}^+$ as follows:

$$\|Q_{<j}^+(\sum_m \chi_{I_m}\phi)\|_{S_{k,j}^{ang}}^2 \lesssim 2^{\frac{j}{2}}\|\sum_m \chi_{I_m}\phi\|_{L^2}^2 \lesssim \sum_m \|\chi_{I_m}\phi\|_{L^\infty L^2}^2$$

(ii) Long intervals, $|I_m| > 2^{-j}$. Then for each m we use a partition of unity adapted to I_m to write

$$1 = \tilde{\chi}_{I_m} + \sum_{l>0} \tilde{\chi}_{I_m}^l$$

where $\tilde{\chi}_{I_m}$ is a smooth cutoff selecting a 2^{-j} neighborhood of I_m, while $\chi_{I_m}^l$ select the region at distance 2^{-j+l} from I_m. Correspondingly, we write

$$Q_{<j}^+ \sum_m \chi_{I_m}\phi = \sum_m \tilde{\chi}_{I_m} Q_{<j}^+(\chi_{I_m}\phi) + \sum_{l>0}\sum_m \tilde{\chi}_{I_m}^l Q_{<j}^+(\chi_{I_m}\phi)$$

Now we estimate each of the sums above. For the first one we use the fact that the bump functions $\tilde{\chi}_{I_m}$ have finite overlapping to write

$$\|\sum_m \tilde{\chi}_{I_m} Q_{<j}^+(\chi_{I_m}\phi)\|_{S_{k,j}^{ang}}^2 \lesssim \sum_m \|\tilde{\chi}_{I_m} Q_{<j}^+(\chi_{I_m}\phi)\|_{S_{k,j}^{ang}}^2 \lesssim \sum_m \|Q_{<j}^+(\chi_{I_m}\phi)\|_{S_{k,j}^{ang}}^2$$

which suffices thanks to part (1). On the other hand, in the second sum, for each l we have at most 2^l overlapping bump functions. So we obtain

$$\|\sum_m \tilde{\chi}_{I_m}^l Q_{<j}^+(\chi_{I_m}\phi)\|_{S_{k,j}^{ang}}^2 \lesssim 2^l \sum_m \|\tilde{\chi}_{I_m}^l Q_{<j}^+(\chi_{I_m}\phi)\|_{S_{k,j}^{ang}}^2 \lesssim 2^l 2^j \sum_m \|\tilde{\chi}_{I_m}^l Q_{<j}^+(\chi_{I_m}\phi)\|_{L^2}^2$$

$$\lesssim 2^l \sum_m 2^{-2Nl}\|\chi_{I_m}\phi\|_{L^\infty L^2}^2$$

which again suffices. Here, at the last stage, we have used the fact that the operator $\tilde{\chi}_{I_m}^l Q_{<j}^+ \chi_{I_m}$ has a $2^{-\frac{j}{2}}2^{-Nl}$ norm from $L^\infty L^2$ to $L^2 L^2$, which is due to the separation of supports of the two cutoff functions. $\qquad\square$

Last but not least, we consider the effect of extension on some of our Strichartz or energy dispersed norms; the role of these norms in our work will be explained in Section 3.3. For an interval I we denote by χ_I^k a generalized cutoff function, which is adapted to the 2^k frequency scale:

$$\chi_I^k(t) = (1 + 2^k \text{dist}(t, I))^{-N}.$$

17

For a function ϕ_I in I we denote by ϕ_I^{ext} its extension as homogeneous waves. Then we have:

Proposition 3.4. *Assume that $|I| \geq 2^{-k}$. Then the following estimates hold for ϕ_I localized at frequency 2^k:*

$$\|\chi_I^k(\phi_I^{ext}, 2^{-k}\partial_t\phi_I^{ext})\|_{L^pL^q} \lesssim \|\phi_I\|_{L^pL^q[I]} + 2^{(\frac{1}{2}-\frac{1}{p}-\frac{4}{q})k}\|\Box\phi_I\|_{L^2L^2[I]}, \qquad (3.16)$$

$$\chi_I^k(t)\|\phi_I^{ext}(t)\|_{ED} \lesssim \|\phi_I\|_{ED[I]}, \qquad (3.17)$$

where (p,q) is any pair of admissible Strichartz exponents on $\mathbb{R}^{1+4}$.

Proof. By rescaling, we can take $k = 0$. It suffices to consider the case when I has the minimal length, i.e., $|I| = 1$; the general case is then easily recovered by applying the same proof to unit-length intervals at each end of I. By translation invariance, we may take $I = [0,1]$.

We first consider the bound (3.16). It suffices to show that for any Strichartz norm L^pL^q and ϕ_I localized at frequency 1 we have

$$\|\chi_I^0(\phi_I^{ext}, \partial_t\phi_I^{ext})\|_{L^pL^q} \lesssim \|\phi_I\|_{L^pL^q[I]} + \|\Box\phi_I\|_{L^2[I]}. \qquad (3.18)$$

By the inhomogeneous Strichartz estimates, this reduces to the case when $\Box\phi = 0$.

We prove this in two steps. First we notice that the Cauchy data at time 0 satisfies

$$\|\phi_I[0]\|_{L^q} \lesssim \|\phi_I\|_{L^pL^q[I]} \qquad (3.19)$$

Equivalently, we have to show that for functions $f_\pm$ localized at frequency 1 we have

$$\|f_\pm\|_{L^q} \lesssim \|e^{it|D|}f_+ + e^{-it|D|}f_-\|_{L^p([0,1];L^q)}$$

We may easily find $t_1 \in [0, 1/3]$ and $t_2 \in [2/3, 1]$ such that the L^q norm of $e^{it_j|D|}f_+ + e^{-it_j|D|}f_-$ is bounded by the right hand side. Then the desired conclusion follows from the linear independence of the symbols $(e^{it_j|D|}, e^{-it_j|D|})$ for $j = 1, 2$.

Secondly, we have the bound

$$\|(\phi, \partial_t\phi)(t)\|_{L^q} \lesssim (1+t^8)^{\frac{1}{2}-\frac{1}{p}}\|(\phi, \partial_t\phi)(0)\|_{L^q}, \qquad 2 \leq q \leq \infty \qquad (3.20)$$

This is trivial for $q = 2$. In the case $q = \infty$, for a fixed $(t, x) \in \mathbb{R}^{1+4}$, by finite speed of propagation we may truncate the initial data for $\phi[0]$ outside a ball of radius $C(1+t)$ without changing $(\phi, \partial_t\phi)(t, x)$. Then the desired bound for $|(\phi, \partial_t\phi)(t, x)|$ follows from Hölder's inequality (to control the energy with the L^∞ norm), the energy estimate and Bernstein's inequality.

Putting together (3.19) and (3.20) we obtain (3.18). Finally, the bound (3.17) follows from (3.20) with $q = \infty$.

$\square$

3.3. **Smallness: energy dispersion and divisibility.** Since our goal is to work with large data MKG-CG solutions, it is crucial to have at our disposal sufficient tools to gain smallness in appropriate settings. One such source of smallness in this article is the *energy dispersion*, which is used as an a-priori bound. Another venue for gaining smallness is to partition the time in finitely many subintervals, on each of which the norm is small. A space-time norm for which this procedure works is said to be *divisible*. In this short subsection, we provide heuristic explanation of both concepts and their use in our context.

We start by discussing the use of energy dispersion as a source of smallness. For application, it is useful to quantify the smallness of the energy dispersion norm $\|\cdot\|_{ED}$ in comparison with the norm $\|\cdot\|_{S^1}$, which is stronger and have the same scaling. We therefore define:

Definition 3.5. For any interval $I \subseteq \mathbb{R}$ and $\varepsilon > 0$, we say that $\phi \in S^1[I]$ is ε-energy dispersed (with respect to the S^1 norm) if

$$\|\phi\|_{ED(I)} \leq \varepsilon \|\phi\|_{S^1[I]}. \tag{3.21}$$

Observe that the S^1 norm is stronger than the ED norm by Bernstein's inequality, i.e.,

$$\|\phi\|_{ED(I)} \lesssim \|\nabla\phi\|_{L^\infty L^2[I]} \lesssim \|\phi\|_{S^1[I]}.$$

Hence the dimensionless quantity $\varepsilon > 0$ can be thought of as measuring the improvement relative to Bernstein's inequality.

Roughly speaking, small energy dispersion improves balanced frequency bilinear interactions. In [23], this improvement was obtained by interpolating the ED norm with the Wolff-Tao bilinear estimate in $L^p L^p$ with $p < 2$. In the present setting, as we have stronger dispersion due to higher dimensionality, we can achieve the same end by simply interpolating the ED norm with Strichartz norms. Indeed, the following linear lemma covers essentially all of our usage of small energy dispersion:

Lemma 3.6. Let $\phi \in S^1[I]$ be ε-energy dispersed. Then for any $k \in \mathbb{Z}$ and any non-sharp pair of Strichartz exponents $(p, q) \in [2, \infty]$ (i.e., $\frac{2}{p} + \frac{3}{q} < \frac{3}{2}$ and $p \neq 2$), we have

$$\sup_k \|(P_k\phi, 2^{-k}P_k\partial_t\phi)\|_{L^p L^q[I]} \leq \varepsilon^{\delta_1} 2^{(2-\frac{1}{p}-\frac{4}{q})k}\|\phi\|_{S^1[I]},$$

where $\delta_1 = \delta_1(p, q) > 0$.

As discussed, this lemma easily follows by interpolating the ED norm with the S_k^{str} component of the S^1 norm, which is possible thanks to the non-sharpness of (q, r). We will often combine this lemma with Hölder's inequality to gain smallness for multilinear estimates.

We now turn to the use of divisibility in our work. The bound (3.13) shows that the N norm is divisible. However, the S^1 norm is not divisible, and this is a source of trouble; see however the result in Theorem 5.1(5). Our workaround is to introduce a weaker norm, denoted DS^1, which collects a subset of the components of the S^1 norm which are divisible. This is defined as follows:

$$\|\phi\|_{DS^1[I]} = \|(|D|^{\frac{1}{6}}\phi, |D|^{-\frac{5}{6}}\partial_t\phi)\|_{L^2 L^6[I]} + \|(|D|^{\frac{5}{6}}\phi, |D|^{-\frac{1}{6}}\partial_t\phi)\|_{L^{10}L^{\frac{30}{13}}[I]} + \||D|^{-\frac{1}{2}}\Box u\|_{L^2 L^2[I]}. \tag{3.22}$$

Precisely, we may include here any divisible Strichartz norm as long as we stay away from the $L^\infty L^2$ endpoint (i.e., the energy). To gain divisibility for A_0, we use the norm

$$\|\nabla A_0\|_{L^2 \dot{H}^{\frac{1}{2}}[I]} = \|(A_0, \partial_t A_0)\|_{L^2 \dot{H}^{\frac{3}{2}} \times L^2 \dot{H}^{\frac{1}{2}}[I]},$$

which is a divisible component of the Y^1 norm.

By Proposition 3.4, we see that the homogeneous wave extension ϕ_I^{ext} of a function ϕ_I in I obeys the bound

$$\|\chi_I^k \phi_I^{ext}\|_{DS^1} \lesssim \|\phi_I\|_{DS^1[I]} \tag{3.23}$$

when ϕ_I is localized at frequency 2^k and $|I| \geq 2^{-k}$.

Our strategy will be to use as much as possible the divisible norms (such as DS^1 or $L^2\dot{H}^{\frac{3}{2}}$) in our bilinear and multilinear estimates, and try to prove smallness for the remainder.

4. The decomposition of the nonlinearity

Recalling the definition of the currents $J_\alpha = -\text{Im}(\phi\overline{D_\alpha\phi})$ we write the MKG-CG system again here as:

$$\Box A_i = \mathcal{P}_i J_x, \qquad (4.1\text{a})$$

$$\Box_A \phi = 0 \qquad (4.1\text{b})$$

The second equation also includes A_0 and $\partial_t A_0$, which are obtained from the elliptic equations

$$\Delta A_0 = J_0, \qquad \Delta\partial_t A_0 = \nabla^i J_i. \qquad (4.2)$$

We now discuss the bounds for each of the components of MKG equation. For the purpose of this section, all analysis is done in a fixed time interval $I = [0, T]$.

We remark that for the most part, bilinear and trilinear estimates for the nonlinearities were already proved in [13] in the context of the small data problem. Our goal here is to understand when and how we can regain smallness in the study of the large data. As discussed in Section 3.3, there are two such sources of smallness:

a) Arising from norm divisibility for either A or ϕ, where a large but divisible norm is made small by selecting a suitable time interval partition. Here we seek to use the DS^1 part of the S^1 norm to measure the bulk of the nonlinearities.

b) Arising from small energy dispersion for ϕ. This is often considered coupled with the additional high modulation bound

$$\|\Box\phi\|_{L^2\dot{H}^{-\frac{1}{2}}} \leq \varepsilon^{\delta_1}\|\phi\|_{S^1} \qquad (4.3)$$

which for MKG-CG solutions is an easy consequence of the ε-energy dispersion; see Theorem 5.2 below.

Two easy ways to gain the two types of estimates above is by using suitable Strichartz estimates. Precisely, for divisibility we need L^pL^q norms with $p < \infty$. On the other hand for energy dispersion we need $p > 2$, as well as non-sharp pairs of exponents (p, q), so that Lemma 3.6 is applicable. Often we can fulfill both at once, and prove the two types of estimates simultaneously.

4.1. The terms A_x.

We decompose A_i into a free and a nonlinear component,

$$A_i = A_i^{free} + A_i^{nl}$$

where

$$\Box A_i^{free} = 0, \qquad A_i^{free}[0] = A_i[0]$$

and

$$\Box A_i^{nl} = \mathcal{P}_i J_x, \qquad A_i^{nl}[0] = 0$$

Given the expression of the currents $J_\alpha = -\text{Im}(\phi\partial_\alpha\bar{\phi}) + A_\alpha|\phi|^2$, we will think of A_x^{nl} given by the above equation as a multilinear expression in ϕ and A, i.e., $A_i^{nl} = \mathbf{A}_i(\phi, \phi, A)$. We can also extend this to a symmetric quadratic form in the first two variables, $\mathbf{A}_i(\phi_1, \phi_2, A)$. We also split it into a quadratic and a cubic part,

$$\mathbf{A}_i(\phi_1, \phi_2, A) = \mathbf{A}_i^2(\phi_1, \phi_2) + \mathbf{A}_x^3(\phi_1, \phi_2, A)$$

$$= -\frac{1}{2}\Box^{-1}\mathcal{P}_i(\phi_1\partial_x\overline{\phi_2} + \overline{\partial_x\phi_1}\phi_2) + \frac{1}{2}\Box^{-1}\mathcal{P}_i(\phi_1\overline{\phi_2}A_x + \overline{\phi_1}\phi_2 A_x),$$

20

where $\Box^{-1}f$ denotes the solution to the inhomogeneous wave equation $\Box u = f$ with $u[0] = 0$. The N bounds we need for $\Box A_x$ are as follows:

Proposition 4.1. *Let ϕ_1, ϕ_2, A be test functions defined on a time interval I containing 0.*
(1) For all admissible frequency envelopes c, d, e we have

$$\|\Box \mathbf{A}_x^2(\phi_1, \phi_2)\|_{N_{cd}[I]} + \|\nabla \mathbf{A}_x^2(\phi_1, \phi_2)\|_{S_{cd}[I]} \lesssim \|\phi_1\|_{S_c^1[I]}\|\phi_2\|_{S_d^1[I]} \tag{4.4}$$

respectively

$$\|\Box \mathbf{A}_x^3(\phi_1, \phi_2, A)\|_{L^1 L^2_{cde}[I]} + \|\nabla \mathbf{A}_x^3(\phi_1, \phi_2, A)\|_{S_{cde}[I]} \lesssim \|\phi_1\|_{DS_c^1[I]}\|\phi_2\|_{DS_d^1[I]}\|A\|_{DS_e^1[I]} \tag{4.5}$$

(2) Further, for each $m > 0$ there is a decomposition

$$\mathbf{A}_x^2(\phi_1, \phi_2) = \mathbf{A}_{x,small}^2(\phi_1, \phi_2) + \mathbf{A}_{x,large}^2(\phi_1, \phi_2)$$

so that we have

$$\|\Box \mathbf{A}_{x,small}^2(\phi_1, \phi_2)\|_{N_{cd}[I]} + \|\nabla \mathbf{A}_{x,small}^2(\phi_1, \phi_2)\|_{S_{cd}[I]} \lesssim 2^{-cm}\|\phi_1\|_{S_c^1[I]}\|\phi_2\|_{S_d^1[I]} \tag{4.6}$$

respectively

$$\|\Box \mathbf{A}_{x,large}^2(\phi_1, \phi_2)\|_{N_{cd}[I]} + \|\nabla \mathbf{A}_{x,large}^2(\phi_1, \phi_2)\|_{S_{cd}[I]} \lesssim 2^{Cm}\|\phi_1\|_{DS_c^1[I]}\|\phi_2\|_{DS_d^1[I]} \tag{4.7}$$

(3) In addition, if ϕ_1 is ε-energy dispersed and satisfies (4.3) then

$$\|\Box \mathbf{A}_x^2(\phi_1, \phi_2)\|_{N_c[I]} + \|\nabla \mathbf{A}_x^2(\phi_1, \phi_2)\|_{S_c[I]} \lesssim \varepsilon^{\delta_1}\|\phi_1\|_{S^1[I]}\|\phi_2\|_{S_c^1[I]} \tag{4.8}$$

and

$$\|\Box \mathbf{A}_x^3(\phi_1, \phi_2, A)\|_{N_{de}[I]} + \|\nabla \mathbf{A}_x^3(\phi_1, \phi_2, A)\|_{S_{de}[I]} \lesssim \varepsilon^{\delta_1}\|\phi_1\|_{S^1[I]}\|\phi_2\|_{S_d^1[I]}\|A\|_{S_e^1[I]} \tag{4.9}$$

Remark 4.2. The $\|\nabla(\cdot)\|_S$ norm bounds follow immediately from the control of $\|\Box(\cdot)\|_N$ thanks to (3.11) and the fact that the initial data vanish for $\mathbf{A}_x^2, \mathbf{A}_x^3$. As we see from (3.3), this norm is slightly weaker than the main 'solution norm' S^1 for high modulations; nevertheless these bounds will prove useful in the proof of the multilinear estimates in Section 8.

Remark 4.3. Given a test function ϕ_1 on I which is ε-energy dispersed and obeys (4.3), the bounds (4.8) and (4.9) still hold with the same right hand sides if we replace ϕ_1 by its frequency projection (e.g., $P_{<k_*}\phi_1$ or $P_{\geq k_*}\phi_1$) on the left hand side. This fact will be evident from the proof. The same remark applies to all the other estimates in this section that rely on ε-energy dispersion.

This proposition is proved in Section 7.

We also state high modulation bounds for $\Box A_x$, which do not require a null structure nor an extra decomposition:

Proposition 4.4. *Let ϕ_1, ϕ_2, A be test functions defined on a time interval I containing 0. For all admissible frequency envelopes c, d, e we have*

$$\|\Box \mathbf{A}_x^2(\phi_1, \phi_2)\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{cd}[I]} \lesssim \|\phi_1\|_{DS_c^1[I]}\|\phi_2\|_{DS_d^1[I]} \tag{4.10}$$

$$\|\Box \mathbf{A}_x^3(\phi_1, \phi_2, A)\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{de}[I]} \lesssim \|\phi_1\|_{DS_c^1[I]}\|\phi_2\|_{DS_d^1[I]}\|A\|_{DS_e^1[I]} \tag{4.11}$$

In addition, if ϕ_1 is ε-energy dispersed, then

$$\|\Box \mathbf{A}_x^2(\phi_1, \phi_2)\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_d[I]} \lesssim \varepsilon^{\delta_1}\|\phi_1\|_{S^1[I]}\|\phi_2\|_{S_d^1[I]} \tag{4.12}$$

$$\|\Box \mathbf{A}_x^3(\phi_1, \phi_2, A)\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{de}[I]} \lesssim \varepsilon^{\delta_1}\|\phi_1\|_{S^1[I]}\|\phi_2\|_{S_d^1[I]}\|A\|_{S_e^1[I]} \tag{4.13}$$

Proof. The whole proposition is a simple consequence of Bernstein's inequality, non-sharp Strichartz estimates and Lemma 3.6. For instance, the $L^2 \dot{H}^{-\frac{1}{2}}$ norm of the Littlewood-Paley piece $P_k \square \mathbf{A}_x^2(\phi_{k_1}, \phi_{k_2})$ (where ϕ_{k_i} is a shorthand for $P_{k_i} \phi_i$) is bounded as follows:

$$\|P_k \square \mathbf{A}_x^2(\phi_{k_1}, \phi_{k_2})\|_{L^2 \dot{H}^{-\frac{1}{2}}[I]} \lesssim \|P_k(\phi_{k_1} \partial_x \overline{\phi}_{k_2})\|_{L^2 \dot{H}^{-\frac{1}{2}}[I]} + \|P_k(\partial_x \overline{\phi}_{k_1} \phi_{k_2})\|_{L^2 \dot{H}^{-\frac{1}{2}}[I]}$$
$$\lesssim 2^{-\delta \max\{|k-k_i|\}} \prod_{i=1,2} \||D|^{\frac{9}{16}} \phi_{k_i}\|_{L^4 L^{\frac{64}{21}}[I]}$$

where the off-diagonal gain arises from applying Bernstein's inequality to the lowest frequency. As $(4, \frac{64}{21})$ is a non-sharp Strichartz estimate, this bound suffices for both (4.10) and (4.12) (via Lemma 3.6). Similarly, for $P_k \square \mathbf{A}_x^3(\phi_{k_1}, \phi_{k_2}, A_{k_3})$ (where $A_{k_3} = P_{k_3} A$), we have

$$\|P_k \square \mathbf{A}_x^3(\phi_{k_1}, \phi_{k_2}, A_{k_3})\|_{L^2 \dot{H}^{-\frac{1}{2}}[I]} \lesssim 2^{-\delta \max\{|k-k_i|\}} \left(\prod_{i=1,2} \||D|^{\frac{1}{6}} \phi_{k_i}\|_{L^6 L^4[I]} \right) \||D|^{\frac{1}{6}} A_{k_3}\|_{L^6 L^4[I]}$$

The argument for the $L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}$ norm is analogous. $\qquad \square$

4.2. **The term A_0.** Here we consider bounds for both A_0 and its time derivative, which are given by (4.2). The first equation can be written in a more explicit form as

$$(-\Delta + |\phi|^2) A_0 = \mathrm{Im}(\phi \partial_t \overline{\phi}) \tag{4.14}$$

which was analyzed earlier in Lemma 2.2. As an immediate corollary of Lemma 2.2 we obtain the following estimate for A_0:

$$\|A_0\|_{L^\infty \dot{H}^1_{c^2}[I]} + \|A_0\|_{L^2 \dot{H}^{\frac{3}{2}}_{c^2}[I]} \lesssim \|\phi\|_{L^\infty \dot{H}^1[I]} \|\phi\|^2_{S^1_c[I]} . \tag{4.15}$$

Given this bound, we return to the equations (4.2) and view them simply as Laplace equations, whose solutions are quadratic expressions in ϕ,

$$A_0 = \mathbf{A}_0(\phi, \phi, A) \qquad \partial_t A_0 = \partial_t \mathbf{A}_0(\phi, \phi, A)$$

which are given by

$$\mathbf{A}_0(\phi, \phi, A_0) = \mathbf{A}_0^2(\phi, \phi) + \mathbf{A}_0^3(\phi, \phi, A_0)$$
$$= -\Delta^{-1} \mathrm{Im}(\phi \partial_t \overline{\phi}) + \Delta^{-1}(\phi \overline{\phi} A_0),$$
$$\partial_0 \mathbf{A}_0(\phi, \phi, A_x) = \partial_0 \mathbf{A}_0^2(\phi, \phi) + \partial_0 \mathbf{A}_0^3(\phi, \phi, A_x)$$
$$= -\Delta^{-1} \partial^j \mathrm{Im}(\phi \partial_j \overline{\phi}) + \Delta^{-1} \partial^j (\phi \overline{\phi} A_j).$$

We also extend these to symmetric quadratic forms in the first two variables ϕ_1, ϕ_2. Our estimates for A_0 and $\partial_0 A_0$ are as follows:

Proposition 4.5. *Let ϕ_1, ϕ_2, A be test functions defined on a time interval I. Let c, d, e be admissible frequency envelopes.*

(1) For any exponent $2 \le p \le \infty$, we have

$$\|\mathbf{A}_0^2(\phi_1, \phi_2)\|_{L^p \dot{H}^{1+\frac{1}{p}}_{cd}[I]} \lesssim \|\phi_1\|_{DS^1_c[I]} \|\phi_2\|_{DS^1_d[I]}$$
$$\|\mathbf{A}_0^3(\phi_1, \phi_2, A_0)\|_{L^p \dot{H}^{1+\frac{1}{p}}_{cde}[I]} \lesssim \|\phi_1\|_{DS^1_c[I]} \|\phi_2\|_{DS^1_d[I]} \|A_0\|_{L^p \dot{H}^{1+\frac{1}{p}}_e[I]}, \tag{4.16}$$

22

$$\|\partial_t \mathbf{A}_0^2(\phi_1, \phi_2)\|_{L^p \dot{H}_{cd}^{\frac{1}{p}}[I]} \lesssim \|\phi_1\|_{DS_c^1[I]} \|\phi_2\|_{DS_c^1[I]}$$

$$\|\partial_t \mathbf{A}_0^3(\phi_1, \phi_2, A_x)\|_{L^p \dot{H}_{cde}^{\frac{1}{p}}[I]} \lesssim \|\phi_1\|_{DS_c^1[I]} \|\phi_2\|_{DS_d^1[I]} \|A_x\|_{DS_e^1[I]} \; . \tag{4.17}$$

(2) In addition, if ϕ_1 is ε-energy dispersed then

$$\|\mathbf{A}_0(\phi_1, \phi_2, A_0)\|_{L^p \dot{H}_c^{1+\frac{1}{p}}[I]} \lesssim \varepsilon^{\delta_1} \|\phi_1\|_{S^1[I]} \|\phi_2\|_{S_c^1[I]} (1 + \|A_0\|_{L^p \dot{H}^{1+\frac{1}{p}}[I]})$$

$$\|\partial_t \mathbf{A}_0(\phi_1, \phi_2, A_x)\|_{L^p \dot{H}_c^{\frac{1}{p}}[I]} \lesssim \varepsilon^{\delta_1} \|\phi_1\|_{S^1[I]} \|\phi_2\|_{S_c^1[I]} (1 + \|A_x\|_{S^1[I]}) \; . \tag{4.18}$$

We omit the proof, as it is similar to Proposition 4.4.

4.3. **The ϕ equation.** We will split the ϕ equation into a leading order paradifferential approximation plus a perturbative part. The paradifferential approximation is given by

$$\Box_A^{p,m} = \Box + 2i \sum_k P_{<k-m} A^\alpha \partial_\alpha P_k \; . \tag{4.19}$$

Here we retain the freedom to choose m arbitrarily large later on. Then the operator $\Box_A$ is written as

$$\Box_A = \Box_A^{p,m} + \mathcal{M}_A^m \tag{4.20}$$

where $\mathcal{M}_A^m = \mathcal{M}_A^{m,2} + \mathcal{M}_{A,A}^{m,3}$ is given by

$$\mathcal{M}_A^{m,2} \psi = 2i \sum_k P_{\geq k-m} A^\alpha \partial_\alpha P_k \psi - i \partial_t A_0 \psi \; ,$$

$$\mathcal{M}_{A,B}^{m,3} \psi = - A^\alpha B_\alpha \psi \; . \tag{4.21}$$

The operator $\mathcal{M}_A^m$ will play a perturbative role in our analysis, just based on the S^1 and $L^2 \dot{H}^{\frac{1}{2}}$ bounds for the coefficients $A_x, \nabla A_0$. Precisely, for its quadratic and cubic parts we have:

Proposition 4.6. *Let A, B, ψ be test functions defined on a time interval I. Let c, d, e be admissible frequency envelopes.*
(1) *The cubic part $\mathcal{M}_{A,B}^{m,3}$ satisfies the bound*

$$\|\mathcal{M}_{A,B}^{m,3} \psi\|_{N_{cde}[I]} \lesssim 2^{Cm} \|(A_x, \nabla A_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|(B_x, \nabla B_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_d[I]} \|\psi\|_{S_e^1} \tag{4.22}$$

where $\|(A_x, \nabla A_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]}$ is a shorthand for $(\|A_x\|_{DS_c^1[I]} + \|\nabla A_0\|_{L^2 \dot{H}_c^{\frac{1}{2}}[I]})$.
(2) *The quadratic part $\mathcal{M}_A^{m,2}$ admits a decomposition*

$$\mathcal{M}_A^{m,2} = \mathcal{M}_{A,small}^{m,2} + \mathcal{M}_{A,large}^{m,2} \tag{4.23}$$

so that we have

$$\|\mathcal{M}_{A,small}^{m,2} \psi\|_{N_{cd}[I]} \lesssim 2^{-cm} \|(A_x, \nabla A_0)\|_{(S^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|\psi\|_{S_d^1[I]} \tag{4.24}$$

while

$$\|\mathcal{M}_{A,large}^{m,2} \psi\|_{N_{cd}[I]} \lesssim 2^{Cm} \|(A_x, \nabla A_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|\psi\|_{S_d^1[I]} \; . \tag{4.25}$$

23

(3) *Further, if ψ is ε-energy dispersed and obeys (4.3), then the quadratic and cubic parts of $\mathcal{M}_A^m$ satisfy*

$$\|\mathcal{M}_A^{m,2}\psi\|_{N_c[I]} \lesssim \left(2^{Cm}\varepsilon^{\delta_1}\|A_x\|_{S_c^1[I]} + \|\nabla A_0\|_{L^2\dot{H}_c^{\frac{1}{2}}[I]}\right)\|\psi\|_{S^1[I]} \tag{4.26}$$

as well as

$$\|\mathcal{M}_{A,B}^{m,3}\psi\|_{N_{cd}[I]} \lesssim 2^{Cm}\varepsilon^{\delta_1}\|(A_x, \nabla A_0)\|_{(S^1\times L^2\dot{H}^{\frac{1}{2}})_c[I]}\|(B_x, \nabla B_0)\|_{(S^1\times L^2\dot{H}^{\frac{1}{2}})_d[I]}\|\psi\|_{S^1[I]} . \tag{4.27}$$

This result is proved in Section 7. We remark the different roles of $\mathcal{M}_{A,small}^{m,2}$ versus $\mathcal{M}_{A,large}^{m,2}$. The first one is small, and thus directly perturbative. The second is not small, but is instead estimated using only a divisible norm of A; thus we can partition time into finitely many intervals where it is small.

Our next goal is to compare the operators $\Box_A^{p,m}$ and $\Box_{A^{free}}^{p,m}$, where we use the convention $A_0^{free} = 0$. We define the bilinear operator $\mathrm{Diff}_A^m\psi$ by

$$\mathrm{Diff}_A^m\psi = \Box_A^{p,m} - \Box = 2i\sum_k P_{<k-m}A^\alpha\partial_\alpha P_k\psi.$$

Hence we have the decomposition

$$\Box_A^{p,m} = \Box_{A^{free}}^{p,m} + \mathrm{Diff}_{A^{nl}}^m$$

For the last term, we no longer use only the S^1 and $L^2\dot{H}^{\frac{1}{2}}$ bounds for A_x and ∇A_0, but instead we rely on the fact that A_x and A_0 come from the equations (4.1a), (4.2). Thus, we replace $\mathrm{Diff}_{A^{nl}}^m\psi$ with the multilinear operator

$$\mathrm{Diff}_{\mathbf{A}}^m(\phi, \phi, A) = 2i\sum_k P_{<k-m}\mathbf{A}^\alpha(\phi, \phi, A)\partial_\alpha P_k .$$

As before, we extend this operator to a symmetric quadratic form in the first two inputs. For the multilinear operator $\mathrm{Diff}_{\mathbf{A}}^m(\phi, \phi, A)\psi$, we have the following estimates:

Proposition 4.7. *Let ϕ_1, ϕ_2, ψ, A be test functions on a time interval I containing 0. Let c, d, e be admissible frequency envelopes.*
(1) The quadratic and cubic parts of the operator $\mathrm{Diff}_{\mathbf{A}}^m(\phi, \phi, A)$ satisfy the bounds

$$\|\mathrm{Diff}_{\mathbf{A}}^{m,2}(\phi_1, \phi_2)\psi\|_{N_f[I]} \lesssim \|\phi_1\|_{S_c^1[I]}\|\phi_2\|_{S_d^1[I]}\|\psi\|_{S_e^1[I]} \tag{4.28}$$

respectively

$$\|\mathrm{Diff}_{\mathbf{A}}^{m,3}(\phi_1, \phi_2, A)\psi\|_{N_f[I]} \lesssim \|(A_x, \nabla A_0)\|_{(DS^1\times L^2\dot{H}^{\frac{1}{2}})[I]}\|\phi_1\|_{DS_c^1[I]}\|\phi_2\|_{DS_d^1[I]}\|\psi\|_{S_e^1[I]} \tag{4.29}$$

where

$$f(k) = e(k)\|c_{\leq k-m}\|_{\ell^2}\|d_{\leq k-m}\|_{\ell^2} . \tag{4.30}$$

(2) Further, for each $m > 0$, $\mathrm{Diff}_{\mathbf{A}}^{m,2}$ admits a decomposition

$$\mathrm{Diff}_{\mathbf{A}}^{m,2} = \mathrm{Diff}_{\mathbf{A},small}^{m,2} + \mathrm{Diff}_{\mathbf{A},large}^{m,2} \tag{4.31}$$

so that $\mathrm{Diff}_{\mathbf{A},small}^{m,2}$ satisfies a better bound,

$$\|\mathrm{Diff}_{\mathbf{A},small}^{m,2}(\phi, \phi)\psi\|_{N_c[I]} \lesssim 2^{-cm}\|\phi\|_{S^1[I]}^2\|\psi\|_{S_c^1[I]} \tag{4.32}$$

while $\mathrm{Diff}^{m,2}_{\mathbf{A},large}$ *is estimated directly in a divisible norm,*

$$\|\mathrm{Diff}^{m,2}_{\mathbf{A},large}(\phi,\phi)\psi\|_{N_c[I]} \lesssim 2^{Cm}\|\phi\|^2_{DS^1[I]}\|\psi\|_{S^1_c[I]} \tag{4.33}$$

This result is proved in Section 8. Again, we remark that the large part is estimated using a divisible norm, which can be made small by subdividing the time interval. We also remark that here we are concerned with unbalanced frequency interactions, so the energy dispersion plays no role.

For the gradient terms in $\square^{p,m}_{A^{free}}$ we only have the following dyadic bound from [13]:

Proposition 4.8. *For a divergence free homogeneous wave A we have the dyadic bound*

$$\|P_k A^j \partial_j P_l \psi\|_{N_l} \lesssim \|P_k A[0]\|_E \|P_l \psi\|_{S^1}, \qquad for\ k < l. \tag{4.34}$$

Due to the lack of ℓ^2 dyadic summation with respect to k in the above bound, the gradient terms in $\square^{p,m}_{A^{free}}$ need to be treated in a nonperturbative manner. This issue was addressed in the small data case in [13] by constructing a microlocal parametrix. Here we adopt the same strategy, but using a different source for the smallness, namely the frequency gap m:

Theorem 4.9. *Let $\square^{p,m}_A$ be the paradifferential gauge-covariant wave operator defined on line (4.19), and suppose that $\square A^{free} = 0$ with $\|A^{free}[0]\|_{\dot{H}^1 \times L^2} \leq E$. If m is sufficiently large, $m \gg_E 1$, then we have the linear bound:*

$$\|\phi\|_{S^1} \lesssim_E \|\phi[0]\|_E + \|\square^{p,m}_{A^{free}}\phi\|_{N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}}. \tag{4.35}$$

Section 9 is devoted to the proof of this result.

Finally, we end this section with estimates that are relevant for high modulation bounds for ϕ. As before, no null structure is necessary.

Proposition 4.10. *Let A, B, ψ be test functions defined on a time interval I. For all admissible frequency envelopes c, d, e, we have*

$$\|\mathcal{M}^{m,2}_A \psi\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{cd}[I]} \lesssim \|(A_x, \nabla A_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|\psi\|_{DS^1_d[I]} \tag{4.36}$$

$$\|\mathcal{M}^{m,3}_{A,B} \psi\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{cde}[I]} \lesssim \|(A_x, \nabla A_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \tag{4.37}$$
$$\times \|(B_x, \nabla B_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_d[I]} \|\psi\|_{DS^1_e[I]} .$$

For every $m > 0$, we also have the bound

$$\|\mathrm{Diff}^m_A \psi\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{cd}[I]} \lesssim \|(A_x, \nabla A_0)\|_{(DS^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|\psi\|_{DS^1_d[I]} \tag{4.38}$$

with an implicit constant independent of m.

In addition, if ψ is ε-energy dispersed, then

$$\|\mathcal{M}^{m,2}_A \psi\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_c[I]} \lesssim \varepsilon^{\delta_1} \|(A_x, \nabla A_0)\|_{(S^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|\psi\|_{S^1[I]} \tag{4.39}$$

$$\|\mathrm{Diff}^m_A \psi\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_c[I]} \lesssim \varepsilon^{\delta_1} \|(A_x, \nabla A_0)\|_{(S^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \|\psi\|_{S^1[I]} \tag{4.40}$$

$$\|\mathcal{M}^{m,3}_{A,B} \psi\|_{(L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_{cd}[I]} \lesssim \varepsilon^{\delta_1} \|(A_x, \nabla A_0)\|_{(S^1 \times L^2 \dot{H}^{\frac{1}{2}})_c[I]} \tag{4.41}$$
$$\times \|(B_x, \nabla B_0)\|_{(S^1 \times L^2 \dot{H}^{\frac{1}{2}})_d[I]} \|\psi\|_{S^1[I]} .$$

Proof. The proof is similar to Proposition 4.4 and 4.5. We sketch the case of the $L^2\dot{H}^{-\frac{1}{2}}$ norm and leave the case of the $L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}}$ norm (which is a simple variant) to the reader.

Compared to the frequency dyadic estimates in the proof of Proposition 4.4, it suffices to note that the following estimates hold:

$$\|P_k(A_{k_1}\partial_t\psi_{k_2})\|_{L^2\dot{H}^{-\frac{1}{2}}} + \|P_k(\partial_t A_{k_1}\psi_{k_2})\|_{L^2\dot{H}^{-\frac{1}{2}}} \lesssim 2^{-\delta\max\{|k-k_i|\}}\|\nabla A_{k_1}\|_{L^2\dot{H}^{\frac{1}{2}}}\||D|^{-\frac{2}{5}}\nabla\psi_{k_2}\|_{L^\infty L^{\frac{5}{2}}}$$

$$\|P_k(A_{k_1}B_{k_2}\psi_{k_3})\|_{L^2\dot{H}^{-\frac{1}{2}}} \lesssim 2^{-\delta\max\{|k-k_i|\}}\|A_{k_1}\|_{L^4\dot{H}^{\frac{5}{4}}}\|B_{k_2}\|_{L^4\dot{H}^{\frac{5}{4}}}\||D|^{\frac{3}{5}}\psi_{k_3}\|_{L^\infty L^{\frac{5}{2}}}$$

where we omitted $[I]$ and used the shorthands $A_k = P_k A_0$, $B_k = P_k B_0$ and $\psi_k = P_k\psi$. The off-diagonal gain is again due to the freedom of choosing where to apply Bernstein's inequality. Moreover, by interpolation with the Sobolev trace theorem, note that

$$\|A_k\|_{L^4\dot{H}^{\frac{5}{4}}} \lesssim \|\nabla A_k\|_{L^2\dot{H}^{\frac{1}{2}}}.$$

Since $(\infty, \frac{5}{2})$ is a non-sharp Strichartz exponent, the above estimates suffice for both divisibility and ε-energy dispersed bounds (via Lemma 3.6). $\qquad\square$

5. THE STRUCTURE OF FINITE S^1 NORM MKG WAVES.

Here we consider an MKG solution (A, ϕ) on a time interval $I = [0, T]$, with finite S^1 norm for (A_x, ϕ). Our main result is an accurate characterization of such maps:

Theorem 5.1. *Let (A, ϕ) be an admissible $C_t\mathcal{H}^1$ solution to the MKG system (1.6) in the Coulomb gauge (1.5) on the time interval $I = [0, T]$ which has energy E and S^1 norm F, i.e., $\|(A_x, \phi)\|_{S^1[I]} \leq F$. Let c be a frequency envelope for the initial data $(A, \phi)[0]$ in the energy space $\dot{H}^1 \times L^2$. Then the following properties hold:*

(1) (Linear well-posedness for $\square_A$) The linear equation

$$\square_A\psi = f, \qquad \psi[0] = \psi_0$$

is well-posed, with bounds

$$\|\psi\|_{S_d^1[I]} \lesssim_F \|\psi[0]\|_{E_d} + \|f\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_d[I]} \tag{5.1}$$

for any admissible frequency envelope d.

(2) (Frequency envelope bound) The solution (A, ϕ) satisfies

$$\|(A_x, \phi)\|_{S_c^1[I]} \lesssim_F 1. \tag{5.2}$$

(3) (Refined Maxwell field bounds) We have

$$\|\square A_i\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c^2}[I]} \lesssim_F 1, \qquad \|\nabla A_0\|_{Y_{c^2}^1[I]} \lesssim_F 1. \tag{5.3}$$

(4) (Refined scalar field bounds) We have

$$\|\square_{A^{free}}\phi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_c[I]} \lesssim_F 1, \tag{5.4}$$

and for each $m > 0$ the following paradifferential estimates hold:

$$\|\square_A^{p,m}\phi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c^2}[I]} + \|\square_{A^{free}}^{p,m}\phi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_c[I]} \lesssim_F 2^{Cm}. \tag{5.5}$$

(5) (Weak divisibility of S^1 norm) There exists a partition $I = \bigcup_{k=1}^K I_k$ with $K \lesssim_F 1$ so that

$$\|(A, \phi)\|_{S^1[I_k]} \lesssim_E 1 \tag{5.6}$$

where the implicit constant is $C(E + E^2)$ times the constant in Theorem 4.9.

Proof. As a preliminary step, we observe that from Lemma 2.2, (4.15), (4.16) and (4.17) we obtain the bound

$$\|\nabla A_0\|_{L^2 \dot{H}^{\frac{1}{2}}_{c^2}[I]} \lesssim_F 1. \tag{5.7}$$

We remark that this bound will later be refined when we prove (3).

(1)). We first prove well-posedness for the equation

$$\Box^{p,m}_A \psi = f, \qquad \psi[0] = (\psi_0, \psi_1)$$

together with the bound

$$\|\psi\|_{S^1_e} \lesssim \|\psi[0]\|_{E_e} + \|f\|_{(N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})_e} \tag{5.8}$$

provided that $m \gg_F 1$. This is done perturbatively, based on the similar result for $\Box^{p,m}_{A^{free}}$ in Theorem 4.9. Using also (5.7), we can split time into $O_F(2^{100Cm})$ intervals I_n so that

$$\|\phi\|_{DS^1[I_n]} + \|A_x\|_{DS^1[I_n]} + \|\nabla A_0\|_{L^2 \dot{H}^{\frac{1}{2}}[I_n]} \lesssim_F 2^{-2Cm} \ .$$

Then within each interval I_n we write the equation above in the form

$$\Box^{p,m}_{A^{free}} \psi = -\mathrm{Diff}^m_{\mathbf{A}} \psi + f = -(\mathrm{Diff}^{m,2}_{\mathbf{A},large} + \mathrm{Diff}^{m,3}_{\mathbf{A}} + \mathrm{Diff}^{m,2}_{\mathbf{A},small}) \psi + f \ .$$

By Propositions 4.7 and 4.10, all the terms on the right are perturbative in $N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}[I_n]$, so if $m \gg_F 1$ then within each such interval we can solve the above equation perturbatively. Reiterating, the global solvability along with (5.8) follows. We note that in this argument the free part of A is reinitialized in each interval I_k. The nonlinear part $A^{nl}_x = \mathbf{A}(\phi, \phi, A_x)$ is also defined separately for each interval.

To get the well-posedness for the $\Box_A$ equation, we repeat the above argument for the expression $\mathcal{M}^m_A$. For the N norm, we apply Proposition 4.6. Then the small part is treated perturbatively by taking $m \gg_F 1$, while for the large part we use again a time interval division in order to gain smallness. For the $L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}$ norm, we use Proposition 4.10 and rely on divisibility for smallness.

(2)). The ϕ bound is a direct consequence of the bound (5.1) applied to ϕ. Then we get the A_x bound from (3.8), (4.4)-(4.5) (for the N norm) and (4.10)-(4.11) (for the $L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}$ norm).

(3)). The A_x bound has been proved in (2), while the desired A_0 estimate follows from (4.15) and Proposition 4.16.

(4)). For the N norm, the bound for $\Box^{p,m}_A \phi$ is a consequence of the estimates (4.22), (4.24) and (4.25) for the components of $\Box^{p,m}_A \phi = -\mathcal{M}^m_A$. For transition to $\Box^{p,m}_{A^{free}} \phi$ we use in addition the bounds (4.28) and (4.29). We can switch back from $\Box^{p,m}_{A^{free}} \phi$ to $\Box_{A^{free}} \phi$ using again the estimates (4.22), (4.24) and (4.25) but for $A = A^{free}$. Finally, for the $L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}$ bound, we use Proposition 4.10 for all parts.

(5)). By Proposition 2.1 and conservation of energy, the linear energy $E_{lin}(A_x[t], \phi[t])$ is bounded by $E + E^2$ uniformly in time. Moreover, the N norm is divisible by (3.13), hence the A part is a direct consequence of (5.3). The similar assertion for ϕ follows similarly from the divisibility of the N norm and the second bound (5.5), since for a fixed $m \gg E$, the $\Box^{p,m}_{A^{free}}$ equation is well-posed in S^1 with implicit constants depending only on E. $\quad\Box$

27

With Theorem 5.1 in hand, we may easily prove the continuation and scattering theorem (Theorem 1.6).

Proof of Theorem 1.6. We start with the continuation result. The idea is to use the frequency envelope bound in Theorem 5.1 to show a uniform lower bound on the energy concentration scale r_c for all $t \in I$, which allows us to apply Theorem 1.4.

By Theorem 5.1, we see that (A, ϕ) obeys the frequency envelope bound

$$\|A_0\|_{Y_{c^2}^1[I]} + \|(A_x, \phi)\|_{S_c^1[I]} \leq \tilde{F}$$

for some $\tilde{F} = O_F(1)$, where $\|c\|_{\ell^2} \lesssim \|(A_x, \phi)[0]\|_{\dot{H}^1 \times L^2}$. In particular, $\lim_{\ell \to \infty} \|c_k\|_{\ell^2(k>\ell)} = 0$. Recall also that both Y_c^1 and S_c^1 control $\|\nabla(\cdot)\|_{L^\infty L_c^2}$. Hence given any small number $\delta > 0$, there exists $\ell \in \mathbb{Z}$ such that the splittings $A = A_{low} + A_{high} := A_{<\ell} + A_{\geq \ell}$ and $\phi = \phi_{low} + \phi_{high} := \phi_{<\ell} + \phi_{\geq \ell}$ obey

$$\|\nabla A_{\mu,high}(t)\|_{L^2} + \|\nabla \phi_{high}(t)\|_{L^2} < \frac{\delta}{10},$$

and by Bernstein's inequality,

$$\|A_{\mu,low}(t)\|_{L^\infty} + \|\nabla A_{\mu,low}(t)\|_{L^\infty} + \|\phi_{low}(t)\|_{L^\infty} + \|\nabla \phi_{low}(t)\|_{L^\infty} \lesssim_{\tilde{F},c} 1.$$

Both bounds are uniform in $t \in I$. Using Hölder's inequality for the low frequency part, we can find $\tilde{r} = \tilde{r}(\tilde{F}, c, \delta) > 0$ such that

$$\|(A_\mu, \partial_t A_\mu)(t)\|_{(\dot{H}^1 \cap L^4) \times L^2(B_{\tilde{r}}(x))} + \|(\phi, \partial_t \phi(t)\|_{(\dot{H}^1 \cap L^4) \times L^2(B_{\tilde{r}}(x))} < \delta$$

for every $t \in I$ and ball $B_{\tilde{r}}(x)$ of radius $\tilde{r}$ and arbitrary center $x \in \mathbb{R}^4$. Recalling the definition (1.11), we see that the energy concentrations scale of the data for (A, ϕ) at time t is uniformly bounded below by $\tilde{r} > 0$, if $\delta > 0$ is chosen sufficiently small depending only on E. Hence by Theorem 1.4, (A, ϕ) can be continued past the endpoints of I as an $C_t \mathcal{H}^1$ admissible solution with appropriate S^1 and Y^1 bounds.

The scattering statement is an easy consequence of (5.3), (5.4), and divisibility of the $N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}$ norm. $\qquad \square$

5.1. MKG waves with small energy dispersion.
Here we continue the analysis above, but add to it the small energy dispersion condition.

Theorem 5.2. *Let (A, ϕ) be an admissible $C_t \mathcal{H}^1$ solution to the MKG system (1.6) in the Coulomb gauge (1.5) on the time interval $I = [0, T]$, which has energy E and S^1 norm F. Suppose furthermore that ϕ is ε-energy dispersed. Then the following properties hold:*

(1) (Elliptic bounds) We have

$$\|\nabla A_0\|_{Y^1[I]} \lesssim_F \varepsilon^{\delta_1} \|\phi\|_{S^1}^2 \tag{5.9}$$

(2) (High modulation bound)

$$\|\Box \phi\|_{L^2 \dot{H}^{-\frac{1}{2}}[I]} + \|\Box \phi\|_{L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}[I]} \lesssim_F \varepsilon^{\delta_1} \|\phi\|_{S^1} \tag{5.10}$$

(3) (Maxwell field bounds) We have

$$\|A_x^{nl}\|_{S^1[I]} + \|\Box A_x\|_{(N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}})[I]} \lesssim_F \varepsilon^{\delta_1} \|\phi\|_{S^1}^2 \tag{5.11}$$

28

(4) (Scalar field bounds) For $m > 0$ we have

$$\|\Box_A^{p,m}\phi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})[I]} \lesssim_F 2^{Cm}\varepsilon^{\delta_1}\|\phi\|_{S^1} \tag{5.12}$$

Proof. (1)). The bound (5.9) follows directly from the estimate (4.18).

(2)). These bounds follow from (4.39)-(4.41).

(3)). The estimate (5.11) follows from (4.8)-(4.9), (4.12)-(4.13) for $\Box A_x^{nl}$ and (3.12).

(4)). The bound (5.12) is a consequence of (4.26)-(4.27), (4.39) and (4.41). $\qquad\square$

6. INDUCTION ON ENERGY

Here we provide the induction on energy argument which gives the proof of our main result in Theorem 1.5. Our induction hypothesis is that the conclusion of the theorem holds up to energy E. Thus we have $F(E)$ and $\epsilon(E)$. Our goal is to show that there exists $c_0 = c_0(E) > 0$ so that the conclusion holds up to energy $E + c_0$. Moreover, we do not allow $c_0(E)$ to depend on $F(E)$, but only on E. The independence of $c_0(E)$ on $F(E)$ allows us to additionally ensure that $c_0(\cdot)$ is a positive non-increasing function on the whole $[0, \infty)$; this property is what makes our induction argument work for all energies. We refer to the beginning of Step 2.3 in the proof of Proposition 6.1 for the precise dependence of c_0 on E. The conclusion is that c_0 needs to be chosen small enough compared to the constant in Theorem 4.9.

To begin with, we observe that it suffices to establish Theorem 1.5 for smooth solutions. Indeed, Theorem 1.4 implies that any admissible $C_t\mathcal{H}^1$ solution (A, ϕ) can be approximated by smooth solutions in the $S^1[J]$ norm (and hence also in the $ED[J]$ norm) for any compact interval J. Thus, we consider smooth data $(A_x[0], \phi[0])$ with energy $E + c_0$, generating a smooth solution (A, ϕ) in $[0, T)$ with $\|\phi\|_{ED(0,T)} \leq \epsilon \ll_E 1$. Then the S^1 norm $\|(A_x, \phi)\|_{S^1(0,t)}$ is a continuous function of time $t \in (0, T)$, satisfying

$$\lim_{t \to 0} \|(A, \phi)\|_{S^1(0,t)} \lesssim E^{1/2}.$$

Hence, in order to prove a uniform bound

$$\|(A_x, \phi)\|_{S^1(0,t)} \leq F \tag{6.1}$$

we can make the bootstrap assumption

$$\|(A_x, \phi)\|_{S^1(0,t)} \leq 2F, \tag{6.2}$$

where F is a positive number to be determined in the proof. By scaling we harmlessly take $t = T$.

Indeed, once we show that (6.1) holds assuming (6.2), a simple continuous induction argument in time implies that $S^1(0, T)$ norm of (A_x, ϕ) is bounded by F. This bound is precisely (1.13) with $F(E + c_0) = F$. Note that the parameter ϵ becomes $\epsilon(E + c_0)$ in Theorem 1.5.

Next, we dispense the easy case when the S^1 norm of ϕ is disproportionally small compared to the overall energy E of (A, ϕ). This procedure allows us to link the small energy dispersion assumption $\|\phi\|_{ED(0,T)} \leq \epsilon$ to the notion of ε-energy dispersion (Definition 3.5) for some $\varepsilon = \varepsilon(\epsilon)$. More precisely, given $\varepsilon > 0$ to be determined, we consider two cases: (i) $\|\phi\|_{S^1(0,T)} \leq$

$\varepsilon E^{1/2}$ or (ii) $\|\phi\|_{S^1(0,T)} > \varepsilon E^{1/2}$. In case (i), a direct application of (4.4)-(4.5), (4.10)-(4.11) and the bootstrap assumption (6.2) gives

$$\|\Box A_x\|_{N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}(0,T)} \lesssim \varepsilon^2 E(1+F).$$

Applying the linear estimate (3.8) and taking ε sufficiently small compared to to F, (6.1) follows directly. Thus we are left with case (ii), in which we may assume that ϕ is ε-energy dispersed (according to Definition 3.5) by taking $\epsilon = \varepsilon^2 E^{1/2}$. Henceforth we eliminate ϵ (which has the dimension of $(energy)^{1/2}$) in favor of the dimensionless parameter ε.

To establish the S^1 bound (6.1) under the assumption that ϕ is ε-energy dispersed on $[0, T]$, we will compare the solution (A, ϕ) with the MKG wave $(\tilde{A}, \tilde{\phi})$ generated by frequency truncated data

$$(\tilde{A}_x[0], \tilde{\phi}[0]) = P_{\leq k^*}(A_x[0], \phi[0])$$

where the *cut frequency* $k^* \in \mathbb{R}$ (hence $P_{\leq k^*}$ is a continuous version of Littlewood-Paley projection) is selected so that $(\tilde{A}, \tilde{\phi})$ has energy E. Note here that we only truncate $\phi[0]$ and $A_j[0]$. The functions A_0 and $\partial_t A_0$, which are also part of the energy, are defined directly from the compatibility conditions (1.7). The fact that such a k^* exists is a consequence of the continuity with respect to k of the A_0 component generated by $P_{\leq k}(A_x[0], \phi[0])$, see Proposition 2.1. We further remark that by part (3) of Proposition 2.1, the energy of both (A, ϕ) and $(\tilde{A}, \tilde{\phi})$ is $\epsilon^{\frac{1}{4}} = O_E(\varepsilon^{1/2})$ close to the corresponding linear energy of $(A_x[0], \phi[0])$, respectively $(\tilde{A}_x[0], \tilde{\phi}[0])$.

We wish to apply the induction hypothesis to obtain an S^1 bound for $(\tilde{A}, \tilde{\phi})$ on $[0, T]$, namely

$$\|(\tilde{A}_x, \tilde{\phi})\|_{S^1(0,T)} \leq F(E). \tag{6.3}$$

For this purpose, we need to know that energy dispersion of $\tilde{\phi}$ is sufficiently small on $[0, T]$. We achieve this smallness by transferring the information for $\|\phi\|_{ED(0,T)}$ to $\|\tilde{\phi}\|_{ED(0,T)}$ by another continuous induction in time.

Indeed, at time $t = 0$, the solution $(\tilde{A}, \tilde{\phi})$ has smooth data and $\|\tilde{\phi}[0]\|_{ED} \leq \varepsilon^2 E^{1/2} \ll \epsilon(E)$. Thus for some short time it will still have energy dispersion $\leq \epsilon(E)$. We claim that $(\tilde{A}, \tilde{\phi})$ extends smoothly up to time T, so that the stronger bound

$$\|\tilde{\phi}\|_{ED(0,t_0)} \leq \frac{1}{2}\epsilon(E) \tag{6.4}$$

holds for all $t_0 \in (0, T]$. We will establish (6.4) under the additional bootstrap assumption

$$\|\tilde{\phi}\|_{ED(0,t_0)} \leq \epsilon(E) \tag{6.5}$$

As before, note that we may take $t_0 = T$ by scaling.

To see how the claim follows from this bootstrap procedure, let T^* be the maximal time $T^* \leq T$ up to which (6.4) holds. Then by our induction hypothesis and Theorem 1.6, the solution $(\tilde{A}, \tilde{\phi})$ extends smoothly past time T^*. Hence (6.5) holds past the time T^* by continuity, therefore (6.4) also holds past time T^* by our claim. This contradicts the maximality of T^* unless $T^* = T$.

To summarize, we have to prove that we can find $c_0 = c_0(E)$, $F \gg_E 1$ and $\varepsilon \ll_F 1$ so that the following statement holds:

Proposition 6.1. *Assume that the MKG waves (A, ϕ), respectively $(\tilde{A}, \tilde{\phi})$, with initial data $(A_x[0], \phi[0])$, respectively $(\tilde{A}_x[0], \tilde{\phi}[0]) = P_{\leq k^*}(A_x[0], \phi[0])$, and energies $E + c_0$, respectively E, are smooth in $[0, T]$ and obey the following hypotheses:*

(i) The S^1 norm of (A_x, ϕ) satisfies (6.2).
(ii) The solution ϕ is ε-energy dispersed (as in Definition 3.5).
(iii) The ED norms of ϕ and $\tilde{\phi}$ obey

$$\|\tilde{\phi}\|_{ED(0,T)} \leq \epsilon(E), \qquad \|\phi\|_{ED(0,T)} \leq \varepsilon^2 E^{1/2}. \tag{6.6}$$

Then the following statements hold:
(1) The S^1 norm of (A_x, ϕ) satisfies (6.1).
(2) The ED norm of $\tilde{\phi}$ satisfies (6.4).

Proof. Unless otherwise stated, all norms below are taken over the time interval $(0, T)$. We will prove the proposition in two steps.

Step 1: The low frequency bound. Here we estimate the difference (B, ψ) given by

$$B = \tilde{A} - A_{<k^*}, \qquad \psi = \tilde{\phi} - \phi_{<k^*}$$

and prove that it satisfies the bound

$$\|(B_x, \psi)\|_{S_{c^*}^1} + \|\nabla B_0\|_{L^2 \dot{H}_{c^*}^{\frac{1}{2}}} \lesssim_F \varepsilon^{\delta_*}, \qquad c^*(k) = 2^{-\delta_0 |k-k^*|} \tag{6.7}$$

One consequence of the ψ bound above, combined with (6.6), is that (6.4) holds.

Before we begin, note that Theorem 5.2 implies the following a-priori bounds:

$$\|\Box \phi\|_{L^2 \dot{H}^{-\frac{1}{2}}} \lesssim_F \varepsilon^{\delta_1} \|\phi\|_{S^1}, \tag{6.8}$$

$$\|\nabla A_0\|_{L^2 \dot{H}^{\frac{1}{2}}} \lesssim_F \varepsilon^{\delta_1} \|\phi\|_{S^1}^2, \tag{6.9}$$

In particular, the bound for ϕ ensures that (4.3) holds, allowing us to apply Propositions 4.1 and 4.6, whereas the bound for ∇A_0 provides smallness in applications of (4.8).

Step 1.1: Bound for B_x. To prove the estimates above, we begin with the bounds for B_x. By definition B_x has zero Cauchy data at time 0, therefore we have

$$B_x = \mathbf{A}_x(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - P_{<k^*} \mathbf{A}_x(\phi, \phi, A)$$
$$= \mathbf{A}_x(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \mathbf{A}_x(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}) + \mathbf{A}_x(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}) - P_{<k^*} \mathbf{A}_x(\phi, \phi, A)$$

In the first difference above we substitute $\tilde{A} = A_{<k^*} + B$ and $\tilde{\phi} = \phi_{<k^*} + \psi$, then use (3.8), (4.4)-(4.5), (4.8)-(4.9), (4.10)-(4.11) and (4.12)-(4.13) (see also Remark 4.3) to obtain

$$\|\mathbf{A}_x(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \mathbf{A}_x(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*})\|_{S_{c^*}^1} \lesssim_F \varepsilon^{\delta_1} \|(B_x, \psi)\|_{S_{c^*}^1} + \|(B_x, \psi)\|_{S_{c^*}^1}^2 + \|(B_x, \psi)\|_{S_{c^*}^1}^3$$

The second difference is localized at frequency $< 2^{k^*+5}$, and all the ϕ factors are ε-energy dispersed. Further, we may rewrite this difference as

$$P_{<k^*+5}\Big(\mathbf{A}_x(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}) - \mathbf{A}_x(\phi_{<k^*-5}, \phi_{<k^*-5}, A_{<k^*-5})\Big)$$
$$- P_{<k^*}\Big(\mathbf{A}_x(\phi, \phi, A) - \mathbf{A}_x(\phi_{<k^*-5}, \phi_{<k^*-5}, A_{<k^*-5})\Big)$$

which shows that of the two or three inputs, at least one has frequency $\geq 2^{k^*-10}$. This input can be measured with the frequency envelope c^* at frequencies below 2^{k^*+10} (the choice of

31

a frequency envelope for frequencies higher than 2^{k^*+10} is less important since the output frequency is at most 2^{k^*+5}). Thus, applying (3.8), (4.8)-(4.9) and (4.12)-(4.13) yields

$$\|\mathbf{A}_x(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}) - P_{<k^*}\mathbf{A}_x(\phi, \phi, A)\|_{S^1_{c^*}} \lesssim_F \varepsilon^{\delta_1} \ .$$

Summing up the last two bounds, we get

$$\|B_x\|_{S^1_{c^*}} \lesssim_F (\varepsilon^{\delta_1} + \|(B_x, \psi)\|^2_{S^1_{c^*}})(1 + \|(B_x, \psi)\|_{S^1_{c^*}}) \tag{6.10}$$

Step 1.2: Bound for B_0. The analysis for B_0 is very similar. Precisely, B_0 solves the equation

$$\begin{aligned}
B_0 &= \mathbf{A}_0(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - P_{<k^*}\mathbf{A}_0(\phi, \phi, A) \\
&= \mathbf{A}_0(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \mathbf{A}_0(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}) + \mathbf{A}_0(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}) - P_{<k^*}\mathbf{A}_0(\phi, \phi, A)
\end{aligned}$$

and the terms on the right can be estimated using (4.16)-(4.18). The same applies for $\partial_t B_0$. We obtain

$$\|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c^*}} \lesssim_F (\varepsilon^{\delta_1} + \|(B_x, \psi)\|^2_{S^1_{c^*}})(1 + \|(B_x, \psi)\|_{S^1_{c^*}} + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c^*}}) \tag{6.11}$$

Step 1.3: Bound for ψ. We now consider ψ, which solves

$$\Box_{\tilde{A}}\psi = -(\Box_{\tilde{A}} - \Box_{A_{<k^*}})\phi_{<k^*} - (\Box_{A_{<k^*}}\phi_{<k^*} - P_{<k^*}\Box_A\phi) \tag{6.12}$$

We start by estimating the right hand side in $(L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c^*}$. For the first difference, we write $\tilde{A} = A_{<k^*} + B$ and observe that at least one input is B (which can be measured using c^*) and $\phi_{<k^*}$ is ε-energy dispersed. Hence by (4.39)-(4.41), we have

$$\begin{aligned}
&\|(\Box_{\tilde{A}} - \Box_{A_{<k^*}})\phi_{<k^*}\|_{(L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c^*}} \\
&\qquad \lesssim_F \varepsilon^{\delta_1}(\|B_x\|_{S^1_{c^*}} + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c^*}})(1 + \|B_x\|_{S^1_{c^*}} + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c^*}}).
\end{aligned} \tag{6.13}$$

For the second difference in (6.12), we claim that the following bound holds:

$$\begin{aligned}
&\|(\Box_{A_{<k^*}}\phi_{<k^*} - P_{<k^*}\Box_A\phi)\|_{(L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c^*}} \\
&\qquad \lesssim_F \varepsilon^{\delta_1}(\|B_x\|_{S^1_{c^*}} + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c^*}})(1 + \|B_x\|_{S^1_{c^*}} + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c^*}}).
\end{aligned} \tag{6.14}$$

To prove this bound, we divide further into the following cases:

(i) At least one of the A frequencies is $> 2^{k^*-10}$. Note that the output frequency is localized to $< 2^{k^*+5}$. Hence by measuring the high frequency input with c^*, using the ε-energy dispersion of ϕ and applying (4.39)-(4.41), we can bound this contribution in $(L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c^*}$ by $\lesssim_F \varepsilon^{\delta_1}$.

(ii) The term $\Box_{A_{<k^*-10}}\phi_{<k^*} - P_{<k^*}(\Box_{A_{<k^*-10}}\phi)$. The contribution of $\phi_{<k^*-5}$ and $\phi_{>k^*+5}$ is zero, so we may assume that ϕ and the output are frequency localized near 2^{k^*}. Then by ε-energy dispersion of ϕ and (4.39)-(4.41), the desired estimate follows.

To estimate the N_{c^*} norm of the right hand side in (6.12), we use a frequency gap parameter m to be chosen later. The first difference in (6.12) is expressed in the form

$$(\Box_{\tilde{A}} - \Box_{A_{<k^*}})\phi_{<k^*} = (\mathcal{M}^m_{\tilde{A}} - \mathcal{M}^m_{A_{<k^*}})\phi_{<k^*} + (\text{Diff}^m_{\mathbf{A}}(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \text{Diff}^m_{\mathbf{A}}(\phi, \phi, A))\phi_{<k^*}$$

In the first term, note that one of the inputs must be B. Then we use the frequency envelope c^* for B, the ε-energy dispersion of $\phi_{<k^*}$, (6.8) and (6.9) via (4.26)-(4.27) to obtain

$$\|(\mathcal{M}_{\tilde{A}}^m - \mathcal{M}_{A_{<k^*}}^m)\phi_{<k^*}\|_{N_{c^*}} \lesssim_F (\varepsilon^{\delta_1}2^{Cm}\|B_x\|_{S_{c^*}^1} + \|\nabla B_0\|_{L^2\dot{H}_{c^*}^{\frac{1}{2}}})(1 + \|B_x\|_{S_{c^*}^1} + \|\nabla B_0\|_{L^2\dot{H}_{c^*}^{\frac{1}{2}}})$$
(6.15)

In the second term, we first replace the argument (ϕ, ϕ, A) by $(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*})$, and estimate the corresponding difference via (4.28)-(4.29) as

$$\|(\text{Diff}_{\mathbf{A}}^m(\phi, \phi, A) - \text{Diff}_{\mathbf{A}}^m(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}))\phi_{<k^*}\|_{N_{c^*}} \lesssim_F 2^{-\delta_0 m}$$
(6.16)

where both the frequency envelope control of N_{c^*} and the gain $2^{-\delta_0 m}$ come from the frequency gap between the difference of the magnetic coefficients $\mathbf{A}(\phi, \phi, A) - \mathbf{A}(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*})$ (which is only used at frequencies below 2^{k^*-m}) and its arguments (ϕ, ϕ, A) (of which at least one must have frequency no smaller than 2^{k^*}; we use c^* to measure this input). Then we are left to establish

$$\|(\text{Diff}_{\mathbf{A}}^m(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \text{Diff}_{\mathbf{A}}^m(\phi_{<k^*}, \phi_{<k^*}, A_{<k^*}))\phi_{<k^*}\|_{N_{c^*}}$$
$$\lesssim_F 2^{-\delta_0 m}(\|(B_x, \psi)\|_{S_{c^*}^1} + \|\nabla B_0\|_{L^2\dot{H}_{c^*}^{\frac{1}{2}}})(1 + \|\psi\|_{S_{c^*}^1}^2)$$
(6.17)

Note that one of the inputs must be (B, ψ). Then (6.17) follows again from (4.28)-(4.29), and using the frequency envelope c^* to measure (B, ψ).

Finally, we still have the second difference in (6.12) in N_{c^*}, for which we claim that

$$\|\Box_{A_{<k^*}}\phi_{<k^*} - P_{<k^*}\Box_A\phi\|_{N_{c^*}} \lesssim_F \varepsilon^{\delta_1}2^{Cm} + 2^{-cm}$$
(6.18)

To see this we write it as

$$(\Box_{A_{<k^*}}\phi_{<k^*} - P_{<k^*}\Box_A\phi) = (\mathcal{M}_{A_{<k^*}}^m\phi_{<k^*} - P_{<k^*}\mathcal{M}_A^m\phi) + [P_{<k^*}, \text{Diff}_A^m]\phi$$

For both differences, note that the output frequency is localized to $< 2^{k^*+5}$. Canceling the like terms in the first difference, we are left with three types of frequency scenarios:

(i) The frequency of one of the A's is at least $k^* - m$. Then we can apply (4.26)-(4.27) and (6.9) to obtain an $\varepsilon^{\delta_1}2^{Cm}$ bound, where we use c^* to measure the high frequency input.

(ii) We have a quadratic term of the form $\partial_t A_{0, <k^*-m}\phi_{[k^*-5, k^*+5]}$, which can be directly estimated by 2^{-cm} using Strichartz bounds and Bernstein's inequality.

(iii) We have a cubic term of the form $A_{<k^*-m}^2\phi_{[k^*-5, k^*+5]}$, which in turn can be estimated directly in L^1L^2 using non-sharp Strichartz estimates, to get an ε^{δ_1} bound.

It remains to consider the commutator term $[P_{<k^*}, \text{Diff}_A^m]\phi$; we claim that the contribution of this term can be estimated by 2^{-cm}. It is clear that only the frequencies close to k^* in ϕ are relevant here, therefore the commutator can be expressed as

$$[P_{<k^*}, A_{<k^*-m}^\alpha]\partial_\alpha\phi_{k^*} = 2^{-k^*}L(\nabla A_{<k^*-m}^\alpha, \partial_\alpha\phi_{k^*})$$

for a bilinear form L with translation invariant integrable kernel. The A_0 term is easy to deal with using the $L^2\dot{H}^{\frac{3}{2}}$ bound for A_0, Strichartz for ϕ and Bernstein's inequality. Thus we are left with the expression $2^{-k^*}L(\nabla A_{<k^*-m}^j, \partial_j\phi_{k^*})$, which has both a null structure and a favorable frequency balance. This we can treat using the bound (7.3) in the beginning of the next section (see also (7.1)). Hence (6.18) follows.

33

Thus, summing up all cases in (6.13)-(6.18), we obtain

$$\|\Box_{\tilde{A}}\psi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c*}} \lesssim_F$$

$$\left(2^{-cm} + 2^{-\delta_0 m} + \varepsilon^{\delta_1}2^{Cm}(1 + \|(B_x,\psi)\|_{S^1_{c*}}) + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c*}}\right)(1 + \|(B_x,\psi)\|^2_{S^1_{c*}} + \|\nabla B_0\|^2_{L^2\dot{H}^{\frac{1}{2}}_{c*}})$$

Optimizing the choice of m, this gives (with $\delta_* \ll \delta_0 \ll c$)

$$\|\Box_{\tilde{A}}\psi\|_{(N\cap L^2\dot{H}^{-\frac{1}{2}}\cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}})_{c*}}$$

$$\lesssim_F \left(\varepsilon^{\delta_*}(1 + \|(B_x,\psi)\|_{S^1_{c*}}) + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c*}}\right)(1 + \|(B_x,\psi)\|^2_{S^1_{c*}} + \|\nabla B_0\|^2_{L^2\dot{H}^{\frac{1}{2}}_{c*}})$$

Recalling that ψ has zero initial data, by Theorem 5.1(1) this implies the estimate

$$\|\psi\|_{S^1_{c*}} \lesssim_F \left(\varepsilon^{\delta_*}(1 + \|(B_x,\psi)\|_{S^1_{c*}}) + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}_{c*}}\right)(1 + \|(B_x,\psi)\|^2_{S^1_{c*}} + \|\nabla B_0\|^2_{L^2\dot{H}^{\frac{1}{2}}_{c*}}). \quad (6.19)$$

Now we can combine this with (6.10) and (6.11), and close to prove (6.7), provided that ε is small enough. We carefully observe here that the smallness of ε depends on F. In turn, we will want later that the choice of F is independent of ε.

Step 2: The high frequency bound. Here we consider the high frequency differences (B^{high}, ψ^{high}) given by

$$B^{high} = A - \tilde{A}, \qquad \psi^{high} = \phi - \tilde{\phi}$$

and prove that they satisfy the S^1 bound

$$\|(B_x^{high}, \psi^{high})\|_{S^1} + \|\nabla B_0\|_{L^2\dot{H}^{\frac{1}{2}}} \lesssim_{F(E)} 1, \quad (6.20)$$

provided that $c_0 = c_0(E)$ is chosen small enough compared to E, but independent of $F(E)$.

Step 2.1: The energy estimate for $(B_x^{high}, \psi^{high})$ and weak divisibility. Here we take the necessary steps to ensure the independence of c_0 on $F(E)$. We first use the energy conservation for (A, ϕ) and $(\tilde{A}, \tilde{\phi})$, together with the low frequency estimates of the previous step, to conclude that the energy norm for (B^{high}, ψ^{high}) stays bounded, i.e.,

$$\|(\nabla B_x^{high}, \nabla \psi^{high})\|^2_{L^\infty L^2} \lesssim c_0 + O_F(\varepsilon^{\frac{1}{4}\delta_*}) \quad (6.21)$$

Indeed, for each $t \in (0, T)$ we have

$$E_{lin}(A, \phi) = E_{lin}(A_{\geq k^*}, \phi_{\geq k^*}) + E_{lin}(A_{<k^*}, \phi_{<k^*}) + \langle A_{\geq k^*}, A_{<k^*}\rangle_{\dot{H}^1\times L^2} + \langle\phi_{\geq k^*}, \phi_{<k^*}\rangle_{\dot{H}^1\times L^2}$$

$$\geq E_{lin}(A_{\geq k^*}, \phi_{\geq k^*}) + E_{lin}(A_{<k^*}, \phi_{<k^*}).$$

where we omitted writing $[t]$ and the subscript x from $A = A_x$. We have used the fact that the operator $P_{<k^*}P_{\geq k^*}$ is non-negative, as it has a non-negative symbol. By Step 1, we know that $(A_{<k^*}, \phi_{<k^*})$ is equal to $(\tilde{A}, \tilde{\phi})$ up to an error of size $O_F(\varepsilon^{\delta_*})$ in S^1_{c*}. Therefore, we have

$$E_{lin}(B^{high}, \psi^{high}) = E_{lin}(A_{\geq k^*}, \phi_{\geq k^*}) + O_F(\varepsilon^{\delta_*})$$

$$\leq E_{lin}(A, \phi) - E_{lin}(A_{<k^*}, \phi_{<k^*}) + O_F(\varepsilon^{\delta_*})$$

$$= E_{lin}(A, \phi) - E_{lin}(\tilde{A}, \tilde{\phi}) + O_F(\varepsilon^{\delta_*}).$$

By Proposition 2.1, recall that $E_{lin}(A, \phi)$ and $E_{lin}(\tilde{A}, \tilde{\phi})$ are $O_E(\varepsilon^{\frac{1}{4}\delta_*})$ close to the corresponding conserved energies $\mathcal{E}[A, \phi]$ and $\mathcal{E}[\tilde{A}, \tilde{\phi}]$, respectively. Hence by the definition of $(\tilde{A}, \tilde{\phi})$, the desired estimate (6.21) follows.

Next, we use the weak divisibility of the S^1 norm in Theorem 5.1(5) to split the time interval I into $O_{F(E)}(1)$ subintervals, on each of which

$$\|(\tilde{A}_x, \tilde{\phi})\|_{S^1[J]} \lesssim_E 1. \tag{6.22}$$

We remark that this bound also relies on the conservation of energy for $(\tilde{A}, \tilde{\phi})$. Due to the uniform bound in (6.21), on each such subinterval J we can reinitialize the data for $(B_x^{high}, \psi^{high})$ and we no longer have any trace of $F(E)$ or $\epsilon(E)$. Instead, $(\tilde{A}_x, \tilde{\phi})$ has S^1 norm $O_E(1)$ and energy dispersion $\tilde{C}_F\varepsilon^{\delta_*}$. Thus, it remains to show that on each J we have the improved bound

$$\|(B^{high}, \psi^{high})\|_{S^1[J]} \lesssim_E 1 \tag{6.23}$$

Then (6.20) would follow by adding the above over $O_{F(E)}(1)$ intervals, using (3.14).

Remark 6.2. Compared to the low frequency estimate in Step 1, here we have a key advantage that we can exploit small energy dispersion for both $\tilde{\phi}$ and ψ^{high}, albeit at the expense of using the S^1 norm of ϕ on the larger interval $(0, T)$. More precisely, Step 1 implies

$$\|\tilde{\phi} - \phi_{<k^*}\|_{S^1(0,T)} + \|\psi^{high} - \phi_{\geq k^*}\|_{S^1(0,T)} \leq \tilde{C}_F\varepsilon^{\delta_*}. \tag{6.24}$$

for some constant $\tilde{C}_F > 1$. Moreover, ϕ is ε-energy dispersed by hypothesis and obeys (6.8) on the large interval $(0, T)$. By Remark 4.3 and a simple extension procedure we may gain $\tilde{C}_F\varepsilon^{\delta_*} + \varepsilon^{\delta_1}\|\phi\|_{S^1(0,T)}$ from $\tilde{\phi}$ or ψ^{high} whenever any of the ε-energy dispersion bounds (4.8)-(4.9), (4.12)-(4.13), (4.18), (4.26)-(4.27), (4.39)-(4.41) are applicable on the smaller interval J. To be precise, one extends all non energy-dispersed inputs of the form ϕ, A_x by homogeneous waves outside J to $(0, T)$ (see Proposition 3.3) and A_0 by a standard Sobolev extension so that $\|\nabla A_0\|_{L^2\dot{H}^{\frac{1}{2}}(0,T)} \lesssim \|\nabla A_0\|_{L^2\dot{H}^{\frac{1}{2}}[J]}$.

Step 2.2: The S^1 bound for B_x^{high}. The bound for B_x^{high} is an easy consequence of small energy dispersion. Indeed, using (6.24), ε-energy dispersion of ϕ, (6.8) as well as the estimates (4.4)-(4.5), (4.8)-(4.9), (4.10)-(4.11) and (4.12)-(4.13) (see also Remark 6.2), we have

$$\|\Box B_x^{high}\|_{N \cap L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}}[J]} \lesssim_F \varepsilon^{\delta_*}, \tag{6.25}$$

where we used the fact that $\delta_* \ll \delta$. Then by the linear estimate (3.8) and (6.21), it follows that

$$\|B^{high}\|_{S^1[J]} \lesssim_E c_0 + O_F(\varepsilon^{\frac{1}{4}\delta_*}) . \tag{6.26}$$

This bound is stronger than what we need for (6.23), but it will be useful in the next step.

Step 2.3: The S^1 bound for ψ^{high}. For ψ^{high}, we claim that a similar bound to (6.25) but with respect to the $\Box_A$ flow holds:

$$\|\Box_A\psi^{high}\|_{N \cap L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}}[J]} \lesssim_F \varepsilon^{\delta_{**}}. \tag{6.27}$$

where $0 < \delta_{**} \ll \delta_*$. Assuming that (6.27) holds, we can conclude (6.23) using the following simple additional bootstrap argument in time. Denoting the initial time in J by t_0, it follows

from (6.21) and continuity of the $S^1[J']$ norm that we have

$$\|\psi^{high}\|_{S^1[J']} \le C_E(c_0 + O_F(\varepsilon^{\frac{1}{4}\delta_*})) \tag{6.28}$$

for a suitably large constant $C_E > 1$ and a sufficiently short interval J' containing t_0. Then to prove (6.28) for $J' = J$, it suffices to establish (6.28) under the bootstrap assumption

$$\|\psi^{high}\|_{S^1[J']} \le 2C_E(c_0 + O_F(\varepsilon^{\frac{1}{4}\delta_*})). \tag{6.29}$$

Choosing c_0 sufficiently small depending on E and $\varepsilon \ll_F 1$, it follows from (6.21), (6.22), (6.26) and (6.29) that

$$\|(A_x, \phi)\|_{S^1[J']} \lesssim_E 1,$$

where the implicit constant is twice that of (6.22). Now applying the linear $\Box_A$ bound (5.1) in Theorem 5.1(1) and enlarging C_E to be larger than the implicit constant in (5.1), the desired estimate (6.28) follows. We remark that the size of c_0 essentially depends on the implicit constant in (5.1), which in turn depends on the constant in Theorem 4.9.

We now turn to the proof of (6.27). We first estimate the $L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}}$ norm, which is easier. Note that

$$\Box_A\psi^{high} = (\Box_{\tilde{A}} - \Box_A)\tilde{\phi}.$$

Exploiting the small energy dispersion of $\tilde{\phi}$ as in Remark 6.2 and applying (4.36)-(4.38), (4.39)-(4.41), as well as (6.7), (6.9) for A_0, $\tilde{A}_0$, we obtain

$$\|\Box_A\psi^{high}\|_{L^2\dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}}\dot{H}^{-\frac{4}{9}}[J]} \lesssim_F \varepsilon^{\delta_*}. \tag{6.30}$$

To bound the N norm in (6.27), we introduce a frequency gap m to be chosen later. Then we write the $\Box_A$ equation for ψ^{high} as follows:

$$\Box_A\psi^{high} = (\Box_{\tilde{A}} - \Box_A)\tilde{\phi} \tag{6.31}$$
$$= (\mathcal{M}_{\tilde{A}}^m - \mathcal{M}_A^m)\tilde{\phi} + (\Box_{\tilde{A}^{free}}^{p,m} - \Box_{A^{free}}^{p,m})\tilde{\phi} + (\text{Diff}_{\mathbf{A}}^m(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \text{Diff}_{\mathbf{A}}^m(\phi, \phi, A))\tilde{\phi},$$

where the decomposition $A = A^{free} + A^{nl}$ is performed with respect to some fixed initial time $t_0 \in J$.

In the first term on the right hand side of (6.31), we take advantage of the small energy dispersion of $\tilde{\phi}$ as in Remark 6.2 and apply (4.22)-(4.24), (4.26)-(4.27), as well as (6.7), (6.9) for A_0, $\tilde{A}_0$ (here it is crucial to use the smallness factor ε^{δ_1} in (6.9)), to obtain

$$\|(\mathcal{M}_{\tilde{A}}^m - \mathcal{M}_A^m)\tilde{\phi}\|_{N[J]} \lesssim_F 2^{Cm}\varepsilon^{\delta_*}.$$

To treat the term $(\Box_{\tilde{A}^{free}}^{p,m} - \Box_{A^{free}}^{p,m})\tilde{\phi}$, we expand

$$(\Box_{\tilde{A}^{free}}^{p,m} - \Box_{A^{free}}^{p,m})\tilde{\phi} = 2i\sum_k (\tilde{A}^{free} - A^{free})_{<k-m}^j \partial_j\tilde{\phi}_k$$

Due to the c^* envelope bound (6.7) in Step 1, note that $\|\tilde{\phi}_k\|_{S^1}$ decays at high frequencies $k > k^*$, whereas $P_{k'}(\tilde{A}^{free} - A^{free})[t_0] = P_{k'}(\tilde{A} - A)[t_0]$ decays at low frequencies $k' < k^*$. Observe moreover that the frequency gap m enforces a separation of at least m between the dyadic frequencies of $\tilde{A}^{free} - A^{free}$ and $\tilde{\phi}$. Combined with the null form estimate (4.34), it follows that

$$\|(\Box_{\tilde{A}^{free}}^{p,m} - \Box_{A^{free}}^{p,m})\tilde{\phi}\|_{N[J]} \lesssim_F 2^{-\delta_0 m}.$$

36

Finally, for the third term in (6.31), we use (4.28) and (4.29). As for the second term, the gain comes from the low frequency bound (6.7) from Step 1. This guarantees that, on one hand, $\tilde{\phi}$ decays at high frequencies $> 2^{k^*}$, and on the other hand the differences $(A - \tilde{A}, \phi - \tilde{\phi})$ decay at low frequencies $< 2^{k^*}$. Again, the frequency gap enforces a frequency separation of at least m, so from (4.28) and (4.29) we obtain

$$\|(\mathrm{Diff}_{\mathbf{A}}^m(\tilde{\phi}, \tilde{\phi}, \tilde{A}) - \mathrm{Diff}_{\mathbf{A}}^m(\phi, \phi, A))\tilde{\phi}\|_{N[J]} \lesssim_F 2^{-\delta_0 m}.$$

Summing up the bounds for the three terms in $\Box_A \psi^{high}$, we conclude that

$$\|\Box_A \psi^{high}\|_{N \cap L^2 \dot{H}^{-\frac{1}{2}} \cap L^{\frac{9}{5}} \dot{H}^{-\frac{4}{9}}[J]} \lesssim_F 2^{Cm} \varepsilon^{\delta_*} + 2^{-\delta_0 m}. \tag{6.32}$$

Optimizing the choice of m, the desired estimate (6.27) follows. $\qquad\square$

7. Bilinear null form estimates

We begin our discussion with the bilinear null form estimates, which play a key role in our analysis. These occur in both equations in the MKG-CG system (1.6). In the ϕ equation we have the expression $A^j \partial_j \phi$, under the Coulomb gauge condition $\partial^j A_j = 0$. We can rewrite this as

$$A^j \partial_j \phi = \partial_k \partial^k \Delta^{-1} A^j \partial_j \phi = Q_{kj}(\partial^k \Delta^{-1} A^j, \phi) \tag{7.1}$$

where Q_{kj} is the standard null form

$$Q_{kj}(u, v) = \partial_k u \partial_j v - \partial_j u \partial_k v$$

In the A equation, on the other hand, we encounter the expression

$$\mathcal{P}_j(\phi \partial_x \bar{\phi}) = \phi \partial_j \bar{\phi} - \partial^k \partial_j \Delta^{-1}(\phi \partial_k \bar{\phi}) = \partial^k \Delta^{-1} Q_{kj}(\phi, \bar{\phi}) \tag{7.2}$$

Thus, it suffices to produce good estimates for the null form Q_{ij}. For that we have

Proposition 7.1. *Let $\mathcal{N}$ be one of the Q_{ij} null form. Then the following bilinear estimates hold:*

$$\|P_j \mathcal{N}(\phi_k, \psi_l)\|_N \lesssim 2^j 2^{-\delta(|j-k|+|j-l|)} \|\phi_k\|_{S^1} \|\psi_l\|_{S^1} \tag{7.3}$$

$$\|Q_{<j-m} P_j \mathcal{N}(Q_{<k-m}\phi_k, Q_{<l-m}\psi_l)\|_N \lesssim 2^j 2^{-\delta m} 2^{C(|j-k|+|k-l|)} \|\phi_k\|_{S^1} \|\psi_l\|_{S^1} \tag{7.4}$$

The first estimate above is the bound (131) in [13]. For the second bound we can harmlessly assume that $|j - k| + |k - l| \ll C$. Then (7.4) is a consequence of the bound (143) in [13].

We remark that the first bound (7.3) easily transfers to an interval I. However, the second one involves modulation localizations, which are inconsistent with interval localizations.

We now use the above null form estimates to conclude the proof of all the remaining results in Section 4, except for Proposition 4.7.

Proof of Proposition 4.1. As in Remark 4.2, it suffices to estimate the $\|\Box(\cdot)\|_N$ norms. We begin with the quadratic part $\mathbf{A}_x^2$. For simplicity, we concentrate on the case when the first two inputs are identical; the general case is a minor extension. We have

$$\Box \mathbf{A}_i^2(\phi, \phi) = \partial_j \Delta^{-1} \mathcal{N}(\phi, \bar{\phi})$$

therefore (4.4) follows by dyadic summation from (7.3).

To prove the more refined bounds for $\mathbf{A}_x^2$ we use a large frequency gap m to first split

$$\Box \mathbf{A}_i^2(\phi, \phi) = \sum_{\max\{|k-k_i|\} \geq m} P_k \partial_j \Delta^{-1} \mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2}) + \sum_{|k-k_i| < m} P_k \partial_j \Delta^{-1} \mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2})$$

The first sum is estimated using (7.3) with a 2^{-cm} constant.

The second sum is essentially diagonal, so it suffices to estimate it for fixed k. For this we consider two cases depending on the relative size of the interval I. The case of short intervals $|I| \leq 2^{-k+m}$ is easy to dispense with, as we have

$$\|P_k \partial_j \Delta^{-1} \mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2})\|_{N[I]} \lesssim 2^{-k} \|\mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2})\|_{L^1 L^2[I]}$$
$$\lesssim 2^{Cm} (2^k |I|)^{\frac{1}{2}} \||D|^{\frac{1}{4}} \phi_{k_1}\|_{L^4[I]} \||D|^{\frac{1}{4}} \phi_{k_2}\|_{L^4[I]}.$$

Here we have a large 2^{Cm} constant, but the Strichartz norms on the right is divisible so this suffices for (4.7). Moreover, since $(4, 4)$ is a non-sharp pair of Strichartz exponents, it will be sufficient for (4.8) too, as we explain below.

We are left with the most interesting case. To summarize, we have $|k - k_j| < m$ and $|I| > 2^{-k+Cm}$. To continue the proof we need to use modulation localizations. In order to be able to do that we extend ϕ outside our interval I by homogeneous waves. Then we decompose

$$P_k \partial_j \Delta^{-1} \mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2}) = Q_{<k-Cm} P_k \partial_j \Delta^{-1} \mathcal{N}(Q_{<k-Cm} \phi_{k_1}, Q_{<k-Cm} \bar{\phi}_{k_2}) + err(k, k_1, k_2)$$

where the error corresponds to at least one modulation larger than $k - Cm$. The first term is estimated using (7.4) with a 2^{-cm} constant. For the error we produce instead a direct bound, with two cases:

(i) High modulation output:

$$\|Q_{>k-Cm} P_k \partial_j \Delta^{-1} \mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2})\|_{N[I]} \lesssim 2^{Cm} 2^{-\frac{3k}{2}} \|\chi_I^k \mathcal{N}(\phi_{k_1}, \bar{\phi}_{k_2})\|_{L^2}$$
$$\lesssim 2^{Cm} \|\chi_I^k |D|^{\frac{1}{4}} \phi_{k_1}\|_{L^4} \|\chi_I^k |D|^{\frac{1}{4}} \phi_{k_2}\|_{L^4}$$
$$\lesssim 2^{Cm} \||D|^{\frac{1}{4}} \phi_{k_1}\|_{L^4[I]} \||D|^{\frac{1}{4}} \phi_{k_2}\|_{L^4[I]}$$

where we used Proposition 3.4 on the last line. Here the relaxed cutoff χ_I^k was inserted in order to account for the fact that the operator $Q_{>k-Cm}$ is nonlocal in time. Its kernel decays rapidly on the $2^{Cm} 2^{-k}$ time scale, and this is exactly the scale captured by χ_I^k. Again the Strichartz norms on the right are both divisible and non-sharp, so this bound suffices for both (4.7) and (4.8) (see below).

(ii) One high modulation input:

$$\|Q_{<k-Cm} P_k \mathcal{N}(Q_{>k-Cm} \phi_{k_1}, \bar{\phi}_{k_2})\|_{N[I]} \lesssim 2^{Cm} \|\chi_I^k \mathcal{N}(Q_{>k-Cm} \phi_{k_1}, \bar{\phi}_{k_2})\|_{L^1 L^2}$$
$$\lesssim 2^{Cm} \|\Box \phi_{k_1}\|_{L^2} \|\chi_I^k \phi_{k_2}\|_{L^2 L^\infty}$$

This suffices for (4.7). To complete the proof of (4.8) we also need to account for the case when ϕ_{k_2} has high modulation. Then we have the following small variation of the previous computation:

$$\|Q_{<k-Cm} P_k \mathcal{N}(\phi_{k_1}, \overline{Q_{>k-Cm} \phi_{k_2}})\|_{N[I]} \lesssim 2^{Cm} \|\chi_I^k \mathcal{N}(\phi_{k_1}, \overline{Q_{>k-Cm} \phi_{k_2}})\|_{L^1 L^2}$$
$$\lesssim 2^{Cm} \|\chi_I^k \phi_{k_1}\|_{L^{\frac{9}{4}} L^\infty} \|\Box \phi_{k_2}\|_{L^{\frac{9}{5}} L^2}$$

where the point is that $(\frac{9}{4}, \infty)$ is a non-sharp Strichartz exponent.

38

To conclude the proof of (4.8) we observe that the above estimates allow us to use the ε-energy dispersion and (4.3) for all the large parts of $\mathbf{A}_x^2$. Hence we obtain a bound of the form

$$\|\Box\mathbf{A}_x^2(\phi_1,\phi_2)\|_{N_c[I]} \lesssim (2^{-cm} + 2^{Cm}\varepsilon^{\delta_1})\|\phi_1\|_{S^1[I]}\|\phi_2\|_{S_c^1[I]}$$

Now (4.8) easily follows by optimizing the choice of m.

Finally we consider the cubic terms $\mathbf{A}_i^3$, which satisfy

$$\Box\mathbf{A}_i^3(\phi,\phi,A_x) = \mathcal{P}(\phi\bar\phi A_x)$$

At the dyadic level, using Bernstein's inequality in a favorable way we obtain

$$\|P_k(\phi_{k_1}\bar\phi_{k_2}A_{k_3})\|_{L^1L^2[I]} \lesssim 2^{-\delta\max|k-k_i|}\||D|^{\frac14}\phi_{k_1}\|_{L^4[I]}\||D|^{\frac14}\phi_{k_2}\|_{L^4[I]}\|A_{k_3}\|_{L^2L^8[I]}$$

All norms on the right are Strichartz norms and are bounded by the S^1 norms, so (4.4) follows. Further, if say ϕ_1 is ε-energy dispersed, then we can bound its non-sharp Strichartz norm L^4 using the energy dispersion at the expense of losing the frequency envelope information, in order to obtain (4.9). $\qquad\Box$

Proof of Proposition 4.6. For the leading part

$$\mathcal{M}_A^{m,main} = 2i\sum_k P_{\geq k-m}A^j\partial_j P_k$$

of $\mathcal{M}_A^m$ we have

$$A^j\partial_j\phi = \mathcal{N}(\partial_k\Delta^{-1}A_j,\phi)$$

To decompose it into a small and a large part we first consider the frequency balance of the two inputs and the output, depending on the frequency gap parameter $m \gg 1$:

$$\mathcal{M}_A^{m,main}\phi = 2i\sum_k P_{\geq k+m}A^j\partial_j P_k\phi + P_{<k-m}(P_{<k-m}A^j\partial_j P_k\phi) + P_{\geq k-m}(P_{[k-m,k+m)}A^j\partial_j P_k\phi)$$

The first two terms are estimated with a favorable 2^{-cm} constant using (7.3), and thus placed in $\mathcal{M}_{A,small}^{m,2}$. It remains to consider the last term. This is essentially diagonal in k, so we can freeze the three frequencies in the allowed range.

Now we consider the size of I. As in the proof of Proposition 4.1 there is one easy case, namely when $|I| \leq 2^{-k+m}$. Dispensing with that, from here on we assume that $|I| > 2^{-k+m}$. The remaining argument uses modulation localizations. To allow for that we extend both A and ϕ outside I as free waves. Then we decompose the last term above as

$$P_{\geq k-m}(P_{[k-m,k+m)}A^j\partial_j P_k\phi) = Q_{<k-Cm}P_{\geq k-m}(Q_{<k-Cm}P_{[k-m,k+m)}A^j\partial_j Q_{<k-Cm}P_k\phi)$$
$$+ \mathcal{M}_{A,large}^{m,main}P_k\phi$$

In the first term, we gain 2^{-cm} by (7.4). Hence this part can be put into $\mathcal{M}_{A,small}^{m,2}$. The remaining part $\mathcal{M}_{A,large}^{m,main}$ contains only terms where all three frequencies are balanced, and at least one modulation is large. But this is estimated exactly as in the proof of Proposition 4.1:

(i) If the output has high modulation, then we bound it in L^2 using using divisible non-sharp Strichartz norms to gain either the divisible bound (4.25), or smallness via ε-energy dispersion as in (4.26).

(ii) If the second input (i.e., ϕ) has high modulation, then we combine the L^2L^∞ bound for A with the L^2 bound for $\Box\phi$.

(iii) If the first input (i.e., A) has high modulation, then we gain both divisibility and smallness via energy dispersion by combining an $L^{\frac{9}{5}}L^2$ bound for $\Box A$ and an $L^{\frac{9}{4}}L^\infty$ bound for ϕ.

We now consider the remaining terms in $\mathcal{M}_A^m$. For the terms $P_{\geq k-m}A_0\partial_t\phi_k$ and $\partial_t A_0\phi$ we estimate

$$\|P_k\sum_{k_2}P_{\geq k-m}A_0\partial_t\phi_{k_2}\|_{L^1L^2[I]} \lesssim 2^{-\delta\max\{|k-k_i|\}}\|P_{k_1}A_0\|_{L^2\dot{H}^{\frac{3}{2}}[I]}\||D|^{-1}\partial_t\phi_{k_2}\|_{L^2L^8[I]}$$

respectively

$$\|P_k(P_{k_1}\partial_t A_0\phi_{k_2})\|_{L^1L^2[I]} \lesssim 2^{-\delta\max\{|k-k_i|\}}\|P_{k_1}\partial_t A_0\|_{L^2\dot{H}^{\frac{1}{2}}[I]}\|\phi_{k_2}\|_{L^2L^8[I]}.$$

Finally, the term $A^\alpha A_\alpha\phi$ is estimated in L^1L^2 with off-diagonal gain using only divisible non-endpoint Strichartz estimates, which suffices. $\qquad\square$

8. Multilinear null form estimates

In this section we discuss directly the bounds for the operator Diff_A^m, and prove Proposition 4.7. The bounds (4.28) and (4.29) were already proved in [13]. The delicate matter is to be able to estimate the bulk of $\text{Diff}_A^{m,2}(\phi,\phi)\psi$ in terms of the divisible norm DS^1 of ϕ. We split our argument into two steps:

(i) First we review the decompositions and the estimates in [13] on the full real line, leading to the proof of (4.28) and (4.29). But we do this in a careful fashion so that we can isolate a bulk part where we get smallness from the frequency gap, and a remaining part where this does not work. For this remaining part we can easily produce a divisible bound. Unfortunately, this last argument uses modulation localizations.
(ii) Secondly, we consider the changes in the previous arguments when the analysis is done on a compact interval I. The challenge here is to be able to accurately estimate the large but divisible part using only information localized to our interval.

8.1. A review of [13].
We decompose Diff_A^m into

$$\text{Diff}_A^m = \mathcal{H}^*\text{Diff}_A^m + (I - \mathcal{H}^*)\text{Diff}_A^m$$

where the operator $\mathcal{H}^*$, introduced in [13], selects the case where $\mathbf{A}$ has high modulation while both the input and the output have small modulation,

$$\mathcal{H}^*\text{Diff}_A^m\psi = \sum_{k_0<k-m}\sum_{j<k_0}Q_{<j-2}(Q_jP_{k_0}\mathbf{A}^\alpha Q_{<j-2}\partial_\alpha\psi_k)$$

The better part $(I - \mathcal{H}^*)\text{Diff}_A^m$ can be still dealt with in a bilinear fashion using the following result:

Proposition 8.1. *We have the bilinear estimate*

$$\|(I - \mathcal{H}^*)\text{Diff}_A^m\psi\|_N \lesssim (\|A_x\|_{\ell^1S^1} + \|\nabla A_0\|_{\ell^1L^2\dot{H}^{\frac{1}{2}}})\|\psi\|_{S^1} . \tag{8.1}$$

Further, we have the low modulation improvement

$$\|(I - \mathcal{H}^*)\text{Diff}_{A^{low}}^m\psi\|_N \lesssim 2^{-cm}(\|A_x\|_{\ell^1S^1} + \|\nabla A_0\|_{\ell^1L^2\dot{H}^{\frac{1}{2}}})\|\psi\|_{S^1} \tag{8.2}$$

where

$$A^{low} = \sum_{k_0} Q_{<k_0-m} P_{k_0} A.$$

The bound (8.1) is the sum of the bounds (54) and (58) in [13]. The bound (8.2) is a corollary of the proof of (54) in [13]; it follows from a similar improvement in the bound (132) in [13], which in turn is tied to the fact that the estimate (143) in [13] is only used in the case when $j_1 < k_2 - m$.

We now turn our attention to the term $\mathcal{H}^*\text{Diff}_{\mathbf{A}}^m$, where it is no longer enough to obtain bounds depending on the above norms of $\mathbf{A}$. Our first tool here is the intermediate norm Z, which has the following properties:

Proposition 8.2. *We have the bilinear estimates:*

$$\|\mathcal{H}^*\text{Diff}_{A_x}^m\psi\|_N \lesssim \|A_x\|_Z\|\psi\|_{S^1} \tag{8.3}$$

respectively

$$\|\mathcal{H}^*\text{Diff}_{A_0}^m\psi\|_N \lesssim \|A_0\|_{\Delta^{-\frac{1}{2}}\square^{\frac{1}{2}}Z+L^1L^\infty}\|\psi\|_{S^1} \tag{8.4}$$

as well as the low modulation improvement

$$\|\mathcal{H}^*\text{Diff}_{A_x^{low}}^m\psi\|_N \lesssim 2^{-cm}\|A_x\|_Z\|\psi\|_{S^1} \tag{8.5}$$

respectively

$$\|\mathcal{H}^*\text{Diff}_{A_0^{low}}^m\psi\|_N \lesssim 2^{-cm}\|A_0\|_{\Delta^{-\frac{1}{2}}\square^{\frac{1}{2}}Z}\|\psi\|_{S^1} . \tag{8.6}$$

These are the bounds (133) and (140) in [13], where the low modulation improvement is again apparent from the proofs.

Combining the estimates (8.1), (8.3), (8.4) and (3.6) we can eliminate the modulation localizations and obtain

Corollary 8.3. *The following estimate holds:*

$$\|\text{Diff}_A^m\psi\|_N \lesssim (\|A_x\|_{\ell^1 S^1} + \|\square A_x\|_{\ell^1 L^1 L^2} + \|\nabla A_0\|_{\ell^1 L^2 \dot{H}^{\frac{1}{2}}} + \|A_0\|_{L^1 L^\infty})\|\psi\|_{S^1} \tag{8.7}$$

Using this estimate, we can dispense with the cubic contributions due to $\mathbf{A}^3 = (\mathbf{A}_0^3, \mathbf{A}_x^3)$. Indeed, combined with the bounds (4.5), (4.11) and (4.16), as well as (3.8) to control $\|\mathbf{A}_x^3\|_{\ell^1 S^1}$, we can use (8.7) to establish (4.29). We remark that the frequency envelope bound in (4.29) is clear from the frequency gap m between the two inputs $\mathbf{A}^3$ and ϕ.

The output of the quadratic part of $\mathbf{A}$ cannot be all dealt with using the Z norm, but a good portion of it is amenable to this strategy. This is described using the operators $\mathcal{H}_{k_0}$ defined by

$$\mathcal{H}_{k_0}\mathbf{A}^2(\phi_{k_1}, \phi_{k_2}) = \sum_{j<k_0} Q_j P_{k_0} \mathbf{A}^2(Q_{<j}\phi_{k_1}, Q_{<j}\phi_{k_2})$$

Precisely, the portion of $\mathbf{A}^2$ which does not have good Z bounds is

$$\mathcal{H}^m\mathbf{A}^2 = \sum_{k_0<k_1-m} \mathcal{H}_{k_0}\mathbf{A}^2(\phi_{k_1}, \phi_{k_2})$$

A key result in [13] is to treat the output of this part in a genuine trilinear fashion, taking advantage of a cancellation between the A_0 and A_x parts, which have otherwise been treated separately. Precisely, we have

41

Proposition 8.4. *For any admissible frequency envelopes c, d, e, we have*

$$\|\mathcal{H}^*\text{Diff}^m_{\mathcal{H}^m\mathbf{A}^2(\phi_1,\phi_2)}\psi\|_{N_f} \lesssim 2^{-cm}\|\phi_1\|_{S^1_c}\|\phi_2\|_{S^1_d}\|\psi\|_{S^1_e}, \tag{8.8}$$

where $f(k)$ is as in (4.30).

For this we refer the reader to the estimate (60) in [13] and its dyadic versions (136)-(138), where the frequency envelope bound and the gain with respect to m are apparent.

Hence it remains to bound

$$\|(I - \mathcal{H}^m)\mathbf{A}^2_x(\phi,\phi)\|_{Z+L^1L^\infty}, \qquad \|(I - \mathcal{H}^m)\mathbf{A}^2_0(\phi,\phi)\|_{\Delta^{-\frac{1}{2}}\square^{\frac{1}{2}}Z+L^1L^\infty}.$$

Considering the dyadic portions

$$P_{k_0}\mathbf{A}^2(\phi_{k_1},\phi_{k_2}),$$

the case of high-high interactions was also discussed in [13]. Precisely, from the bounds (134) and (141) in [13] we have

Proposition 8.5. *For $k_0 < k_1 - C$, we have the dyadic bound*

$$\|(I - \mathcal{H}^m)P_{k_0}\mathbf{A}^2_x(\phi_{k_1},\phi_{k_2})\|_Z + \|(I - \mathcal{H}^m)P_{k_0}\mathbf{A}^2_0(\phi_{k_1},\phi_{k_2})\|_{\Delta^{-\frac{1}{2}}\square^{\frac{1}{2}}Z}$$
$$\lesssim 2^{-\delta|k_0-k_1|}\|\phi_{k_1}\|_{S^1}\|\phi_{k_2}\|_{S^1}. \tag{8.9}$$

This suggests that we should decompose $\mathbf{A}^2$ into a *high $\times$ high $\to$ low* portion and a better reminder. We will be more accurate and set

$$\mathbf{A}^2 = \sum_{k_1 \geq k_0+m} P_{k_0}\mathbf{A}^2(\phi_{k_1},\phi_{k_2}) + \sum_{k_0-m<k_{1,2}<k_0+m} P_{k_0}\mathbf{A}^2(\phi_{k_1},\phi_{k_2}) + \sum_{k_{min}<k_0-m} P_{k_0}\mathbf{A}^2(\phi_{k_1},\phi_{k_2})$$
$$:= \mathbf{A}^{2,hh(m)} + \mathbf{A}^{2,med(m)} + \mathbf{A}^{2,hl(m)},$$

where $k_{min} = \min\{k_1, k_2\}$. Note that no modulation localizations are present here.

We first handle the part $\mathbf{A}^{2,hh(m)}$. Recall from the proofs of (4.4), (4.10), (4.16) and (4.17) that there is a bound with an off-diagonal decay of the form

$$\|P_{k_0}\mathbf{A}^2_x(\phi_{k_1},\phi_{k_2})\|_{S^1} + \|P_{k_0}\mathbf{A}^2_0(\phi_{k_1},\phi_{k_2})\|_{L^2\dot{H}^{\frac{3}{2}}} \lesssim 2^{-\delta(|k_0-k_1|+|k_0-k_1|)}\|\phi_{k_1}\|_{S^1}\|\phi_{k_2}\|_{S^1} \tag{8.10}$$

when $k_0 < k_1 - C$.

Combining the bounds (8.8), (8.9) with (8.3), (8.4) for the $\mathcal{H}^*$ portion and (8.10) with (8.1) for the $I - \mathcal{H}^*$ portion, we obtain the following:

Corollary 8.6. *For any admissible frequency envelopes c, d, e, we have*

$$\|\text{Diff}^m_{\mathbf{A}^{2,hh(m)}}(\phi_1,\phi_2)\psi\|_{N_f} \lesssim 2^{-cm}\|\phi_1\|_{S^1_c}\|\phi_2\|_{S^1_d}\|\psi\|_{S^1_e} \tag{8.11}$$

where $f(k)$ is as in (4.30).

Again, no modulation localizations are present here.

The remaining parts of $\mathbf{A}^2$ have no contributions from $\mathcal{H}^m\mathbf{A}^2$, so we will estimate them entirely using the Z norm or the simpler L^1L^∞ bound. The latter suffices in the case of $\mathbf{A}_0$, whose dyadic pieces are readily bounded by

$$\|P_{k_0}\mathbf{A}^2_0(\phi_{k_1},\phi_{k_2})\|_{L^1L^\infty} \lesssim 2^{-2k_0}\|P_{k_0}\Delta\mathbf{A}^2_0(\phi_{k_1},\phi_{k_2})\|_{L^1L^\infty}$$
$$\lesssim 2^{-2(k_0-k_{max})-\frac{1}{2}|k_1-k_2|}\prod_{j=1,2}\|(|D|^{-\frac{1}{2}}\phi_{k_j},|D|^{-\frac{3}{2}}\partial_t\phi_{k_j})\|_{L^2L^\infty} \tag{8.12}$$

42

where $k_{max} = \max\{k_1, k_2\}$. In both $\mathbf{A}_0^{2,hl(m)}$ and $\mathbf{A}_0^{2,med(m)}$, note that we have $k_{max} \leq k_0 + m$ for (say) $m \geq 3$. Moreover, the above dyadic bound sums up easily due to the off-diagonal decay. Using the embedding $L^1L^2 \subset N$, we easily obtain

Corollary 8.7. *For any admissible frequency envelopes c, d, e, we have*

$$\|\mathrm{Diff}^m_{\mathbf{A}_0^{2,hl(m)}}(\phi_1, \phi_2)\psi\|_{N_f} + \|\mathrm{Diff}^m_{\mathbf{A}_0^{2,med(m)}}(\phi_1, \phi_2)\psi\|_{N_f} \lesssim 2^{Cm}\|\phi_1\|_{DS_c^1}\|\phi_2\|_{DS_d^1}\|\psi\|_{S_e^1} \quad (8.13)$$

where $f(k)$ is as in (4.30).

We now consider the contributions of $\mathbf{A}_x^{2,hl(m)}$ and $\mathbf{A}_x^{2,med(m)}$. Our first tool is due to the estimates (134) and (135) in [13], which give

Proposition 8.8. *The following estimate holds:*

$$\|P_{k_0}\mathbf{A}_x^2(\phi_{k_1}, \phi_{k_2})\|_Z \lesssim 2^{C|k_0-k_{max}|}2^{-\delta|k_1-k_2|}\|\phi_{k_1}\|_{S^1}\|\phi_{k_2}\|_{S^1} \quad (8.14)$$

This gives a gain for the high-low portion of A_x. Hence in combination with (8.3), (8.4), (8.9) for the $\mathcal{H}^*$ portion and (8.1), (8.10) for the $I - \mathcal{H}^*$ portion, we obtain a result with no modulation localizations:

Corollary 8.9. *For any admissible frequency envelopes c, d, e, we have*

$$\|\mathrm{Diff}^m_{\mathbf{A}^{2,hl(m)}}(\phi_1, \phi_2)\psi\|_{N_f} \lesssim 2^{-cm}\|\phi_1\|_{S_c^1}\|\phi_2\|_{S_d^1}\|\psi\|_{S_e^1} \quad (8.15)$$

where $f(k)$ is as in (4.30).

Finally, it remains to consider the contribution of $\mathbf{A}_x^{2,med(m)}$. There the estimate (8.14) suffices for the bound (4.29), but provides no divisible norm estimate. To summarize, we are left with the case

$$k_0 - m < k_1, k_2 < k_0 + m.$$

Here we can take advantage of the low modulation decay in (8.2) and (8.5) to obtain

Corollary 8.10. *The following bound holds for large enough C and k_0, k_1, k_2 as above:*

$$\|\mathrm{Diff}^m_{\mathbf{A}_x^{2,med(m),low}}(\phi_{k_1}, \phi_{k_2})\psi_k\|_N \lesssim 2^{-cm}\|\phi_{k_1}\|_{S^1}\|\phi_{k_2}\|_{S^1}\|\psi_k\|_{S^1} \quad (8.16)$$

where

$$\mathbf{A}_x^{2,med(m),low} = \sum_{k_0} Q_{<k_0-m}P_{k_0}\mathbf{A}_x^{2,med(m)}\ .$$

Thus we can restrict ourselves to high modulations in $\mathbf{A}_x$, i.e.,

$$\mathbf{A}_x^{2,med(m),high}(\phi_1, \phi_2) = \sum_{k_0} P_{k_0}Q_{>k_0-Cm}\mathbf{A}_x^{2,med(m)}(\phi_1, \phi_2).$$

For this part, we can use the L^1L^∞ norm. Precisely, each dyadic piece obeys the estimate

$$\|P_{k_0}Q_{>k_0-Cm}\mathbf{A}_x^2(\phi_{k_1}, \phi_{k_2})\|_{L^1L^\infty} \lesssim 2^{-2k_0+Cm}\|\Box\mathbf{A}_x^2(\phi_{k_1}, \phi_{k_2})\|_{L^1L^\infty}$$

$$\lesssim 2^{Cm}\||D|^{-\frac{1}{2}}\phi_{k_1}\|_{L^2L^\infty}\||D|^{-\frac{1}{2}}\phi_{k_2}\|_{L^2L^\infty}\ .$$

Recall that we are in the scenario $k_0 - m < k_1, k_2 < k_0 + m$. Combined with the embeddings $L^1L^2 \subset N$ and $\nabla S^1 \subset L^\infty L^2$, we obtain

$$\|\mathrm{Diff}^m_{\mathbf{A}_x^{2,med(m),high}(\phi_1, \phi_2)}\psi\|_{N_f} \lesssim 2^{Cm}\|\phi\|_{DS_c^1}\|\phi\|_{DS_d^1}\|\psi_k\|_{S_e^1}$$

43

where c, d, e are any admissible frequency envelopes and f is as in (4.30). Thus the proof of Proposition 4.7 is concluded on the entire real line.

8.2. Interval localized bounds.

Here we seek to prove the result of Proposition 4.7 in a time interval I. Due to the paradifferential nature of the operator $\mathrm{Diff}_{\mathbf{A}}^m$, we can fix the frequency 2^k of the input ψ and simply estimate the expression $\mathrm{Diff}_{\mathbf{A}}^m \psi_k$. For $\mathbf{A}$ we consider its components successively:

a) The cubic terms $\mathbf{A}^3$. Here we simply extend $\mathbf{A}^3$ outside I as a homogeneous wave, and then use the bound (8.7). By Propositions 4.1, 4.4 and 4.5, we know that $\mathbf{A}^3$ is entirely estimated by divisible norms.

b) The contributions of $\mathbf{A}^{2,hh(m)}$ and $\mathbf{A}_x^{2,hl(m)}$. Here we extend ϕ outside I as a homogeneous wave, and then apply (8.11), respectively (8.15).

c) The contributions of $\mathbf{A}_0^{2,med(m)}$ and $\mathbf{A}_0^{2,hl(m)}$. These are estimated directly via (8.13); no extensions are necessary.

d) The contribution of $\mathbf{A}_x^{2,med(m)}$. This is the part where the divisible bound is more difficult to gain. In what follows, we simply write $\mathbf{A}_x = \mathbf{A}_x^2$. To review, we have to estimate the expression

$$\|P_{k_0}\mathbf{A}^i(\phi_{k_1}, \phi_{k_2})\partial_i \psi_k\|_N$$

where the frequency balance is

$$k_0 < k - m, \qquad k_0 - m < k_1, k_2 < k_0 + m .$$

This is where the length of the time interval I plays a role. Comparing it to k_0, we distinguish two scenarios:

(i) Short time intervals, $|I| \leq 2^{-k_0+m}$. Then we have a direct estimate,

$$
\begin{aligned}
\|P_{k_0}\mathbf{A}^i(\phi_{k_1}, \phi_{k_2})\partial_i \psi_k\|_{N[I]} &\lesssim \|P_{k_0}\mathbf{A}^i(\phi_{k_1}, \phi_{k_2})\partial_i \psi_k\|_{L^1 L^2[I]} \\
&\lesssim |I|^{\frac{1}{2}}\|P_{k_0}\mathbf{A}_x(\phi_{k_1}, \phi_{k_2})\|_{L^2 L^\infty[I]}\|\nabla_x \psi_k\|_{L^\infty L^2[I]} \\
&\lesssim 2^{Cm}\||D|^{-\frac{1}{2}}P_{k_0}\mathbf{A}_x(\phi_{k_1}, \phi_{k_2})\|_{L^2 L^\infty[I]}\|\nabla_x \psi_k\|_{L^\infty L^2[I]} .
\end{aligned}
$$

Summing over k_0, k_1, k_2 and recalling the definition of the $S_{k_0}^{str} \subset S_{k_0}$, we obtain

$$\|\mathrm{Diff}_{\mathbf{A}_x^{2,med(m)}(\phi_1,\phi_2)}^m \psi_k\|_{N[I]} \lesssim 2^{Cm}\|\nabla P_{<k-m}\mathbf{A}_x^{2,med(m)}(\phi_1, \phi_2)\|_{\ell^1 S[I]}\|\psi_k\|_{S^1[I]},$$

The right hand side can be controlled by Proposition 4.1. The splitting into small and large parts is then achieved by using the corresponding statements (4.6)-(4.7) for $\mathbf{A}_x^2$. To be precise, (4.6)-(4.7) apply to the full operator $\mathbf{A}_x^2$. Nevertheless, $\mathbf{A}_x^2 - \mathbf{A}_x^{2,med(m)}$ gains 2^{-cm} by (8.10), and thus this difference can be put into the 'small' part.

(ii) Long time intervals, $|I| > 2^{-k_0+m}$. This is the difficult case. Our proof here involves modulation localizations, so we need to consider appropriate extensions of $\mathbf{A}_x$ and ψ_k. Since ψ_k is an independent variable, for it we can simply use the canonical extension as homogeneous waves. For $\mathbf{A}_x$, instead, we extend its arguments ϕ_1 and ϕ_2 as homogeneous waves.

The bound (8.16) suffices for low modulations of $\mathbf{A}^{2,med(m)}$, therefore it suffices to estimate

$$\sum_{k_0 < k-m} \|Q_{>k_0-Cm} P_{k_0} \mathbf{A}_x(\phi_{k_1}, \phi_{k_2}) \nabla_x \psi_k\|_{N[I]} \lesssim$$

$$\sum_{k_0 < k-m} \|Q_{>k_0-Cm} P_{k_0} \mathbf{A}_x(\phi_{k_1}, \phi_{k_2})\|_{L^1 L^\infty[I]} \|\nabla_x \psi_k\|_{L^\infty L^2}$$

To estimate the localized $L^1 L^\infty[I]$ norm we write

$$Q_{>k_0-Cm} P_{k_0} \mathbf{A}_x(\phi_{k_1}, \phi_{k_2}) = Q_{>k_0-Cm} P_{k_0} \Box^{-1}\Big(\frac{1}{2}\mathcal{P}_x \mathrm{Im}(\phi_{k_1} \nabla_x \overline{\phi}_{k_2} + \nabla_x \overline{\phi}_{k_1} \phi_{k_2})\Big)$$

$$= 2^{-k_0+2Cm} L(\phi_{k_1}, \phi_{k_2})$$

where L is a bilinear translation invariant form whose kernel is localized near 0 on the 2^{-k_0} scale in space-time. This allows us to estimate the tails outside I as follows:

$$\|Q_{>k_0-Cm} P_{k_0} \mathbf{A}_x(\phi_{k_1}, \phi_{k_2})\|_{L^1 L^\infty[I]} \lesssim 2^{-k_0} 2^{Cm} \|\chi_I^{k_0} \phi_{k_1}\|_{L^2 L^\infty} \|\chi_I^{k_0} \phi_{k_2}\|_{L^2 L^\infty}.$$

Since k_1 and k_2 are close to k_0, we conclude using Proposition 3.4 that

$$\|Q_{>k_0-Cm} P_{k_0} \mathbf{A}_x(\phi_{k_1}, \phi_{k_2})\|_{L^1 L^\infty[I]} \lesssim 2^{Cm} \|\phi_{k_1}\|_{DS^1[I]} \|\phi_{k_2}\|_{DS^1[I]}$$

which is the sought after divisible bound. The proof of Proposition 4.7 is concluded.

9. The paradifferential parametrix

The goal of this section is to prove Theorem 4.9. Instead of producing an exact solution operator, it is easier to produce parametrix with small errors. Then the exact solution is obtained in a straightforward iterative fashion. The result we produce here is as follows:

Theorem 9.1. *Let A_x be a Coulomb magnetic potential solving the free wave equation with energy E^2, and let $m > 5$. Consider any finite energy initial data (ϕ_0, ϕ_1) localized in frequency ≈ 1, and a source $f \in N$ which is localized in frequency ≈ 1 and modulation $\lesssim 1$. Then there exists an approximate solution ϕ so that*

$$\|\phi\|_{S_0} \lesssim_E \|(\phi_0, \phi_1)\|_{L^2 \times L^2} + \|f\|_{N_0},$$
$$\|\phi[0] - (\phi_0, \phi_1)\|_{L^2 \times L^2} + \|\Box_A^{p,m}\phi - f\|_{N_0} \lesssim_E 2^{-cm}(\|(\phi_0, \phi_1)\|_{L^2 \times L^2} + \|f\|_{N_0}). \tag{9.1}$$

We remark that the frequency support of the approximate solution ϕ is only slightly larger compared to ϕ_0, ϕ_1 and f; it is essentially also localized at frequency ≈ 1 and modulation $\lesssim 1$. After choosing m sufficiently large, Theorem 9.1 directly implies Theorem 4.9; see [13, Proof of Theorem 6.3].

The definition of our parametrix is identical to the one used for the small data problem in [13], which was based on [19]. The main difference is in the source of smallness for the errors. In [13] this comes from the smallness of the energy of A. Here, we rely instead on the frequency gap m, which must be large in terms of the energy E.

The parametrix is constructed using pseudodifferential operators with rough symbols. Given a symbol $a(t, x, \tau, \xi)$, its left- and right-quantizations are denoted $a(t, x, D)$ and $a(D, y, s)$, respectively. We also use the standard convention $D_\mu = \frac{1}{i}\partial_\mu$.

To prove the theorem it suffices to consider initial data $\phi[0]$ and source f with frequency localization in ≈ 1, and construct the approximate solution ϕ with a similar localization. Thus we work with the unit-frequency localized paradifferential magnetic wave operator

$$\Box_A^{p,m} = \Box + 2iA_{<-m}^j \partial_j. \tag{9.2}$$

where A solves the free wave equation $\Box A = 0$ with initial data $A[0] \in \dot{H}_x^1 \times L_x^2$.

Given an additional small angular localization parameter $0 < \sigma < 1/2$, we construct a parametrix for (9.2) as follows. For $\xi \in \mathbb{R}^4$ we define

$$\omega := \frac{\xi}{|\xi|}, \qquad L_\pm^\omega := \partial_t \pm \omega \cdot \nabla_x, \qquad \Delta_{\omega^\perp} := \Delta - (\omega \cdot \nabla_x)^2.$$

Note that

$$\Box = -L_+^\omega L_-^\omega + \Delta_{\omega^\perp}.$$

Define the angular sector projection $\Pi_{>\theta}^\omega$ by the formula

$$\mathcal{F}[\Pi_{>\theta}^\omega f](\xi) := \left(1 - \eta(\frac{\angle(\xi,\omega)}{\theta})\right)\left(1 - \eta(\frac{\angle(-\xi,\omega)}{\theta})\right)\widehat{f}(\xi).$$

It is important to note that if f is real, then so is $\Pi_{>\theta}^\omega f$. We also define

$$\Pi_\theta^\omega := \Pi_{>\frac{\theta}{2}}^\omega - \Pi_{>\theta}^\omega, \qquad \Pi_{\leq\theta}^\omega := 1 - \Pi_{>\theta}^\omega.$$

For each $\ell \leq 0$, we define $\psi_{\ell,\pm}$ to be

$$\psi_{\ell,\pm}(t,x,\xi) := \mp L_\pm^\omega \Delta_{\omega^\perp}^{-1} \Pi_{>2^{\sigma\ell}}^\omega(\omega \cdot P_\ell A) \tag{9.3}$$

The full phase $\psi_\pm$ is then defined to be

$$\psi_\pm(t,x,\xi) = \psi_{<-m,\pm}(t,x,\xi) := \sum_{\ell<-m} \psi_{\ell,\pm}(t,x,\xi). \tag{9.4}$$

Note that we have

$$L_\mp^\omega \psi_\pm = \mp \sum_{\ell<-m} \Pi_{>2^{\sigma\ell}}^\omega(\omega \cdot P_\ell A) \tag{9.5}$$

In other words, $\psi_\pm$ represent roughly the output of the integration of the (bulk of the) magnetic potential A along light rays. Here we exclude the output of small angle interactions, which is on one hand perturbative, and on the other hand would yield a bad dependence of $\psi_\pm$ on ξ. This is akin to symbol smoothing for rough pdo's.

We use the pseudodifferential gauge transform

$$e_{<0}^{-i\psi_\pm}(t,x,D) := (S_{<0}e^{-i\psi_\pm})(t,x,D)$$

where $S_{<0}$ is taken with respect to the (t,x) variables of the symbol. Its dual is

$$e_{<0}^{i\psi_\pm}(D,y,s)$$

As the symbol is independent of the time Fourier variable $\xi_0 = \tau$, we see that the left and right quantizations with respect to t are the same, i.e.,

$$e_{<0}^{i\psi_\pm}(D,y,s) = e_{<0}^{i\psi_\pm}(t,D,y).$$

46

The operators $e_{<0}^{-i\psi_\pm}(t,x,D)$, respectively $e_{<0}^{i\psi_\pm}(D,y,s)$ are used on the left and on the right in order to conjugate the paradifferential operator $\Box_A^{p,m}$ to the d'Alembertian $\Box$. Precisely, our parametrix is given by

$$
\begin{aligned}
\phi(t,x) = {}& e_{<0}^{-i\psi_\pm}(t,x,D)|D|^{-1}e^{\pm it|D|}e_{<0}^{i\psi_\pm}(D,y,0)(|D|\phi_0 \pm \phi_1) \\
& + \frac{1}{2}\int_0^t e_{<0}^{-i\psi_\pm}(t,x,D)|D|^{-1}e^{\pm i(t-s)|D|}e_{<0}^{i\psi_\pm}(D,y,s)Q_\pm f(s)ds
\end{aligned}
\tag{9.6}
$$

To show that the above parametrix satisfies the bounds in Theorem 9.1 we need the following mapping properties for the operators $e_{<0}^{-i\psi_\pm}(t,x,D)$, respectively $e_{<0}^{i\psi_\pm}(D,y,s)$:

Theorem 9.2. *For $m > 0$, let $\psi_\pm$ be defined as in (9.4). Then the following mapping properties hold with $Z \in \{N_0, L^2, N_0^*\}$, with implicit constants which depend on the energy E of A:*

(1) (Boundedness)

$$
e_{<0}^{\pm i\psi_\pm{}'}(t,x,D): \quad Z \to Z
\tag{9.7}
$$

(2) (Dispersive estimates)

$$
e_{<0}^{\pm i\psi_\pm{}'}(t,x,D): \quad S_0^\sharp \to S_0
\tag{9.8}
$$

(3) (Derivative bounds)

$$
(\nabla e_{<0}^{\pm i\psi_\pm})(t,x): Z \to 2^{-m}Z
\tag{9.9}
$$

(4) (Approximate unitarity of $e^{i\psi_\pm}$ on L_x^2) For each $t \in \mathbb{R}$, we have

$$
e_{<0}^{-i\psi_\pm}(t,x,D)e_{<0}^{i\psi_\pm}(t,D,y) - I: L_x^2 \to 2^{-(1-\delta_0)m}L_x^2
\tag{9.10}
$$

(5) (Approximate unitarity of $e^{i\psi_\pm}$ on N) We have

$$
e_{<0}^{-i\psi_\pm}(t,x,D)e_{<0}^{i\psi_\pm}(D,y,s) - I: N_0 \to 2^{-\delta_1 m}N_0
\tag{9.11}
$$

(6) (Parametrix error estimate) We have

$$
e_{<0}^{-i\psi_\pm}(t,x,D)\Box - \Box_A^p e_{<0}^{-i\psi_\pm}(t,x,D): S_{0,\pm}^\sharp \to 2^{-\delta_2 m}N_{0,\pm}
\tag{9.12}
$$

Remark 9.3. The small constants $\sigma, \delta_0, \delta_1, \delta_2$ and δ are now *different* from those used in the earlier part of the paper. They are chosen in the following logical order: σ, δ_0, δ_1, δ_2. On the other hand, we reserve the symbol $\delta > 0$ for a free small number, whose value may vary depending on the usage.

This result mirrors Theorem 3 in [13], with the key difference that the smallness is now due to the frequency gap parameter m. Assuming these bounds, the conclusion of Theorem 9.1 follows in the same way as in [13].

To prove the above theorem, we may directly borrow the estimates from [13] which do not involve smallness, namely (9.7) and (9.8). The implicit constant in these estimates will now depend on the energy E of A. The remainder of the section is devoted to the proof of the new bounds (9.9), (9.10), (9.11) and (9.12).

9.1. **Review of decomposability calculus.** Here we give a brief review of the notion of *decomposable symbols* developed in [19, 12, 13], which provides a convenient way to keep track of mixed $L_t^q L_x^r$-type bounds. The particular version we use is from [12, 13].

Given $\theta \in 2^{\mathbb{Z}_-}$, where $\mathbb{Z}_-$ denotes the set of nonpositive integers, consider a covering of the unit sphere $\mathbb{S}^3 = \{\xi : |\xi| = 1\} \subseteq \mathbb{R}^4$ by solid angular caps of the form $\{\xi \in \mathbb{S}^3 : |\phi - \frac{\xi}{|\xi|}| < \theta\}$ with uniformly finite overlaps. We enumerate these caps by the centers $\phi \in \mathbb{S}^3$, and denote by $\{b_\theta^\phi(\xi)\}_\phi$ the associated smooth partition of unity on $\mathbb{S}^3$.

Consider a smooth symbol $c(t, x; \xi)$ which is homogeneous of degree zero in ξ, i.e., depends only on the angular variable $\omega := \frac{\xi}{|\xi|}$. We say that $c(t, x; \xi)$ is *decomposable in* $L_t^q L_x^r$ (where $1 \leq q, r \leq \infty$) if there exists an expansion $c = \sum_{\theta \in 2^{\mathbb{Z}_-}} c^{(\theta)}$ such that

$$\sum_{\theta \in 2^{\mathbb{Z}_-}} \|c^{(\theta)}\|_{D_\theta L_t^q L_x^r} < \infty, \tag{9.13}$$

where

$$\|c^{(\theta)}\|_{D_\theta L_t^q L_x^r} := \Big\| \Big(\sum_{k=0}^{40} \sum_\phi \sup_\omega \|b_\theta^\phi(\omega) \theta^k \partial_\xi^{(k)} c^{(\theta)}\|_{L_x^r}^2 \Big)^{\frac{1}{2}} \Big\|_{L_t^q}. \tag{9.14}$$

We denote the class of such symbols by $DL_t^q L^r$. For $c \in DL_t^q L_x^r$, we define the norm $\|c\|_{DL_t^q L_x^r}$ by taking the infimum of (9.13) over all possible decompositions $c = \sum_{\theta \in 2^{\mathbb{Z}_-}} c^{(\theta)}$.

The class $DL_t^q L_x^r$ provides a convenient framework for establishing $L_t^q L_x^r$-type estimates for pseudo-differential operators arising from products of symbols. The following lemma collects the key properties that we need.

Lemma 9.4. *The following statements concerning the class $DL_t^q L_x^r$ hold.*
(1) For any symbols $c \in DL^{q_1} L^{r_1}$ and $d \in DL^{q_2} L^{r_2}$, its product obeys the Hölder-type bound

$$\|cd\|_{DL_t^q L_x^r} \lesssim \|c\|_{DL_t^{q_1} L_x^{r_1}} \|d\|_{DL_t^{q_2} L_x^{r_2}}$$

where $1 \leq q_1, q_2, q, r_1, r_2, r \leq \infty$, $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{q}$ and $\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r}$.
(2) Let $a(t, x; \xi)$ be a smooth symbol whose left-quantization $a(t, x; D)$ satisfies the fixed time bound

$$\sup_t \|a(t, x; D)\|_{L_x^2 \to L_x^2} \leq A.$$

Then for any symbol $c \in DL_t^q L_x^r$, we have the space-time bounds

$$\|(ac)(t, x; D)\|_{L_t^{q_1} L_x^2 \to L_t^{q_2} L_x^{r_2}} \lesssim A \|c\|_{DL_t^q L_x^r}$$

where $1 \leq q_1, q_2, q, r_2, r \leq \infty$, $\frac{1}{q_1} + \frac{1}{q} = \frac{1}{q_2}$ and $\frac{1}{2} + \frac{1}{r} = \frac{1}{r_2}$. An analogous statement holds in the case of right-quantization.

For a proof, see [12, Chapter 10] and [13, Lemma 7.1].

We borrow another lemma from [13], which relates the product of quantized operators with the product of the corresponding symbols within the framework of decomposable symbols.

Lemma 9.5. *Let $a(t, x; \xi)$, $b(t, x; \xi)$ be smooth symbols, where we assume furthermore that a is homogeneous of degree zero in ξ. Then we have*

$$\|a(t, x; D)b(t, x; D) - (ab)(t, x; D)\|_{L_t^{q_0} L_x^2 \to L_t^q L_x^2}$$
$$\lesssim \|(\partial_\xi a)(t, x; D)\|_{DL_t^{q_2} L_x^\infty} \|(\partial_x b)(t, x; D)\|_{L_t^{q_0} L_x^2 \to L_t^{q_1} L_x^2}$$

48

where $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}$. *An analogous statement holds in the case of right-quantization.*

For a proof, see [13, Lemma 7.2].

9.2. **Symbol bounds for ψ.** We first consider the size and regularity of the dyadic pieces of $\psi_{k,\pm}$, namely

$$\psi_{k,\pm}^{(\theta)}(t, x, \xi) := (\Pi_\theta^\omega \psi_{k,\pm})(t, x, \xi).$$

Given the symbol dependence on the angle, it is useful to keep in mind that the size of $\psi_{k,\pm}^{(\theta)}(t, x, \xi)$ is roughly given by

$$\psi_{k,\pm}^{(\theta)} \approx 2^{-k}\theta^{-2}\Pi_\theta^\omega(\omega \cdot A_k).$$

We borrow the following decomposability estimates for the symbol $\psi_{k,\pm}^{(\theta)}$ from [13]:

Lemma 9.6 (Decomposability estimates [13, Section 7.3]). *For $\frac{2}{q} + \frac{3}{r} \leq \frac{3}{2}$, we have*

$$\|(\psi_{k,\pm}^{(\theta)}, 2^{-k}\nabla\psi_{k,\pm}^{(\theta)})\|_{DL_t^q L_x^r} \lesssim 2^{-(\frac{1}{q}+\frac{4}{r})k}\theta^{\frac{1}{2}-\frac{2}{q}-\frac{3}{r}}E, \tag{9.15}$$

Moreover, for any $\beta \geq 0$, we have

$$\|\partial_\xi^\beta \Pi_\theta^\omega(\omega \cdot A_k)\|_{DL_t^q L_x^r} \lesssim 2^{(1-\frac{1}{q}-\frac{4}{r})k}\theta^{\frac{5}{2}-\frac{2}{q}-\frac{3}{r}-\beta}E. \tag{9.16}$$

In particular, for $q > 4$,

$$\|\partial_x^{\alpha-1}\nabla\psi_{k,\pm}\|_{DL_t^q L_x^\infty} \lesssim 2^{(\alpha-\frac{1}{q})k}E. \tag{9.17}$$

Remark 9.7. From the decomposability bound (9.17) with $q = \infty$, (9.9) follows easily.

We also collect here additional symbol bounds which are cruder but useful for estimating oscillatory kernels:

Lemma 9.8 (Symbol bound for $\psi_\pm^{(\theta)}$ [13, Section 7.3]). *The following symbol bounds hold.*
(1) For any $\alpha, \beta \geq 0$ and $2 \leq q \leq \infty$ we have

$$\|\partial_x^{\alpha-1}\nabla\partial_\xi^\beta\psi_{k,\pm}^{(\theta)}\|_{L_t^q L_{x,\xi}^\infty} \lesssim 2^{(-\frac{1}{q}+\alpha)k}\theta^{\frac{1}{2}-\frac{2}{q}-\beta}E. \tag{9.18}$$

When $\alpha = 0$, we interpret the expression on the left hand side as $\partial_\xi^\beta\psi_{k,\pm}^{(\theta)}$.
(2) For $q > 4$ and $1 \leq \beta \leq \sigma^{-1}(1 - \frac{1}{q})$, we have

$$\|\partial_x^\alpha\partial_\xi^\beta\psi_{k,\pm}\|_{L_t^q L_{x,\xi}^\infty} \lesssim 2^{(-\frac{1}{q}+\alpha)k}2^{\sigma(\frac{1}{2}-\frac{2}{q}-\beta)k}E. \tag{9.19}$$

(3) For $1 \leq \beta \leq \sigma^{-1}$, we have

$$|\partial_\xi^\beta(\psi_\pm(t, x, \xi) - \psi_\pm(t, y, \xi))| \lesssim \langle x - y \rangle^{\sigma(\beta-\frac{1}{2})}E. \tag{9.20}$$

49

9.3. **Fixed-time L^2 bounds.** Here we prove (9.10). For later use, we prove the following stronger result:

Proposition 9.9. *For sufficiently small $\sigma > 0$, there exists $\delta_0 > 0$ such that the following is true: For every $\ell, k \leq 0$ with $\ell + C \leq k$, we have*

$$\|(e^{-i\psi_{<\ell,\pm}}_{<k}(t,x,D)e^{i\psi_{<\ell,\pm}}_{<k}(D,y,t) - 1)P_0\|_{L^2_x \to L^2_x} \lesssim_E 2^{(1-\delta_0)\ell} + 2^{10(\ell-k)}. \qquad (9.21)$$

where the constant is independent of k, ℓ.

We remind the reader that $e^{i\psi_\pm}_{<0}(D,y,s) = e^{i\psi_\pm}_{<0}(t,D,y)$, since the symbol is independent of $\tau = \xi_0$. In particular, this pseudodifferential operator makes sense on every fixed time slice. Note that (9.10) follows by taking $k = 0$ and noting that $\psi_\pm = \psi_{<-m,\pm}$.

To begin the proof of Proposition 9.9, we prove a closely related estimate which does not involve space-time Littlewood-Paley projections for $e^{i\psi_{<\ell,\pm}}$.

Lemma 9.10. *Let $\ell \leq 0$ and $a(D)$ be a multiplier such that $a(\xi)$ is a smooth bump function adapted to $\{|\xi| \lesssim 1\}$. Then we have*

$$\|e^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e^{i\psi_{<\ell,\pm}}(D,y,t) - a(D)\|_{L^2_x \to L^2_x} \lesssim_E 2^{(1-\delta_0)\ell}. \qquad (9.22)$$

Furthermore, for any $k \in \mathbb{R}$ we have

$$\|e^{-i\psi_{<\ell,\pm}}(t,x,D)P_0\|_{L^2_x \to L^2_x} \lesssim_E 1, \qquad (9.23)$$

$$\|e^{-i\psi_{<\ell,\pm}}_{<k}(t,x,D)P_0\|_{L^2_x \to L^2_x} \lesssim_E 1. \qquad (9.24)$$

Proof. We first reduce (9.23) and (9.24) to proving (9.22). By a TT^* argument, (9.23) is equivalent to L^2_x boundedness of $e^{-i\psi_{<\ell,\pm}}(t,x,D)P_0^2 e^{i\psi_{<\ell,\pm}}(D,y,s)$, which follows from (9.22) and the L^2_x boundedness of $a(D) = P_0^2$. Next, note that

$$e^{-i\psi_{<\ell,\pm}}_{<k}(t,x,\xi) = \int e^{-i\psi_{<\ell,\pm}}((t,x) - z, \xi)2^{5k}m(2^k z)\,d^{1+4}z$$

where $m(z)$ is the kernel for $S_{<0}$. As the hypotheses for (9.23) are obviously invariant under translations, the left- (and also right-) quantization of each $e^{-i\psi_{<\ell,\pm}}((t,x) - z, \xi)$ obeys the same bound as (9.23). Therefore, by the rapid decay of $m(\cdot)$, (9.24) follows.

The proof of (9.22) is an easy consequence of non-stationary phase, thanks to the fact that $\sigma > 0$ can be taken arbitrarily small. The kernel of the operator in (9.22) is given by

$$K_1(t,x,y) := C \int (e^{i(\psi_{<\ell,\pm}(t,x,\xi) - \psi_{<\ell,\pm}(t,y,\xi))} - 1)a(\xi)e^{i\xi \cdot (x-y)}\,d^4\xi$$

$$= C \int \int_0^1 \Psi_\pm e^{i\rho\Psi_\pm}(t,x,t,y,\xi)a(\xi)e^{i\xi \cdot (x-y)}\,d\rho\,d^4\xi$$

where

$$\Psi_\pm(t,x,s,y,\xi) := \psi_{<\ell,\pm}(t,x,\xi) - \psi_{<\ell,\pm}(s,y,\xi).$$

We divide into two cases, namely when $|x - y| \lesssim 2^{-\delta\ell}$ and $|x - y| \gtrsim 2^{-\delta\ell}$.

Case 1: $|x - y| \lesssim 2^{-\delta\ell}$. In this case, using (9.18) with $q = \infty$ and $|\alpha| = 1$ and $\beta = 0$ for each frequency ℓ' and summing up in $\ell' < \ell$, we obtain

$$|\Psi_\pm(t,x,t,y,\xi)| \lesssim 2^{(1-\delta)\ell}E.$$

Since $\Psi_\pm$ is real-valued and $\operatorname{supp} a \subseteq \{|\xi| \lesssim 1\}$, it easily follows that

$$|K_1(t,x,y)| \lesssim 2^{(1-\delta)\ell} E \qquad \text{for } |x-y| \lesssim 2^{-\delta\ell}. \tag{9.25}$$

Case 2: $|x-y| \gtrsim 2^{-\delta\ell}$. Here we integrate by parts in ξ for N-times and use the bound (9.20). Then we obtain

$$|K_1(t,x,y)| \lesssim_{\sigma,N,E} \frac{1}{|x-y|^{(1-\sigma)N+\frac{1}{2}\sigma}} \qquad \text{for } |x-y| \gtrsim 2^{-\delta\ell}. \tag{9.26}$$

Combining Cases 1 and 2, it follows that

$$\sup_x \int |K_1(t,x,y)|\, d^4y + \sup_y \int |K_1(t,x,y)|\, d^4x \lesssim_E 2^{(1-\delta_0)\ell}$$

if σ, δ are small enough and N is sufficiently large. Estimate (9.22) now follows. $\qquad\square$

Next, we borrow a lemma from [13], which is useful for handling $e_k^{i\psi_{<\ell,\pm}}$ when $k > \ell$.

Lemma 9.11. *For $\ell + C \le k$ and every $t \in \mathbb{R}$, we have*

$$\|e_k^{-i\psi_{<\ell,\pm}}(t,x,D)P_0\|_{L_x^2 \to L_x^2} \lesssim_E 2^{10(\ell-k)}. \tag{9.27}$$

Furthermore, for $1 \le q \le p \le \infty$ and $\ell + C \le k$, we have

$$\|e_k^{-i\psi_{<\ell,\pm}}(t,x,D)P_0\|_{L_t^p L_x^2 \to L_t^q L_x^2} \lesssim_E 2^{(\frac{1}{p}-\frac{1}{q})\ell} 2^{10(\ell-k)}. \tag{9.28}$$

These estimates also hold for $e_k^{-i\psi_{<\ell,\pm}}(D,y,s)$.

Remark 9.12. The specific factor 10 in the gain $2^{10(\ell-k)}$ is irrelevant, but it is important to note that this number is much bigger than 1. This will be very useful in our proof of (9.11), where we will use this factor to dominate smaller factors. In fact, a variant of the proof below allows us to make this gain as large as we want, by making the implicit constant larger.

Proof. Consider frequency projections $S_k^{(1)}, \ldots, S_k^{(6)}, \widetilde{S}_k^{(1)}, \ldots, \widetilde{S}_k^{(5)}$, which obey the same bounds as S_k and furthermore satisfy

$$S_k^{(1)} := S_k, \quad S_k^{(i)} = 2^{-2k}\widetilde{S}_k^{(i)}(\partial_t^2 + \Delta), \quad \widetilde{S}_k^{(i)} = \widetilde{S}_k^{(i)} S_k^{(i+1)}$$

for $i = 1, \ldots, 5$. Thanks to the assumption $\ell + C \le k$, we may write at the level of symbols

$$\begin{aligned}
e_k^{-i\psi_{<\ell,\pm}} &= 2^{-2k}\widetilde{S}_k^{(1)}(\partial_t^2 + \Delta)e^{-i\psi_{<\ell,\pm}} \\
&= 2^{-2k}\widetilde{S}_k^{(1)}(-2i\Delta\psi_{<\ell,\pm} - |\nabla\psi_{<\ell,\pm}|^2)S_k^{(2)}e^{-i\psi_{<\ell,\pm}} \\
&= \cdots = 2^{-10k}\prod_{j=1}^{5}\left[\widetilde{S}_k^{(j)}(-2i\Delta\psi_{<\ell,\pm} - |\nabla\psi_{<\ell,\pm}|^2)\right]e^{-i\psi_{<\ell,\pm}}
\end{aligned}$$

Here we used the fact that $\psi_{<\ell,\pm}(t,x,\xi)$ solves the free wave equation $\partial_t^2\psi_{<\ell,\pm} = \Delta\psi_{<\ell,\pm}$ for each ξ, since A does. Disposing of the nested projections $\widetilde{S}_k^{(j)}$ by translation invariance, using the decomposability bound (9.17) and L_x^2 boundedness of $e^{-i\psi_{<\ell,\pm}}(t,x,D)P_0$, the desired estimate follows. $\qquad\square$

We are now ready to prove Proposition 9.9.

Proof of Proposition 9.9. Thanks to the frequency localization of the symbol $e_{<k}^{i\psi_{<\ell,\pm}}(s,y,\xi)$, note that we can harmlessly put in a multipler $a(D)$ whose symbol is a smooth bump function adapted to $\{|\xi| \lesssim 1\}$. The operator in (9.21) therefore equals

$$(e_{<0}^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e_{<0}^{i\psi_{<\ell,\pm}}(D,y,s) - a(D))P_0.$$

For the purpose of proving (9.21), we can safely dispose P_0 on the right. Next, note that

$$e^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e^{i\psi_{<\ell,\pm}}(D,y,s) - e_{<k}^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e_{<k}^{i\psi_{<\ell,\pm}}(D,y,s)$$
$$= e_{\geq k}^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e^{i\psi_{<\ell,\pm}}(D,y,s) + e_{<k}^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e_{\geq k}^{i\psi_{<\ell,\pm}}(D,y,s).$$

By Lemma 9.11 and (9.23), the operators on the right hand side obey

$$\|e_{\geq k}^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e^{i\psi_{<\ell,\pm}}(D,y,s)\|_{L_x^2 \to L_x^2} \lesssim_E 2^{10(\ell-k)},$$
$$\|e_{<k}^{-i\psi_{<\ell,\pm}}(t,x,D)a(D)e_{\geq k}^{i\psi_{<\ell,\pm}}(D,y,s)\|_{L_x^2 \to L_x^2} \lesssim_E 2^{10(\ell-k)}.$$

Combining these bounds with (9.22) established in Lemma 9.10, (9.21) follows. $\qquad\square$

9.4. **Space-time bounds.** Here we establish (9.11). More precisely, we will show that:

Proposition 9.13. *For $\sigma > 0$ sufficiently small, there exists $\delta_1 > 0$ such that the following holds: For $j \leq C$, we have*

$$\|Q_j[e_{<0}^{-i\psi_\pm}(t,x,D)e_{<0}^{i\psi_\pm}(D,y,s) - 1]P_0 Q_{<0}\|_{N^* \to X_\infty^{0,1/2}} \lesssim_E 2^{-\delta_1 m}. \qquad (9.29)$$

The estimate (9.29) proves the $X_\infty^{0,1/2}$ part of (9.11). Note that the $L_t^\infty L_x^2$ portion of (9.11) follows immediately from (9.10).

To ease the notation, we omit writing $\pm$ in $\psi_\pm$. Also, we omit the dependence of the constants on E. It will be convenient to define the compound symbols

$$\Psi(t,x,s,y,\xi) := \psi(t,x,\xi) - \psi(s,y,\xi),$$
$$\Psi_{<\ell}(t,x,s,y,\xi) := \psi_{<\ell}(t,x,\xi) - \psi_{<\ell}(s,y,\xi).$$

The symbol Ψ_ℓ is defined in the obvious way.

Given a compound symbol $a(t,x,s,y,\xi)$, we define the double space-time frequency projection

$$a_{\ll k}(t,x,s,y,\xi) := S_{<k}^{t,x}S_{<k}^{s,y}a(t,x,s,y,\xi),$$

where $S_{<k}^{t,x}$ is the space-time frequency projection applied to (t,x), etc. Therefore, according to our conventions,

$$e_{\ll k}^{-i\Psi}(t,x,D,y,s) = e_{<k}^{-i\psi}(t,x,D)e_{<k}^{i\psi}(D,y,s).$$

We begin with a lemma for frequency localizing the gauge transform $e^{-i\Psi_{<\ell}}$, which will be used several times in our argument.

Lemma 9.14. *For $2 \leq q \leq \infty$ and $\ell + C \leq k \leq 0$, we have*

$$\|(e_{\ll C}^{-i\Psi_{<\ell}} - e_{\ll k}^{-i\Psi_{<\ell}})P_0\|_{L_t^p L_x^2 \to L_t^q L_x^2} \lesssim 2^{(\frac{1}{p}-\frac{1}{q})\ell}2^{10(\ell-k)}. \qquad (9.30)$$

Proof. Proceeding as in the last part of the proof of Proposition 9.9, we may write

$$
(e_{\ll C}^{-i\Psi_{<\ell}} - e_{\ll k}^{-i\Psi_{<\ell}})P_0
$$
$$
= [e_{<C}^{-i\psi_{<\ell}}(t,x,D)a(D)e_{<C}^{i\psi_{\leq \ell}}(D,y,s) - e_{<k}^{-i\psi_{<\ell}}(t,x,D)a(D)e_{<k}^{i\psi_{\leq \ell}}(D,y,s)]P_0
$$
$$
= [e_{k\leq \cdot <C}^{-i\psi_{<\ell}}(t,x,D)a(D)e_{<C}^{i\psi_{\leq \ell}}(D,y,s) + e_{<k}^{-i\psi_{<\ell}}(t,x,D)a(D)e_{k\leq \cdot <C}^{i\psi_{\leq \ell}}(D,y,s)]P_0
$$

where $a(\xi)$ is a smooth bump function adapted to $\{|\xi| \lesssim 1\}$. Then (9.30) follows from Lemma 9.11. $\qquad\square$

We are now ready to prove Proposition 9.13.

Proof of Proposition 9.13. We proceed in several steps. Let $\delta > 0$ be a small number to be determined later.

Step 1: High modulation input. For $j' \geq j - C$, we claim that

$$
\|Q_j[e_{\ll 0}^{-i\Psi}(t,x,D,y,s) - 1]P_0 Q_{j'}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m} 2^{\frac{1}{2}(j-j')}. \tag{9.31}
$$

Using the $X_\infty^{0,1/2}$ portion of N_0^*, (9.31) follows from

$$
\|Q_j[e_{\ll 0}^{-i\Psi} - 1]P_0 Q_{j'}\|_{L_{t,x}^2 \to L_{t,x}^2} \lesssim 2^{-(1-\delta_0)m}.
$$

Since $Q_j, Q_{j'}$ are easily disposable, this estimate follows easily from (9.10).

Step 2: Low modulation input, $-\frac{1}{2}m \leq j \leq C$. In this step, we take care of the easy case $-\frac{1}{2}m \leq j \leq C$. Under this assumption, we claim that

$$
\|Q_j[e_{\ll 0}^{-i\Psi}(t,x,D,y,s) - 1]P_0 Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m}. \tag{9.32}
$$

Note that

$$
Q_j[e_{\ll j-C}^{-i\Psi} - 1]P_0 Q_{<j-C} = 0
$$

by modulation localization. Using the $L_t^\infty L_x^2$ portion of N_0^*, it suffices to prove

$$
\|Q_j[e_{\ll 0}^{-i\Psi} - e_{\ll j-C}^{-i\Psi}]P_0 Q_{<j-C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-4m} 2^{-\frac{1}{2}j}.
$$

Since $Q_j, Q_{<j-C}$ are disposable on $L_{t,x}^2$ and $L_t^\infty L_x^2$, this estimate follows from (9.30) and the fact that $\Psi = \Psi_{<-m}$.

Step 3: Low modulation input, $j < -\frac{1}{2}m$, main decomposition. Henceforth, we consider the case $j < -\frac{1}{2}m$. The goal of Steps 3–6 is to establish

$$
\|Q_j[e_{\ll 0}^{-i\Psi}(t,x,D,y,s) - e_{\ll 0}^{-i\Psi_{<j-\delta m}}(t,x,D,y,s)]P_0 Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m}. \tag{9.33}
$$

At the level of symbols, we begin by writing

$$
e^{-i\Psi} - e^{-i\Psi_{<j-\delta m}} = -i\int_{\ell \geq j-\delta m} \Psi_\ell\, e^{-i\Psi_{<j-\delta m}}\, d\ell - \iint_{\ell \geq \ell' \geq j-\delta m} \Psi_\ell \Psi_{\ell'}\, e^{-i\Psi_{<j-\delta m}}\, d\ell' d\ell
$$
$$
+ i\iiint_{\ell \geq \ell' \geq \ell'' \geq j-\delta m} \Psi_\ell \Psi_{\ell'} \Psi_{\ell''}\, e^{-i\Psi_{<\ell''}}\, d\ell'' d\ell' d\ell
$$
$$
=: \mathcal{L} + \mathcal{Q} + \mathcal{C}.
$$

We treat $\mathcal{L}$, $\mathcal{Q}$ and $\mathcal{C}$ in Steps 4, 5, and 6, respectively.

Step 4: Low modulation input, $j < -\frac{1}{2}m$, contribution of $\mathcal{L}$. In this step, we prove

$$\|Q_j \mathcal{L}_{\ll 0}(t, x, D, y, s) P_0 Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m}. \tag{9.34}$$

We further decompose $\mathcal{L}$ as follows. We first separate out the low frequency part of the gauge transform, then decompose according to the frequency of Ψ_ℓ (depending on whether ℓ is higher or comparable to j), and finally replace the gauge transform by 1:

$$\mathcal{L} = -i \int_{\ell \geq j - \delta m} \Psi_\ell \left(e^{-i\Psi_{<j-\delta m}} - e^{-i\Psi_{\ll j-C}} \right) d\ell \tag{9.35}$$

$$ -i \int_{\ell \geq j + 10\delta m} \Psi_\ell \, e^{-i\Psi_{<j-\delta m}}_{\ll j-C} \, d\ell \tag{9.36}$$

$$ -i \int_{j - \delta m \leq \ell \leq j + 10\delta m} \Psi_\ell \left(e^{-i\Psi_{<j-\delta m}}_{\ll j-C} - 1 \right) d\ell \tag{9.37}$$

$$ -i \int_{j - \delta m \leq \ell \leq j + 10\delta m} \Psi_\ell \, d\ell \tag{9.38}$$

$$ =: \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \mathcal{L}_4.$$

We treat the contribution of $\mathcal{L}_1, \ldots, \mathcal{L}_4$ separately.

Step 4.1: Contribution of $\mathcal{L}_1$. For $\mathcal{L}_1$, the double frequency localization $(\mathcal{L}_1)_{\ll 0}$ and the fact that $\ell < -m < 0$ allow us to write (at the level of symbols)

$$(\mathcal{L}_1)_{\ll 0} = \left(-i \int_{\ell \geq j - \delta m} \Psi_\ell \left(e^{-i\Psi_{<j-\delta m}}_{\ll C} - e^{-i\Psi_{<j-\delta m}}_{\ll j-C} \right) d\ell \right)_{\ll 0}.$$

As the rest of the argument for $\mathcal{L}_1$ will be translation invariant, we can easily dispose the double frequency localization $(\cdot)_{\ll 0}$. We are now reduced to proving

$$\|\Psi_\ell \left(e^{-i\Psi_{<j-\delta m}}_{\ll C} - e^{-i\Psi_{<j-\delta m}}_{\ll j-C} \right) P_0\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{-(10-\frac{1}{2})\delta m} 2^{\frac{1}{6}(j-\delta m - \ell)}$$

for $\ell \geq j - \delta m$. This estimate follows from the decomposability bound (9.17) with $q = 6$ and (9.30) with $(p, q) = (\infty, \frac{1}{3})$.

Step 4.2: Contribution of $\mathcal{L}_2$. Before we begin, note that the double frequency localization $(\cdot)_{\ll 0}$ does nothing to $\mathcal{L}_2$, $\mathcal{L}_3$ and $\mathcal{L}_4$, thanks to their frequency localization properties. Therefore, we drop $(\cdot)_{\ll 0}$ from now on.

In this step, the desired gain in m will be obtained from $\ell \geq j + 10\delta m$, and we do not exploit the difference structure in Ψ_ℓ. In fact, in order to apply decomposability bounds, we divide $\Psi_\ell(t, x, s, y, \xi) = \psi_\ell(t, x, \xi) - \psi_\ell(s, y, \xi)$ and treat each term separately. Here we only consider the case $\psi_\ell(t, x, \xi)$; the argument for the other case is analogous.

Thanks to the frequency localization of $e^{i\psi_{<j-\delta m}}_{<j-C}(s, y, \xi)$, the contribution of $\psi_\ell(t, x, \xi)$ in the integrand in (9.36) equals

$$Q_j (\psi_\ell e^{-i\psi_{<j-\delta m}}_{<j-C})(t, x, D) \widetilde{P}_0 \widetilde{Q}_{<j-C} e^{i\psi_{j-\delta m}}_{<j-C}(D, y, s) P_0 Q_{<j-C}$$

where $\widetilde{P}_0 \widetilde{Q}_{<j-C}$ is a slightly enlarged version of $P_0 Q_{<j-C}$. Then by the frequency localization of $e^{-i\psi_{<j-\delta m}}_{<j-C}(t, x, \xi)$, the modulation of the output forces an angular separation between the

54

spatial frequency of $\psi_\ell(t,x,\xi)$ and ξ of the size $\sim 2^{-\frac{1}{2}(\ell-j)_+}$. Therefore, the preceding operator equals

$$Q_j(\Pi^\omega_{>2^{-\frac{1}{2}(\ell-j)_+}}\psi_\ell e^{-i\psi_{<j-C}^{-\delta m}})(t,x,D)e^{i\psi_{<j-C}^{-\delta m}}(D,y,s)P_0Q_{<j-C},$$

where we dropped $\widetilde{P}_0\widetilde{Q}_{<j-C}$ as it is of no more use. Using fixed-time L^2_x boundedness of $e^{i\psi_{<j-C}^{-\delta m}}P_0$ and the decomposability bound (9.15) summed over $\theta \gtrsim 2^{-\frac{1}{2}(\ell-j)_+}$, it follows that

$$\|Q_j(\Pi^\omega_{>2^{-\frac{1}{2}(\ell-j)_+}}\psi_\ell e^{-i\psi_{<j-C}^{-\delta m}})(t,x,D)e^{i\psi_{<j-C}^{-\delta m}}(D,y,s)P_0Q_{<j-C}\|_{L^\infty_t L^2_x \to L^2_{t,x}}$$
$$\lesssim 2^{-\frac{1}{2}j}2^{\frac{1}{4}(j-\ell)} \tag{9.39}$$

for $\ell \geq j$. Now integrating this bound over $\ell \geq j + 10\delta m$, we obtain a gain of $2^{-\frac{5}{2}\delta m}$ from the factor $2^{\frac{1}{4}(j-\ell)}$, which is acceptable.

Step 4.3: Contribution of $\mathcal{L}_3$. For $\mathcal{L}_3$, we make use of the difference structure in the phase $\Psi_{<j-\delta m}$, but not for the symbol Ψ_ℓ. Thus we again only consider $\psi_\ell(t,x,\xi)$. We remind the reader that in this case, the ℓ-integral in (9.37) is taken over $j - \delta m \leq \ell \leq j + 10\delta m$.

Proceeding as in Step 4.2, the contribution of this term in the integrand for (9.37) equals

$$\widetilde{\mathcal{L}}_3(\ell) := Q_j(\Pi^\omega_{>2^{-\frac{1}{2}(\ell-j)_+}}\psi_\ell(e^{-i\Psi_{\ll j-C}^{<j-\delta m}}-1))(t,x,D,y,s)P_0Q_{<j-C}.$$

Then proceeding as in the proof of (9.39), but using (9.21) instead of mere boundedness of $e^{-i\Psi_{\ll j-C}^{<j-\delta m}}$, and integrating over $j - \delta m \leq \ell \leq j + 10\delta m$, we obtain

$$\|\int_{j-\delta m \leq \ell \leq j+10\delta m}\widetilde{\mathcal{L}}_3(\ell)\,d\ell\|_{L^\infty_t L^2_x \to L^2_{t,x}} \lesssim 2^{-\frac{1}{2}j}2^{\frac{1}{2}\delta m}(2^{(1-\delta_0)(j-\delta m)}+2^{-10\delta m}).$$

This bound is good if $\delta > 0$ is sufficiently small, since $j < -\frac{1}{2}m$.

Step 4.4: Contribution of $\mathcal{L}_4$. As in the previous step, the ℓ-integral in (9.38) is taken over $j - \delta m \leq \ell \leq j + 10\delta m$. Here we make use of the difference structure of the symbol Ψ_ℓ. The modulation localization properties again allow us to write

$$Q_j\Psi_\ell(t,x,D,y,s)P_0Q_{<j-C} = \sum_{\theta \gtrsim 2^{-\frac{1}{2}(\ell-j)_+}} Q_j(\psi_\ell^{(\theta)}(t,x,D)-\psi_\ell^{(\theta)}(D,y,s))P_0Q_{<j-C}.$$

As usual, we can harmlessly put in an operator $a(D)$ which is a slightly enlarged version of P_0. We now claim that for $2 \leq q \leq \infty$, the following bound holds:

$$\|\psi_\ell^{(\theta)}(t,x,D)a(D)-a(D)\psi_\ell^{(\theta)}(D,y,s)\|_{L^\infty_t L^2_x \to L^q_t L^2_x} \lesssim 2^{-\frac{1}{q}\ell}2^\ell\theta^{-C} \tag{9.40}$$

To prove this bound, we compute the kernel of $\psi_\ell^{(\theta)}(t,x,D)a(D)-a(D)\psi_\ell^{(\theta)}(D,y,s)$.

$$K_2(t,x,y) = C\int (\psi_\ell^{(\theta)}(t,x,\xi)-\psi_\ell^{(\theta)}(t,y,\xi))a(\xi)e^{i\xi\cdot(x-y)}\,d^4\xi$$
$$= C\int\int_0^1 (x-y)\cdot(\partial_x\psi_\ell^{(\theta)})(t,\rho x+(1-\rho)y,\xi)a(\xi)e^{i\xi\cdot(x-y)}\,d\rho\,d^4\xi$$
$$= C\int\int_0^1 \partial_\xi\cdot[(\partial_x\psi_\ell^{(\theta)})(t,\rho x+(1-\rho)y,\xi)a(\xi)]e^{i\xi\cdot(x-y)}\,d\rho\,d^4\xi.$$

Integrating by parts in ξ several times and using the symbol bound (9.18), we obtain a kernel bound which implies (9.40).

55

Applying (9.40) with $q = 2$, it follows that

$$\|Q_j(\psi_\ell^{(\theta)}(t, x, D) - \psi_\ell^{(\theta)}(D, y, s))P_0 Q_{<j-C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}\ell} 2^\ell \theta^{-C}.$$

Summing over $\theta \gtrsim 2^{-\frac{1}{2}(\ell-j)_+}$ and integrating over $j - \delta m \leq \ell \leq j + \delta m$, we arrive at

$$\| \int_{j-\delta m \leq \ell \leq j+10\delta m} Q_j \Psi_\ell(t, x, D, y, s) P_0 Q_{<j-C} \, d\ell \|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^j 2^{C\delta m}.$$

This is good for sufficiently small $\delta > 0$, as $j < -\frac{1}{2}m$. This completes the proof of (9.34).

Step 5: Low modulation input, $j < -\frac{1}{2}m$, contribution of $\mathcal{Q}$. Here we prove

$$\|Q_j \mathcal{Q}_{\ll 0}(t, x, D, y, s) P_0 Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m}. \tag{9.41}$$

As in Step 4, we begin by further decomposing $\mathcal{Q}$:

$$\mathcal{Q} = -\iint_{\ell \geq \ell' \geq j-\delta m} \Psi_\ell \Psi_{\ell'} \left(e^{-i\Psi_{<j-\delta m}} - e^{-i\Psi_{\ll j-C}^{<j-\delta m}} \right) d\ell' d\ell \tag{9.42}$$

$$- \iint_{\substack{\ell \geq \ell' \geq j-\delta m \\ \ell \geq j+10\delta m}} \Psi_\ell \Psi_{\ell'} \, e^{-i\Psi_{\ll j-C}^{<j-\delta m}} \, d\ell' d\ell \tag{9.43}$$

$$- \iint_{j-\delta m \leq \ell' \leq \ell \leq j+10\delta m} \Psi_\ell \Psi_{\ell'} \left(e^{-i\Psi_{\ll j-C}^{<j-\delta m}} - 1 \right) d\ell' d\ell \tag{9.44}$$

$$- \iint_{j-\delta m \leq \ell' \leq \ell \leq j+10\delta m} \Psi_\ell \Psi_{\ell'} \, d\ell' d\ell \tag{9.45}$$

$$=: \mathcal{Q}_1 + \mathcal{Q}_2 + \mathcal{Q}_3 + \mathcal{Q}_4.$$

We treat each of these terms below.

Step 5.1: Contribution of $\mathcal{Q}_1$. Proceeding as in Step 4.1, we have

$$(\mathcal{Q}_1)_{\ll 0} = \left(-\iint_{\ell \geq \ell' \geq j-\delta m} \Psi_\ell \Psi_{\ell'} \left(e^{-i\Psi_{\ll C}^{<j-\delta m}} - e^{-i\Psi_{\ll j-C}^{<j-\delta m}} \right) d\ell' d\ell \right)_{\ll 0}$$

and the outer $(\cdot)_{\ll 0}$ can be disposed by translation invariance as before. Next, by (9.17) (with $q = 6$ for Ψ_ℓ, $q = \infty$ for $\Psi_{\ell'}$) and (9.30) with $(p, q) = (\infty, 3)$, we have

$$\|\Psi_\ell \Psi_{\ell'} \left(e^{-i\Psi_{\ll C}^{<j-\delta m}} - e^{-i\Psi_{\ll j-C}^{<j-\delta m}} \right)\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{-(10-\frac{1}{2})\delta m} 2^{\frac{1}{6}(j-\delta m-\ell)}.$$

Integrating over $\ell \geq \ell' \geq j - \delta m$, we see the desired gain of $2^{-(10-\frac{1}{2})\delta m}$.

Step 5.2: Contribution of $\mathcal{Q}_2$. As in Steps 4.2, 4.3 and 4.4, $(\cdot)_{\ll 0}$ does nothing to $\mathcal{Q}_2, \mathcal{Q}_3, \mathcal{Q}_4$, and therefore can be removed. Also, in this step we split $\Psi_\ell(t, x, s, y, \xi) = \psi_\ell(t, x, \xi) - \psi_\ell(s, y, \xi)$ and handle only the contribution of $\psi_\ell(t, x, \xi)\psi_{\ell'}(t, x, \xi)$, as the argument for the other parts is the same.

As in Step 4.2, the contribution of $\psi_\ell(t, x, \xi)\psi_{\ell'}(t, x, \xi)$ in the integrand in (9.43) equals

$$Q_j(\psi_\ell \psi_{\ell'} e_{<j-C}^{-i\psi_{<j-\delta m}})(t, x, D) \widetilde{P}_0 \widetilde{Q}_{<j-C} e_{<j-C}^{i\psi_{<j-\delta m}}(D, y, s) P_0 Q_{<j-C} \, .$$

We first split

$$\psi_{\ell'}(t, x, \xi) = \Pi_{>2^{-\frac{1}{2}(\ell'-j)_+ - C'}}^\omega \psi_{\ell'}(t, x, \xi) + \Pi_{\leq 2^{-\frac{1}{2}(\ell'-j)_+ - C'}}^\omega \psi_{\ell'}(t, x, \xi).$$

56

The first term is good, as we already see an angular separation. For the contribution of the second term, we can apply an argument similar to Step 4.2 to conclude that there is an angular separation between the spatial frequency of $\psi_\ell(t,x,\xi)$ and ξ of size $\sim 2^{-\frac{1}{2}(\ell-j)_+}$. Therefore, the preceding operator equals

$$Q_j(\psi_\ell \Pi^\omega_{>2^{-\frac{1}{2}(\ell'-j)_+-C'}} \psi_{\ell'} e^{-i\psi_{<j-C}^{<j-\delta m}})(t,x,D)e^{i\psi_{<j-C}^{<j-\delta m}}(D,y,s)P_0 Q_{<j-C}$$

$$+ Q_j(\Pi^\omega_{\gtrsim 2^{-\frac{1}{2}(\ell-j)_+}} \psi_\ell \Pi^\omega_{\leq 2^{-\frac{1}{2}(\ell'-j)_+-C'}} \psi_{\ell'} e^{-i\psi_{<j-C}^{<j-\delta m}})(t,x,D)e^{i\psi_{<j-C}^{<j-\delta m}}(D,y,s)P_0 Q_{<j-C}$$

$$=: \widetilde{\mathcal{Q}}_{2,1}(\ell,\ell') + \widetilde{\mathcal{Q}}_{2,2}(\ell,\ell')$$

For $\widetilde{\mathcal{Q}}_{2,1}$ we use $\psi_\ell \in DL_t^6 L_x^\infty$ and $\Pi^\omega_{\geq 2^{\frac{1}{2}(j-\ell')-C'}} \psi_{\ell'} \in DL_t^3 L_x^\infty$, and vice versa for $\widetilde{\mathcal{Q}}_{2,2}$; see (9.15), (9.17). We also use fixed-time L_x^2 boundedness of $e^{i\psi_{<j-C}^{<j-\delta m}} P_0$ in both cases. Then

$$\|\widetilde{\mathcal{Q}}_{2,1}(\ell,\ell')\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{-\frac{4}{3}\delta m} 2^{\frac{1}{6}(j+10\delta m-\ell)} 2^{\frac{1}{3}(j-\delta m-\ell')}$$

$$\|\widetilde{\mathcal{Q}}_{2,2}(\ell,\ell')\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{-\frac{19}{6}\delta m} 2^{\frac{1}{3}(j+10\delta m-\ell)} 2^{\frac{1}{6}(j-\delta m-\ell')}.$$

which are good once integrated over $\{\ell \geq \ell' \geq j - \delta m\} \cap \{\ell \geq j + 10\delta m\}$.

Step 5.3: Contribution of $\mathcal{Q}_3$. We again only consider $\psi_\ell(t,x,\xi)\psi_{\ell'}(t,x,\xi)$. Proceeding as in the previous step, the contribution of this term in the integrand in (9.44) equals

$$Q_j(\psi_\ell \Pi^\omega_{\gtrsim 2^{-\frac{1}{2}(\ell'-j)_+}} \psi_{\ell'}(e^{-i\Psi_{\ll j-C}^{<j-\delta m}}-1))(t,x,D,y,s)P_0 Q_{<j-C}$$

$$+ Q_j(\Pi^\omega_{\gtrsim 2^{-\frac{1}{2}(\ell-j)_+}} \psi_\ell \Pi^\omega_{\lesssim 2^{-\frac{1}{2}(\ell'-j)_+}} \psi_{\ell'}(e^{-i\Psi_{\ll j-C}^{<j-\delta m}}-1))(t,x,D,y,s)P_0 Q_{<j-C}$$

$$=: \widetilde{\mathcal{Q}}_{3,1}(\ell,\ell') + \widetilde{\mathcal{Q}}_{3,2}(\ell,\ell').$$

We proceed as in Step 5.2, but replace the use of L_x^2 boundedness of $e^{i\psi_{<j-C}^{<j-\delta m}} P_0$ by (9.21). Integrating these bounds over $j - \delta m \leq \ell' \leq \ell \leq j + 10\delta m$, we obtain

$$\left\| \int_{j-\delta m \leq \ell' \leq \ell \leq j+10\delta m} \widetilde{\mathcal{Q}}_{3,1}(\ell,\ell')\,d\ell d\ell' \right\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{\frac{1}{2}\delta m}(2^{(1-\delta_0)(j-\delta m)} + 2^{-10\delta m})$$

$$\left\| \int_{j-\delta m \leq \ell' \leq \ell \leq j+10\delta m} \widetilde{\mathcal{Q}}_{3,2}(\ell,\ell')\,d\ell d\ell' \right\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{\frac{1}{2}\delta m}(2^{(1-\delta_0)(j-\delta m)} + 2^{-10\delta m}).$$

Taking $\delta > 0$ sufficiently small and using the fact that $j < -\frac{1}{2}m$, the desired gain in m follows.

Step 5.4: Contribution of $\mathcal{Q}_4$. Proceeding as in Steps 5.2 and 5.3 for every possible contribution of

$$(\psi_\ell(t,x,\xi) - \psi_\ell(s,y,\xi))(\psi_{\ell'}(t,x,\xi) - \psi_{\ell'}(s,y,\xi))$$

and recombining the expressions, it follows that

$$Q_j \Psi_\ell \Psi_{\ell'}(t,x,D,y,s)P_0 Q_{<j-C} = Q_j \Psi_\ell \Psi_{\ell'}^{(\gtrsim 2^{-\frac{1}{2}(\ell'-j)_+})}(t,x,D,y,s)P_0 Q_{<j-C}$$

$$+ Q_j \Psi_\ell^{(\gtrsim 2^{-\frac{1}{2}(\ell-j)_+})} \Psi_{\ell'}^{(\lesssim 2^{-\frac{1}{2}(\ell'-j)_+})}(t,x,D,y,s)P_0 Q_{<j-C}$$

$$=: \widetilde{\mathcal{Q}}_{4,1}(\ell,\ell') + \widetilde{\mathcal{Q}}_{4,2}(\ell,\ell')$$

where

$$\Psi_\ell = \Psi_\ell^{(>\theta)} + \Psi^{(\leq\theta)}, \qquad \Psi_\ell^{(>\theta)}(t,x,s,y,\xi) := \Pi_{>\theta}^\omega \psi_\ell(t,x,\xi) - \Pi_{>\theta}^\omega \psi_\ell(\xi,y,s).$$

Using (9.40) with $q = 3$ and summing up in $\theta \gtrsim 2^{-\frac{1}{2}(\ell'-j)_+}$, we obtain

$$\|\Psi_{\ell'}^{(\gtrsim 2^{-\frac{1}{2}(\ell'-j)_+})}(t,x,D,y,s)P_0\|_{L_t^\infty L_x^2 \to L_t^3 L_x^2} \lesssim 2^{-\frac{1}{3}\ell'} 2^{\ell'} 2^{\frac{C}{2}(\ell'-j)_+}.$$

By the decomposability bound (9.17) with $q = 6$ for Ψ_ℓ, it follows that

$$\|\widetilde{\mathcal{Q}}_{4,1}(\ell,\ell')\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{\ell'} 2^{\frac{1}{6}(j-\ell)} 2^{C(\ell'-j)_+}.$$

Then integrating over $j - \delta m \leq \ell' \leq \ell \leq j + 10\delta m$, we arrive at

$$\| \int_{j-\delta m \leq \ell' \leq \ell \leq j+10\delta m} \widetilde{\mathcal{Q}}_{4,1}(\ell,\ell')\, d\ell d\ell'\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^j 2^{C\delta m}$$

which is acceptable for $\delta > 0$ sufficiently small, since $j < -\frac{1}{2}m$. The term $\widetilde{\mathcal{Q}}_{4,2}$ is treated similarly, with the roles of Ψ_ℓ and $\Psi_{\ell'}$ swapped. This completes the proof of (9.41).

Step 6: Low modulation input, $j < -\frac{1}{2}m$, contribution of $\mathcal{C}$. In this step, we establish

$$\|Q_j \mathcal{C}_{\ll 0}(t,x,D,y,s)P_0 Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m}. \tag{9.46}$$

This step is easier than Steps 4 and 5, as we do not need to get the angle separation to apply the decomposability bound (9.15); instead, we can use (9.17). Thanks to this fact, the gauge transform need not be as finely localized in frequency as $\mathcal{L}$ and $\mathcal{Q}$. Accordingly, we make the following decomposition:

$$\mathcal{C} = i \iiint_{\ell \geq \ell' \geq \ell'' \geq j-\delta m} \Psi_\ell \Psi_{\ell'} \Psi_{\ell''} \left(e^{-i\Psi_{<\ell''}} - e_{\ll -C}^{-i\Psi_{<\ell''}} \right) d\ell'' d\ell' d\ell \tag{9.47}$$

$$+ i \iiint_{\substack{\ell \geq \ell' \geq \ell'' \geq j-\delta m \\ \ell \geq j+10\delta m}} \Psi_\ell \Psi_{\ell'} \Psi_{\ell''} e_{\ll -C}^{-i\Psi_{<\ell''}} d\ell'' d\ell' d\ell \tag{9.48}$$

$$+ i \iiint_{j-\delta m \leq \ell'' \leq \ell' \leq \ell \leq j+10\delta m} \Psi_\ell \Psi_{\ell'} \Psi_{\ell''} \left(e_{\ll -C}^{-i\Psi_{<\ell''}} - 1 \right) d\ell'' d\ell' d\ell \tag{9.49}$$

$$+ i \iiint_{j-\delta m \leq \ell'' \leq \ell' \leq \ell \leq j+10\delta m} \Psi_\ell \Psi_{\ell'} \Psi_{\ell''}\, d\ell'' d\ell' d\ell \tag{9.50}$$

$$=: \mathcal{C}_1 + \mathcal{C}_2 + \mathcal{C}_3 + \mathcal{C}_4.$$

We treat $\mathcal{C}_1, \ldots, \mathcal{C}_4$ separately.

Step 6.1: Contribution of $\mathcal{C}_1$. Proceeding as in Steps 4.1 and 5.1, it follows that

$$(\mathcal{C}_1)_{\ll 0} = \left(i \iiint_{\ell \geq \ell' \geq \ell'' \geq j-\delta m} \Psi_\ell \Psi_{\ell'} \Psi_{\ell''} \left(e_{\ll C}^{-i\Psi_{<\ell''}} - e_{\ll -C}^{-i\Psi_{<\ell''}} \right) d\ell'' d\ell' d\ell \right)_{\ll 0}$$

where the outer $(\cdot)_{\ll 0}$ may be easily disposed by translation invariance. Moreover, we have

$$\|\Psi_\ell \Psi_{\ell'} \Psi_{\ell''} \left(e_{\ll C}^{-i\Psi_{<\ell''}} - e_{\ll -C}^{-i\Psi_{<\ell''}} \right)\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j} 2^{\frac{1}{3}(j-\delta m-\ell'')} 2^{\frac{1}{6}(j-\delta m-\ell)} 2^{10\ell''} 2^{\frac{1}{2}\delta m}$$

by (9.17) and (9.30). Integrating over $j - \delta m \leq \ell'' \leq \ell' \leq \ell \leq -m$, this is acceptable.

58

Step 6.2: Contribution of $\mathcal{C}_2$. As before, by frequency localization properties, the double frequency projection $(\cdot)_{\ll 0}$ leaves $\mathcal{C}_2$, $\mathcal{C}_3$ and $\mathcal{C}_4$ unchanged. Using (9.17) with $q = 6$ for every factor of ψ and L_x^2 boundedness of $e_{\ll -C}^{-i\Psi_{\leq \ell''}}$, it follows that

$$\|Q_j\Psi_\ell\Psi_{\ell'}\Psi_{\ell''}e_{\ll -C}^{-i\Psi_{\leq \ell''}}P_0Q_{<j-C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j}2^{\frac{1}{6}(j-\ell)}2^{\frac{1}{6}(j-\ell')}2^{\frac{1}{6}(j-\ell'')}.$$

Integrating over $\{\ell \geq \ell' \geq \ell'' \geq j - \delta m\} \cap \{\ell \geq j + 10\delta m\}$, this is good.

Step 6.3: Contribution of $\mathcal{C}_3$. Here we use (9.17) with $q = 6$ for every factor of ψ and (9.21). Then we have

$$\|Q_j\Psi_\ell\Psi_{\ell'}\Psi_{\ell''}(e_{\ll -C}^{-i\Psi_{\leq \ell''}} - 1)P_0Q_{<j-C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j}2^{\frac{1}{6}(j-\ell)}2^{\frac{1}{6}(j-\ell')}2^{\frac{1}{6}(j-\ell'')}2^{(1-\delta_0)\ell''}$$

Integrating over $\{j - \delta m \leq \ell'' \leq \ell' \leq \ell \leq j + 10\delta m\}$ and using the fact that $j < -\frac{1}{2}m$, we obtain the desired gain in m.

Step 6.4: Contribution of $\mathcal{C}_4$. Summing up (9.40) with $q = 6$ in $\theta \gtrsim 2^{\sigma k}$, we obtain

$$\|\psi_\ell(t, x, D)a(D) - a(D)\psi_\ell(D, y, s)\|_{L_t^\infty L_x^2 \to L_t^6 L_x^2} \lesssim 2^{-\frac{1}{6}\ell}2^{(1-C\sigma)\ell}, \tag{9.51}$$

where $a(\xi)$ is any smooth bump function adapted to $\{|\xi| \lesssim 1\}$. Applying the decomposability bound (9.17) twice with $q = 6$, it follows that

$$\|Q_j\Psi_\ell\Psi_{\ell'}\Psi_{\ell''}P_0Q_{<j-C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j}2^{\frac{1}{6}(j-\ell)}2^{\frac{1}{6}(j-\ell')}2^{\frac{1}{6}(j-\ell'')}2^{(1-C\sigma)\ell}.$$

We integrate this over $\{j - \delta m \leq \ell'' \leq \ell' \leq \ell \leq j + 10\delta m\}$. Since $j < -\frac{1}{2}m$, the desired gain in m follows provided that $\sigma > 0$ is sufficiently small.

Step 7: Low modulation input, $j < -\frac{1}{2}m$, low frequency phase. To establish (9.29), it is only left to prove

$$\|Q_j[e_{\ll 0}^{-i\Psi_{<j-\delta m}}(t, x, D, y, s) - 1]P_0Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-\delta_1 m}. \tag{9.52}$$

Since

$$Q_j[e_{\ll j-C}^{-i\Psi_{<j-\delta m}} - 1]P_0Q_{<j-C} = 0$$

by modulation localization, it suffices to establish

$$\|Q_j[e_{\ll 0}^{-i\Psi_{<j-\delta m}}(t, x, D, y, s) - e_{\ll j-C}^{-i\Psi_{<j-\delta m}}]P_0Q_{<j-C}\|_{N^* \to X_\infty^{0,1/2}} \lesssim 2^{-(10+\frac{1}{2})\delta m}.$$

Proceeding as in Step 2, this estimate is reduced to

$$\|[e_{\ll 0}^{-i\Psi_{<j-\delta m}}(t, x, D, y, s) - e_{\ll j-C}^{-i\Psi_{<j-\delta m}}]P_0\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}j}2^{-(10+\frac{1}{2})\delta m}.$$

The last estimate follows from (9.30). $\qquad\square$

9.5. **Parametrix error estimate.** Here we prove (9.12). The argument here is essentially the same as in [13].

Step 1: Decomposition of the parametrix error. At the level of left-quantized operators, we compute

$$\Box_A^p e_{<0}^{-i\psi_\pm}(t,x,D) - e_{<0}^{-i\psi_\pm}(t,x,D)\Box$$

$$= 2(\partial^\mu e^{-i\psi_\pm})_{<0}\partial_\mu + (\Box e^{-i\psi_\pm})_{<0} + 2iA_{<-m}^\ell e_{<0}^{-i\psi_\pm}\partial_\ell + 2iA_{<-m}^\ell(\partial_\ell e^{-i\psi_\pm})_{<0}$$

$$= 2(\omega \cdot \partial_x\psi_\pm e^{-i\psi_\pm})_{<0}|D| - 2(\omega \cdot A_{<-m}e_{<0}^{-i\psi_\pm})|D| - 2(\partial_t\psi_\pm e^{-i\psi_\pm})_{<0}D_t$$

$$- (\partial^\mu\psi_\pm\partial_\mu\psi_\pm e^{-i\psi_\pm})_{<0} + 2A_{<-m}^\ell(\partial_\ell\psi_\pm e^{-i\psi_\pm})_{<0},$$

where we are using the shorthand $\omega = \xi/|\xi|$. This computation can be justified simply by using the direct definition of left-quantization, or by using the symbol calculus as in [13]. On the last line, we used the fact that $\Box\psi_\pm(t,x,\xi) = 0$ as $\Box A = 0$.

To see the cancellation between $A_{<-m}$ and $L_\mp^\omega\psi_\pm$, we add and subtract $2(\pm\partial_t\psi_\pm - \omega \cdot A_{<-m}e^{-i\psi_\pm})_{<0}|D|$. Then we can write

$$\Box_A^p e_{<0}^{-i\psi_\pm} - e_{<0}^{-i\psi_\pm}\Box = -2\Big(\big(\pm\partial_t\psi_\pm - \omega \cdot \partial_x\psi_\pm + \omega \cdot A_{<-m}\big)e^{-i\psi_\pm}\Big)_{<0}|D|$$

$$- 2(\partial_t\psi_\pm e^{-i\psi_\pm})_{<0}(D_t \mp |D|)$$

$$- ((-\partial_t\psi_\pm\partial_t\psi_\pm + \partial_x\psi_\pm \cdot \partial_x\psi_\pm)e^{-i\psi_\pm})_{<0}$$

$$+ 2A_{<-m} \cdot (\partial_x\psi_\pm e^{-i\psi_\pm})_{<0}$$

$$+ 2\big([\omega \cdot A_{<-m}, S_{<0}]e^{-i\psi_\pm}\big)(t,x,D)|D|$$

$$=: \mathrm{Diff}_1 + \mathrm{Diff}_2 + \mathrm{Diff}_3 + \mathrm{Diff}_4 + \mathrm{Diff}_5.$$

Step 2: Estimate for Diff_1. Being highest order, this is a-priori the most dangerous term. This is precisely the point where we need $\sigma > 0$. In this step we prove

$$\|\mathrm{Diff}_1 P_0\|_{N^* \to N} \lesssim 2^{-\frac{1}{2}\sigma m} + 2^{-m}. \tag{9.53}$$

Step 2.1: Preliminary reduction. By (9.5), it follows that

$$\mathrm{Diff}_1 = -2\Big(\sum_{k<-m} \Pi_{\leq 2^{\sigma k}}^\omega(\xi \cdot A_k)e^{-i\psi_\pm}\Big)_{<0}(t,x,D).$$

Note that $e^{-i\psi_\pm}$ can be replaced by $e_{<C}^{-i\psi_\pm}$ by the frequency localization of A. The outer $(\cdot)_{<0}$ can be easily disposed by translation invariance. Therefore, it suffices to consider

$$\mathcal{E}_1 := 2\sum_{k<-m} (\Pi_{\leq 2^{\sigma k}}^\omega(\xi \cdot A_k)e_{<C}^{-i\psi_\pm})(t,x,D)P_0$$

Step 2.2: Reduction to bilinear estimate. Our next order of business is to remove $e_{<C}^{-i\psi_\pm}$. For this purpose, consider the operator

$$\mathcal{E}_2 := 2\sum_{k<-m} (\Pi_{\leq 2^{\sigma k}}^\omega(\xi \cdot A_k))(t,x,D)e_{<C}^{-i\psi_\pm}(t,x,D)P_0$$

We claim that

$$\|\mathcal{E}_1 - \mathcal{E}_2\|_{L_t^\infty L_x^2 \to L_t^1 L_x^2} \lesssim 2^{-m}. \tag{9.54}$$

This estimate contributes the term 2^{-m} in (9.53), as $N^* \subseteq L_t^\infty L_x^2$ and $L_t^1 L_x^2 \subseteq N$.

Thanks to frequency localization of $e^{-i\psi_\pm}_{<C}$, we can harmlessly insert an operator $a(D)$ between the two pseudodifferential operators in $\mathcal{E}_2$, where $a(\xi)$ is a smooth bump function adapted to $\{|\xi| \sim 1\}$. Then by Lemma 9.5, it follows that

$$\|\mathcal{E}_1 - \mathcal{E}_2\|_{L_t^\infty L_x^2 \to L_t^1 L_x^2} \lesssim \sum_{k<-m} \|\partial_\xi(\Pi^\omega_{\leq 2^{\sigma k}}(\xi \cdot A_k) a(\xi))\|_{DL_t^2 L_x^\infty} \|(-i\partial_x \psi_\pm e^{-i\psi_\pm})_{<C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2}$$

Note that

$$\partial_\xi(\Pi^\omega_{\leq 2^{\sigma k}}(\xi \cdot A_k) a(\xi)) = (\frac{\xi}{|\xi|} a(\xi) + |\xi| \partial_\xi a(\xi)) \Pi^\omega_{\leq 2^{\sigma k}}(\omega \cdot A_k) + |\xi| a(\xi) \partial_\xi(\Pi^\omega_{\leq 2^{\sigma k}}(\omega \cdot A_k))$$

The factors involving only ξ can easily be removed as they are bounded. Invoking (9.16) and summing over $\theta \lesssim 2^{\sigma k}$ and $k < -m$, it follows that

$$\sum_{k<-m} \|\partial_\xi(\Pi^\omega_{\leq 2^{\sigma k}}(\xi \cdot A_k) a(\xi))\|_{DL_t^2 L_x^\infty} \lesssim \sum_{k<-m} 2^{\frac{1}{2}k} 2^{\frac{1}{2}\sigma k} \lesssim 2^{-\frac{1}{2}(1+\sigma)m}. \tag{9.55}$$

On the other hand, summing (9.15) over $\theta \gtrsim 2^{\sigma k}$ and $k \leq -m$, we obtain

$$\|\nabla \psi_\pm\|_{DL_t^2 L_x^\infty} \lesssim 2^{-\frac{1}{2}(1-\sigma)m}. \tag{9.56}$$

Now replacing $e^{-i\psi_\pm}$ by $e^{-i\psi_\pm}_{<2C}$, removing the outer $(\cdot)_{<C}$ by translation invariance as usual and using (9.56), we obtain

$$\|(\partial_x e^{-i\psi_\pm})_{<C}\|_{L_t^\infty L_x^2 \to L_{t,x}^2} \lesssim 2^{-\frac{1}{2}(1-\sigma)m}.$$

Combining (9.55) and (9.56), estimate (9.54) follows.

Step 2.3: Bilinear estimate. It is now only left to treat $\mathcal{E}_2$. Note that the operator $e^{-i\psi_\pm}_{<C}(t,x,D)P_0$ can be easily removed at this point, as it is bounded on N_0^*. Therefore, it suffices to show

$$\| \sum_{k<-m} \Pi^\omega_{\leq 2^{\sigma k}} A_k(t,x,D) \cdot \partial_x \widetilde{P}_0 \|_{N^* \to N} \lesssim 2^{-\frac{1}{2}\sigma m}$$

where $\widetilde{P}_0$ is a slightly enlarged version of P_0.

We now divide $A_k(t,x)$ and the input to $\pm$-waves (i.e., supported in $\{\pm\tau > 0\}$). For concreteness, consider the case when both are $+$-waves; other cases can be handled similarly. Recall that $\Pi^\omega_{\leq 2^{\sigma k}}$ localizes A_k into angular sectors of size $\sim 2^{\sigma k}$ centered at $\omega = \frac{\xi}{|\xi|}$ (close-angle) and $-\omega$ (far-angle). Therefore, by a Whitney-type decomposition in angles, it suffices to consider the sum

$$\sum_{k<-m} \sum_{\ell<\sigma k} \sum_{\substack{\phi,\phi' \\ \text{dist}(\phi,\phi')\sim 2^\ell}} (P_k P_\ell^\phi A) \cdot \partial_x \widetilde{P}_0 P_\ell^{\phi'} + \sum_{k<-m} \sum_{\substack{\phi,\phi' \\ \text{dist}(\phi,\phi')\sim 1}} (P_k P_{\sigma k}^\phi A) \cdot \partial_x \widetilde{P}_0 P_{\sigma k}^{\phi'},$$

where the first sum corresponds to the close-angle interaction, and the second sum corresponds to the far-angle interaction.

We begin by treating the close-angle interaction. We split this sum into two cases, depending on whether the input modulation is $> k + 2\ell - C$ or otherwise.

Step 2.3.1: Close-angle, high modulation input. By the sharp $L_t^2 L_x^6$ Strichartz estimate and Bernstein, we have

$$\|P_k P_\ell^\phi A\|_{L_t^2 L_x^\infty} \lesssim 2^{\frac{1}{2}k} 2^{\frac{1}{2}\ell}. \tag{9.57}$$

We estimate the output in $L_t^1 L_x^2$ and the input in $X_\infty^{0,1/2}$, using (9.57) for $P_k P_\ell^\phi A$. Note that, thanks to the null structure in $A \cdot \partial_x$, we also gain a factor of 2^ℓ. Using ℓ^2 summability in angles for A and the input, we obtain

$$\sum_{\substack{\phi,\phi' \\ \mathrm{dist}(\phi,\phi') \sim 2^\ell}} \|(P_k P_\ell^\phi A) \cdot \partial_x \widetilde{P}_0 P_\ell^{\phi'} Q_{>k+2\ell-C}\|_{X_\infty^{0,1/2} \to L_t^1 L_x^2} \lesssim 2^{\frac{1}{2}\ell}.$$

Summing over $\ell < \sigma k$ and then $k < -m$, the desired gain of $2^{-\frac{1}{2}\sigma m}$ follows.

Step 2.3.2: Close-angle, low modulation input. In this case, by elementary geometry of the cone, the output modulation is $\sim 2^{k+2\ell}$. Placing the output in $X_1^{0,-1/2}$ and the input in $L_t^\infty L_x^2$, the numerology is the same as in Step 2.3.1 and we obtain a gain of $2^{-\frac{1}{2}\sigma m}$.

Step 2.3.3: Far-angle. We proceed as in the case of close-angle interaction, this time splitting the input into $Q_{>k-C} + Q_{\leq k-C}$. In this case we do not gain from the null structure, but obtain the desired gain $2^{\frac{1}{2}\sigma k}$ from (9.57).

Step 3: Estimate for Diff_2. Here we need to use the $S_\pm^\sharp$ norm. We claim that

$$\|\mathrm{Diff}_2\|_{S_\pm^\sharp \to N} \lesssim 2^{-m}.$$

This estimate follows from the obvious mapping property

$$D_t \mp |D| : S_\pm^\sharp \to N.$$

and estimate (9.9).

Step 4: Estimate for Diff_3 and Diff_4. Again, we replace $e^{-i\psi_\pm}$ by $e_{<C}^{-i\psi_\pm}$, and dispose the outer $(\cdot)_{<0}$ by translation invariance. Summing up (9.15) in $\theta \gtrsim 2^{\sigma k}$, we have

$$\|\nabla \psi_{k,\pm}\|_{DL_t^2 L_x^\infty} \lesssim 2^{\frac{1}{2}(1-\sigma)k} \|A[0]\|_{\dot{H}_x^1 \times L_x^2}$$

On the other hand, since $A_k = P_k A$ is independent of ξ, it follows from Strichartz that

$$\|A_k\|_{DL_t^2 L_x^\infty} \lesssim \|A_k\|_{L_t^2 L_x^\infty} \lesssim 2^{\frac{1}{2}k} \|A_k[0]\|_{\dot{H}_x^1 \times L_x^2}$$

Then by decomposability and L_x^2 boundedness of $e_{<C}^{-i\psi_\pm}$, it follows that

$$\|\mathrm{Diff}_3 + \mathrm{Diff}_4\|_{L_t^\infty L_x^2 \to L_t^1 L_x^2} \lesssim 2^{-(1-\sigma)m}$$

which is enough.

Step 5: Estimate for Diff_5. For each component, the commutator may be written as

$$[A_{<-m}, S_{<0}](\phi) = L(\nabla A_{<-m}, \phi)$$

where L is a translation invariant bilinear operator with an integrable kernel. Using this expression, we now proceed as in Step 1. Summation in $k < -m$ is now possible thanks to the extra derivative ∇, and we obtain

$$\|\mathrm{Diff}_5\|_{N^* \to N} \lesssim 2^{-m}.$$

Combining Steps 1–5, estimate (9.12) follows.

References

1. Hajer Bahouri and Patrick Gérard, *High frequency approximation of solutions to critical nonlinear wave equations*, Amer. J. Math. **121** (1999), no. 1, 131–175. MR 1705001 (2000i:35123)

2. Scipio Cuccagna, *On the local existence for the Maxwell-Klein-Gordon system in R^{3+1}* Comm. PDE **24** (1999), no. 5-6, 851–867

3. Douglas M. Eardley and 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 649158 (83e:35106a)

4. ———, *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 649159 (83e:35106b)

5. Markus Keel, Tristan Roy and Terence Tao, *Global well-posedness of the Maxwell-Klein-Gordon equation below the energy norm* Discrete Contin. Dyn. Syst. 30 (2011), no. 3, 573–621.

6. Carlos E. Kenig and Frank Merle, *Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case*, Invent. Math. **166** (2006), no. 3, 645–675. MR 2257393 (2007g:35232)

7. ———, *Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation*, Acta Math. **201** (2008), no. 2, 147–212. MR 2461508 (2011a:35344)

8. Sergiu Klainerman and Matei Machedon, *On the Maxwell-Klein-Gordon equation with finite energy*, Duke Mathematical Journal (1994).

9. Sergiu Klainerman and Daniel Tataru, *On the optimal local regularity for Yang-Mills equations in $\mathbf{R^{4+1}}$*, Journal of the American Mathematical Society (1999).

10. Herbert Koch, Daniel Tataru, and Monica Vişan, *Dispersive equations and nonlinear waves*, Springer, 2014.

11. Joachim Krieger and Wilhelm Schlag, *Concentration compactness for critical wave maps*, EMS Publishing House, 2009.

12. Joachim Krieger and Jacob Sterbenz, *Global Regularity for the Yang-Mills Equations on High Dimensional Minkowski Space*, arXiv.org (2005).

13. Joachim Krieger, Jacob Sterbenz, and Daniel Tataru, *Global well-posedness for the Maxwell-Klein Gordon equation in 4+1 dimensions. Small energy*, arXiv.org (2012), Duke Mathematical Journal, to appear.

14. Andrew Lawrie and Sung-Jin Oh, *A refined threshold theorem for (1+2)-dimensional wave maps into surfaces*, preprint (2015).

15. Matei Machedon and Jacob Sterbenz, *Almost optimal local well-posedness for the $(3+1)$-dimensional Maxwell-Klein-Gordon equations*, Journal of the American Mathematical Society (2004).

16. Vincent Moncrief, *Global existence of Maxwell-Klein-Gordon fields in $(2+1)$-dimensional spacetime*, J. Math. Phys. **21** (1980), no. 8, 2291–2296. MR 579231 (82c:81089)

17. Sung-Jin Oh and Daniel Tataru, *Local well-posedness of the $(4+1)$-dimensional Maxwell-Klein-Gordon equation at energy regularity*, Annals of PDE, to appear, (2015)

18. ———, *Global well-posedness and scattering of the $(4+1)$-dimensional Maxwell-Klein-Gordon equation*, Inventiones, to appear (2015).

19. Igor Rodnianski and Terence Tao, *Global regularity for the Maxwell-Klein-Gordon equation with small critical Sobolev norm in high dimensions*, Comm. Math. Phys. **251** (2004), no. 2, 377–426. MR 2100060 (2005i:35256)

20. Sigmund Selberg, *Almost optimal local well-posedness of the Maxwell-Klein-Gordon equations in 1 + 4 dimensions*, Comm. Partial Differential Equations **27** (2002), no. 5-6, 1183–1227. MR 1916561 (2003f:35247)

21. Sigmund Selberg and Achenef Tesfahun, *Finite-energy global well-posedness of the Maxwell-Klein-Gordon system in Lorenz gauge*, Communications in Partial Differential Equations (2010).

22. Jacob Sterbenz, *Global regularity and scattering for general non-linear wave equations II. (4+1) dimensional Yang-Mills equations in the Lorentz gauge* Amer. J. of Math. **129** (2007), no. 3, 611–664

23. Jacob Sterbenz and Daniel Tataru, *Energy dispersed large data wave maps in 2 + 1 dimensions*, Comm. Math. Phys. **298** (2010), no. 1, 139–230. MR 2657817 (2011g:58045)

24. _____, *Regularity of wave-maps in dimension* $2 + 1$, Comm. Math. Phys. **298** (2010), no. 1, 231–264. MR 2657818 (2011h:58026)

25. Terence Tao, *Global Regularity of Wave Maps II. Small Energy in Two Dimensions*, Communications in Mathematical Physics (2001).

26. _____, *Global regularity of wave maps III. Large energy from* R^{1+2} *to hyperbolic spaces*, arXiv.org (2008).

27. _____, *Global regularity of wave maps IV. Absence of stationary or self-similar solutions in the energy class*, arXiv.org (2008).

28. _____, *Global regularity of wave maps V. Large data local wellposedness and perturbation theory in the energy class*, arXiv.org (2008).

29. _____, *Global regularity of wave maps VI. Abstract theory of minimal-energy blowup solutions*, arXiv.org (2009).

30. _____, *Global regularity of wave maps VII. Control of delocalised or dispersed solutions*, arXiv.org (2009).

31. Daniel Tataru, *On global existence and scattering for the wave maps equation*, Amer. J. Math. **123** (2001), no. 1, 37–77. MR 1827277 (2002c:58045)

Department of Mathematics, UC Berkeley, Berkeley, CA, 94720
E-mail address: sjoh@math.berkeley.edu

Department of Mathematics, UC Berkeley, Berkeley, CA, 94720
E-mail address: tataru@math.berkeley.edu