CONCENTRATION COMPACTNESS FOR THE CRITICAL MAXWELL-KLEIN-GORDON EQUATION

JOACHIM KRIEGER AND JONAS LÜHRMANN

Abstract. We prove global regularity, scattering and a priori bounds for the energy critical Maxwell-Klein-Gordon equation relative to the Coulomb gauge on $\mathbb{R}^{1+n}$-dimensional Minkowski space. The proof is based upon a modified Bahouri-Gérard profile decomposition [1] and a concentration compactness/rigidity argument by Kenig-Merle [5], following the method developed by the first author and Schlag [10] in the context of critical wave maps. This is a preliminary version of the final paper.

CONTENTS

1. Introduction 1
2. Function spaces and technical preliminaries 6
3. Microlocalized magnetic wave equation 9
4. Breakdown criterion 28
5. A concept of weak evolution 28
6. How to arrive at the minimal energy blowup solution 42
7. Concentration compactness step 43
8. Towards ruling out a minimal blowup solution with the compactness property 99
9. Technical details for ruling out a minimal blowup solution 136
References 147

1. Introduction

The Maxwell-Klein-Gordon system on Minkowski space-time $\mathbb{R}^{1+n}$, $n \geq 1$, is a classical field theory for a complex scalar field $\phi : \mathbb{R}^{1+n} \to \mathbb{C}$ and a connection 1-form $A_\alpha : \mathbb{R}^{1+n} \to \mathbb{R}$ for $\alpha = 0, 1, \ldots, n$. Defining the covariant derivative

$$D_\alpha = \partial_\alpha + iA_\alpha$$

and the curvature 2-form

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

the formal Lagrangian action functional of the Maxwell-Klein-Gordon system is given by

$$\int_{\mathbb{R}^{1+n}} \left( \frac{1}{4} F_{\alpha\beta} F^{\alpha\beta} + \frac{1}{2} D_\alpha \phi D^\alpha \overline{\phi} \right) dx \, dt,$$

References

BSCGI0\L 157694.
where the Einstein summation convention is in force and Minkowski space $\mathbb{R}^{1+n}$ is endowed with the standard metric $\text{diag}(-1, +1, \ldots, +1)$. Then the Maxwell-Klein-Gordon equations are the associated Euler-Lagrange equations

$$
\begin{align*}
\partial^\beta F_{\alpha\beta} &= \text{Im}(\phi D_\alpha \overline{\phi}), \\
\Box \phi &= 0,
\end{align*}
$$

where $\Box_A = D_\alpha D^\alpha$ is the covariant d’Alembertian. The system has two important features. First, it enjoys the gauge invariance $A_\alpha \mapsto A_\alpha - \partial_\alpha \gamma$, $\phi \mapsto e^{i\gamma} \phi$ for any suitably regular function $\gamma : \mathbb{R}^{1+n} \to \mathbb{R}$. Second, it is Lorentz invariant. Moreover, the system admits a conserved energy

$$
E(A, \phi) := \int_{\mathbb{R}^n} \left( \frac{1}{4} \sum_{\alpha,\beta} F^2_{\alpha\beta} + \frac{1}{2} \sum_{\alpha} |D_\alpha \phi|^2 \right) dx.
$$

Given that the system of equations (1.1) is invariant under the scaling transformation

$$
A_\alpha(t, x) \mapsto \lambda A_\alpha(\lambda t, \lambda x), \quad \phi(t, x) \mapsto \lambda \phi(\lambda t, \lambda x) \quad \text{for} \quad \lambda > 0,
$$

one distinguishes between the energy sub-critical case corresponding to $n \leq 3$, the energy critical case for $n = 4$, and the energy super-critical case for $n \geq 5$.

Imposing the Coulomb condition $\sum_{j=1}^n \partial_j A_j = 0$ for the spatial components of the connection form $A$, the Maxwell-Klein-Gordon system decouples into a system of wave equations for the dynamical variables $(A_j, \phi)$, $j = 1, \ldots, n$, coupled to an elliptic equation for the temporal component $A_0$,

$$
\begin{align*}
\Box A_j &= -\mathcal{P}_j \text{Im}(\phi D_\alpha \overline{\phi}), \\
\Box A_\phi &= 0, \\
\Delta A_0 &= -\text{Im}(\phi \partial_t \overline{\phi}),
\end{align*}
$$

where $\mathcal{P}$ is the standard projection onto divergence free vector fields.

We observe that in the formulation (MKG-CG), the components $(A_j, \phi)$, $j = 1, \ldots, n$, implicitly completely describe $(A_\alpha, \phi)$, since the missing component $A_0$ is uniquely determined by the elliptic equation

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

For this reason, we will mostly work in terms of the dynamical variables $(A_x, \phi)$, it being understood that required bounds on $A_0$ can be directly inferred from (1.3). In particular, to describe initial data for (MKG-CG), we will use the notation $A_j[0] := (A_j, \partial_j A_j)(0, \cdot)$ and $\phi[0] := (\phi, \partial_t \phi)(0, \cdot)$. Often, we will simply denote these by $(A_x, \phi)[0]$.

The present work will give a complete analysis of the energy critical case $n = 4$. More precisely, we implement an analysis closely analogous to the one by the first author and Schlag [10] in the context of critical wave maps in order to prove existence, scattering and a priori bounds for large global solutions to (MKG-CG). Moreover, we establish a concentration compactness phenomenon, which describes a kind of “atomic decomposition” of sequences of solutions of bounded energy.

To formulate our main result, we introduce the following notion of admissible data for the evolution problem (MKG-CG) on $\mathbb{R}^{1+4}$. 

---

\[\text{CONCENTRATION COMPACTNESS FOR THE CRITICAL MKG EQUATION}\]
Definition 1.1. We call $C^\infty$-smooth data $(A_x, \phi)[0]$ admissible, provided $A_x[0]$ satisfy the Coulomb condition and $\phi[0]$ as well as all spatial curvature components $F_{jk}[0]$ are Schwartz class. Moreover, we require that for $j = 1, \ldots, 4$,

$$|A_j[0](x)| \leq \langle x \rangle^{-3} \text{ as } |x| \to \infty.$$  

In particular, admissible data are of class $H^s(\mathbb{R}^4) \times H^{s-1}(\mathbb{R}^4)$ for any $s \geq 1$ and thus, the local existence and uniqueness theory developed by Selberg [20] applies to them. One can easily verify that as long as the solution exists in the sense of [20], and hence in the smooth sense, it will be admissible on fixed time slices. The above notion of admissible data therefore leads to a natural concept of solution to work with, and we call such solutions admissible. Our main theorem can then be stated as follows.

Theorem 1.2. Consider the evolution problem (MKG-CG) on $\mathbb{R}^{1+4}$. There exists a function

$$K : (0, \infty) \to (0, \infty)$$

with the following property. Let $(A_x, \phi)[0]$ be an admissible Coulomb class data set such that the corresponding full set of components $(A_x, \phi)$ has energy $E$. Then there exists a global in time admissible solution $(A, \phi)$ to (MKG-CG) with initial data $(A_x, \phi)[0]$ that satisfies for any $\frac{1}{q} + \frac{3}{2r} \leq \frac{3}{4}$ with $2 \leq q \leq \infty$, $2 \leq r < \infty$, $\gamma = 2 - \frac{q}{r} - \frac{3}{r}$ the following a priori bound

$$(1.4) \quad \left\| \left( (-\Delta)^{-\frac{1}{2}} \nabla_{t,x} A_x, (-\Delta)^{-\frac{1}{2}} \nabla_{t,x} \phi \right) \right\|_{L^q_t L^r_x(\mathbb{R} \times \mathbb{R}^4)} \leq C_r K(E).$$

The solution scatters in the sense that there exist finite energy free waves $f_j$ and $g$, $\Box f_j = \Box g = 0$, such that for $j = 1, \ldots, 4$,

$$\lim_{t \to +\infty} \left\| \nabla_{t,x} A_j(t, \cdot) - \nabla_{t,x} f_j(t, \cdot) \right\|_{L^2_t(\mathbb{R}^4)} = 0, \quad \lim_{t \to +\infty} \left\| \nabla_{t,x} \phi(t, \cdot) - \nabla_{t,x} g(t, \cdot) \right\|_{L^2_t(\mathbb{R}^4)} = 0,$$

and analogously with different free waves for $t \to -\infty$.

In fact, we will prove the significantly stronger a priori bound

$$\left\| (A_x, \phi) \right\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq K(E),$$

where the precise definition of the $S^1$ norm will be introduced in Section 2.

Very recently, a proof of the global regularity and scattering affirmations in the preceding theorem was obtained by Oh-Tataru [16] [17] [18], following the method developed by Sterbenz-Tataru [21] [22] in the context of critical wave maps. The results and methods of our work were announced by us at a conference on 20 January 2015.

1.1. Historical remarks. The existence of global smooth solutions to the Maxwell-Klein-Gordon equation in $n = 3$ space dimensions follows from seminal work of Eardley-Moncrief [2] [3] and the important refined well-posedness theory of Klainerman-Machedon [6]. The latter work introduced many new techniques into the study of (MKG-CG) such as exploitation of null structures to get improved local well-posedness. Almost optimal local well-posedness of (MKG-CG) in $n = 3$ space dimensions was obtained in the work of Machedon-Sterbenz [13]. It revealed a deep trilinear null structure in the system that is also of crucial importance in our work. The corresponding almost optimal local well-posedness result for (MKG-CG) on $\mathbb{R}^{1+4}$ was proved by Selberg [20], which we rely on in this paper. A similar result to the one in [20], but for a model type system had been obtained earlier by Klainerman-Tataru [8].

Small energy global well-posedness of the energy critical Maxwell-Klein-Gordon equation on $\mathbb{R}^{1+4}$ was established in joint work of the first author with Sterbenz and Tataru [12]. This work
also provides the functional analytic framework that we will draw on and constantly refer to in this paper. The main ingredient in [12] was an approximate parametrix construction for wave equations involving a magnetic potential interaction term. This followed closely the template provided in the work by Rodnianski-Tao [19] on global regularity of (MKG-CG) for small critical Sobolev norm data in space dimensions \( n \geq 6 \). The key novelty in [12] was the realization that this parametrix is in fact compatible with the complicated spaces introduced for critical wave maps in [26,27], [24,25]. [9] and that these spaces play a pivotal role in the energy critical small data global existence theory for (MKG-CG) on \( \mathbb{R}^{1+4} \), especially in light of the deep null structure revealed in [13].

We also mention that for the closely related Yang-Mills equation finite energy global well-posedness in (energy sub-critical) \( n = 3 \) space dimensions was proved by Klainerman-Machedon [7]. A different proof was later obtained by Oh [14, 15], using the Yang-Mills heat flow. Global regularity for the Yang-Mills system for small critical Sobolev data for \( n \geq 6 \) was established by the first author and Sterbenz [11].

For energy critical problems, it is a standard strategy to attempt to prove global regularity for large energies by reducing to a small energy global existence result via finite speed of propagation and exclusion of an energy concentration scenario. Such a method worked well, for example, for the critical defocusing nonlinear wave equation \( \square u = u^5 \) on \( \mathbb{R}^{1+3} \) and for the radial critical wave maps on \( \mathbb{R}^{1+2} \). At this point, with the exception of some special problems, it appears that a general large data result cannot be inferred by using the small data result as a black box, but instead requires a more or less complete re-working of the small data theory. See, for example, the works on the critical large data wave maps [21, 22], [10], [23]. Our approach here is to implement a similar strategy as the one by the first author and Schlag [10] for critical wave maps, which consists of essentially two steps: First, a novel “covariant” Bahouri-Gérard procedure to take into account the non-negligible influence of low on high frequencies in the magnetic potential interaction term. Second, an implementation of a variant of a concentration compactness/rigidity argument by Kenig-Merle, following more or less the sequence of steps in [5]. As the latter was introduced in the context of a scalar wave equation, and we are considering a complex nonlinearly linked system, we believe that the implementation of this step for the energy critical Maxwell-Klein-Gordon equation is also of interest in its own right.

1.2. Overview of the paper. The purpose of this paper is to prove Theorem 1.2. In fact, we will derive a significantly stronger version, namely the existence of a function \( K : (0, \infty) \to (0, \infty) \) with

\[
\left\| (A_x, \phi) \right\|_{S^1} \leq K(E), \quad E = E(A, \phi)
\]

for any admissible solution \( (A, \phi) \), where the precise definition of the \( S^1 \) norm and its time localized versions will be given in Section 2 and Definition 4.1. The purpose of this norm is to control the regularity of the solutions \( (A, \phi) \).

Beginning the argument at this point, we assume that the existence of such a function fails for some finite energy level. Thus, the set of energies

\[
\mathcal{E} := \left\{ E > 0 : \sup_{(A, \phi) \text{admissible}} \left\| (A_x, \phi) \right\|_{S^1} = +\infty \right\}
\]

is non-empty. Hence, it has a (necessarily positive [12]) infimum, which we call \( E_{\text{crit}} \),

\[
E_{\text{crit}} := \inf \mathcal{E}.
\]
By definition we can then find a sequence of admissible solutions \((A_n, \phi_n)\) such that
\[
E(A_n, \phi_n) \to E_{\text{crit}}, \quad \lim_{n \to \infty} \| (A_{\lambda, n}, \phi_n) \|_{S^1} = +\infty.
\]
As in \cite{10}, we call such a sequence an \textit{essentially singular sequence}. The goal of this paper is to rule out the existence of such an object. This will be accomplished in broad strokes by the following two steps.

1. Extracting an energy class minimal blowup solution via a modified Bahouri-Gérard procedure, which consists of an inductive sequence of low-frequency approximations and a profile extraction process taking into account the effect of the magnetic potential interaction. Here we closely follow the procedure introduced by the first author and Schlag \cite{10}, but we have to subtly diverge from the profile extraction process there to correctly capture the asymptotic evolution of the atomic components. See Section 7.5 and, in particular, the choice \(\Box_{A_n, \text{true}}\) to model the correct asymptotic evolution, as opposed to the operator \(\Box_{A_n, \text{low}}\) used in \cite{10}. We note that the heart of the modified Bahouri-Gérard procedure resides in Section 7.

2. Ruling out the minimal blowup solution by essentially following the method of Kenig-Merle \cite{5} with an additional ingredient introduced in \cite{10}, namely a Vitali covering argument. However, the construction of the correct virial type identities appears significantly more complicated here than for wave maps, in particular the identity in self-similar coordinates in Section 8.4, where we use the trick of working in the Cronstrom gauge to simplify the computations.

The rigorous execution of this program requires a lot of technical preparations, which are carried out in the sections leading up to Section 7. We now give an overview of the structure of this paper.

- In Section 2 we lay out the functional framework following \cite{12}.
- In Section 3 we prove key estimates for the linear magnetic wave equation \(\Box_{A} u = f\).
- In Section 4 we state the property of the \(S^1\) norm as a regularity controlling device.
- In Section 5 we show how to unambiguously locally evolve Coulomb energy class data \((A_x, \phi)(0)\) via approximation by smoothed data and truncation in physical space (to reduce to the admissible setup). Here one needs to pay close attention to the fact that solutions to (MKG-CG) do not obey as good a perturbation theory with respect to the \(S^1\) spaces as, say, critical wave maps (in a suitable gauge), due to a low frequency divergence. Hence, one needs to be very careful about the correct choice of smoothing, using low-frequency truncations of the data. Moreover, to ensure the existence of an energy class local evolution of Coulomb energy class data \((A_x, \phi)(0)\) on a non-trivial time slice around \(t = 0\), we need to prove uniform \(S^1\) norm bounds for the approximations, which we accomplish similarly to the procedure in \cite{10} via localization in physical space, see Proposition 5.2. We also introduce the concept of the “lifespan” of such an energy class solution and the definition of its \(S^1\) norm.
- In Section 6 we then state that energy class data \((A_x, \phi)(0)\) obtained as the limit of the data of an essentially singular sequence, which will be the outcome of the modified Bahouri-Gérard procedure, lead to a singular solution in the sense that
\[
\sup_{J \subset I} \|(A_x, \phi)\|_{S^1(J \times \mathbb{R}^4)} = +\infty,
\]
where \(I\) denotes its lifespan. The proof of this as well as a number of further technical assertions will be relegated to Section 7.4.
In Section 7 we carry out the modified Bahouri-Gérard procedure. In Section 7.1 and Section 7.2 we extract the “small” and “large” frequency atoms mimicking closely the procedure in [10]. Then we show in Section 7.3 how the lowest frequency “non-atomic” part of the low frequency approximation induction can be globally evolved with good $S^1$-bounds. In Section 7.4 we prove several technical assertions that all use the core perturbative result from Step 3 of the proof of Proposition 7.4. In Section 7.5 we add the first “large” frequency atom by extracting frequency profiles and invoking the induction on energy hypothesis, i.e. that all profiles have energy strictly less than $E_{\text{crit}}$. The end result of the modified Bahouri-Gérard procedure is obtained in Section 7.6, see in particular Theorem 7.26. We then have a minimal blowup solution, denoted by $(\mathcal{A}_\infty, \Phi_\infty)$, with the required compactness property.

In Section 8 and Section 9 we prove the non-existence of such a minimal blowup solution. More precisely, in Section 8.1 we show that if a Lorentz transformed solution, when placed in the Coulomb gauge, satisfies good $S^1$-bounds, then so does the original solution. Together with a computation of the effect of a Lorentz transformation on the energy in Section 8.2, this will then be used in Section 9 to rigorously infer a vanishing momentum condition for a minimal blowup solution $(\mathcal{A}_\infty, \Phi_\infty)$. In Section 8.3, we collect several conservation and virial identities, which are mostly summarized in Proposition 8.3. These will be used in Section 9 to reduce to a self-similar blowup scenario for $(\mathcal{A}_\infty, \Phi_\infty)$. To prepare for the latter, we derive a virial functional in self-similar coordinates in Section 8.4. In Section 9 all loose ends are tied up to rule out the minimal blowup solution $(\mathcal{A}_\infty, \Phi_\infty)$, which then finally shows that an essentially singular sequence cannot exist, thereby proving Theorem 1.2.

We remark that we will often abuse notation and denote the spatial components $A_x$ of the connection form simply by $A$.

2. Function spaces and technical preliminaries

We will be working with the same function spaces that were used for the small energy global well-posedness result for the MKG-CG system [12] together with their time-localized versions. In this section we recall their definitions very briefly. For a more detailed discussion of these spaces we refer to Section 3 in [12] and [25], [21], [10].

We will mainly use three function spaces $N, N^*, \text{ and } S$. Their dyadic subspaces $N_k, N_k^*$ and $S_k$ satisfy

$$N_k = \mathcal{L}^1_1 L^2_1 + X^{0, -\frac{1}{2}}, \quad N_k^* = \mathcal{L}^\infty_1 L^2_1 \cap X^{0, \frac{1}{2}}, \quad X^{0, \frac{1}{2}} \subseteq S_k \subseteq N_k^*.$$ 

Then we have

$$\|F\|_N^2 = \sum_{k \in \mathbb{Z}} \|P_k F\|_{N_k}^2, \quad \|F\|_{N^*}^2 = \sum_{k \in \mathbb{Z}} \|P_k F\|_{N_k^*}^2.$$ 

The space $S_k$ is defined by

$$\|\phi\|^2_S = \|\phi\|_{S^\text{str}}^2 + \|\phi\|_{S^\text{ang}}^2 + \|\phi\|_{X^{0, \frac{1}{2}}_{1, \infty}}^2,$$

where

$$S^\text{str} = \bigcap_{r \geq 1} \mathcal{L}^{\frac{1}{2} + \frac{3}{2}} L^r_1 L^r_1,$$

$$\|\phi\|_{S^\text{ang}}^2 = \sum_{l > 0} \|P^{\text{ang}}_{Q < k + 2l} (\phi)\|_{S^\text{ang}_k(l)}^2,$$

and

$$\|\phi\|_{X^{0, \frac{1}{2}}_{1, \infty}}^2 = \sup_{l > 0} \sum_{\omega} \|P^{\omega}_{Q < k + 2l} (\phi)\|_{X^{0, \frac{1}{2}}_{1, \infty}(l)}^2.$$
and the angular sector norms $S^\omega_k(l)$ are defined below.

To introduce the angular sector norms $S^\omega_k(l)$ we first define the plane wave space

$$\|\phi\|_{PW_k^\omega(l)} = \inf_{\phi = \sum_{l=0}^\infty \phi_{\omega,l}} \int_{|\omega - \omega'| \leq 2l} \|\phi_{\omega,l}'\|_{L^2_{\text{rad}} L^\infty_{\omega'(\omega)}} \, d\omega'$$

and the null energy space

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

where the norms are with respect to $\ell^\omega_\omega = t \pm \omega \cdot x$ and the transverse variable, while $\nabla_\omega$ denotes spatial differentiation in the $(\ell^\omega_\omega)^\perp$ plane. We now set

$$\|\phi\|^2_{S^\omega_k(l)} = \|\phi\|^2_{S^\omega_k} + 2^{-2k}\|\phi\|^2_{NE} + 2^{-3k} \sum_{\pm} \|Q^\pm \phi\|^2_{PW_k^\omega(l)}$$

and

$$\|\phi\|^2_{S^\omega_0} = \sum_{k \in \mathbb{Z}} \|\nabla_{t,x} P_k \phi\|^2_{S_k} + \|\square \phi\|^2_{L^1_{\omega'} H_k^{-\frac{1}{2}}}$$

and the higher derivative norms

$$\|\phi\|_{S^\omega_N} := \|\nabla_{t,x}^{N-1} \phi\|_{S^\omega_1}, \quad N \geq 2.$$

Moreover, we introduce

$$\|u\|^2_{S^\omega_k} = \|\nabla_{t,x} u\|_{L^\infty_{\omega'} L^2_{\omega}} + \|\square u\|_{N_k} .$$

On occasion we need to separate the two characteristic cones ($\tau = \pm \xi$). To this end we define

$$\begin{align*}
N_{k,\pm}, & \quad N_k = N_{k,\pm} \cap N_{k,-} \\
S_{k,\pm}^\#, & \quad S_k^\# = S_{k,\pm}^\# + S_{k,-}^\# \\
N_{k,\pm}^*, & \quad N_k^* = N_{k,\pm}^* + N_{k,-}^* .
\end{align*}$$

We will also use an auxiliary space of $L^1_{\omega'} L^\infty_{x}$ type,

$$\|\phi\|_{Z_k} = \sum_k \|P_k \phi\|_{Z_k}, \quad \|\phi\|_{Z_k} = \sup_{l \in \mathbb{C}} \sum_{\omega} 2^l \|Q^\omega_{P_k} \phi\|^2_{L^1_{\omega'} L^\infty_{x}} .$$

Finally, to control the component $A_0$, we define

$$\|A_0\|^2_{Y^1} = \|\nabla_{t,x} A_0\|^2_{L^\infty_{\omega'} L^2_{\omega}} + \|A_0\|^2_{L^2_{\omega'} H_{1/2}} + \|\partial_t A_0\|_{L^2_{\omega'} H_{1/2}}$$

and the higher derivative norms

$$\|A_0\|_{Y^N} = \|\nabla_{t,x}^{N-1} A_0\|_{Y^1}, \quad N \geq 2.$$

We will need to work with time-localized versions of the $S_k$ and $N_k$ spaces. For any compact interval $I \subset \mathbb{R}$ and $k \in \mathbb{Z}$, we define

$$\|\phi\|_{S_k(I \times \mathbb{R}^4)} := \inf_{\psi_t = \phi_t} \|P_k \psi\|_{S_k(I \times \mathbb{R}^4)} .$$
with $\psi$ and $\tilde{\psi}$ Schwartz functions. Analogously, we define $N_k(I \times \mathbb{R}^d)$.

The following lemma shows that the $S_k$ and $Z_k$ spaces are compatible with time cutoffs. We will frequently use this fact without further mentioning.

**Lemma 2.1.** Let $\chi_I$ be a smooth cutoff to a time interval $I \subset \mathbb{R}$. Then it holds for all $k \in \mathbb{Z}$ that

$$\|P_k(\chi_I \phi)\|_{S_k(\mathbb{R} \times \mathbb{R}^d)} \leq \|P_k \phi\|_{S_k(\mathbb{R} \times \mathbb{R}^d)}$$

and

$$\|P_k(\chi_I \phi)\|_{Z_k(\mathbb{R} \times \mathbb{R}^d)} \leq \|P_k \phi\|_{Z_k(\mathbb{R} \times \mathbb{R}^d)}$$

**Proof.** (Outline) This is obvious for the Strichartz type norms. It remains to show it for the $X^{0,\frac{1}{2}}_\infty$ and $S^{ang}_k$ components. We start with the former. For fixed $j \in \mathbb{Z}$, we have

$$Q_j(\chi_I \phi) = Q_j(Q_j + o(1)) (\chi_I Q_{\leq j-C} \phi) + Q_j(\chi_I Q_{> j-C}(\phi)).$$

Using the bound

$$\|Q_j + o(1)(\chi_I)\|_{L^2_x} \leq 2^{-\frac{1}{2}j},$$

we obtain

$$2^{\frac{j}{2}}\|Q_j(Q_j + o(1)) (\chi_I P_k Q_{\leq j-C} \phi)\|_{L^2_{x,t}} \leq 2^{\frac{j}{2}}\|Q_j + o(1)(\chi_I)\|_{L^2_x} \|P_k Q_{\leq j-C} \phi\|_{L^2_{x,t}} \leq \|P_k \phi\|_{X^{0,\frac{1}{2}}_\infty}.$$  

Moreover, we find

$$2^{\frac{j}{2}}\|Q_j(\chi_I P_k Q_{> j-C}(\phi))\|_{L^2_{x,t}} \leq 2^{\frac{j}{2}}\|\chi_I\|_{L^\infty_x} \|P_k Q_{> j-C}(\phi))\|_{L^2_{x,t}} \leq \|P_k \phi\|_{X^{0,\frac{1}{2}}_\infty}.$$  

Thus, we have

$$\|P_k \chi_I \phi\|_{X^{0,\frac{1}{2}}_\infty} \leq \|P_k \phi\|_{S_k}.$$  

Next, we consider the $S^{ang}_k$ component, which is given by

$$\|\phi\|_{S^{ang}_k} = \sup_{l < 0} \|P_l^a Q_{< k+2l} \phi\|_{S^{ang}_k(l)}^2.$$  

We write

$$P_l^a Q_{< k+2l}(\chi_I \phi) = P_l^a Q_{< k+2l}(\chi_I Q_{< k+2l+C} \phi) + P_l^a Q_{< k+2l}(\chi_I Q_{\geq k+2l+C} \phi)$$

Then the first term on the right hand side is bounded by

$$\|P_l^a Q_{< k+2l}(\chi_I Q_{< k+2l+C} \phi)\|_{S^{ang}_k(l)} \leq \|P_l^a Q_{< k+2l+C} \phi\|_{S^{ang}_k(l)},$$

where we have used the fact that the operator $P_l^a Q_{< k+2l}$ is dispersive. For the second term above, we use that

$$\sum_\omega \|P_l^a Q_{< k+2l}(\chi_I Q_{\geq k+2l+C} \phi)\|_{S^{ang}_k(l)}^2 \leq \|P_l Q_{< k+2l}(\chi_I Q_{\geq k+2l+C} \phi)\|_{X^{0,\frac{1}{2}}_l}^2 \leq \|\phi\|_{X^{0,\frac{1}{2}}_l}^2.$$  

For the $Z_k$ space, fix a scale $l < 0$ and consider the expression

$$\sum_\omega 2^l \|P_l^a Q_{k+2l}(\chi_I \phi)\|_{L^1_x L^\infty_t}^2.$$  

Write

$$P_l^a Q_{k+2l}(\chi_I \phi) = P_l^a Q_{k+2l}(Q_{< k+2l-C}(\chi_I) \phi) + P_l^a Q_{k+2l}(Q_{\geq k+2l-C}(\chi_I) \phi).$$
For the first term on the right hand side, we have
\[ \left\| P^\mu_i Q_{k+2l} (Q_{\geq k+2l} - c(\chi) \phi) \right\|_{L^1_t L^2_x} \lesssim \left\| P^\mu_i Q_{k+2l} + O(1) \phi \right\|_{L^1_t L^2_x}, \]
which leads to an acceptable contribution. For the second term on the right hand side, we use
\[ \left\| P^\mu_i Q_{k+2l} (Q_{\geq k+2l} - c(\chi) \phi) \right\|_{L^1_t L^2_x} \lesssim \left\| Q_{\geq k+2l} - c(\chi) \right\|_{L^1_t} 2^j 2^{j+\frac{k}{2}} \left\| P^\mu_i \phi_k \right\|_{L^2_t L^2_x} \]
It follows that
\[ 2^j \left\| P^\mu_i Q_{k+2l} (Q_{\geq k+2l} - c(\chi) \phi) \right\|_{L^1_t L^2_x} \lesssim (2^{-\frac{k}{2}} \left\| P^\mu_i \phi_k \right\|_{L^2_t L^2_x}), \]
which can be square-summed over \( \omega \), see (9) in [12]. \( \square \)

3. Microlocalized magnetic wave equation

In this section we assume that the spatial components of the connection form \( A \) are solutions to the linear wave equation \( \Box A^j = 0 \) on \( \mathbb{R} \times \mathbb{R}^4 \) for \( j = 1, \ldots, 4 \) and that \( A \) is in Coulomb gauge. We define the magnetic wave operator
\[ \Box_A^\mu = \Box + 2i \sum_{k \in \mathbb{Z}} P_{\leq k} \Box A^j P_k \partial_j. \]
The goal of this section is to derive the following linear estimate for the magnetic wave operator \( \Box_A^\mu \).

**Theorem 3.1.** Suppose that \( \Box A^j = 0 \) on \( \mathbb{R} \times \mathbb{R}^4 \) for \( j = 1, \ldots, 4 \) and that \( A \) is in Coulomb gauge. For all \( f \in \mathcal{N}(\mathbb{R} \times \mathbb{R}^4) \) and \( (g, h) \in H^1_4(\mathbb{R}^4) \times L^2_4(\mathbb{R}^4) \), the solution to the magnetic wave equation
\[ \Box_A^\mu \phi = f \text{ on } \mathbb{R} \times \mathbb{R}^4, \]
\[ (\phi, \phi_t)|_{t=0} = (g, h) \]
exists globally and satisfies
\[ \|\phi\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq C \left( \|g\|_{H^1_4} + \|h\|_{L^2_4} + \|f\|_{(\mathcal{N} \cap \mathbf{U}^{L^2_4(\mathbb{R}^4)}_4)} \right), \]
where the constant \( C > 0 \) depends only on \( \|\nabla_{x,A} f\|_{L^2_4} \) and grows at most polynomially in \( \|\nabla_{x,A} f\|_{L^2_4} \).

**Proof.** By time reversibility it suffices to prove the existence of the solution \( \phi \) on the time interval \( [0, \infty) \). Let \( \varepsilon > 0 \) be a sufficiently small constant to be fixed later. We may cover the time interval \( [0, \infty) \) by finitely many consecutive closed intervals \( I_1, \ldots, I_J \) with the following properties. The number of intervals \( J \) depends only on \( \|\nabla_{x,A} f\|_{L^2_4} \) and \( \varepsilon \), the intervals \( I_j \) overlap at most two at a time, consecutive intervals have intersection with non-empty interior and \( [0, \infty) = \bigcup_{j=1}^{\infty} I_j \). Most importantly, the intervals \( I_j \) are chosen such that a finite number of suitable space-time norms of the magnetic potential \( A \) that will be specified later are less than \( \varepsilon \) uniformly on all intervals \( I_j \).

We first construct suitable local solutions to the magnetic wave equation (3.2) on the intervals \( I_j \). The precise statement is summarized in the following theorem. Its proof will be given further below and is based on a parametrix construction. The accuracy of the parametrix crucially relies on the above mentioned smallness of suitable space-time norms of the magnetic potential \( A \) on the intervals \( I_j \). We use the notation \( I_j = [T_j^{(l)}, T_j^{(r)}] \) for the left and right endpoints of \( I_j \).

**Theorem 3.2.** Let \( f \in \mathcal{N}(\mathbb{R} \times \mathbb{R}^4) \) and \( (\tilde{g}, \tilde{h}) \in H^1_4(\mathbb{R}^4) \times L^2_4(\mathbb{R}^4) \). For \( j = 1, \ldots, J \) there exists a solution \( \phi^{(j)} \in S^1(\mathbb{R} \times \mathbb{R}^4) \) to
\[ \Box_A^\mu \phi^{(j)} = f \text{ on } I_j \times \mathbb{R}^4, \]
\[ (\phi^{(j)}, \phi_t^{(j)} )|_{t=T_j^{(l)}} = (\tilde{g}, \tilde{h}) \]
in the sense that \( \| \chi_I (\nabla^P_A \phi^{(j)} - f) \|_{N(\mathbb{R} \times \mathbb{R}^4)} = 0 \) for a sharp cutoff \( \chi_I \) to the time interval \( I_j \). Moreover, it holds that

\[
\| \phi^{(j)} \|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq C \left( \| \tilde{g} \|_{H^1} + \| \tilde{h} \|_{L^2_x} + \| f \| \right)_{(N \cap f \subset L^2_x H^{\frac{1}{2}}(\mathbb{R} \times \mathbb{R}^4))},
\]

where the constant \( C > 0 \) depends only on \( \| \nabla_{t,x} A \|_{L^2_x} \).

Finally, we obtain the solution \( \phi \) to the magnetic wave equation (3.2) on \( [0, \infty) \times \mathbb{R}^4 \) by patching together suitable local solutions on the intervals \( I_j \). Given \( (g, h) \in H^1(\mathbb{R}^4) \times L^2(\mathbb{R}^4) \) and \( f \in N(\mathbb{R} \times \mathbb{R}^4) \), Theorem 3.2 yields a solution \( \phi^{(1)} \in S^1(\mathbb{R} \times \mathbb{R}^4) \) on \( I_1 = [0, T_1^{(r)}) \) to

\[
\Box^P_A \phi^{(1)} = f \text{ on } I_1 \times \mathbb{R}^4,
(\phi^{(1)}, \phi_t^{(1)})|_{t=0} = (g, h).
\]

Next, we obtain a solution \( \phi^{(2)} \in S^1(\mathbb{R} \times \mathbb{R}^4) \) on \( I_2 = [T_2^{(l)}, T_2^{(r)}] \) to

\[
\Box^P_A \phi^{(2)} = f \text{ on } I_2 \times \mathbb{R}^4,
(\phi^{(2)}, \phi_t^{(2)})|_{t=T_2^{(l)}} = (\phi^{(1)}(T_2^{(l)}), \phi_t^{(1)}(T_2^{(l)}))
\]

where we recall that \( I_1 \cap I_2 \neq \emptyset \) with \( T_2^{(l)} < T_1^{(r)} \). We proceed analogously for the remaining intervals \( I_3, \ldots, I_J \). By uniqueness, we must have \( \phi^{(j)}|_{I_j \cap I_{j+1}} = \phi^{(j+1)}|_{I_j \cap I_{j+1}} \) for \( j = 1, \ldots, J - 1 \). We choose a smooth partition of unity \( \{ \chi_j \} \) subordinate to the cover \( \{ I_j \} \) such that \( \text{supp}(\chi_j) \subset I_j \) and \( \text{supp}(\chi_j) \subset (I_{j-1} \cap I_j) \cup (I_j \cap I_{j+1}) \). We then define

\[
\phi = \sum_{j=1}^J \chi_j \phi^{(j)}.
\]

Since we have \( \chi_j' + \chi_{j+1}' = 0 \) on \( I_j \cap I_{j+1} \) for \( j = 1, \ldots, J - 1 \), it follows that

\[
\sum_{j=1}^J \chi_j' \phi^{(j)} = 0 \text{ on } \mathbb{R} \times \mathbb{R}^4
\]

and hence,

\[
\nabla_{t,x} \sum_{j=1}^J \chi_j \phi^{(j)} = \sum_{j=1}^J \chi_j \nabla_{t,x} \phi^{(j)} \text{ on } \mathbb{R} \times \mathbb{R}^4.
\]

Similarly, we find that

\[
\Box \sum_{j=1}^J \chi_j \phi^{(j)} = \sum_{j=1}^J \chi_j \Box \phi^{(j)}.
\]
Using Lemma 2.1 and estimate (3.5), we thus conclude that
\[ \|\phi\|_{L^1_t(\mathbb{R}^4)} = \left\| \sum_{j=1}^J \chi_j \phi^{(j)} \right\|_{L^1_t(\mathbb{R}^4)} \leq \sum_{j=1}^J \|\phi^{(j)}\|_{L^1_t(\mathbb{R}^4)} \]
\[ \leq C(\|\nabla_x A\|_{L^2_t}) \left( \sum_{j=1}^J \|\phi^{(j)}(T_j^{(t)}(f))\|_{H^1_x} + \|\partial_t \phi^{(j)}(T_j^{(t)}(f))\|_{L^2_x} + \|f\|_{N(\mathbb{R}^4)} \right) \]
\[ \leq C(J)C(\|\nabla_x A\|_{L^2_t})\left(\|g\|_{H^1_x} + \|\tilde{h}\|_{L^2_x} + \|f\|_{N(\mathbb{R}^4)} \right). \]
Since \( J \) depends only on the size of \( \|\nabla_x A\|_{L^2_t} \) and \( \varepsilon \), we obtain the desired estimate (3.3). \( \square \)

We proceed with the proof of Theorem 3.2.

**Proof of Theorem 3.2** We begin by considering for every \( k \in \mathbb{Z} \) the frequency localized problem
\[ \Box_{A_k}^{p} \phi^{(j)}_k = f_k \text{ on } I_j \times \mathbb{R}^4, \]
(3.6)
\[ (\phi^{(j)}_k, \partial_t \phi^{(j)}_k)_{|=T_j^{(t)}} = (\tilde{g}_k, \tilde{h}_k), \]
where \( \Box_{A_k}^{p} = \Box + 2iP_{k-1}C\partial_t \). Let \( \chi_{I_j} \) denote a sharp cutoff to the time interval \( I_j \). We first want to construct an approximate solution \( \phi^{(j)}_{app,k} \) to (3.6) that satisfies
\[ \|\phi^{(j)}_{app,k}\|_{L^1_t(\mathbb{R}^4)} \leq C \left( \|\tilde{g}_k\|_{H^1_x} + \|\tilde{h}_k\|_{L^2_x} + \|f_k\|_{N(\mathbb{R}^4)} \right) \]
and
\[ \|\phi^{(j)}_{app,k}(T_j^{(t)}(f)) - \tilde{g}_k\|_{H^1_x} + \|\partial_t \phi^{(j)}_{app,k}(T_j^{(t)}(f)) - \tilde{h}_k\|_{L^2_x} + \|\chi_{I_j}(\Box_{A_k}^{p} \phi^{(j)}_{app,k} - f_k)\|_{(N(\mathbb{R}^4) \cap L^2_tH^1_x)^4(\mathbb{R}^4)} \]
\[ \leq \varepsilon \left( \|\tilde{g}_k\|_{H^1_x} + \|\tilde{h}_k\|_{L^2_x} + \|f_k\|_{(N(\mathbb{R}^4) \cap L^2_tH^1_x)^4(\mathbb{R}^4)} \right). \]
(3.8)

To this end we split
\[ f_k = f_{k}^{hyp} + f_{k}^{ell}, \]
where \( f_{k}^{hyp} \) is supported in the region \( ||\tau|| - ||\xi|| \leq 2^k \). We note that it holds that
\[ \|\Box^{-1} f_{k}^{ell}\|_{L^1_t(\mathbb{R}^4)} \leq \|f_{k}^{ell}\|_{N(\mathbb{R}^4)}. \]
(3.9)

Theorem 3.3 below then yields an approximate solution \( \phi^{(j)}_{app,k} \) to
\[ \Box_{A_k}^{p} \phi^{(j)}_{app,k} = f_{k}^{hyp} \text{ on } I_j \times \mathbb{R}^4, \]
(3.10)
\[ (\phi^{(j)}_{app,k}, \partial_t \phi^{(j)}_{app,k})_{|=T_j^{(t)}} = (\tilde{g}_k, \tilde{h}_k) - ((\Box^{-1} f_{k}^{ell})(T_j^{(t)}), (\partial_t \Box^{-1} f_{k}^{ell})(T_j^{(t)}) \]
that satisfies
\[ \|\phi^{(j)}_{app,k}\|_{L^1_t(\mathbb{R}^4)} \leq \|\tilde{g}_k\|_{H^1_x} + \|\tilde{h}_k\|_{L^2_x} + \|f_k\|_{N(\mathbb{R}^4)} \]
(3.11)
and
\[
\|\tilde{\phi}_{app,k}^{(j)}(T_j^{(i)}) - (\tilde{g}_k - (\Box^{-1} f_{k}^{(j)}))\|_{L^2} + \|\partial_t \tilde{\phi}_{app,k}^{(j)}(T_j^{(i)}) - (\tilde{h}_k - (\Box^{-1} f_{k}^{(j)}))\|_{L^2_x}
\]
(3.12)
\[
+ \|\chi_l A_{k}^j \tilde{\phi}_{app,k}^{(j)}(T_j^{(i)}) - f_{k}^{(j)}\|_{N_0(\mathbb{R}^{3})}
\]
\[
\lesssim \epsilon(\|\tilde{g}_k\|_{L^2_x} + \|\tilde{h}_k\|_{L^2_x} + \|f_k\|_{N_0(\mathbb{R}^{3})})
\]

We remark that because of scaling invariance Theorem [3.3] below is only formulated for the case \( k = 0 \). Next we set
\[
\phi_{app,k}^{(j)} = \tilde{\phi}_{app,k}^{(j)} + (\Box^{-1} f_{k}^{(j)})
\]
and find that
\[
\|\chi_l (\Box_{\mathbb{A}} A_\mathbb{A}_{k}^{j}) \phi_{app,k}^{(j)}(T_j^{(i)}) - f_k\|_{(N_0 \cap L^2_x H^\frac{1}{2}) (\mathbb{R}^{3})}
\]
(3.13)
\[
\lesssim \chi_l A_{k}^j \|P_k \partial_t (\Box^{-1} f_{k}^{(j)})\|_{(N_0 \cap L^2_x H^\frac{1}{2}) (\mathbb{R}^{3})}
\]
\[
\lesssim \epsilon(\|\Box^{-1} f_{k}^{(j)}\|_{L^1(\mathbb{R}^{3})})
\]
\[
\lesssim \|f_k\|_{N_0(\mathbb{R}^{3})}
\]

Here we used that the intervals \( I_j \) can be chosen such that uniformly for all \( j = 1, \ldots, J \),
\[
\|\mathcal{A}\|_{L^2_t L^2_x(I_j \times \mathbb{R}^3)} \leq \epsilon
\]
and thus,
\[
\|\chi_l A_{k}^j P_k \partial_t (\Box^{-1} f_{k}^{(j)})\|_{(N_0 \cap L^2_x H^\frac{1}{2}) (\mathbb{R}^{3})}
\]
\[
\lesssim \|\chi_l A_{k}^j \|_{L^2_t L^2_x(H^{\frac{1}{2}})} \|P_k \nabla_x (\Box^{-1} f_{k}^{(j)})\|_{(L^2_x \cap L^\infty_x H^{\frac{1}{2}})}
\]
\[
\lesssim \epsilon(\|\Box^{-1} f_{k}^{(j)}\|_{L^1(\mathbb{R}^{3})})
\]
\[
\lesssim \|f_k\|_{N_0(\mathbb{R}^{3})}
\]

From (3.9), (3.11), (3.12), and (3.13) it now follows immediately that \( \phi_{app,k}^{(j)} \) is an approximate solution to (3.6) that satisfies the estimates (3.7) and (3.8).

Finally, we reassemble the approximate solutions \( \phi_{app,k}^{(j)} \) to the frequency localized problems (3.6) to a full approximate solution \( \phi_{app} = \sum_{k \in \mathbb{Z}} \phi_{app,k}^{(j)} \) to (3.4) satisfying
\[
\|\phi_{app}(T_j^{(i)}) - \tilde{g}\|_{L^2} + \|\partial_t \phi_{app}(T_j^{(i)}) - \tilde{h}\|_{L^2} + \|\chi_l (\Box A \phi_{app} - f)\|_{N_0 \cap L^2_x H^\frac{1}{2}}
\]
\[
\lesssim \epsilon(\|\tilde{g}\|_{L^2} + \|\tilde{h}\|_{L^2} + \|f\|_{N_0 \cap L^2_x H^\frac{1}{2}})
\]
and
\[
\|\phi_{app}\|_{L^1(\mathbb{R}^{3})} \lesssim \|\tilde{g}\|_{L^2} + \|\tilde{h}\|_{L^2} + \|f\|_{N_0 \cap L^2_x H^\frac{1}{2}}
\]

Applying this procedure iteratively to the successive errors, we obtain an exact solution \( \phi^{(j)} \) to (3.4) satisfying (3.5).

We now turn to the heart of the matter, namely the construction of the approximate solutions to the frequency localized magnetic wave equations.
Theorem 3.3. Let $(\tilde{g}, \tilde{h}) \in H^1_t(L^2_\xi(\mathbb{R}^4) \times L^2_\xi(\mathbb{R}^4))$ and $\tilde{f} \in N(\mathbb{R} \times \mathbb{R}^4)$. Assume that $\tilde{f}, \tilde{g}, \tilde{h}$ are frequency localized at $|\xi| \sim 1$ and that $\tilde{f}$ is localized at modulation $|\tau| - |\xi| \lesssim 1$. For $j = 1, \ldots, J$ there exists an approximate solution $\tilde{\phi}_{app}^{(j)}$ to
\begin{equation}
\Box_{A,0}^\rho \phi = \tilde{f} \text{ on } I_j \times \mathbb{R}^4,
\end{equation}
in the sense that
\begin{equation}
\|\tilde{\phi}_{app}^{(j)}\|_{\mathcal{S}_0(\mathbb{R} \times \mathbb{R}^4)} \lesssim \|\tilde{g}\|_{L^2_\xi} + \|\tilde{h}\|_{L^2_\xi} + \|\tilde{f}\|_{N_0(\mathbb{R} \times \mathbb{R}^4)}
\end{equation}
and
\begin{equation}
\|\tilde{\phi}_{app}^{(j)}(T^j) - \tilde{g}\|_{L^2_\xi} + \|\tilde{\phi}_{app}^{(j)}(T^j) - \tilde{h}\|_{L^2_\xi} + \|\chi_j(\Box_{A,0}^\rho \tilde{\phi}_{app}^{(j)} - \tilde{f})\|_{N_0(\mathbb{R} \times \mathbb{R}^4)}
\lesssim e(\|\tilde{g}\|_{L^2_\xi} + \|\tilde{h}\|_{L^2_\xi} + \|\tilde{f}\|_{N_0(\mathbb{R} \times \mathbb{R}^4)}),
\end{equation}
where $\chi_j$ denotes a sharp cutoff to the time interval $I_j$.

Proof. In order to prove estimates and construct a parametrix for the frequency localized magnetic wave equation (3.14) we adapt the scheme in Section 6 of [12] to our time-localized setting. We will use frequency localized renormalization operators $e^{-i\psi\xi}(t, x, D)$ and $e^{+i\psi\xi}(D, y, s)$, where $P(x, D)$ denotes the left quantization and $P(D, y)$ the right quantization of a pseudodifferential operator $P$ and where the subscript $< 0$ denotes the space-time frequency localization of the symbol at frequencies $\ll 1$. For the definition of the phase correction $\psi_\pm$ in the renormalization operator $e^{+i\psi\xi}(D, y, s)$ we need to introduce some notation.

For any $\zeta \in \mathbb{R}^4 \setminus \{0\}$ we set
\[ \omega = \frac{\zeta}{|\zeta|}, \quad L^\omega_\pm := \pm \partial_t + \omega \cdot \nabla_x, \quad \Delta^\omega := \Delta - (\omega \cdot \nabla_x)^2. \]
Moreover, for any $\omega \in S^3$ and any angle $0 < \theta \lesssim 1$, we define the sector projection $\Pi^{\omega}_\theta$ in frequency space by the formula
\[ \Pi^{\omega}_\theta f(\zeta) := \left(1 - \eta\left(\frac{\zeta(\omega \cdot \zeta)}{\theta}\right)\right)\left(1 - \eta\left(\frac{\zeta(-\omega \cdot \zeta)}{\theta}\right)\right)f(\zeta), \]
where $\eta(y)$ is a bump function on $\mathbb{R}$ which equals 1 when $|y| < \frac{4}{3}$ and vanishes for $|y| > 1$, and $\zeta(\omega, \omega)$ is the angle between $\zeta$ and $\omega$. Thus, $\Pi^{\omega}_\theta$ restricts $f$ smoothly (except at the frequency origin) to the sector of frequencies $\zeta$ whose angle with both $\omega$ and $-\omega$ is $\geq \theta$. Similarly, we define the Fourier multipliers $\Pi^\omega_\theta$, $\Pi^{\omega}_0$ and $\Pi^{\omega}_{\theta > \theta_2}$.

Let $C_1, C_2 > 0$ be constants to be chosen sufficiently large later on depending on the size of $\|\nabla_{t,x} A\|_{L^2_\xi}$ and let $\sigma > 0$ be a constant to be chosen sufficiently small. We then define the phase correction $\psi_\pm$ by
\begin{equation}
\psi_\pm(t, x, \zeta) = \sum_{-C_2 < k < 0} L^\omega_\pm \Delta^{-1} \Pi^{\omega}_{2^1 \xi_1 > 2^2 \xi_2 < 2^3 \xi_3} A_k \cdot \omega + \sum_{k < -C_2} L^\omega_\pm \Delta^{-1} \Pi^{\omega}_{2^1 \xi_1 > 2^2 \xi_2 > 2^3 \xi_3} A_k \cdot \omega.
\end{equation}
Note that the first sum effectively only starts at $k \leq -\frac{C_2}{\sigma}$. See Section 6 in [12] for a motivation for such a choice of phase correction. We emphasize that this phase slightly differs from the one used in [12], because for intermediate frequencies $-C_2 \leq k < 0$ the high angles are cut off.
We define the approximate solution $\tilde{\phi}_{app}^{(j)}$ to (3.14) by

$$
\tilde{\phi}_{app}^{(j)}(t) = \frac{1}{2} \sum_{\pm} \sum_{c} \int_{-\infty}^{\infty} e^{i\phi(t, x, D)} e^{i\phi(t, y, T^{(l)}_j)} |D| \left| e^{i\phi} \right| \left( D, y, s \right)(-i\tilde{f}) \right),
$$

where

$$
K_j \tilde{f}(t) = \int_{T_j}^{t} e^{i(t-s)|D|} \tilde{f}(s) d s.
$$

In order to prove the estimates (3.15) and (3.16) we establish the following crucial time-localized mapping properties of the renormalization operator $e^{i\phi t}$.

**Theorem 3.4.** For $j = 1, \ldots, J$, the frequency localized renormalization operators have the following mapping properties with $Z \in \{N_0(\mathbb{R} \times \mathbb{R}^4), L^2_0(\mathbb{R}^4), N^*_0(\mathbb{R} \times \mathbb{R}^4)\}$,

\begin{align}
(3.18) \quad & \chi I \cdot e^{i\phi t} : Z \rightarrow Z, \\
(3.19) \quad & \chi I \cdot \partial_t e^{i\phi t} : Z \rightarrow eZ, \\
(3.20) \quad & \chi I \cdot (e^{i\phi t} + \phi(t, x, D)e^{i\phi t}) : Z \rightarrow eZ, \\
(3.21) \quad & \chi I \cdot (e^{i\phi t} + \Delta_p \phi(t, x, D)) : N^*_0(\mathbb{R} \times \mathbb{R}^4) \rightarrow eN_0(\mathbb{R} \times \mathbb{R}^4), \\
(3.22) \quad & \chi I \cdot e^{i\phi t} : S^*_{0}(\mathbb{R} \times \mathbb{R}^4) \rightarrow S_0(\mathbb{R} \times \mathbb{R}^4),
\end{align}

where $\chi I$ denotes a sharp cutoff to the time interval $I_j$. In the estimates (3.15) and (3.16), the operator $e^{i\phi t}$ and, respectively $\partial_t e^{i\phi t}$, stands for both left and right quantization.

The estimates (3.15) and (3.16) then follow by adapting the manipulations in the proof of Theorem 4 in [12] to our time-localized setting.

The remainder of this section is devoted to the proof of Theorem 3.4. To this end we will adapt the general scheme of Sections 7 – 11 in [12] to our large data setting. The accuracy of the approximate solution $\tilde{\phi}_{app}^{(j)}$ relies on the error estimates (3.19), (3.20) and (3.21). While in [12] the small energy assumption can be used to achieve smallness in the corresponding error estimates, we have to argue more carefully here, using the high angle cut-off in the definition of the phase correction and smallness of suitable space-time norms of $A$ on sufficiently small time intervals, namely the intervals $I_j$.

### 3.1. Decomposable function spaces.

We begin by reviewing the notion of decomposable function spaces and estimates from [19], [11], and [12].

Let $c(t, x, D)$ be a pseudodifferential operator whose symbol $c(t, x, \xi)$ is homogeneous of degree 0 in $\xi$. Assume that $c$ has a representation

$$
c = \sum_{\theta \in 2^{-N} \mathbb{Z}^d} c^{(\theta)}.
$$

Let $1 \leq q, r \leq \infty$. For every $\theta \in 2^{-N} \mathbb{Z}^d$, we define

$$
\|c^{(\theta)}\|_{D_{0}(L^q_0(L^r_0(\mathbb{R} \times \mathbb{R}^4)))} = \left( \sum_{l=0}^{100} \sum_{l=0}^{100} \sup_{\omega \in \Gamma_\theta} \| \hat{b}_{\gamma}(\omega)(\theta \nabla_{\xi})^{l} c^{(\theta)} \|_{L^r_0(\mathbb{R}^4)}^{2} \right)^{1/2},
$$

where
where \( \{ \Gamma^p \}_\theta \) is a uniformly finitely overlapping covering of \( S^3 \) by caps of diameter \( \sim \theta \) and \( \{ b^p \}_\theta \) is a smooth partition of unity subordinate to the covering \( \{ \Gamma^p \}_\theta \). Then we define the decomposable norm

\[
\|c\|_{D(L^q_s L_t^r)(\mathbb{R} \times \mathbb{R}^4)}^\ast = \inf_{c = \sum c^{(\theta)} \sum_{\theta \in 2^{-i}}} \|c^{(\theta)}\|_{D(L^q_s L_t^r)(\mathbb{R} \times \mathbb{R}^4)}.
\]

We will repeatedly use the following decomposable estimates.

**Lemma 3.5** ([12] Lemma 7.1). Let \( P(t, x, D) \) be a pseudodifferential operator with symbol \( p(t, x, \xi) \). Suppose that \( P \) satisfies the fixed-time estimate

\[
\sup_{t \in \mathbb{R}} \|P(t, x, D)\|_{L^2_x \to L^2_t} \leq 1.
\]

Let \( 1 \leq q, q_1, q_2, r, r_1 \leq \infty \) such that \( \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \) and \( \frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_2} \). For any symbol \( c(t, x, \xi) \in D(L^{q_1}_t L^{r_1}_x)(\mathbb{R} \times \mathbb{R}^4) \) that is zero homogeneous in \( \xi \), we have

\[
\|c(t, x, D)\phi\|_{L^q_t L^r_x(\mathbb{R} \times \mathbb{R}^4)} \leq \|c\|_{D(L^{q_1}_t L^{r_1}_x)(\mathbb{R} \times \mathbb{R}^4)} \|\phi\|_{L^{q_2}_t L^{r_2}_x(\mathbb{R} \times \mathbb{R}^4)}.
\]

By duality we obtain decomposable estimates for right quantizations.

**Lemma 3.6.** Let \( P \) be a pseudodifferential operator with symbol \( p(t, x, \xi) \). Suppose that \( P \) satisfies the fixed-time estimate

\[
\sup_{t \in \mathbb{R}} \|P(t, x, D)\|_{L^2_x \to L^2_t} \leq 1.
\]

Let \( 1 \leq q < \infty \) and \( 1 \leq q_1, q_2 \leq \infty \) such that \( \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \). For any symbol \( c(t, x, \xi) \in D(L^{q_1}_t L^{r_1}_x)(\mathbb{R} \times \mathbb{R}^4) \) that is zero homogeneous in \( \xi \), the right-quantized operator \((\mathcal{T} \mathcal{P})(D, y, t)\) has the following mapping property

\[
\|\mathcal{T} \mathcal{P}(D, y, t)\phi\|_{L^q_t L^r_x(\mathbb{R} \times \mathbb{R}^4)} \leq \|c\|_{D(L^{q_1}_t L^{r_1}_x)(\mathbb{R} \times \mathbb{R}^4)} \|\phi\|_{L^{q_2}_t L^{r_2}_x(\mathbb{R} \times \mathbb{R}^4)}.
\]

**Proof.** Let \( 1 < q' \leq \infty \) be the the conjugate exponent to \( q \) and define \( \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} \). By duality, Hölder’s inequality and Lemma 3.5, we have

\[
\|\mathcal{T} \mathcal{P}(D, y, t)\phi\|_{L^q_t L^r_x} = \sup_{\|\phi\|_{L^{q_2}_t L^{r_2}_x} \leq 1} \langle \phi, (\mathcal{T} \mathcal{P})(D, y, t)\phi \rangle
\]

\[
= \sup_{\|\phi\|_{L^{q_2}_t L^{r_2}_x} \leq 1} \langle c(t, x, D)\psi, \phi \rangle
\]

\[
\leq \sup_{\|\phi\|_{L^{q_2}_t L^{r_2}_x} \leq 1} \|c(t, x, D)\psi\|_{L^{q_1}_t L^{r_1}_x} \|\phi\|_{L^{q_2}_t L^{r_2}_x}
\]

\[
\leq \sup_{\|\phi\|_{L^{q_2}_t L^{r_2}_x} \leq 1} \|c\|_{D(L^{q_1}_t L^{r_1}_x)} \|\psi\|_{L^{q_1}_t L^{r_1}_x} \|\phi\|_{L^{q_2}_t L^{r_2}_x}
\]

\[
\leq \|c\|_{D(L^{q_1}_t L^{r_1}_x)} \|\phi\|_{L^{q_2}_t L^{r_2}_x}.
\]

From [11] Lemma 10.2] we have the following Hölder-type estimate for decomposable norms

\[
(3.23) \quad \prod_{i=1}^m c_i \|D(L^{q_i}_t L^{r_i}_x) \| \leq \prod_{i=1}^m \|c_i\|_{D(L^{q_i}_t L^{r_i}_x)},
\]

where \( m \in \mathbb{N}, 1 \leq q, r, q_i, r_i \leq \infty \) for \( i = 1, \ldots, m \) and \( \frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{q_3} + \cdots + \frac{1}{q_m} \).
3.2. Some symbol bounds for phases. Recall that the magnetic potential $A$ is assumed to be supported at frequencies $\leq 1$. For any integer $k < 0$ and any dyadic angle $0 < \theta \leq 1$, we use the notation
$$\psi_k^{(0)}(t, x, \xi) = L_A^\omega A_k \cdot \omega$$
and
$$\psi_{ck} = \sum_{l < k} P_l \psi_z.$$ 

**Lemma 3.7.** For any $t, s \in \mathbb{R}$, $x, y \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$ and any integer $k < 0$, it holds that
$$|\psi_k(t, x, \xi) - \psi_k(s, y, \xi)| \leq (2^{-C_1/2} + 2^{-C_2}) |\nabla_{t,x} A(0)|_{L^2_t}(t - s) + |x - y|.$$ 

Moreover, we have for any multi-index $\alpha \in \mathbb{N}_0^d$ with $1 \leq |\alpha| \leq \sigma^{-1}$ that
$$|\nabla_{t,x}^\alpha (\psi(t, x, \xi) - \psi(s, y, \xi))| \leq (t - s)^{\sigma (|\alpha| - 1)} |\nabla_{t,x} A(0)|_{L^2_t}.$$ 

**Proof.** For any $t \in \mathbb{R}, x \in \mathbb{R}^d$, $\xi \in \mathbb{R}^d$ and any integer $k < 0$, we obtain that
$$|\psi_k^{(0)}(t, x, \xi)| \leq \|L_A^\omega A_k \cdot \omega\|_{L^\infty_t} \leq (\theta^2)^{1/2} \|L_A^\omega A_k \cdot \omega\|_{L^\infty_t} \leq \theta^{3/2} 2^k \|\nabla_{t,x} A(0)|_{L^2_t} \leq \theta^{1/2} |\nabla_{t,x} A_k|_{L^2_t},$$
where we used Bernstein’s inequality, the Coulomb gauge of $A$ and that $|\Delta_{\omega}^{-1}(\xi)| \sim 2^{-2k} \theta^2$ on the frequency support of $\Pi^0 A_k$. Similarly, we find
$$|\nabla_{t,x}^\alpha \psi_k^{(0)}(t, x, \xi)| \leq 2^k \theta^{1/2} |\nabla_{t,x} A_k|_{L^2_t}.$$ 

Thus, we have
$$|\psi_k(t, x, \xi) - \psi_k(s, y, \xi)|$$
$$\leq \sum_{k \leq -C_2} \sum_{2^k \leq s, t < 2^k} |\psi_k^{(0)}(t, x, \xi) - \psi_k^{(0)}(s, y, \xi)| + \sum_{k \leq -C_2} \sum_{2^k \leq s, t < 2^k} |\psi_k^{(0)}(t, x, \xi) - \psi_k^{(0)}(s, y, \xi)|$$
$$\leq \left( \sum_{k \leq -C_2} \sum_{2^k \leq s, t < 2^k} 2^k \theta^{1/2} + \sum_{k \leq -C_2} \sum_{2^k \leq s, t < 2^k} 2^k \theta^{1/2} \right)^{1/2} |\nabla_{t,x} A_k|_{L^2_t}(|x - y| + |t - s|)$$
$$\leq (2^{-C_1/2} + 2^{-C_2}) |\nabla_{t,x} A_k|_{L^2_t}(|x - y| + |t - s|).$$

We now turn to the proof of (3.25). To this end we note that differentiating with respect to $\xi$ yields $\theta^{-1}$ factors, i.e. for any $\alpha \in \mathbb{N}_0^d$ it holds that
$$|\nabla_{\xi}^\alpha \psi_k^{(0)}(t, x, \xi)| \leq \theta^{3/2 - |\alpha|} |\nabla_{t,x} A_k|_{L^2_t}$$
and
$$|\nabla_{t,x} \nabla_{\xi}^\alpha \psi_k^{(0)}(t, x, \xi)| \leq 2^k \theta^{3/2 - |\alpha|} |\nabla_{t,x} A_k|_{L^2_t},$$
For any $1 \leq |\alpha| \leq \sigma^{-1}$ and $l < 0$ we then obtain
$$|\nabla_{\xi}^\alpha (\psi_k(t, x, \xi) - \psi_k(s, y, \xi))|$$
$$\leq \sum_{k \leq -C_2} \sum_{2^k \leq s, t < 2^k} 2^k \theta^{3/2 - |\alpha|} |\nabla_{t,x} A_k|_{L^2_t}(|x - y| + |t - s|) + \sum_{k \leq -C_2} \sum_{2^k \leq s, t < 2^k} \theta^{3/2 - |\alpha|} |\nabla_{t,x} A_k|_{L^2_t}$$
$$\leq 2^{l(1 - \sigma (|\alpha| - 3/2))} |\nabla_{t,x} A_k|_{L^2_t}(|x - y| + |t - s|) + 2^{-\sigma (|\alpha| - 3/2)} |\nabla_{t,x} A_k|_{L^2_t}.$$
Optimizing the choice of \( l < 0 \) we find that
\[
|\nabla_\xi^l(\psi_\pm(t, x, \xi) - \psi_\pm(s, y, \xi))| \leq \langle |t - s| + |x - y| \rangle^{-\frac{1}{2}} |\nabla_{t,x}A|_{L^2_x}.
\]
\[
\square
\]

We will frequently use the following bounds on decomposable norms of the phase \( \psi_\pm \).

**Lemma 3.8 (\[12\] Lemma 7.3).** Let \( 2 \leq q, r \leq \infty \) with \( \frac{2}{q} + \frac{3}{r} \leq \frac{3}{2} \). For any integer \( k < 0 \) and any dyadic angle \( \theta \in 2^{-N} \) the component \( \psi_k(\theta) = L^\omega_\pm \Delta^\omega_\pm 1 \Pi^\omega_\theta A_k \cdot \omega \) satisfies
\[
(3.26) \quad \left\| (\psi_k^{(t)}, 2^{-k} \nabla_{t,x_0} \psi_k^{(t)}) \right\|_{D_k^0 L^2_x L^r_t}(\mathbb{R} \times \mathbb{R}^4) \leq 2^{-\left(\frac{1}{4} + \frac{3}{r} + \frac{3}{2} - \frac{1}{q} - \frac{1}{2} \right)} \left\| \nabla_{t,x} A \right\|_{L^2_x}.
\]

### 3.3. Oscillatory integral estimates

In order to prove the mapping properties in Theorem 3.4 we need pointwise kernel bounds for operators of the form
\[
T_a = e^{-i \varphi_y}(t, x, D) a(D) e^{i \varphi_x(D, y, s)},
\]
where \( a \) is localized at frequency \( |\xi| \sim 1 \). The kernel of \( T_a \) is given by the oscillatory integral
\[
K_a(t, x; y, s) = \int_{\mathbb{R}^d} e^{-i \varphi_x(t, x_0, \xi)} e^{i \varphi_y(D, y_0, s)} a(\xi) d\xi,
\]
where \( a \) is a smooth bump function with support on the annulus \( |\xi| \sim 1 \).

**Lemma 3.9.** For any \( t, s \in \mathbb{R}, x, y \in \mathbb{R}^4 \) and any integer \( 1 \leq N \leq \sigma^{-1} \), we have
\[
(3.27) \quad |K_a(t, x; s, y)| \leq \frac{\left\| \nabla_{t,x} A \right\|_{L^2_x}}{\langle |x - y| \rangle^N(1-\sigma)}
\]
and
\[
(3.28) \quad |K_a(t, x; t, y) - \tilde{a}(x - y)| \leq \min(\langle 2^{-C_1/2} + 2^{-C_2} \rangle, \frac{1}{\langle |x - y| \rangle^N(1-\sigma)}}) \left\| \nabla_{t,x} A \right\|_{L^2_x}.
\]

Moreover, it holds that
\[
(3.29) \quad |K_a(t, x; s, y)| \leq \langle t - s \rangle^{-\frac{1}{2}} \langle |t - s| - |x - y| \rangle^{-N} \left\| \nabla_{t,x} A \right\|_{L^2_x}.
\]

**Proof.** If \( |x - y| \leq 1 \) we obtain by taking absolute values that
\[
|K_a(t, x; s, y)| \leq 1.
\]
If instead \( |x - y| \gg 1 \), we use (3.25) and integrate by parts repeatedly to obtain for any \( 1 \leq N \leq \sigma^{-1} \) that
\[
|K_a(t, x; t, y)| \leq \frac{\left\| \nabla_{t,x} A \right\|_{L^2_x}}{|x - y|^{N(1-\sigma)}}.
\]
This proves (3.27). In order to show (3.28), we integrate by parts repeatedly for \( |x - y| \gg 1 \), while for \( |x - y| \leq 1 \), we use (3.24) to estimate
\[
|K_a(t, x; t, y) - \tilde{a}(x - y)| \leq \int_{\mathbb{R}^d} |e^{i \varphi_x(t, x_0, \xi)} - 1| |a(\xi)| d\xi,
\]
\[
\leq \int_{\mathbb{R}^d} |\psi_x(t, x, \xi) - \psi_x(t, y, \xi)| |a(\xi)| d\xi,
\]
\[
\leq (2^{-C_1/2} + 2^{-C_2}) \left\| \nabla_{t,x} A \right\|_{L^2_x} |x - y|.
\]
Finally, the proof of (3.29) proceeds along the lines of Proposition 6(a) in [12]. We only have to argue a bit more that away from the cone for sufficiently large $|t - s|$ the phase is still non-degenerate. But this is because away from the cone
\[
\left| -i \nabla_{\xi} \psi_{\pm}(t, x, \xi) - \psi_{\pm}(s, y, \xi) + i(t - s) \frac{\xi}{|\xi|} + i(x - y) \right|
\geq c(|t - s| + |x - y|) - C \|\nabla_{t,x} A\| \langle |t - s| + |x - y| \rangle^{\frac{1}{2} + \sigma}
\]
and we choose $0 < \sigma \ll 1$ sufficiently small. □

To deal with the frequency localized operators $e^{\pm i\psi_{\pm}}(t, x, D)$ and $e^{i\psi_{\pm}}(D, y, s)$, we need to produce similar estimates for the kernel $K_{a, <0}$ of the operator
\[
T_{a, <0} = e^{\pm i\psi_{\pm}}(t, x, D) a(D) e^{\pm i(t - s)|D|} e^{i\psi_{\pm}}(D, y, s).
\]
Noting that the frequency localized symbol $e^{\pm i\psi_{\pm}}$ can be represented as
\[
e^{\pm i\psi_{\pm}} = \int_{\mathbb{R}^{1+4}} m(z) e^{\pm iT_{\psi_{\pm}} z} dz,
\]
where $m(z)$ is an integrable bump function on the unit scale and $T_{z}$ denotes space-time translation in the direction $z \in \mathbb{R}^{1+4}$, the transition to these frequency localized operators can be made just as in Proposition 7 in [12]. We obtain the following estimates for $K_{a, <0}$.

**Lemma 3.10.** For any $t, s \in \mathbb{R}, x, y \in \mathbb{R}^{4}$ and any integer $1 \leq N \leq \sigma^{-1}$, we have
\[
|K_{a, <0}(t, x; t, y)| \leq \frac{\|\nabla_{t,x} A\|_{L^{2}_{x,y}}}{\langle |x - y| \rangle^{N(1 - \sigma)}}
\]
and
\[
|K_{a, <0}(t, x; t, y) - a(x - y)| \leq \min\left\{2^{-C_{1}/2} + 2^{-C_{2}}, \frac{1}{|x - y|^{N(1 - \sigma)}} \right\} \|\nabla_{t,x} A\|_{L^{2}_{x,y}}^{2}.
\]
Moreover, it holds that
\[
|K_{a, <0}(t, x; s, y)| \leq (t - s)^{-\frac{1}{2}} |t - s|^{-\frac{1}{2}} |x - y|^{-N} \|\nabla_{t,x} A\|_{L^{2}_{x,y}}^{2}.
\]

### 3.4. Fixed-time $L^{2}_{x}$ estimates in Theorem 3.4

In this subsection we prove the fixed-time $L^{2}_{x}$ estimates in Theorem 3.4 using the above oscillatory integral estimates. To obtain a small factor $\varepsilon$ in the estimates (3.18) and (3.19), we additionally have to fix the constants $C_{1}, C_{2} > 0$ in the definition of the phase correction $\psi_{\pm}$ sufficiently large.

**Lemma 3.11.** For any $t \in \mathbb{R}$, we have
\[
\left\| e^{\pm i\psi_{\pm}}(t, x, D) P_{0} \phi \right\|_{L^{2}_{x}} \leq \|\nabla_{t,x} A\|_{L^{2}_{x,y}} \|P_{0} \phi\|_{L^{2}_{x}}
\]
and
\[
\left\| e^{\pm i\psi_{\pm}}(D, y, t) P_{0} \phi \right\|_{L^{2}_{x}} \leq \|\nabla_{t,x} A\|_{L^{2}_{x,y}} \|P_{0} \phi\|_{L^{2}_{x}}
\]

**Proof.** The claim follows immediately from the kernel bound (3.30) and a $TT^{*}$-argument. □

**Lemma 3.12.** For any $\varepsilon > 0$ the constants $C_{1}, C_{2} > 0$ in the definition of the phase correction $\psi_{\pm}$ can be chosen sufficiently large (depending on the size of $\varepsilon^{-1}$ and $\|\nabla_{t,x} A\|_{L^{2}_{x,y}}$) such that we have
\[
\left\| \left(\nabla_{t,x} e^{\pm i\psi_{\pm}}(t, x, D) P_{0} \phi \right) \right\|_{L^{2}_{x}} \leq \varepsilon \|P_{0} \phi\|_{L^{2}_{x}}.
\]
Proof. Using Lemma 3.5 and (3.33) we obtain that
\[
\left\| (\nabla_{t,x} e^{-i\theta_x}) (t, x, D) P_0 \phi \right\|_{L^2_x} \leq \left\| (\nabla_{t,x} \psi_\pm) e^{-i\theta_x} (t, x, D) P_0 \phi \right\|_{L^2_x}
\]
\[
\leq \sum_{-C_1 \leq k < 0} \sum_{2^{\theta_k} < 2^{-C_1}} \left\| (\nabla_{t,x} \psi_k^{(t)}) e^{-i\theta_x} (t, x, D) P_0 \phi \right\|_{L^2_x} + \sum_{k < -C_2} \sum_{2^{\theta_k} < |\theta|} \left\| (\nabla_{t,x} \psi_k^{(t)}) e^{-i\theta_x} (t, x, D) P_0 \phi \right\|_{L^2_x}
\]
\[
\leq \sum_{-C_1 \leq k < 0} \sum_{2^{\theta_k} < 2^{-C_1}} \left\| (\nabla_{t,x} \psi_k^{(t)}) \|D_\beta\|_{L^\infty_y L^\infty_t} \|\nabla_{t,x} A\|_{L^2_x} \right\|_{L^2_x} P_0 \phi \right\|_{L^2_x} + \sum_{k < -C_2} \sum_{2^{\theta_k} < |\theta|} \left\| (\nabla_{t,x} \psi_k^{(t)}) \|D_\beta\|_{L^\infty_y L^\infty_t} \|\nabla_{t,x} A\|_{L^2_x} \right\|_{L^2_x} P_0 \phi \right\|_{L^2_x}
\]
\[
\leq \left( \sum_{-C_1 \leq k < 0} \sum_{2^{\theta_k} < 2^{-C_1}} 2^k \theta^{1/2} + \sum_{k < -C_2} \sum_{2^{\theta_k} < |\theta|} 2^k \theta^{1/2} \right) \left\| (\nabla_{t,x} A)^2 \right\|_{L^2_x} \|P_0 \phi\|_{L^2_x}
\]
\[
\leq (2^{-C_1/2} + 2^{-C_2}) \left\| (\nabla_{t,x} A)^2 \right\|_{L^2_x} \|P_0 \phi\|_{L^2_x},
\]
from which the assertion follows. \(\square\)

Lemma 3.13. For any \( \epsilon > 0 \) the constants \( C_1, C_2 > 0 \) in the definition of the phase correction \( \psi_\pm \) can be chosen sufficiently large (depending on the size of \( \epsilon^{-1} \) and \( \|\nabla_{t,x} A\|_{L^2_x} \)) such that we have
\[
(3.36) \quad \| (e^{-i\theta_x} (t, x, D) e^{i\theta_x} (D, y, t) - 1) P_0 \phi \|_{L^2_x} \leq \epsilon \|P_0 \phi\|_{L^2_x}.
\]

Proof. The integral kernel of
\[
(e^{-i\theta_x} (t, x, D) e^{i\theta_x} (D, y, t) - 1) \delta(D)
\]
is given by \( K_{a,0}(t, x; t, y) - \tilde{\delta}(x-y) \). Using (3.31) we find that
\[
\sup_y \int_{\mathbb{R}^4} |K_{a,0}(t, x; t, y) - \tilde{\delta}(x-y)| \, dx
\]
\[
\leq \int_{\mathbb{R}^4} \min\{(2^{-C_1/2} + 2^{-C_2}) \frac{1}{|x|^{N(1-\sigma)}}, \frac{1}{|x|^{N(1-\sigma)}}\} \|\nabla_{t,x} A\|_{L^2_x} \, dx
\]
\[
\leq \inf_{R > 0} \left\{ (2^{-C_1/2} + 2^{-C_2}) R^4 + \frac{1}{R^{N(1-\sigma)}} \right\} \|\nabla_{t,x} A\|_{L^2_x}.
\]
Choosing \( C_1, C_2 > 0 \) sufficiently large depending on the size of \( \epsilon^{-1} \) and \( \|\nabla_{t,x} A\|_{L^2_x} \), we obtain that
\[
\sup_y \int_{\mathbb{R}^4} |K_{a,0}(t, x; t, y) - \tilde{\delta}(x-y)| \, dx \leq \epsilon
\]
and similarly for \( \sup_x \int_{\mathbb{R}^4} |K_a(t, x; t, y) - \tilde{\delta}(x-y)| \, dy \). The assertion then follows from Schur’s lemma. \(\square\)

Remark 3.14. The fixed-time \( L^2 \) bounds from Lemma 3.11, Lemma 3.12 and Lemma 3.13 in fact hold for the operators \( e^{2i\theta_x} (t, x, D) e^{-i\theta_x} (t, x, D) \) and \( e^{2i\theta_x} (t, x, D) \) for any \( k, l < 0 \). The proofs in this and the previous subsection can be easily adapted to obtain this assertion.
3.5. Modulation localized estimates. All implicit constants in this subsection may depend on the size of $\|\nabla_{t,x}A\|_{L^2_t}$.

**Proposition 3.15.** For any $\varepsilon > 0$ the intervals $I_j$ can be chosen such that uniformly for all $j = 1, \ldots, J$ and all integers $k \leq k' \pm O(1) < 0$, it holds that

$$
\left\| Q_k(\chi I, e^{\pm \imath \theta_k \cdot \cdot \cdot + t, x, D}) P_0 \phi \right\|_{L^2_t L^2_x(\mathbb{R} \times \mathbb{R}^d)} \leq \varepsilon 2^k 2^{\delta(k-k')} \| P_0 \phi \|_{N_0^\varepsilon(\mathbb{R} \times \mathbb{R}^d)}.
$$

In the proof of Proposition 3.15 we will use the following result whose proof will be given later.

**Lemma 3.16.** Let $1 \leq q \leq p \leq \infty$. For any $\varepsilon > 0$ the intervals $I_j$ can be chosen such that uniformly for all $j = 1, \ldots, J$ and all integers $k + C \leq l \leq 0$, the following operator bound holds

$$
\left\| \chi I(t, x, D) \right\|_{L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^d)} \leq \varepsilon 2^k 2^{\delta(k-k')}.
$$

**Proof of Proposition 3.15.** In the following we denote an interval $I_j$ just by $I$ and $e^{\pm \imath \theta_k}$ stands for the left quantization $e^{\pm \imath \theta_k} = \chi I(t, x, D)$.

We first reduce to the case $k = k' \pm O(1)$. To this end we will use that Proposition 9 and Lemma 10 in [12] hold without the $\varepsilon$ smallness gain also for large energies. We split

$$
Q_k(\chi I, e^{\pm \imath \theta_k \cdot \cdot \cdot + t, x, D}) P_0 \phi = Q_k(Q_{\leq k-c}(\chi I) e^{\pm \imath \theta_k \cdot \cdot \cdot + t, x, D}) P_0 \phi + Q_k(Q_{k-c,k+c}(\chi I) e^{\pm \imath \theta_k \cdot \cdot \cdot + t, x, D}) P_0 \phi.
$$

(3.38)

For the first term we obtain

$$
2^k \left\| Q_k(Q_{\leq k-c}(\chi I) e^{\pm \imath \theta_k} P_0 \phi) \right\|_{L^2_t L^2_x} = \frac{1}{2^k} \left\| Q_k(Q_{\leq k-c}(\chi I) Q_{k-c,0}(\chi I) e^{\pm \imath \theta_k} P_0 \phi) \right\|_{L^2_t L^2_x} \leq 2^k \left\| Q_{k-c,0}(\chi I) e^{\pm \imath \theta_k} P_0 \phi \right\|_{L^2_t L^2_x} \leq 2^\delta 2^{\delta(k-k')} \| P_0 \phi \|_{N_0^\varepsilon},
$$

where we used Proposition 9 from [12]. We estimate the second term from (3.38) by

$$
2^k \left\| Q_k(Q_{k-c,k+c}(\chi I) e^{\pm \imath \theta_k} P_0 \phi) \right\|_{L^2_t L^2_x} \leq 2^k \left\| Q_{k-c,k+c}(\chi I) \right\|_{L^2_t L^2_x} \| e^{\pm \imath \theta_k} P_0 \phi \|_{L^2_t L^2_x} \leq 2^k 2^{-\frac{k'}{k}} \| e^{\pm \imath \theta_k} P_0 \phi \|_{L^2_t L^2_x}.
$$

Using continuous Littlewood-Paley resolutions to decompose the group element we have

$$
e^{\pm \imath \theta_k} = e^{\pm \imath \theta_k e^{\pm \imath \theta_0 \cdot \cdot \cdot}} \pm i \int_{t > k-C} S \left( \psi e^{\pm \imath \theta_0 \cdot \cdot \cdot} \right) dt.
$$

By Lemma 10 in [12] and the decomposable estimates (3.26) we find

$$
\| e^{\pm \imath \theta_k} P_0 \phi \|_{L^2_t L^2_x} \leq \| e^{\pm \imath \theta_k} e^{\pm \imath \theta_0 \cdot \cdot \cdot} P_0 \phi \|_{L^2_t L^2_x} + \int_{t > k-C} \| S \left( \psi e^{\pm \imath \theta_0 \cdot \cdot \cdot} \right) P_0 \phi \|_{L^2_t L^2_x} dt

\leq 2^{-\frac{k'}{k}} \| P_0 \phi \|_{L^2_t L^2_x} + \int_{t > k-C} \| \psi \| \| P_0 \phi \|_{L^2_t L^2_x} dt

\leq 2^{-\frac{k'}{k}} \| P_0 \phi \|_{L^2_t L^2_x} + \int_{t > k-C} 2^{-\frac{k}{} k} \| \nabla_{t,x} A \|_{L^2_t L^2_x} \| P_0 \phi \|_{L^2_t L^2_x} dt

\leq 2^{-\frac{k'}{k}} \| P_0 \phi \|_{N_0^\varepsilon}.
$$
Thus, the second term from (3.38) is bounded by
\[ 2^{\frac{j}{2}} \| Q_k(Q_{j-k+C} (\chi_l) e^{\pm i\phi_x} P_0 \phi)) \|_{L^2_t L^2_x} \lesssim 2^{\frac{k}{2}(k-k')} \| P_0 \phi \|_{N_0^k}. \]

To estimate the third term from (3.38) we first use Bernstein in time to obtain
\[ 2^{\frac{j}{2}} \| Q_k(Q_{j-k+C} (\chi_l) e^{\pm i\phi_x} P_0 \phi)) \|_{L^2_t L^2_x} \leq 2^{\frac{j}{2}} \sum_{j \geq k+C} \| Q_k(Q_j(\chi_l) Q_{j+O(1)} e^{\pm i\phi_x} P_0 \phi)) \|_{L^2_t L^2_x} \]
\[ \lesssim 2^{k} \sum_{j \geq k+C} \| Q_j(\chi_l) \|_{L^2} \| Q_{j+O(1)} e^{\pm i\phi_x} P_0 \phi \|_{L^2_t L^2_x} \]
\[ \lesssim 2^{k} \sum_{j = k+C}^{k'+C} 2^{-\frac{j}{2}} \| Q_{j+O(1)} e^{\pm i\phi_x} P_0 \phi \|_{L^2_t L^2_x} + 2^{k} \sum_{j \geq k'+C} 2^{-\frac{j}{2}} \| Q_{j+O(1)} e^{\pm i\phi_x} P_0 \phi \|_{L^2_t L^2_x}. \]

For the first sum we use Proposition 9 from [12], for the second sum we first note that there is no modulation interference since \( k' < j - C \) and then use the fixed-time \( L^2_t \rightarrow L^2_x \) estimate for \( e^{\pm i\phi_x} \).

Hence,
\[ \lesssim 2^{k} \sum_{j = k+C}^{k'+C} 2^{-\frac{j}{2}} 2^{\frac{j}{2}(j-k')} \| \nabla_{\ell,x} A \|_{L^2_t} \| P_0 \phi \|_{N_0^k} + 2^{k} \sum_{j \geq k'+C} 2^{-\frac{j}{2}} \| P_0 \phi \|_{X_{\infty}^{\frac{1}{2}}}. \]
\[ \lesssim (2^{\delta(j-k')} + 2^{k-k'}) \| P_0 \phi \|_{N_0^k}. \]

Putting things together we find that
\[ 2^{\frac{j}{2}} \| Q_k(\chi_l e^{\pm i\phi_x} P_0 \phi) \|_{L^2_t L^2_x} \lesssim 2^{\delta(k-k')} \| P_0 \phi \|_{N_0^k} \]
and for sufficiently large \(|k| \gg |k'|\) we therefore trivially gain a smallness factor \( \varepsilon \) from \( 2^{\frac{j}{2}(k-k')} \).

We are thus reduced to the case \( k = k' \pm O(1) \) and it remains to show that
\[ 2^{\frac{j}{2}} \| Q_k(\chi_l e^{\pm i\phi_x} P_0 \phi) \|_{L^2_t L^2_x} \lesssim \varepsilon \| P_0 \phi \|_{N_0^k}. \]

As in the proof of Proposition 9 in [12] we expand the untruncated group element
\[ e^{\pm i\phi} = e^{\pm i\phi_{\ell-x} \cdot C} + i \int_{\ell > k-C} \psi(x) e^{i \phi_{\ell-x} \cdot C} \, dl - \int_{\ell > k-C} \psi(x) e^{i \phi_{\ell-x} \cdot C} \, dl' \, dl \]
\[ + i \int_{\ell, \ell' > k-C} \psi(x) \psi(x') e^{i \phi_{\ell-x} \cdot C} \, dl' \, dl \]
\[ = Z + L + Q + C \]
and estimate each of these terms separately.

**Zero order term Z:** From Lemma 3.16 we immediately obtain that
\[ \| Q_k(\chi_l e^{\pm i\phi_{\ell-x} \cdot C} P_0 \phi) \|_{L^2_t L^2_x} \lesssim \varepsilon 2^{\frac{j}{2}} \| P_0 \phi \|_{N_0^k}. \]

**Linear term L:** We have to show that
\[ \| Q_k(\chi_l \int_{\ell > k-C} S_k(\psi(x) e^{i \phi_{\ell-x} \cdot C} P_0 \phi) \, dl) \|_{L^2_t L^2_x} \lesssim \varepsilon 2^{\frac{j}{2}} \| P_0 \phi \|_{N_0^k}. \]
To this end we decompose $\psi_l$ into a small and a large angular part

$$\psi_l = \sum_{2^{c_1} < \theta < 2^{c_3}} \psi_l^{(0)} + \sum_{2^{-c_1} \leq \theta \leq 1} \psi_l^{(0)}.$$  

In order to bound the small angular part we split

$$\chi_l = \mathcal{Q}_{2k-\mathcal{C}}(\chi_l) + \mathcal{Q}_{\mathcal{C}}(\chi_l).$$

Using Lemma 3.5 we estimate the first term by

$$\left\| \mathcal{Q}(\mathcal{Q}_{2k-\mathcal{C}}(\chi_l) \int_{l \geq \mathcal{C}} S_k \left( \sum_{2^{c_1} < \theta < 2^{c_3}} (\psi_l^{(0)} e^{\pm i \phi \cdot k \cdot c}) \mathcal{P}_0 \phi \right) dt \right\|_{L^2_x L^2_t}\leq \left\| \mathcal{Q}_{2k-\mathcal{C}}(\chi_l) \int_{l \geq \mathcal{C}} S_k \left( \sum_{2^{c_1} < \theta < 2^{c_3}} \left\| \psi_l^{(0)} \right\|_{\mathcal{D}_k(L^4_x L^\infty_t)} \left\| e^{\pm i \phi \cdot k \cdot c} \mathcal{P}_0 \phi \right\|_{L^4_x L^\infty_t} dt \right\|_{L^2_x L^2_t} \leq 2^{-\frac{k}{2}} \int_{l \geq \mathcal{C}} \sum_{2^{c_1} < \theta < 2^{c_3}} \left\| \nabla_{\mathcal{A}} A \right\|_{L^2_x} \left\| \mathcal{P}_0 \phi \right\|_{L^4_x L^\infty_t} \right. \left\| \mathcal{P}_0 \phi \right\|_{N_0^\infty} \right. \right.$$  

For the second term we have

$$\mathcal{Q}(\mathcal{Q}_{\mathcal{C}}(\chi_l) \int_{l \geq \mathcal{C}} S_k \left( \sum_{2^{c_1} < \theta < 2^{c_3}} (\psi_l^{(0)} e^{\pm i \phi \cdot k \cdot c}) \mathcal{P}_0 \phi \right) dt) = \mathcal{Q}(\mathcal{Q}_{\mathcal{C}}(\chi_l) \mathcal{Q}_{k+O(1)} \int_{l \geq \mathcal{C}} S_k \left( \sum_{2^{c_1} < \theta < 2^{c_3}} (\psi_l^{(0)} e^{\pm i \phi \cdot k \cdot c}) \mathcal{P}_{k+O(1)} \phi \right) dt) \right.$$

Then since $\psi_l^{(0)}$ is a free wave, we can write this as

$$\mathcal{Q}(\mathcal{Q}_{\mathcal{C}}(\chi_l) \mathcal{Q}_{k+O(1)} \int_{l \geq \mathcal{C}} S_k \left( \sum_{2^{c_1} < \theta < 2^{c_3}} (\psi_l^{(0)} e^{\pm i \phi \cdot k \cdot c}) \mathcal{Q}_{k+O(1)} \mathcal{P}_0 \phi \right) dt) \right.$$

and estimate by

$$\left\| \mathcal{Q}(\mathcal{Q}_{\mathcal{C}}(\chi_l) \mathcal{Q}_{k+O(1)} \int_{l \geq \mathcal{C}} S_k \left( \sum_{2^{c_1} < \theta < 2^{c_3}} (\psi_l^{(0)} e^{\pm i \phi \cdot k \cdot c}) \mathcal{Q}_{k+O(1)} \mathcal{P}_0 \phi \right) dt) \right\|_{L^2_x L^2_t} \leq \int_{l \geq \mathcal{C}} \sum_{2^{c_1} < \theta < 2^{c_3}} \left\| \psi_l^{(0)} \right\|_{\mathcal{D}_k(L^4_x L^\infty_t)} \left\| e^{\pm i \phi \cdot k \cdot c} \mathcal{Q}_{k+O(1)} \mathcal{P}_0 \phi \right\|_{L^4_x L^\infty_t} \right. \left\| \mathcal{P}_0 \phi \right\|_{N_0^\infty} \right. \leq 2^{-\frac{k}{2}} \int_{l \geq \mathcal{C}} \sum_{2^{c_1} < \theta < 2^{c_3}} \left\| \nabla_{\mathcal{A}} A \right\|_{L^2_x} \left\| \mathcal{Q}_{k+O(1)} \mathcal{P}_0 \phi \right\|_{L^4_x L^\infty_t} \right. \left\| \mathcal{P}_0 \phi \right\|_{N_0^\infty} \right. \right.$$  

Here we used Lemma 3.5 the fixed-time $L^2_x \to L^2_x$ estimate for $e^{\pm i \phi \cdot k \cdot c}$ and then Bernstein in time.
The large angular part in (3.39) has to be estimated more carefully. Noting that the symbol localization $S_k(\psi_1e^{\pm i\theta \cdot \phi_{-k}})$ can be represented as

$$
S_k(\psi_1e^{\pm i\theta \cdot \phi_{-k}}) = \int_{\mathbb{R}^{1+4}} m_k(z)(T_x\psi_T)e^{\pm iT_x\psi_{-k}} \, dz,
$$

where $m_k$ is an integrable bump function at scale $k$ and $T_x$ denotes translation in space-time direction $z \in \mathbb{R}^{1+4}$. We derive the following key estimate for the large angular part

$$
\left\| Q_k(\chi_I \int_{b_k-C} \int_{\mathbb{R}^{1+4}} m_k(z) \left( \sum_{2^{-3} \leq \theta \leq 1} (T_x\psi_T) \right) e^{\pm iT_x\psi_{-k}} P_0 \phi \, dz \, dt \right\|_{L^2_t L^4_z} \leq C_3^{1/2} 2^{-k/2} \left\| \left( \int_{b_k-C} \int_{\mathbb{R}^{1+4}} |m_k(z)| \sum_{2^{-3} \leq \theta \leq 1} \sup_{\omega \in \Gamma_{\theta}} \left( 2^{-\frac{5}{2}} \| b_0(\omega) \Pi^0_\theta \nabla_{x,A} \|_{L^2_x} \right)^2 \, dz \, dt \right\|_{L^2_t L^4_z}^2 \right. 
$$

This estimate can be proven by carefully opening up the proof of the decomposable estimates in Lemma 3.5. We emphasize that uniformly for all integers $k < 0$, the quantity

$$
\left\| \left( \int_{b_k-C} \int_{\mathbb{R}^{1+4}} |m_k(z)| \sum_{2^{-3} \leq \theta \leq 1} \sup_{\omega \in \Gamma_{\theta}} \left( 2^{-\frac{5}{2}} \| b_0(\omega) \Pi^0_\theta \nabla_{x,A} \|_{L^2_x} \right)^2 \, dz \, dt \right\|_{L^2_t L^4_z}^2 \right. 
$$

is bounded by $\| \nabla_{x,A} \|_{L^2_z}$ by Strichartz estimates.

By first fixing $C_3 > 0$ sufficiently large and then suitably choosing the intervals $I_j$, the estimate of the linear term $L$ follows.

**Quadratic and cubic terms $Q$ and $C$:** Using the above ideas these can be estimated similarly. We omit the details.

**Proof of Lemma 3.16** As in [12, Lemma 10] we write the symbol as

$$
S_{\ell}e^{\pm i\theta \cdot \phi_{-k}} = (\pm i)^5 2^{-5l} \prod_{r=1}^{5} [S_{\ell}^{(r)} \partial_t \psi_{-k}] e^{\pm i\theta \cdot \phi_{-k}},
$$

where the product denotes a nested multiplication by $S_{\ell} \partial_t \psi_{-k}$ for a series of frequency cutoffs $S_{\ell}^{(r+1)} S_{\ell}^{(r)} = S_{\ell}^{(r)} \approx S_{\ell}$ with expanding widths. Then we have

$$
S_{\ell}^{(r)} \partial_t \psi_{-k} = \int_{\mathbb{R}^{1+4}} m_{\ell}^{(r)}(z_{\ell}) (T_{z_{\ell}} \partial_t \psi_{-k}) \, dz_{\ell},
$$

where $m_{\ell}^{(r)}$ is an integrable bump function at scale $l$ and $T_{z_{\ell}}$ denotes translation in space-time direction $z_{\ell} \in \mathbb{R}^{1+4}$. The claim now reduces to proving that the intervals $I_j$ can be chosen such that
uniformly for \( j = 1, \ldots, J \) and all integers \( k \leq l - C \), it holds that

\[
\left\| \chi_I \left( \prod_{r=1}^{5} \int_{\mathbb{R}^{1+4}} m_i^{(r)}(z_r)(T_z \partial_t \psi_k(t)) e^{i\phi \cdot \xi}(t, x, D) \right) \right\|_{L^q_t L^p_x L^2} \leq e^{2^{4l}2^{1/2-l}k^k}.
\]

To this end we show that the intervals \( I_j \) can be chosen such that uniformly for \( j = 1, \ldots, J \), all integers \( k \leq l - C \) and all integers \( k_1, \ldots, k_5 < k \), we have the operator bound

\[
\left\| \chi_I \left( \prod_{r=1}^{5} \int_{\mathbb{R}^{1+4}} m_i^{(r)}(z_r)(T_z \partial_t \psi_k(t)) e^{i\phi \cdot \xi}(t, x, D) \right) \right\|_{L^q_t L^p_x L^2} \leq e^{2^{1-\frac{k}{q}}2^{1-\frac{k}{q}}k_1 \cdots 2^{1-\frac{k}{q}}k_5},
\]

where \( \frac{1}{q} = \frac{1}{q} - \frac{1}{p} \). By summing over dyadic frequencies, the estimate (3.41) then follows.

In order to prove (3.42), we split \( \partial_t \psi_k(t) \) into a small and a large angular part

\[
\partial_t \psi_k(t) = \sum_{2^{a_k} < \theta_i < 2^{a_k+1}} \partial_t \psi_k(t) + \sum_{2^{a_k} \leq \theta_i \leq 1} \partial_t \psi_k(t)
\]

for some constant \( C_3 > 0 \) to be chosen sufficiently large later in the proof. We estimate the small angular part using Hölder-type estimates for decomposable function spaces (3.24) and the bounds (3.23) for the phase,

\[
\left\| \chi_I \left( \sum_{2^{a_k} < \theta_i < 2^{a_k+1}} \int_{\mathbb{R}^{1+4}} m_i^{(1)}(z_1)(T_z \partial_t \psi_k(t)) d z_1 \right) \left( \prod_{r=2}^{5} \int_{\mathbb{R}^{1+4}} m_i^{(r)}(z_r)(T_z \partial_t \psi_k(t)) d z_r \right) e^{i\phi \cdot \xi}(t, x, D) \phi \right\|_{L^q_t L^p_x L^2} \leq \left( \sum_{2^{a_k} < \theta_i < 2^{a_k+1}} \int_{\mathbb{R}^{1+4}} |m_i^{(1)}(z_1)||T_z \partial_t \psi_k(t)||_{L^q_t L^p_x L^2} d z_1 \right) \left( \prod_{r=2}^{5} \int_{\mathbb{R}^{1+4}} |m_i^{(r)}(z_r)||T_z \partial_t \psi_k(t)||_{L^q_t L^p_x L^2} d z_r \right) \leq 2^{C_3} \left( \frac{1}{\frac{k}{q}+\frac{1}{3}} \right)^{\frac{1}{2}} \left( \frac{1}{\frac{k}{q}+\frac{1}{3}} \right)^{\frac{1}{2}} \left( \frac{1}{\frac{k}{q}+\frac{1}{3}} \right)^{\frac{1}{2}} |\nabla x A| \right\|_{L^2}^5 \phi \|_{L^p_t L^2}.
\]

Here we dropped the time cutoff \( \chi_I \) and used the space-time translation invariance of the decomposable function spaces. For the large angular part we establish the crucial estimate

\[
\left\| \chi_I \left( \sum_{2^{a_k} \leq \theta_i \leq 1} \int_{\mathbb{R}^{1+4}} m_i^{(1)}(z_1)(T_z \partial_t \psi_k(t)) d z_1 \right) \left( \prod_{r=2}^{5} \int_{\mathbb{R}^{1+4}} m_i^{(r)}(z_r)(T_z \partial_t \psi_k(t)) d z_r \right) e^{i\phi \cdot \xi}(t, x, D) \phi \right\|_{L^q_t L^p_x L^2} \leq 2^{2^{-\frac{k}{q}}k_1 \cdots 2^{1-\frac{k}{q}}k_5 |\nabla x A| \|_{L^2}^{1/2} \times \left( \sum_{2^{a_k} \leq \theta_i \leq 1} m_i^{(1)}(z_1) \sum_{2^{-C_3} \leq \theta_i \leq 1} \sup_{\Gamma_{\theta_i}^{\omega} \in \Omega^{1/2}} \left( \frac{1}{\frac{k}{q}+\frac{1}{3}} \right)^{\frac{1}{2}} |\nabla x A| \|_{L^2}^{1/2} \right)^{1/2} \right\|_{L^q_t L^p_x L^2},
\]

where \( r_0 \geq 2 \) is such that the exponent pair \((\bar{q}, r_0)\) is sharp wave admissible. This estimate can be proven by carefully opening up the proof of Lemma 3.5 and of Hölder-type estimates for decomposable function spaces (3.23).
Noting that uniformly for all integers $k_1 \leq l - C$, the quantity
\[
\left\| \left( \int_{E^{1+4}_{1}} \left| m_{1}^{(1)}(z_{1}) \right| \sum_{2^{-j_{1}} \leq z_{1} \leq 1} \sum_{\Gamma_{\ell_{1}} \in \mathcal{E}_{\Gamma_{\ell_{1}}}} \sup_{2^{-j_{1}} \leq z_{1} \leq 1} \left( 2^{j_{1} \frac{k_{1}}{4} + \frac{k_{1}}{2} - 2k_{1}} \left| b_{0}^{(1)}(\omega) \Pi_{0}^{\text{su}} \nabla_{l_{1}} A_{l_{1}} A_{l_{1}} \right| \right)^{2} dz_{1} \right)^{1/2} \right\|_{L^{2}(\mathbb{R})}
\]
is bounded by $\|\nabla_{l_{1}} A \|_{L^{2}}$ by Strichartz estimates, the assertion follows by first choosing $C_{3} > 0$ sufficiently large and then suitably choosing the intervals $I_{j}$.

Proof. The proof proceeds analogously to the one of Proposition 3.15 using Lemma 3.6 in place of Lemma 3.5. 

3.6. Proof of the $N_{0} \to N_{0}$ and $N_{0}^{*} \to N_{0}^{*}$ bounds (3.18) for $\chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}$.

Proposition 3.18. For $j = 1, \ldots, J$ it holds that
\[
\left\| \chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(t, x, D) P_{0} \right\|_{L^{2}_{N_{0}}(\mathbb{R} \times \mathbb{R}^{d})} \leq \left\| P_{0} \right\|_{N_{0}(\mathbb{R} \times \mathbb{R}^{d})}
\]
and
\[
\left\| \chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0} \right\|_{N_{0}(\mathbb{R} \times \mathbb{R}^{d})} \leq \left\| P_{0} \right\|_{N_{0}(\mathbb{R} \times \mathbb{R}^{d})}.
\]

Proof. We begin with the proof of (3.44). To simplify the notation we denote an interval $I_{j}$ just by $I$ in what follows. If $\phi$ is an $L_{1}^{I} L_{2}^{s}$ atom, the claim follows immediately from the fixed-time $L_{2}^{s} \to L_{2}^{s}$ estimate for $e^{z_{k_{1}} \psi_{k_{1}}}(t, x, D)$. The key point is therefore to show that if $\phi$ is an $X_{1}^{0, -\frac{1}{2}}$ atom at modulation $k$, then we have
\[
\left\| \chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(t, x, D) Q_{k} P_{0} \right\|_{L^{2}_{N_{0}}} \lesssim 2^{-\frac{1}{2}k} \| P_{0} \|_{L^{2}_{N_{0}}}.
\]
By duality, this is equivalent to proving
\[
\left\| Q_{k}(\chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}) \right\|_{L^{2}_{N_{0}}} \lesssim 2^{-\frac{1}{2}k} \| P_{0} \|_{N_{0}^{*}}.
\]
As in [12] Proposition 9.1 we now write
\[
Q_{k}(\chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}) = Q_{k}(\chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}) + Q_{k}(\chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0})
\]
\[
= Q_{k}(Q_{c_{k}-c}(\chi_{l_{1}}) e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}) + Q_{k}(Q_{c_{k}-c}(\chi_{l_{1}}) e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0})
\]
\[
+ Q_{k}(\chi_{l_{1}} e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}).
\]
For the first term we obtain
\[
2^{-\frac{1}{2}k} \left\| Q_{k}(Q_{c_{k}-c}(\chi_{l_{1}}) e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}) \right\|_{L^{2}_{N_{0}}} \lesssim \| P_{0} \|_{L^{2}_{N_{0}}} \leq \| P_{0} \|_{N_{0}^{*}},
\]
because the output modulation directly transfers to $\phi$. We estimate the second term by
\[
2^{-\frac{1}{2}k} \left\| Q_{k}(Q_{c_{k}-c}(\chi_{l_{1}}) e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0}) \right\|_{L^{2}_{N_{0}}} \leq 2^{\frac{1}{2}k} \left\| Q_{c_{k}-c}(\chi_{l_{1}}) e^{z_{k_{1}} \psi_{k_{1}}}(D, y, t) P_{0} \right\|_{L^{2}_{N_{0}}} \leq \| P_{0} \|_{N_{0}^{*}},
\]
where we used that \( \|Q_{\geq k}c(\chi_i)\|_{L^2_x} \lesssim 2^{-\frac{1}{2}k} \). To deal with the last term we use Proposition \[3.17\]

The proof of \ref{prop:3.45} works similarly using Proposition \[3.15\].

In a similar vein we obtain the following \(N_0^* \to N_0^*\) bounds.

**Proposition 3.19.** For \( j = 1, \ldots, J \) it holds that
\begin{align}
&\|\chi_I e^{\varphi_{\leq 0}}(t, x, D)P_0\phi\|_{N_0^*} \leq \|P_0\phi\|_{N_0^*} \\
&\text{and}
\end{align}

\[3.46\]

\[3.47\]

**3.7. Proof of the \(N_0 \to \varepsilon N_0\) and \(N_0^* \to \varepsilon N_0^*\) bounds (3.19) for \(\chi_I \partial_t e^{\varphi_{\leq 0}}\).**

**Proposition 3.20.** For any \( \varepsilon > 0 \) the intervals \(I_j\) can be chosen such that uniformly for all \( j = 1, \ldots, J \) it holds that
\begin{align}
&\|\chi_I \partial_t e^{\varphi_{\leq 0}}(t, x, D)P_0\phi\|_{N_0^*} \leq \varepsilon \|P_0\phi\|_{N_0^*} \\
&\text{and}
\end{align}

\[3.48\]

\[3.49\]

**Proof.** We proceed as in the proof of Proposition \ref{prop:3.18} using the \(L^2_{\varepsilon} \to \varepsilon L^2_{\varepsilon}\) bound for \(\partial_t e^{\varphi_{\leq 0}}\) and that we have for \( k \leq k' \) \( O(1) \),
\[3.50\]

\[3.51\]

for both left and right quantization. The latter estimate can be proven similarly to the proof of Proposition \ref{prop:3.15}.

**Proposition 3.21.** For any \( \varepsilon > 0 \) the intervals \(I_j\) can be chosen such that uniformly for all \( j = 1, \ldots, J \) it holds that
\begin{align}
&\|\chi_I \partial_t e^{\varphi_{\leq 0}}(t, x, D)P_0\phi\|_{N_0^*} \leq \varepsilon \|P_0\phi\|_{N_0^*} \\
&\text{and}
\end{align}

\[3.52\]

\[3.53\]

**3.8. Proof of the renormalization error estimate (3.20).**

**Proposition 3.22.** For any \( \varepsilon > 0 \) the intervals \(I_j\) can be chosen such that uniformly for all \( j = 1, \ldots, J \) we have
\begin{align}
&\|\chi_I (e^{\varphi_{\leq 0}}(t, x, D) e^{\varphi_{\leq 0}}(D, y, t) - 1)P_0\phi\|_{N_0^*} \leq \varepsilon \|P_0\phi\|_{N_0^*} \\
&\text{and}
\end{align}

\[3.54\]

\[3.55\]

**Proof.** We prove the \(N_0^* \to \varepsilon N_0^*\) estimate \[3.53\]. The bound \[3.52\] then follows by duality. The \(L^\infty_{\varepsilon}L^2_{\varepsilon}\) part of \[3.52\] follows immediately from the fixed-time \(L^2_{\varepsilon} \to \varepsilon L^2_{\varepsilon}\) estimate \[3.36\]. The \(X^{0, \frac{1}{2}}\) part reduces to showing that we can choose the intervals \(I_j\) such that uniformly for \( j = 1, \ldots, J \) and all \( k \in \mathbb{Z} \),
\[3.56\]

\[3.57\]
We use the notation
\[ R_{ck} = e^{-i\phi_x}(t, x, D)e^{i\phi_y}(D, y, t) \]
to write
\[ Q_k(\chi_I, (R < a - 1)P_0\phi) = Q_k(\chi_I, (R < a - 1)Q_{k\text{-}c\text{-}c}P_0\phi) + Q_k(\chi_I, (R < a - 1)Q_{\text{c}-c\text{-}c}P_0\phi) + Q_k(\chi_I, (R < a - R_{k\text{-}c\text{-}c})Q_{\text{c}-c\text{-}c}P_0\phi). \]  
(3.54)

Using the fixed-time \( L^2_{\alpha} \) estimate (3.36) for \( (R < a - 1) \), we bound the first term in (3.54) by
\[ 2^{2k}\left\| Q_k(\chi_I, (R < a - 1)Q_{\text{c}-c\text{-}c}P_0\phi) \right\|_{L^2_{\alpha}} \leq 2^{2k}\left\| (R < a - 1)Q_{\text{c}-c\text{-}c}P_0\phi \right\|_{L^2_{\alpha}} \leq 2^{2k}\epsilon\left\| Q_{\text{c}-c\text{-}c}P_0\phi \right\|_{L^2_{\alpha}} \leq \epsilon\|P_0\phi\|_{X_{\alpha, \epsilon}}. \]

To estimate the second term in (3.54) we observe that we have
\[ Q_k(\chi_I, (R < a - c\text{-}c)Q_{\text{c}-c\text{-}c}P_0\phi) = Q_k((Q_{k-c-c}(\chi_I))(R < a - c\text{-}c)Q_{\text{c}-c\text{-}c}P_0\phi) \]
and hence by the fixed-time \( L^2_{\alpha} \) estimate for \( (R < a - c\text{-}c) \),
\[ 2^{2k}\left\| Q_k(\chi_I, (R < a - c\text{-}c)Q_{\text{c}-c\text{-}c}P_0\phi) \right\|_{L^2_{\alpha}} \leq 2^{2k}\left\| Q_{k-c-c}(\chi_I) \right\|_{L^2_{\alpha}} \left\| (R < a - 1)Q_{\text{c}-c\text{-}c}P_0\phi \right\|_{L^2_{\alpha}} \leq \epsilon\|P_0\phi\|_{X_{\alpha, \epsilon}}. \]

Finally, we expand the third term in (3.54) as follows
\[ Q_k(\chi_I, (R < a - R_{k-c\text{-}c})Q_{\text{c}-c\text{-}c}P_0) = Q_k(\chi_I, (e^{\phi_x}(t, x, D) - e^{i\phi_x}(t, x, D))e^{i\phi_y}(D, y, t)Q_{\text{c}-c\text{-}c}P_0\phi \]
\[ + Q_k(\chi_I, e^{\phi_x}(t, x, D)(e_{\text{c}-c\text{-}c}(D, y, t) - e^{i\phi_x}(D, y, t))Q_{\text{c}-c\text{-}c}P_0\phi). \]

To handle the first term in the above expansion we use Proposition 3.15 and the \( N^*_{\alpha} \rightarrow N^*_{\alpha} \) estimate (3.47) for \( e^{i\phi_x}(D, y, t) \) to find that
\[ 2^{2k}\left\| Q_k(\chi_I, (e^{\phi_x}(t, x, D) - e^{i\phi_x}(t, x, D))e^{i\phi_y}(D, y, t)Q_{\text{c}-c\text{-}c}P_0\phi \right\|_{L^2_{\alpha}} \]
\[ \leq \sum_{k' = k - c}^{k'} 2^{2k}\left\| Q_k(\chi_I, e^{\phi_x}(t, x, D)e_{\text{c}-c\text{-}c}(D, y, t)Q_{\text{c}-c\text{-}c}P_0\phi \right\|_{L^2_{\alpha}} \]
\[ \leq \sum_{k' = k - c}^{k'} 2^{2k'}Q_k(\chi_I, e_{\text{c}-c\text{-}c}(D, y, t)Q_{\text{c}-c\text{-}c}P_0\phi) \]
\[ \leq \epsilon\|P_0\phi\|_{X_{\alpha, \epsilon}}. \]

Observing that
\[ Q_k(\chi_I, e^{i\phi_x}(t, x, D)(e_{\text{c}-c\text{-}c}(D, y, t) - e^{i\phi_y}(D, y, t))Q_{\text{c}-c\text{-}c}P_0\phi \]
\[ = Q_k(e_{\text{c}-c\text{-}c}(t, x, D)Q_{k+O(1)}(\chi_I, (e^{i\phi_x}(D, y, t) - e^{i\phi_y}(D, y, t))Q_{\text{c}-c\text{-}c}P_0\phi). \]

we estimate the second term analogously using the fixed-time \( L^2_{\alpha} \) estimate for \( e^{i\phi_x}(t, x, D) \)
and Proposition 3.17.

3.9. Proof of the renormalization error estimate (3.21). This estimate can be proven by adapting the proof in [12, Section 10.2] to our large data setting using similar ideas as above. The additional errors generated by the high-angle cutoff for intermediate frequencies in the definition of the phase correction \( \psi \) can be controlled by divisibility of suitable space-time norms of \( A \).
3.10. **Proof of the dispersive estimate** (3.22). Since the $S$ space is compatible with time localizations by Lemma 2.1, the dispersive estimate (3.22) follows immediately from the estimate (83) in [12].

### 4. Breakdown criterion

**Definition 4.1.** Let $T_0, T_1 > 0$. For any admissible solution $(A, \phi)$ to the MKG-CG system on $(-T_0, T_1) \times \mathbb{R}^4$, we define

$$
\| (A, \phi) \|_{S^1((-T_0, T_1) \times \mathbb{R}^4)} := \sup_{0 < T < T_0, 0 < T' < T_1} \left( \sum_{j=1}^{4} \| A_j \|_{S^1([-T,T'] \times \mathbb{R}^4)} + \| \phi \|_{S^1([-T,T'] \times \mathbb{R}^4)} \right)^\frac{1}{2}.
$$

We establish the following blowup criterion for admissible solutions to the MKG-CG system.

**Proposition 4.2.** Let $(-T_0, T_1)$ be the maximal interval of existence of an admissible solution $(A, \phi)$ to the MKG-CG system. If $\| (A, \phi) \|_{S^1((-T_0, T_1) \times \mathbb{R}^4)} < \infty$, then it must hold that $T_0 = T_1 = \infty$.

The idea of the proof of Proposition 4.2 is to establish an a priori bound on a subcritical norm

$$
\sup_{t \in (-T_0, T_1)} \sum_{j=1}^{4} \| A_j(t) \|_{H^s_t \times H^{s-1}_x} + \| \phi(t) \|_{H^s_t \times H^{s-1}_x} < \infty
$$

for some $s > 1$. By the local well-posedness result [20] for the MKG-CG system it then follows that the solution can be smoothly extended beyond the time interval $(-T_0, T_1)$. To this end, we will use Tao’s device of frequency envelopes. For sufficiently small $\sigma > 0$ we define for all $k \in \mathbb{Z}$,

$$
c_k := \left( \sum_{l \in \mathbb{Z}} 2^{-\sigma |k-l|} \left( \sum_{j=1}^{4} \| P_l A_j(0) \|_{H^s_t \times L^2_x}^2 + \| P_l \phi(0) \|_{H^s_t \times L^2_x}^2 \right) \right)^{\frac{1}{2}}.
$$

Proposition 4.2 then is a consequence of the following result.

**Proposition 4.3.** Let $(-T_0, T_1)$ be the maximal interval of existence of an admissible solution $(A, \phi)$ to the MKG-CG system. If $\| (A, \phi) \|_{S^1((-T_0, T_1) \times \mathbb{R}^4)} < \infty$, there exists $C = C(\| (A, \phi) \|_{S^1((-T_0, T_1) \times \mathbb{R}^4)}) < \infty$ such that for all $k \in \mathbb{Z}$,

$$
\| P_k A \|_{S^1((-T_0, T_1) \times \mathbb{R}^4)} + \| P_k \phi \|_{S^1((-T_0, T_1) \times \mathbb{R}^4)} \leq C c_k.
$$

**Proof.** A sketch of the proof is given in Subsection 7.4 \(\square\)

### 5. A concept of weak evolution

In order to implement the contradiction argument after the concentration compactness step, we have to define the notion of a solution to the MKG-CG system that is merely of energy class. In the context of critical wave maps in [10] this is achieved by first approximating an energy class datum by Schwartz class data in the energy topology. One then defines the energy class solution as a suitable limit of the associated Schwartz class solutions. Using perturbation theory, one shows that this limit is well-defined and independent of the approximating sequence.

For the MKG-CG system we have to argue more carefully, because it appears that the strong perturbative step in the context of the critical wave maps in [10] is not available due to a low frequency divergence. However, the problem with evolving irregular data is really a “high frequency issue” and it appears that truncating high frequencies away does not lead to the same problems as a general perturbative step. More concretely, consider Coulomb energy class data at time $t = 0$. By
truncating in frequency, we can assume that the frequency support of either input is at $|\xi| \leq K$ for some $K > 0$. Then the problem becomes to show that we can add high-frequency perturbations to the data, i.e. supported in frequency space at $|\xi| > K$ at time $t = 0$, and to obtain a perturbed global evolution.

**Proposition 5.1.** Let $(A, \phi)$ be an admissible solution to the MKG-CG system on $[-T_0, T_1] \times \mathbb{R}^4$ for some $T_0, T_1 > 0$. Assume that $(A, \phi)[0]$ have frequency support at $|\xi| \leq K$ for some $K > 0$ and that $\|A\|_{L^1([-T_0, T_1] \times \mathbb{R}^4)} = L < \infty$. Then there exists $\delta_1(L) > 0$ with the following property: Let $(A + \delta A, \phi + \delta \phi)$ be any other admissible solution to the MKG-CG system defined locally around $t = 0$ such that

$$E(\delta A, \delta \phi)(0) = \delta_0 < \delta_1(L)$$

and such that $(\delta A, \delta \phi)[0]$ have frequency support at $|\xi| > K$. Then $(A + \delta A, \phi + \delta \phi)$ extends to an admissible solution to the MKG-CG system on the whole time interval $[-T_0, T_1]$ and satisfies

$$\|(A + \delta A, \phi + \delta \phi)\|_{L^1([-T_0, T_1] \times \mathbb{R}^4)} \leq \tilde{L}(L, \delta_0).$$

Moreover, we have

$$\|(\delta A, \delta \phi)\|_{L^1([-T_0, T_1] \times \mathbb{R}^4)} \to 0$$

as $\delta_0 \to 0$.

**Proof.** A sketch of the proof is given in Subsection 7.4. \qed

The above high-frequency perturbation result suggests that we could define the MKG-CG evolution of energy class Coulomb data as a suitable limit of the evolutions of low frequency approximations of the energy class data. More precisely, for Coulomb data $(A, \phi)[0] \in \dot{H}^1_t \times L^2_x$, we pick a sequence of smoothings $(A_n, \phi_n)[0]$ by truncating the frequency support of $(A, \phi)[0]$ so that

$$\lim_{n \to 0} (A_n, \phi_n)[0] = (A, \phi)[0]$$

in the sense of $\dot{H}^1_t \times L^2_x$. Here the rather technical issue appears whether there exists a smooth (local) solution $(A_n, \phi_n)$ to the MKG-CG system with initial data $(A_n, \phi_n)[0]$. The hypothesis $(A, \phi)[0] \in \dot{H}^1_t \times L^2_x$ does not guarantee that $A(0)$ and $\phi(0)$ are $L^2$ integrable in the low frequencies. For this reason we cannot directly invoke the local well-posedness result [20] to obtain a smooth local solution. The natural way around this is to localize in physical space. This will be explained in more detail in Subsection 5.2 below.

For each smooth local solution $(A_n, \phi_n)$ to the MKG-CG system with initial data $(A_n, \phi_n)[0]$ we then define

$$I_n := \bigcup_{\tilde{I} \in \mathcal{I}_n} \tilde{I},$$

where

$$\mathcal{I}_n := \{ \tilde{I} \subset \mathbb{R} \text{ open interval with } 0 \in \tilde{I} : \sup_{J \subset \tilde{I}, J \text{ closed}} \|(A_n, \phi_n)|_{\dot{S}^1(J \times \mathbb{R}^4)} < \infty \}. $$

We call $I_n$ the maximal lifespan of the solution $(A_n, \phi_n)$.

In order to define a canonical evolution of Coulomb energy class data, we have to show that the low frequency approximations $(A_n, \phi_n)$ exist on some joint time interval and satisfy uniform $S^1$ norm bounds there.

**Proposition 5.2.** Let $(A, \phi)[0]$ be Coulomb energy class data and let $\{(A_n, \phi_n)[0]\}_n$ be a sequence of smooth low frequency truncations of $(A, \phi)[0]$ such that

$$\lim_{n \to \infty} (A_n, \phi_n)[0] = (A, \phi)[0]$$
in the sense of $\dot{H}^1_x \times L^2_x$. Denote by $(A_n, \phi_n)$ the smooth solutions to the MKG-CG system with initial data $(A_n, \phi_n)[0]$ and with maximal intervals of existence $I_n$. Then there exists a time $T_0 \equiv T_0(A, \phi) > 0$ such that $[-T_0, T_0] \subset I_n$ for all sufficiently large $n$ and

$$\lim_{n \to \infty} \sup \| (A_n, \phi_n) \|_{S^1((-T_0, T_0) \times \mathbb{R}^4)} \leq C(A, \phi),$$

where $C(A, \phi) > 0$ is a constant that depends only on the energy class data $(A, \phi)[0]$.

**Proof.** The proof is given in Subsection 5.1 below.

Using Proposition 5.1 and Proposition 5.2, we may introduce the following notion of energy class solutions to the MKG-CG system that we outlined above.

**Definition 5.3.** Let $(A, \phi)[0]$ be Coulomb energy class data and let $\{(A_n, \phi_n)[0]\}_n$ be a sequence of smooth low frequency truncations of $(A, \phi)[0]$ such that

$$\lim_{n \to \infty} (A_n, \phi_n)[0] = (A, \phi)[0]$$

in the sense of $\dot{H}^1_x \times L^2_x$. We denote by $(A_n, \phi_n)$ the smooth solutions to the MKG-CG system with initial data $(A_n, \phi_n)[0]$ and define $I = (-T_0, T_1) = \cup \tilde{I}$ to be the union of all open time intervals $\tilde{I}$ containing 0 with the property that

$$\sup_{J \subset I, J \text{ closed}} \inf_{n \to \infty} \| (A_n, \phi_n) \|_{S^1(J \times \mathbb{R}^4)} < \infty.$$ #1

Then we define the MKG-CG evolution of $(A, \phi)[0]$ on $I \times \mathbb{R}^4$ to be

$$(A, \phi)[t] := \lim_{n \to \infty} (A_n, \phi_n)[t], \quad t \in I,$$

where the limit is taken in the energy topology. We call $I$ the maximal lifespan of $(A, \phi)$. For any closed interval $J \subset I$, we set

$$\| (A, \phi) \|_{S^1(J \times \mathbb{R}^4)} := \lim_{n \to \infty} \| (A_n, \phi_n) \|_{S^1(J \times \mathbb{R}^4)}.$$ #2

We obtain the following characterization of the maximal lifespan $I$ of an energy class solution.

**Lemma 5.4.** Let $(A, \phi), (A_n, \phi_n)$ and $I$ be as in Definition 5.3. Assume in addition that $I \neq (-\infty, \infty)$. Then

$$\sup_{J \subset I, J \text{ closed}} \lim_{n \to \infty} \inf \| (A_n, \phi_n) \|_{S^1(J \times \mathbb{R}^4)} = \infty.$$

We call an energy class solution $(A, \phi)$ with maximal lifespan $I$ singular, if either $I \neq \mathbb{R}$, or if $I = \mathbb{R}$ and

$$\sup_{J \subset I, J \text{ closed}} \| (A, \phi) \|_{S^1(J \times \mathbb{R}^4)} = \infty.$$ #3

5.1. **Proof of Proposition 5.2.** A natural idea is to localize the data $(A_n, \phi_n)[0]$ in physical space to ensure smallness of the energy and to then try to “patch together” the local solutions obtained from the small energy global well-posedness result [12]. The problem is that the MKG-CG system does not have the finite speed of propagation property due to non-local terms in the equation for the magnetic potential $A$. To overcome this difficulty, we exploit that the Maxwell-Klein-Gordon system enjoys gauge invariance.

We first describe how we suitably localize the data $(A_n, \phi_n)[0]$ in physical space to obtain admissible Coulomb data with small energy that can be globally evolved by [12]. Let $\chi \in C^\infty_c(\mathbb{R}^4)$ be a
smooth cutoff function with support in the ball $B(0, \frac{3}{4})$ and such that $\chi \equiv 1$ on $B(0, \frac{5}{4})$. For $x_0 \in \mathbb{R}^4$ and $r_0 > 0$, we set $\chi_{\{x-x_0\leq r_0\}}(x) := \chi(\frac{x-x_0}{r_0})$. Then we define
\begin{equation}
\gamma_0(0, \cdot) := -\Delta^{-1} \partial_j \chi_{\{x-x_0\leq r_0\}}(\cdot) A_n^j(0, \cdot)
\end{equation}
and for $j = 1, \ldots, 4$,
\begin{equation}
\bar{A}_{n,j}(0, \cdot) := \chi_{\{x-x_0\leq r_0\}}(\cdot) A_{n,j}(0, \cdot) - \partial_j \gamma_n(0, \cdot).
\end{equation}
We determine $\bar{A}_{n,0}(0, \cdot)$ as the solution to the elliptic equation
\begin{equation}
\Delta \bar{A}_{n,0} = -\text{Im}(\chi_{\{x-x_0\leq r_0\}} \phi_n \chi_{\{x-x_0\leq r_0\}} \partial_\phi \phi_n) + |\chi_{\{x-x_0\leq r_0\}} \phi_n|^2 A_{n,0}
\end{equation}
on $\mathbb{R}^4$, where $\phi_n$ and $A_{n,0}$ are evaluated at time $t = 0$. We note that $\bar{A}_n$ is in Coulomb gauge. Then we set
\begin{equation}
\partial_t \gamma_n(0, \cdot) := A_{n,0}(0, \cdot) - \bar{A}_{n,0}(0, \cdot)
\end{equation}
and define $\partial_t \bar{A}_{n,j}(0, \cdot)$ for $j = 1, \ldots, 4$, first just on $B(x_0, \frac{5}{4} r_0)$, by setting
\begin{equation}
\partial_t \bar{A}_{n,j}(B(x_0, \frac{5}{4} r_0))(0, \cdot) := (\partial_t A_{n,j}(0, \cdot) - \partial_j \gamma_n(0, \cdot))|_{B(x_0, \frac{5}{4} r_0)}.
\end{equation}
We observe that $\Delta(A_{n,0}(0, \cdot) - \bar{A}_{n,0}(0, \cdot)) = 0$ on $B(x_0, \frac{5}{4} r_0)$ by the definition of $\bar{A}_{n,0}(0, \cdot)$. Thus, the data $\partial_t \bar{A}_{n,j}|_{B(x_0, \frac{5}{4} r_0)}(0, \cdot)$ satisfy the Coulomb compatibility condition $\partial_j(\partial_t \bar{A}_{n,j}(0, \cdot)) = 0$ on $B(x_0, \frac{5}{4} r_0)$. Using [4, Proposition 2.1], we extend $(\partial_t \bar{A}_{n,j}(0, \cdot)|_{B(x_0, \frac{5}{4} r_0)}$ to the whole of $\mathbb{R}^4$ while maintaining the Coulomb compatibility condition and such that $\|\partial_t \bar{A}_{n,j}\|_{L^2(\mathbb{R}^4)} \leq \|\partial_t \bar{A}_{n,j}\|_{L^2(B(x_0, \frac{5}{4} r_0))}$. Finally, we define
\begin{equation}
\bar{\phi}_n(0, \cdot) := e^{\partial_t \gamma_n(0, \cdot)} \chi_{\{x-x_0\leq r_0\}}(\cdot) \phi_n(0, \cdot)
\end{equation}
and
\begin{equation}
\partial_t \bar{\phi}_n(0, \cdot) := i \partial_t \gamma_n(0, \cdot) e^{\partial_t \gamma_n(0, \cdot)} \chi_{\{x-x_0\leq r_0\}}(\cdot) \phi_n(0, \cdot) + e^{\partial_t \gamma_n(0, \cdot)} \chi_{\{x-x_0\leq r_0\}}(\cdot) \partial_t \phi_n(0, \cdot).
\end{equation}
In the next lemma we prove that by choosing $r_0 > 0$ sufficiently small, we can ensure that the Coulomb data $(\bar{A}_n, \bar{\phi}_n)[0]$ have small energy for all sufficiently large $n$. Here we exploit that the convergence $(A_n, \phi_n)[0] \rightarrow (A, \phi)[0]$ in the energy topology as $n \rightarrow \infty$ implies a uniform non-concentration property of the energy of the data $(A_n, \phi_n)[0]$. We denote by $\varepsilon_0 > 0$ the small energy threshold of the small energy global well-posedness result [12] for the MKG-CG system.

**Lemma 5.5.** Let $(\bar{A}_n, \bar{\phi}_n)$ be defined as in (5.1) – (5.7). Given $\varepsilon_0 > 0$ there exists $r_0 > 0$ such that uniformly for all $x_0 \in \mathbb{R}^4$ and for all sufficiently large $n$, it holds that
\begin{equation}
E(\bar{A}_n, \bar{\phi}_n) < \varepsilon_0.
\end{equation}

**Proof.** We start with the components $\bar{A}_{n,j}$. Suppressing that $A_n$ is evaluated at time $t = 0$, we have for $j = 1, \ldots, 4$, that
\begin{equation}
\|\nabla \bar{A}_{n,j}\|_{L^2(\mathbb{R}^4)}^2 \leq \|\nabla \chi_{\{x-x_0\leq r_0\}} A_{n,j}\|_{L^2(\mathbb{R}^4)}^2 + \|\nabla \partial_j \gamma_n\|_{L^2(\mathbb{R}^4)}^2
\end{equation}
\begin{equation}
\leq \sum_{i=1}^4 \|\nabla \chi_{\{x-x_0\leq r_0\}} A_{n,i}\|_{L^2(\mathbb{R}^4)}^2
\end{equation}
\begin{equation}
\leq \sum_{i=1}^4 \frac{1}{r_0^2} \int_{B(x_0, \frac{5}{4} r_0)} |A_{n,i}(x)|^2 \, dx + \int_{B(x_0, \frac{5}{4} r_0)} |\nabla \chi A_{n,i}(x)|^2 \, dx
\end{equation}
\begin{equation}
\leq \sum_{i=1}^4 \left( \int_{B(x_0, \frac{5}{4} r_0)} |A_{n,i}(x)|^4 \, dx \right)^{1/2} + \int_{B(x_0, \frac{5}{4} r_0)} |\nabla \chi A_{n,i}(x)|^2 \, dx.
\end{equation}
Next we note that we can pick \( r_0 > 0 \) such that we have for the energy class data \( A \) that
\[
\sup_{x_0 \in \mathbb{R}^4} \sum_{j=1}^{4} \int_{B(x_0, \frac{3}{4}r_0)} |\nabla_x A_j(x)|^2 \, dx + \int_{B(x_0, \frac{3}{4}r_0)} |A_j(x)|^4 \, dx < \epsilon_0.
\]
Since \( A_n \to A \) in \( H^1_0(\mathbb{R}^4) \) as \( n \to \infty \), we also obtain for sufficiently large \( n \) that
\[
\sup_{x_0 \in \mathbb{R}^4} \sum_{j=1}^{4} \int_{B(x_0, \frac{3}{4}r_0)} |\nabla_x A_{n,j}(x)|^2 \, dx + \int_{B(x_0, \frac{3}{4}r_0)} |A_{n,j}(x)|^4 \, dx < \epsilon_0.
\]
From (5.8) we conclude that \( \|\nabla_x \tilde{A}_n, j\|^2_{L^2(\mathbb{R}^4)} \leq \epsilon_0. \) In a similar manner we argue that \( r_0 > 0 \) can be picked such that for all sufficiently large \( n \) we also have
\[
\sum_{j=1}^{4} \|\partial_t \tilde{A}_{n,j}\|^2_{L^2(\mathbb{R}^4)} + \|\nabla_x \tilde{A}_{n,0}\|^2_{L^2(\mathbb{R}^4)} + \|\nabla_x \bar{\phi}_n\|^2_{L^2(\mathbb{R}^4)} \leq \epsilon_0
\]
and hence,
\[
E(\tilde{A}_n, \bar{\phi}_n) \leq \epsilon_0.
\]
\[
\square
\]
By Lemma 5.5 we can pick \( r_0 > 0 \) such that the data \((\tilde{A}_n, \bar{\phi}_n)[0]\) can be globally evolved for sufficiently large \( n \) by the small energy global well-posedness result [12] and we obtain global \( S^1 \) norm bounds on their evolutions \((\tilde{A}_n, \bar{\phi}_n)\). For \( t > 0 \) we then define
\[
\tilde{\gamma}_n(t, \cdot) := A_{n,0}(t, \cdot) - \tilde{A}_{n,0}(t, \cdot),
\]
which implies that
\[
\tilde{\gamma}_n(t, \cdot) = \tilde{\gamma}_n(0, \cdot) + \int_0^t (A_{n,0}(s, \cdot) - \tilde{A}_{n,0}(s, \cdot)) \, ds.
\]
Our next goal is to relate the evolutions \((\tilde{A}_n, \bar{\phi}_n)\) and \((A_n, \phi_n)\) on the light cone
\[
K_{x_0, \frac{3}{4}r_0} = \{(t, x) : 0 \leq t < \frac{3}{4}r_0, |x - x_0| < \frac{3}{4}r_0 - t\}
\]
over the ball \( B(x_0, \frac{3}{4}r_0) \). These identities will be the key ingredient to recover \( S^1 \) norm bounds for \((A_n, \phi_n)\) from those of \((\tilde{A}_n, \bar{\phi}_n)\).

Lemma 5.6. Let \((\tilde{A}_n, \bar{\phi}_n)\) and \(\tilde{\gamma}_n\) be defined as in (5.1) - (5.7) and (5.9) - (5.10) such that \(E(\tilde{A}_n, \bar{\phi}_n) < \epsilon_0.\) For all sufficiently large \( n \) it holds that
\[
\tilde{A}_{n,j} = A_{n,j} - \partial_j \tilde{\gamma}_n \text{ on } K_{x_0, \frac{3}{4}r_0}
\]
for \( j = 1, \ldots, 4 \) and that
\[
\bar{\phi}_n = e^{i\tilde{\gamma}_n} \phi_n \text{ on } K_{x_0, \frac{3}{4}r_0}.
\]
Proof. To simplify the notation we omit the subscript \( n \). Using the equations that \((A, \phi), (\tilde{A}, \bar{\phi}), \) and \(\tilde{\gamma}\) satisfy, we obtain that
\[
\square \tilde{A}_j = -\operatorname{Im}(\bar{\phi} D_j \bar{\phi}) + \partial_j \Delta^{-1} \partial^j \operatorname{Im}(\bar{\phi} D_i \bar{\phi}) \text{ on } \mathbb{R}^4 \times \mathbb{R}^4
\]
and
\[
\square (A_j - \partial_j \tilde{\gamma}) = -\operatorname{Im}(\phi D_j \phi) + \partial_j \Delta^{-1} \partial^j \operatorname{Im}(\phi D_i \phi) - \partial_j \int_0^t \left\{\operatorname{Im}(\phi D_i \phi) - \operatorname{Im}(\bar{\phi} D_i \bar{\phi})\right\} \, ds \text{ on } K_{x_0, \frac{3}{4}r_0}.
\]
where we use the notation $\tilde{D}_\alpha = \partial_\alpha + iA_\alpha$. Next we introduce the quantities
\[ B_j = \tilde{A}_j - (A_j - \partial_j \gamma) \]
and
\[ \psi = \tilde{\phi} - e^{i\gamma} \phi. \]

From (5.11) and (5.12) we infer that
\[ \Box B_j = \text{Im}(\phi D_j \phi) - \text{Im}(\phi D_j \phi) - \partial_j \int_0^t \left\{ \text{Im}(\phi D_j \phi) - \text{Im}(\phi D_j \phi) \right\} ds \text{ on } K_{x_0, \frac{4}{\epsilon_0}}. \]
The first two terms in the above equation can be rewritten as
\[ \text{Im}(\phi D_j \phi) - \text{Im}(\phi D_j \phi) = B_j \phi^2 - \text{Im}(\psi(\partial_j + i\tilde{A}_j)(\psi + e^{i\gamma} \phi)) - \text{Im}(e^{i\gamma} \phi(\partial_j + i\tilde{A}_j)\psi) \]
and similarly we obtain for the last term that
\[ \text{Im}(\phi D_j \phi) - \text{Im}(\phi D_j \phi) = -\text{Im}(\psi(\partial_t + i\tilde{A}_0)(\psi + e^{i\gamma} \phi)) - \text{Im}(e^{i\gamma} \phi(\partial_t + i\tilde{A}_0)\psi). \]
We conclude that the wave equation for $B_j$ on the light cone $K_{x_0, \frac{4}{\epsilon_0}}$ is of the schematic form
\[
\Box B_j = f_1 B_j + f_2 |\psi|^2 + f_3 \psi + f_4 \overline{\psi} + f_5 (\partial_j \psi) + f_6 (\partial_j \overline{\psi})
+ \partial_j \int_0^t \left\{ f_7 |\psi|^2 + f_8 \psi + f_9 \overline{\psi} + f_{10} (\partial_j \psi) + f_{11} (\partial_j \overline{\psi}) \right\} ds,
\]
where $f_1, \ldots, f_{11}$ are smooth functions on $K_{x_0, \frac{4}{\epsilon_0}}$. To obtain a wave equation for $\psi$, we note that $B_0 = \tilde{A}_0 - (A_0 - \partial_t \gamma) = 0$ by construction and write
\[
0 = \Box \tilde{\phi} - e^{i\gamma} \Box \phi = \Box \tilde{\phi}(\psi + e^{i\gamma} \phi) - e^{i\gamma} \Box \phi
= \Box \tilde{\phi} + \Box_{B+A} \phi - \Box_{B+A} (e^{i\gamma} \phi) - e^{i\gamma} \Box \phi
= \Box \tilde{\phi} + \Box_B (e^{i\gamma} \phi) - \Box (e^{i\gamma} \phi) - 2B^j (A_j - \partial_j \gamma) e^{i\gamma} \phi + \Box_{A-B} (e^{i\gamma} \phi) - e^{i\gamma} \Box \phi
= \Box \tilde{\phi} + i(\partial^j B_j)(e^{i\gamma} \phi) + 2iB^j \partial_j (e^{i\gamma} \phi) - B^j B_j (e^{i\gamma} \phi) - 2B^j (A_j - \partial_j \gamma) e^{i\gamma} \phi.
\]
Thus, $\psi$ satisfies a wave equation on the light cone $K_{x_0, \frac{4}{\epsilon_0}}$ of the schematic form
\[
\Box \psi = f \psi + f_2 \partial^\alpha \psi + g_j B^j + g \partial_j B^j + h(\partial_j B^j),
\]
where $f, f_\alpha, g, g_j, h$ are smooth functions on $K_{x_0, \frac{4}{\epsilon_0}}$. Since $B[0]$ and $\psi[0]$ vanish on $B(x_0, \frac{5}{4})$ by our choice of the initial data $(\tilde{A}, \tilde{\phi})[0]$, we conclude from (5.13) and (5.14) by a standard energy argument that indeed
\[ \tilde{A}_j = A_j - \partial_j \gamma \text{ on } K_{x_0, \frac{5}{4}} \]
and
\[
\tilde{\phi} = e^{i\gamma} \phi \text{ on } K_{x_0, \frac{5}{4}}.
\]
It is clear that given $\epsilon_0 > 0$, there exists $R > 0$ such that for all sufficiently large $n$, it holds that
\[
E(\chi_{||x||>R}(\cdot)A_n(0, \cdot), \chi_{||x||>R}(\cdot)\phi_n(0, \cdot)) < \epsilon_0.
\]

For our later purposes we have to localize the initial data \((A_n, \phi_n)\) outside the large ball \(B(0, R)\) in a scaling invariant way. For any \(x_i \in \mathbb{R}^4\) with \(|x_i| \sim 2R 2^m\) for some \(m \in \mathbb{N}\), we set \(r_i := 2R 2^{m-1}\). Then we define

\[
\gamma_n^{(i)}(0, \cdot) := \Delta^{-1} \partial_j (\chi_{|x-x_i| \leq r_i}(\cdot) A_n^j(0, \cdot))
\]

and for \(j = 1, \ldots, 4\),

\[
\bar{A}_n^{(i)}(0, \cdot) := \chi_{|x-x_i| \leq r_i}(\cdot) A_n, j(0, \cdot) - \partial_j \gamma_n^{(i)}(0, \cdot).
\]

We define \(A_n^{(i)}(0, \cdot), \partial_t \bar{A}_n^{(i)}(0, \cdot), \bar{A}_n(0, \cdot)\) analogously to (5.5)–(5.7) and \(\gamma_n^{(i)}(t, \cdot), \partial_t \gamma_n^{(i)}(t, \cdot)\) for \(t > 0\) analogously to (5.9)–(5.10). Similarly to Lemma 5.5 and Lemma 5.6 we obtain

**Lemma 5.7.** Given \(\varepsilon_0 > 0\) there exists \(R > 0\) such that the initial data \((\bar{A}_n^{(i)}, \bar{\phi}_n^{(i)})\) defined as above satisfy for all sufficiently large \(n\) that

\[
E(\bar{A}_n^{(i)}, \bar{\phi}_n^{(i)}) < \varepsilon_0.
\]

and

**Lemma 5.8.** For all sufficiently large \(n\) it holds that

\[
\bar{A}_n^{(i)} = A_n, j - \partial_j \gamma_n^{(i)} \text{ on } K_{s_i, \frac{2}{3} r_i}
\]

for \(j = 1, \ldots, 4\), and that

\[
\bar{\phi}_n^{(i)} = e^{\gamma_n^{(i)}} \phi_n \text{ on } K_{s_i, \frac{2}{3} r_i},
\]

where \(K_{s_i, \frac{2}{3} r_i} := \{(t, x) : 0 \leq t < \frac{2}{3} r_i, |x - x_i| < \frac{2}{3} r_i - t\}\).

We now begin with the proof of Proposition 5.2 where we suitably “patch together” the small energy global evolutions constructed above.

**Proof of Proposition 5.2** By time reversibility, it suffices to only prove the statement in forward time. We pick \(r_0 > 0\) sufficiently small and \(R > 0\) sufficiently large according to Lemma 5.5 and Lemma 5.7. Then we cover the ball \(B(0, 2R) \subset \mathbb{R}^4\) by the supports of finitely many cutoffs \(\chi_{|x-x_i| \leq r_i}\) with \(r_l = r_0 = r_0 = 1, \ldots, L\) for some \(L \in \mathbb{N}\). We divide the complement \(B(0, 2R)^c\) of the ball \(B(0, 2R)\) into dyadic annuli \(A_m := \{x \in \mathbb{R}^4 : 2R 2^{m-1} < |x| \leq 2R 2^m\}, m \in \mathbb{N}\), and cover each \(A_m\) by the supports of finitely many suitable cutoffs \(\chi_{|x-x_i| \leq r_i}(\cdot)\) with \(|x| \sim 2R 2^m\) and \(r_1 \sim 2R 2^{m-1}\). This can be carried out in such a way that \(\{\text{supp}(\chi_{|x-x_i| \leq r_i})\}_{i=1}^\infty\) is a uniformly finitely overlapping covering of \(\mathbb{R}^4\). We denote by \((\bar{A}_n^{(i)}, \bar{\phi}_n^{(i)})\) the associated global solutions to MKG-CG with small energy data given by Lemma 5.5 and Lemma 5.7. Fix \(0 < T_0 < r_0\) such that

\[
[0, T_0] \times \mathbb{R}^4 \subset \bigcup_{l=1}^\infty K_{s_l, \frac{2}{3} r_l}.
\]

Then Lemma 5.6 and Lemma 5.8 imply that the evolutions \((A_n, \phi_n)\) exist on the time interval \([0, T_0]\) uniformly for all sufficiently large \(n\). The covering of \(\mathbb{R}^4\) by the supports of the cutoffs \(\chi_{|x-x_i| \leq r_i}(\cdot)\) can be done in such a way that there exists a uniformly finitely overlapping, smooth partition of unity \(\{\chi_l\}_{l \in \mathbb{N}} \subset C^\infty_c(\mathbb{R} \times \mathbb{R}^4)\),

\[
1 = \sum_{l=1}^\infty \chi_l \text{ on } [0, T_0] \times \mathbb{R}^4,
\]

so that each \(\chi_l\) satisfies \(K_{s_l, r_l} \subset \text{supp}(\chi_l) \subset K_{s_l, \frac{2}{3} r_l}\).
In order to obtain uniform $S^1$ norm bounds on the evolutions $A_n$ on $[0, T_0] \times \mathbb{R}^4$, we define for $i, j = 1, \ldots, 4$ and $l \in \mathbb{N}$, the curvature tensors

$$F_{n,ij} = \partial_i A_{n,j} - \partial_j A_{n,i}$$

and

$$F_n^{(l)}_{n,ij} = \partial_i \tilde{A}_{n,j}^{(l)} - \partial_j \tilde{A}_{n,i}^{(l)}.$$  

From Lemma 5.6 and Lemma 5.8 we conclude that

$$F_{n,ij} = \sum_{l=1}^{\infty} \chi_l F_{n,ij} = \sum_{l=1}^{\infty} \chi_l F_n^{(l)}_{n,ij} \text{ on } [0, T_0] \times \mathbb{R}^4.$$  

Using the Coulomb gauge, we find for $j = 1, \ldots, 4$ that

$$A_{n,j} = \Delta^{-1} \partial^j F_{n,ij} = \sum_{l=1}^{\infty} \Delta^{-1} \partial^j (\chi_l F_n^{(l)}_{n,ij}) \text{ on } [0, T_0] \times \mathbb{R}^4.$$  

By an almost orthogonality argument, we obtain for $j = 1, \ldots, 4$ that

$$\|A_{n,j}\|_{S^1([0,T_0] \times \mathbb{R}^4)} = \left\| \sum_{l=1}^{\infty} \Delta^{-1} \partial^j (\chi_l F_n^{(l)}_{n,ij}) \right\|_{S^1([0,T_0] \times \mathbb{R}^4)}^{1/2} \leq C(A, \phi) \left( \sum_{l=1}^{\infty} \|\Delta^{-1} \partial^j (\chi_l F_n^{(l)}_{n,ij})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}^2 \right)^{1/2} \leq C(A, \phi) \left( \sum_{l=1}^{\infty} \|\Delta^{-1} \nabla_j (\chi_l \nabla \tilde{A}_n^{(l)})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}^2 \right)^{1/2}.$$  

Next we want to invoke the following multiplier bound for $S^1$ norms that will be proven at the end of this subsection.

**Lemma 5.9.** Let $\chi \in C^\infty(\mathbb{R} \times \mathbb{R}^4)$ satisfy

$$\max_{k=0,1,\ldots,4} \|\nabla^k \chi\|_{L^r_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq D \text{ for all } 1 \leq q, r \leq \infty$$

for some $D > 0$. Then there exists a constant $C > 0$ independent of $\chi$ such that for all $\psi \in S^1(\mathbb{R} \times \mathbb{R}^4)$, it holds that

$$\|\Delta^{-1} \nabla_j (\chi \nabla \psi)\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq CD\|\psi\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}.$$  

By scaling invariance of the $S^1$ norm and the scaling invariant setup of the partition of unity $\{\chi_l\}_{l \in \mathbb{N}}$, we are in a position to apply Lemma 5.9 uniformly for all multipliers $\chi_l$ to estimate the right-hand side of (5.16) by

$$C(A, \phi) \left( \sum_{l=1}^{\infty} \|\tilde{A}_n^{(l)}\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}^2 \right)^{1/2}.$$  

By the small data global well-posedness result [12], this is in turn bounded by

$$C(A, \phi) \left( \sum_{l=1}^{\infty} \|
abla_t \phi_n^{(l)}(0)\|_{L^2_t L^2_x(\mathbb{R}^4)}^2 + \|
abla_t \tilde{A}_n^{(l)}(0)\|_{L^2_t L^2_x(\mathbb{R}^4)}^2 \right)^{1/2}.$$  

It remains to square sum over the $\dot{H}^1_\chi \times L^2_\chi$ norms of the initial data $(\phi_n^{(l)}, \tilde{A}_n^{(l)}(0))$, which we defer to the end of the proof.
To deduce uniform $S^1$ norm bounds for the evolutions $\phi_n$ on $[0, T_0] \times \mathbb{R}^4$, we use Lemma 5.6 and Lemma 5.8 to write

$$\phi_n = \sum_{i=1}^{\infty} \chi_i \phi_n = \sum_{i=1}^{\infty} \chi_i e^{-r_n(i)} \phi_n(i)$$

on $[0, T_0] \times \mathbb{R}^4$.

By an almost orthogonality argument, we have

$$\|\phi_n\|_{S^1([0,T_0] \times \mathbb{R}^4)} \leq C(A, \phi) \left( \sum_{i=1}^{\infty} \|\chi_i e^{-r_n(i)} \phi_n(i)\|_{S^1([0,T_0] \times \mathbb{R}^4)}^2 \right)^{1/2}. \tag{5.19}$$

Here it is not immediate how to obtain $S^1$ norm bounds for $\chi_i e^{-r_n(i)} \phi_n(i)$, because $\gamma_n(i)$ implicitly depends on the unknown quantity $\phi_n$, since we defined in (5.10) for $t > 0$

$$\gamma_n(i)(t, \cdot) = \gamma_n(i)(0, \cdot) + \int_0^t (\Delta A_{t,0}(s, \cdot) - \tilde{A}(i)(s, \cdot)) \, ds$$

and we have $\Delta A_{t,0} = - \Im(\phi_n D_{i} \phi_n)$. We will overcome this difficulty by exploiting that $\partial_t \gamma_n(i)$ is a harmonic function on every fixed-time slice of $K(x_i, \frac{\gamma}{\sqrt{r}} t)$ in view of Lemma 5.6 respectively Lemma 5.8 and its definition

$$\partial_t \gamma_n(i)(t, \cdot) = A_{n,0}(t, \cdot) - \tilde{A}_{n,0}(t, \cdot) \text{ for } t \geq 0.$$ It therefore enjoys the interior derivative estimates for harmonic functions on every fixed-time slice of $K(x_i, \frac{\gamma}{\sqrt{r}} t)$

The partition of unity (5.15) was chosen in such a way that the cutoff functions $\chi_i$ satisfy for all $(t, x) \in \text{supp}(\chi_i)$ that $B(x, \frac{\gamma}{2} r_i) \subset K_{x_i, \frac{\gamma}{\sqrt{r}} t_i} \cap \{t\} \times \mathbb{R}^4$. Thus, for all integers $k \geq 0$ and $(t, x) \in \text{supp}(\chi_i)$ we obtain from the interior derivative estimates for harmonic functions that

$$\left| \chi_i \nabla_x \partial_t \gamma_n(i) \right| \leq \frac{C(k)}{r_i^{1+k}} \|\partial_t A_{n,0}(t, \cdot)\|_{L^2(B(x, \frac{\gamma}{\sqrt{r}}))} \leq \frac{C(k)}{r_i^{1+k}} \|\partial_t A_{n,0}(t, \cdot)\|_{L^2(\mathbb{R}^4)} \leq \frac{C(k)}{r_i^{1+k}} E(A, \phi)^{1/2}.$$ We conclude that

$$\bigl(5.20\bigr) \quad r_i^{1+k} \|\chi_i \nabla_x \partial_t \gamma_n(i)\|_{L^\infty L^2(\mathbb{R}^4)} \leq C(k, A, \phi).$$ Similarly, we observe that by Lemma 5.6,

$$\partial_t^2 \gamma_n(i)(t, \cdot) = \partial_t A_{n,0}(t, \cdot) - \tilde{A}_{n,0}(t, \cdot)$$

is harmonic on every fixed-time slice of $K(x_i, \frac{\gamma}{\sqrt{r}} t_0)$. The interior derivative estimates for harmonic functions then yield for all integers $k \geq 0$ and $(t, x) \in \text{supp}(\chi_i)$ that

$$\left| \chi_i \nabla_x \partial_t^2 \gamma_n(i)(t, x) \right| \leq \frac{C(k)}{r_i^{2+k}} \|\partial_t A_{n,0}(t, \cdot) - \tilde{A}_{n,0}(t, \cdot)\|_{L^2(B(x, \frac{\gamma}{\sqrt{r}}))} \leq \frac{C(k)}{r_i^{2+k}} \|\partial_t A_{n,0}(t, \cdot) - \tilde{A}_{n,0}(t, \cdot)\|_{L^2(\mathbb{R}^4)}.$$ Since we have

$$\|\partial_t A_n(t, \cdot)\|_{L^2(\mathbb{R}^4)} \leq \|\nabla_x \partial_t A_n(t, \cdot)\|_{L^2} \leq \sum_{i=1}^{4} \|\text{Im}(\phi_n D_{i} \phi_n)\|_{L^2} \leq \sum_{i=1}^{4} \|\phi_n\|_{L^2(\mathbb{R}^4)} \|D_{i} \phi_n\|_{L^2} \leq E(A, \phi)$$
and analogously for $\|\partial_t A_n(t, \cdot)\|_{L^2_t(\mathbb{R}^4)}$, it follows that

\begin{equation}
I_{k}^{2+k}\|\chi_1 \nabla^k_{\gamma_n} \gamma_n(i)\|_{L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq C(k, A, \phi).
\end{equation} 

(5.21)

Next, we note that $\gamma_n^{(i)}(0, \cdot)$ as defined in (5.1) is harmonic on the ball $B(x_l, \frac{\sqrt{3}}{r_l})$. As before, the interior derivative estimates for harmonic functions give for all integers $k \geq 0$ that

\begin{equation}
I_{k}^{2+k}\|\chi_1 \nabla^k \gamma_n(0, \cdot)\|_{L^p_t L^q_x(\mathbb{R}^4)} \leq C(k, A, \phi).
\end{equation} 

(5.22)

For all $(t, x) \in \text{supp}(\chi_l)$ we then obtain from

\begin{equation}
\gamma_n^{(i)}(t, x) = \gamma_n^{(i)}(0, x) + \int_0^t \partial_t \gamma_n^{(i)}(s, x) \, ds
\end{equation} 

that

\begin{equation}
I_{k}^{2+k}\|\chi_1 \nabla^k \gamma_n(0, \cdot)\|_{L^p_t L^q_x(\mathbb{R}^4)} \leq C(k, A, \phi).
\end{equation} 

(5.23)

From (5.20) – (5.23) we conclude that for all integers $k \geq 0$ there exists a constant $C(k, A, \phi) > 0$, depending only on $k$ and the energy class data $(A, \phi)$, such that for all sufficiently large $n$, for all $l \in \mathbb{N}$ and for all $1 \leq q, r \leq \infty$,

\begin{equation}
\max_{m=1,2} r_{l}^{k+m}\|\nabla^k \partial_t \psi_l(x_l e^{-t} \chi^{(i)}_l)\|_{L^p_t L^q_x(\mathbb{R}^d)} \leq C(k, A, \phi).
\end{equation} 

(5.24)

Similarly to Lemma 5.9, we also have the following multiplier bound for the $S^1$ norm.

**Lemma 5.10.** Let $\chi \in C^\infty(\mathbb{R} \times \mathbb{R}^4)$ satisfy

\begin{equation}
\max_{k=0,1,2} \max_{m=0,1,2} \|\nabla^k \partial_t \chi\|_{L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq D \text{ for all } 1 \leq q, r \leq \infty
\end{equation} 

for some $D > 0$. Then there exists a constant $C > 0$ independent of $\chi$ such that for all $\psi \in S^1(\mathbb{R} \times \mathbb{R}^4)$,

\begin{equation}
\|\chi \psi\|_{L^1_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq C D \|\psi\|_{L^1_t L^q_x(\mathbb{R} \times \mathbb{R}^4)}.
\end{equation} 

In view of (5.24), the scaling invariance of the $S^1$ norm and the scaling invariant setup of the partition of unity $\{\chi_l\}_{l \in \mathbb{N}}$, we can apply Lemma 5.10 uniformly for all multipliers $\chi_l$ to estimate the right hand side of (5.19) by

\begin{equation}
C(A, \phi)^{\frac{1}{2}} \left(\sum_{l=1}^{\infty} \|\tilde{\phi}_n^{(i)}\|_{L^2(\mathbb{R}^4)}^2\right)^{1/2}.
\end{equation} 

(5.25)

By the small energy global well-posedness result [12], this is in turn bounded by

\begin{equation}
C(A, \phi)^{\frac{1}{2}} \left(\sum_{l=1}^{\infty} \|\nabla_l \tilde{\phi}_n^{(i)}(0)\|_{L^2}^2 + \|\nabla_l \tilde{A}_n^{(i)}(0)\|_{L^2}^2\right)^{1/2}.
\end{equation} 

(5.26)

It remains to square sum over the $H^1_t \times L^2_x$ norms of the initial data $(\tilde{\phi}_n^{(i)}, \tilde{A}_n^{(i)})(0)$ in (5.18) and in (5.27). Here we have, for example, from the definition

\begin{equation}
\tilde{\phi}_n^{(i)}(0, \cdot) := e^{i r_n^{(i)}(0)} \chi_{|x-x_l| \leq r_l}() \psi_l(0, \cdot)
\end{equation} 


that
(5.28)
\[
\sum_{l=1}^{\infty} \int_{\mathbb{R}^4} |\nabla_x \tilde{y}_n^{(l)}(0, x)|^2 \, dx \leq \sum_{l=1}^{\infty} \int_{\mathbb{R}^4} \left( |\nabla_x \tilde{y}_n^{(l)}(0, x)|^2 |\chi_{|x-x_l| \leq 4r}|(x)|^2 + |\nabla_x \chi_{|x-x_l| \leq 4r}(x)|^2 |\phi_n(0, x)|^2 \right) \, dx \\
+ \sum_{l=1}^{\infty} \int_{\mathbb{R}^4} |\chi_{|x-x_l| \leq 4r}(x)|^2 |\nabla_x \phi_n(0, x)|^2 \, dx.
\]
By the construction of the partition of unity, we have uniformly for all \(l \in \mathbb{N}\) and \(x \in \mathbb{R}^4\) that
\[
|\nabla_x \chi_{|x-x_l| \leq 4r}(x)|^2 \leq \frac{C(A, \phi)}{|x|^2}
\]
and, using also (5.22), that
\[
|\nabla_x \tilde{y}_n^{(l)}(0, x)|^2 |\chi_{|x-x_l| \leq 4r}(x)|^2 \leq \frac{C(A, \phi)}{|x|^2}.
\]
By Hardy’s inequality and the uniform finite overlap of the supports of the cutoffs \(\chi_{|x-x_l| \leq 4r}(\cdot)\), we conclude that (5.28) is bounded by
\[
C(A, \phi)\parallel \nabla_x \phi_n(0, \cdot)\parallel_{L^2_k}^2 \leq C(A, \phi)E(A, \phi)
\]
uniformly for all sufficiently large \(n\). Proceeding similarly with the other terms in (5.27), we finally obtain that (5.28) is bounded by \(C(A, \phi)E(A, \phi)\) uniformly for all sufficiently large \(n\). This finishes the proof of Proposition 5.2.

It remains to prove Lemma 5.9 and Lemma 5.10. We only give the proof of Lemma 5.10, the other one being similar.

**Proof of Lemma 5.10** We have to prove that for any \(\psi \in S^1(\mathbb{R} \times \mathbb{R}^4)\),
\[
\|\chi \psi\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} = \left( \sum_{k \in \mathbb{Z}} \|P_k \nabla_{l \times} (\chi \psi)\|^2_{L^2_k(\mathbb{R} \times \mathbb{R}^4)} + \|\nabla (\chi \psi)\|^2_{L^1_k L^{4/3}_x(\mathbb{R} \times \mathbb{R}^4)} \right)^{1/2} \leq CD\|\psi\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}.
\]
To this end we will constantly invoke the assumed space-time bounds (5.25) for the multiplier \(\chi\).

We first consider the \(S^1_k\) component of the \(S^1\) norm. Here we denote by \((q, r)\) any wave-admissible exponent pair, i.e. satisfying \(2 \leq q, r \leq \infty\) and \(\frac{2}{q} + \frac{3}{r} = \frac{1}{2}\). For any \(k \in \mathbb{Z}\) we have
(5.29) \[
\|P_k \nabla_{l \times} (\chi \psi)\|_{S^1_k} \leq \|P_k((\nabla_{l \times} \chi) \psi)\|_{S^1_k} + \|P_k(\chi \nabla_{l \times} \psi)\|_{S^1_k}
\]
and begin with estimating the first term on the right hand side of (5.29). If \(k \leq 0\), we obtain by the Bernstein and Sobolev inequalities uniformly for all \((q, r)\) that
\[
2^{(\frac{1}{4} + \frac{1}{q} - 2)k} \|P_k((\nabla_{l \times} \chi) \psi)\|_{L^q_{l \times} L^2_k} \leq 2^{\frac{1}{4}} 2^{2k} \|\nabla_{x \times} \chi\|_{L^{q\times}_l L^2_k} \|\psi\|_{L^{q\times}_l L^2_k} \leq 2^{2k} \|\nabla_{x \times} \chi\|_{L^{q\times}_l L^2_k} \|\psi\|_{L^{q\times}_l L^2_k} \leq 2^{2k} \|\nabla_{x \times} \chi\|_{L^{q\times}_l L^2_k} \|\psi\|_{S^1}.
\]
Here we used that
\[
\|\nabla_x \psi\|_{L^{q\times}_l L^2_k} \leq \left( \sum_{k \in \mathbb{Z}} \|P_k \nabla_{x \times} \psi\|^2_{L^{q\times}_l L^2_k} \right)^{1/2} \leq \left( \sum_{k \in \mathbb{Z}} \|P_k \nabla_{x \times} \psi\|^2_{S^1_k} \right)^{1/2} \leq \|\psi\|_{S^1}.
\]
If \( k > 0 \), we have
\[
2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} \\
\leq 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((P_{<k-C}(\nabla_{t,x}\chi))\psi)\|_{L^2_t L^4_x} + 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((P_{\leq k-C}(\nabla_{t,x}\chi))P_{k+O(1)}\psi)\|_{L^2_t L^4_x} \\
\leq 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_{>k-C}\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\psi\|_{L^2_t L^4_x} + 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_{\leq k-C}\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|P_{k+O(1)}\psi\|_{L^2_t L^4_x} \\
\leq 2^{\frac{1}{k}}2^{-k}\|\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\nabla_{t,x}\psi\|_{L^2_t L^4_x} + 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_{\leq k-C}\nabla_{t,x}\chi\|_{L^2_t L^4_x}2^{-k}\|P_{k+O(1)}\nabla_{t,x}\psi\|_{L^2_t L^4_x} \\
\leq 2^{-\frac{1}{k}}\|\nabla_{t,x}^2\partial_t \chi\|_{L^2_t L^4_x}\|\psi\|_{L^1_t} + \|\nabla_{t,x}\chi\|_{L^2_t L^4_x}2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_{k+O(1)}\nabla_{t,x}\psi\|_{L^2_t L^4_x},
\]
where we used the reverse Bernstein inequality
\[
\|P_{>k-C}\nabla_{t,x}\chi\|_{L^2_t L^4_x} \leq 2^{-k}\|\nabla_{t,x}\chi\|_{L^2_t L^4_x}.
\]
Square-summing over \( k \in \mathbb{Z} \) yields the desired bound. We continue with the second term on the right hand side of (5.29). If \( k \leq 0 \), we use Bernstein’s inequality to bound
\[
2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} \leq 2^{\frac{1}{k} - 2k}\|\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\nabla_{t,x}\psi\|_{L^2_t L^4_x} \leq 2^{2k}\|\chi\|_{L^2_t L^4_x}\|\psi\|_{L^1_t}.
\]
For \( k > 0 \) we find
\[
2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} \\
\leq 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((P_{>k-C}\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} + 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_k((P_{\leq k-C}\chi)P_{k+O(1)}\nabla_{t,x}\psi)\|_{L^2_t L^4_x} \\
\leq 2^{\frac{1}{k}}\|P_{>k-C}\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\nabla_{t,x}\psi\|_{L^2_t L^4_x} + 2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_{\leq k-C}\chi\|_{L^2_t L^4_x}\|P_{k+O(1)}\nabla_{t,x}\psi\|_{L^2_t L^4_x} \\
\leq 2^{-\frac{1}{k}}\|\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\psi\|_{L^1_t} + \|\chi\|_{L^2_t L^4_x}2^{\frac{1}{k} + \frac{3}{k} - 2k}\|P_{k+O(1)}\nabla_{t,x}\psi\|_{L^2_t L^4_x}.
\]
The desired bound again follows after square-summing over \( k \in \mathbb{Z} \).

Next we consider the \( X^{0,1}_\infty \) component of the \( S_k \) norm. For any \( k \in \mathbb{Z} \) we have
\[
(5.30) \quad \|P_k\nabla_{t,x}(\chi)\psi\|_{X^{0,\frac{1}{k}}_\infty} \leq \|P_k((\nabla_{t,x}\chi)\psi)\|_{X^{0,\frac{1}{k}}_\infty} + \|P_k(\chi\nabla_{t,x}\psi)\|_{X^{0,\frac{1}{k}}_\infty}.
\]
We start with the first term on the right hand side of (5.30). If \( k \leq 0 \), we split into a small and a large modulation term
\[
(5.31) \quad P_k((\nabla_{t,x}\chi)\psi) = P_kQ_{\leq 0}((\nabla_{t,x}\chi)\psi) + P_kQ_{>0}((\nabla_{t,x}\chi)\psi).
\]
We easily estimate the small modulation term using Bernstein’s inequality,
\[
\|P_kQ_{\leq 0}((\nabla_{t,x}\chi)\psi)\|_{X^{0,\frac{1}{k}}_\infty} \leq \|P_k((\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} \leq 2^{2k}\|\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\psi\|_{L^\infty_t L^4_x} \\
\leq 2^{2k}\|\nabla_{t,x}\chi\|_{L^2_t L^4_x}\|\psi\|_{L^1_t}.
\]
To estimate the large modulation term we consider for any \( j > 0 \),
\[
2^{\frac{j}{2}}\|P_kQ_j((\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} \leq 2^{\frac{j}{2}}\|P_kQ_j((P_{>j-C}(\nabla_{t,x}\chi)\psi)\|_{L^2_t L^4_x} \\
+ 2^{\frac{j}{2}}\|P_kQ_j((P_{\leq j-C}Q_{>j-C}(\nabla_{t,x}\chi)\psi))\|_{L^2_t L^4_x} \\
+ 2^{\frac{j}{2}}\|P_kQ_j((P_{\leq j-C}Q_{\leq j-C}(\nabla_{t,x}\chi)Q_{j+O(1)}\psi))\|_{L^2_t L^4_x}.
\]
We bound the first term using the reverse Bernstein inequality,

$$2^{\frac{j}{2}} \| P_k Q_j ((P_{j+C}(\nabla_t, \dot{x} \chi))) \|_{L_t^2 L_x^2} \leq 2^{2k} 2^{\frac{j}{2}} \| (P_{j+C}(\nabla_t, \dot{x} \chi)) \|_{L_t^2 L_x^4} \leq 2^{2k-\frac{j}{2}} \| \nabla_x \nabla_t \dot{x} \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1}.$$  

For the second term on the right hand side of (5.32) we obtain from a reverse Bernstein estimate in time that

$$2^{\frac{j}{2}} \| P_k Q_j ((P_{j-C} Q_{j-C}(\nabla_t, \dot{x} \chi))) \|_{L_t^2 L_x^2} \leq 2^{2k-\frac{j}{2}} \| \nabla_x \nabla_t \dot{x} \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1}.$$  

The third term on the right hand side of (5.32) can be estimated via a Littlewood-Paley trichotomy

$$2^{\frac{j}{2}} \| P_k Q_j ((P_{\leq j-C} Q_{\leq j-C}(\nabla_t, \dot{x} \chi))) \|_{L_t^2 L_x^2} \leq \sum_{l \geq k-C} 2^{\frac{j}{2}} \| (P_{l+C}(\nabla_t, \dot{x} \chi)) \|_{L_t^2 L_x^2} \| P_{l+C} Q_j \|_{L_t^2 L_x^4} \| \psi \|_{S^1} + \sum_{l \geq k-C} 2^{\frac{j}{2}} \| (P_{l-C}(\nabla_t, \dot{x} \chi)) \|_{L_t^2 L_x^4} \| \psi \|_{S^1}.$$  

(5.33)

We bound the high-high case by

$$\sum_{l \geq k+C} 2^{\frac{j}{2}} \| P_{l+C} Q_j \|_{L_t^2 L_x^2} \| P_{l+C} Q_j \|_{L_t^2 L_x^4} \| \psi \|_{S^1} \leq 2^{\frac{j}{2}} \| \nabla_t \nabla_x \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1}.$$  

The high-low case is estimated by

$$\sum_{l \geq k+C} 2^{\frac{j}{2}} \| P_{l+C} Q_j \|_{L_t^2 L_x^2} \| P_{l+C} Q_j \|_{L_t^2 L_x^4} \| \psi \|_{S^1} \leq 2^{\frac{j}{2}} \| \nabla_t \nabla_x \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1}.$$  

and the low-low case by

$$2^{\frac{j}{2}} \| P_{\leq k-C} Q_{\leq j-C}(\nabla_t, \dot{x} \chi) \|_{L_t^2 L_x^2} \| P_{k+C} Q_{\leq j-C}(\nabla_t, \dot{x} \chi) \|_{L_t^2 L_x^4} \| \psi \|_{S^1} \leq 2^k \| \nabla_t \nabla_x \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1}.$$  

Thus, we obtain the following estimate for the large modulation term in (5.31),

$$\| P_k Q_{\geq 0}(\nabla_t, \dot{x} \chi) \psi \|_{X_0^{\frac{1}{2}}} \leq 2^{2k} \| \nabla_t^2 \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1} + 2^k \| \nabla_t \nabla_x \chi \|_{L_t^2 L_x^4} \| \psi \|_{S^1} + \| \nabla_t \nabla_x \chi \|_{L_t^2 L_x^4} \| P_{k+C} Q_{\leq j-C}(\nabla_t, \dot{x} \chi) \|_{X_0^{\frac{1}{2}}}.$$  

If $k > 0$ in the first term on the right hand side of (5.30), we again split into a small and a large modulation term,

$$P_k(\nabla_t, \dot{x} \chi) \psi = P_k Q_{\leq k}(\nabla_t, \dot{x} \chi) \psi + P_k Q_{> k}(\nabla_t, \dot{x} \chi) \psi.$$  

(5.34)
We can immediately dispose of the small modulation term as follows

\[
\left\| P_k Q_j ((\nabla_t x \chi) \psi) \right\|_{L^{2}_t L^\infty_x} \lesssim 2^{\frac{j}{2}} \left\| P_k ((\nabla_t x \chi) \psi) \right\|_{L^{2}_t L^\infty_x}^{\frac{1}{2}} \lesssim 2^{-k} \left( \left\| \nabla_x \nabla_t x \chi \right\|_{L^2_t L^1_x} \left\| \psi \right\|_{L^\infty_t L^1_x} + \left\| \nabla_x \chi \right\|_{L^2_t L^1_x} \left\| \nabla_x \psi \right\|_{L^\infty_t L^1_x} \right). 
\]

To treat the large modulation term, we find that for any \( j > k \),

\[
2^{-k} \left\| P_j Q_j ((\nabla_t x \chi) \psi) \right\|_{L^{2}_t L^\infty_x} \leq 2^{\frac{k}{2}} \left\| P_j Q_j ((P_{> j \cdot C} (\nabla_t x \chi)) \psi) \right\|_{L^{2}_t L^\infty_x} + 2^{\frac{j}{2}} \left\| P_j Q_j ((P_{\leq j \cdot C} Q_j \cdot c (\nabla_t x \chi)) \psi) \right\|_{L^{2}_t L^\infty_x} + 2^{\frac{j}{2}} \left\| P_j Q_j ((P_{\leq j \cdot C} Q_j \cdot c (\nabla_t x \chi)) Q_j \cdot O(1) \psi) \right\|_{L^{2}_t L^\infty_x}. 
\]

(5.35)

We estimate the first term by

\[
2^{-k} \left\| P_{> j \cdot C} \nabla_t x \chi \right\|_{L^2_t L^1_x} \left\| \psi \right\|_{L^\infty_t L^1_x} \lesssim 2^{-\frac{k}{2}} \left\| \nabla_x \nabla_t x \chi \right\|_{L^2_t L^1_x} \left\| \nabla_x \psi \right\|_{L^\infty_t L^1_x} \lesssim 2^{-\frac{k}{2}} \left\| \nabla_x \nabla_t x \chi \right\|_{L^2_t L^1_x} \left\| \psi \right\|_{C^1}.
\]

The second term on the right hand side of (5.35) is bounded by

\[
2^{-k} \left\| P_{\leq j \cdot C} Q_{> j \cdot C} \nabla_t x \chi \right\|_{L^2_t L^1_x} \left\| \psi \right\|_{L^\infty_t L^1_x} \lesssim 2^{-\frac{k}{2}} \left\| \nabla_t x \partial_\chi \right\|_{L^2_t L^1_x} \left\| \nabla_x \psi \right\|_{L^\infty_t L^1_x} \lesssim 2^{-\frac{k}{2}} \left\| \nabla_t x \partial_\chi \right\|_{L^2_t L^1_x} \left\| \psi \right\|_{C^1}.
\]

Using a Littlewood-Paley trichotomy we obtain the following estimate of the third term in (5.35)

\[
2^{-k} \left( \left\| \nabla_x \nabla_t x \chi \right\|_{L^\infty_t L^1_x} + \left\| \nabla^2_x \nabla_t x \chi \right\|_{L^\infty_t L^1_x} \right) \left\| \psi \right\|_{C^1} + \left\| \nabla_t x \chi \right\|_{L^\infty_t L^1_x} \left\| \nabla_x \psi \right\|_{C^1} \lesssim \left\| \nabla_t x \chi \right\|_{L^\infty_t L^1_x} \left\| \psi \right\|_{C^1}.
\]

Finally, square-summing over \( k \in \mathbb{Z} \) gives the desired bound for the first term on the right hand side of (5.30). The second term on the right hand side of (5.30) can be handled similarly.

It remains to treat the \( S^\text{ang}_k \) component of the \( S_k \) norm, which is given by

\[
\left\| P_k (\chi \psi) \right\|_{S^\text{ang}_k}^2 = \sup_{l \geq 0} \sum_{\omega} \left\| P_k P^\omega_l Q_{\cdot k+2l} (\chi \psi) \right\|_{S^\text{ang}_l}^2.
\]

Since \( \| \cdot \|_{S^\text{ang}_l} \) involves square sums over rectangular regions in frequency space, the factor \( \chi \) may prevent us from passing this localization to the input \( \psi \). This issue can only occur when the factor \( \chi \) is at spatial frequency \( \gtrsim 2^{k+2l} \). We therefore split into

\[
P_k P^\omega_l Q_{\cdot k+2l} (\chi \psi) = P_k P^\omega_l Q_{\cdot k+2l} ((P_{> k+2l} \chi) \psi) + P_k P^\omega_l Q_{\cdot k+2l} ((P_{\leq k+2l} \chi) \tilde{P}_k \tilde{P}^\omega_l \tilde{Q}_{\cdot k+2l} \psi).
\]

The first term can be estimated proceeding similarly as before, while for the second term we use the disposability of the operator \( P_k P^\omega_l Q_{\cdot k+2l} \).

Finally, we consider the \( \ell^1 L^2_t H^{-\frac{1}{2}}_x \) component of the \( S^1 \) norm. Here we have to estimate

\[
\left\| \Box (\chi \psi) \right\|_{\ell^1 L^2_t H^{-\frac{1}{2}}_x} \lesssim \sum_{k \in \mathbb{Z}} 2^{-\frac{k}{2}} \left\| P_k \Box (\chi \psi) \right\|_{L^2_t L^1_x},
\]

where

\[
\Box (\chi \psi) = (\Box \chi) \psi + 2 \partial_\alpha \partial_\alpha \Box \chi \psi + \chi \Box \psi.
\]
Using the multiplier bounds \((5.25)\), we estimate the first two terms by placing \(\psi\) into \(L^4_t L^4_x\), respectively. \(\tilde{\partial}^a \psi\) into \(L^\infty_t L^2_x\). For the last term we use a Littlewood-Paley trichotomy. \(\square\)

### 5.2. Localizing in physical space

In this subsection we consider Coulomb data \((A, \phi)[0] \in H^s_t \times H^{s-1}_x\) for all \(s \geq 1\). We show that there exists \(T > 0\) and a \(C^\infty\) solution of the same regularity class on each time slice of the space-time slab \([-T, T] \times \mathbb{R}^4\) satisfying the required \(S^1\) norm bound

\[
\|(A, \phi)\|_{S^1([-T, T] \times \mathbb{R}^4)} < \infty.
\]

To this end we fix a large \(R_0 > 1\). For each \(R \geq R_0\), we consider a cutoff \(\chi_R \in C^\infty_c(\mathbb{R}^4)\) that equals 1 on the ball \(B_R(0)\) and has support in a dilate of \(B_R(0)\). In the previous Subsection \(5.1\) we demonstrated that upon writing

\[
\tilde{A}_R := \chi_R A - \nabla \chi_R
\]

for the spatial components of a new connection form \(\tilde{A}_R\), where

\[
\chi_R = \Delta^{-1} \partial_j (\chi_R A^j),
\]

there is a way to pick the remaining data \(\tilde{\partial} A_R(0)\) and \(\tilde{\phi}_R[0]\), so that the corresponding data are all of class \(H^{1+}_s \times L^0_x\) and of Coulomb class. Thus, we obtain local solutions with these data from the local well-posedness result \([20]\). We can also arrange that \(\tilde{\phi}_R[0]\) is supported within the ball of radius \(\frac{R}{10}\) centered at the origin. It is then also easy to verify that

\[
(\tilde{A}_R, \tilde{\phi}_R)[0] \to (A, \phi)[0] \text{ as } R \to \infty
\]

with respect to the \(H^s_t \times H^{s-1}_x\) topology for any \(s \geq 1\). Moreover, the argument in the previous subsection together with Proposition \(4.3\) implies that these solutions extend of class \(H^s_t \times H^{s-1}_x\) to a space-time slab \([-T, T] \times \mathbb{R}^4\), where \(T > 0\) is independent of \(R \geq R_0\). It then remains to check that the corresponding local solutions on \([-T, T] \times \mathbb{R}^4\), call them again \((\tilde{A}_R, \tilde{\phi}_R)\), converge with respect to the \(S^1\) norm. This will essentially follow from the perturbation theory developed later on in the key Step 3 of the proof of Proposition \(7.4\). The following proposition can be proved.

**Proposition 5.11.** The sequence \([(\tilde{A}_R, \tilde{\phi}_R)]_{R \geq R_0}\) converges in \(S^1([-T, T] \times \mathbb{R}^4)\) as \(R \to \infty\). The limit is also of class \(H^s_t \times H^{s-1}_x\) for all \(s \geq 1\) on each time slice of \([-T, T] \times \mathbb{R}^4\), hence of class \(C^\infty\), and a smooth solution to the MKG-CG system on \([-T, T] \times \mathbb{R}^4\) with initial data \((A, \phi)[0]\).

**Proof.** A sketch of the proof is given in Subsection \(7.4\). \(\square\)

### 6. How to arrive at the minimal energy blowup solution

In this section we address another delicate issue arising due to the difficulties with the perturbation theory for the MKG-CG system. Assume that \((A_n, \phi_n)\) is an “essentially singular sequence” of admissible solutions to the MKG-CG system that converges at time \(t = 0\) in the energy topology to a Coulomb energy class data pair \((A, \phi)[0]\) with \(E(A, \phi) = E_{\text{crit}}\),

\[
\lim_{n \to \infty} (A_n, \phi_n)[0] = (A, \phi)[0].
\]

Using the concept of MKG-CG evolution for energy class data from Definition \(5.3\), we obtain an energy class solution \((A, \phi)\) with maximal lifespan \(I\). We then want to infer that

\[
\sup_{I \subseteq (J \in \text{closed})} \|(A, \phi)\|_{S^1(I \times \mathbb{R}^4)} = \infty,
\]

while by construction it holds that \(E(A, \phi) = E_{\text{crit}}\). In view of Lemma \(5.4\) it suffices to consider the case \(I = \mathbb{R}\). The problem here is that while we have

\[
\lim_{n \to \infty} \|(A_n, \phi_n)\|_{S^1(I \times \mathbb{R}^4)} = \infty,
\]
where $I_n$ are suitably chosen time intervals, we cannot use an immediate perturbative argument to obtain (6.1) as is possible for wave maps in [10]. The reason comes from the fact that the $(A_n, \phi_n)$ may have non-negligible low-frequency components. Nevertheless, we obtain the following result.

**Proposition 6.1.** Let $(A_n, \phi_n)$ be an essentially singular sequence of admissible solutions to the MKG-CG system. Assume that

$$\lim_{n \to \infty} (A_n, \phi_n)[0] = (A, \phi)[0]$$

in the energy topology for some Coulomb energy class data pair $(A, \phi)[0]$. Let $I$ be the maximal lifespan of the MKG-CG evolution $(A, \phi)$ of this data pair given by Definition [5.3]. Then it holds that

$$\sup_{J \subset I, J \text{closed}} \| (A, \phi) \|_{S^1(J \times \mathbb{R}^4)} = \infty.$$

**Proof.** A sketch of the proof can be found in Subsection [7.4].

We shall later on need certain variations of the preceding proposition.

**Corollary 6.2.** Let $\{(A_n, \phi_n)[0]\}_{n \in \mathbb{N}}$ and $(A, \phi)[0]$ be Coulomb energy class data such that

$$\lim_{n \to \infty} (A_n, \phi_n)[0] = (A, \phi)[0]$$

in the energy topology and let $I$ be the maximal lifespan of the MKG-CG evolution of $(A, \phi)[0]$. If $J \subset I$ is a compact time interval, then it holds that

$$\limsup_{n \to \infty} \| (A_n, \phi_n) \|_{S^1(J \times \mathbb{R}^4)} < \infty.$$

This entails the following important corollary.

**Corollary 6.3.** Let $\{(A_n, \phi_n)[0]\}_{n \in \mathbb{N}} \subset H^1_x \times L^2_x$ be a compact subset of Coulomb energy class data. Then there exists an open interval $I_*$ centered at $t = 0$ with the property that

$$I_* \subset I_n$$

for all $n \in \mathbb{N}$, where $I_n$ denotes the maximal lifespan of the MKG-CG evolutions of $(A_n, \phi_n)[0]$ given by Definition [5.3].

**Proof.** We argue by contradiction. Assume that there exists a subsequence $\{(A_{n_k}, \phi_{n_k})[0]\}_{k \in \mathbb{N}}$ for which at least one of the lifespan endpoints of the associated MKG-CG evolutions converges to $t = 0$. Passing to a further subsequence, which we again denote by $\{(A_{n_k}, \phi_{n_k})[0]\}_{k \in \mathbb{N}}$, we may assume that

$$\lim_{k \to \infty} (A_{n_k}, \phi_{n_k})[0] = (A, \phi)[0]$$

in the energy topology for some Coulomb energy class data $(A, \phi)[0]$. The contradiction now follows from Corollary [6.2].

### 7. Concentration compactness step

#### 7.1. General considerations.

We begin by sorting out the relationship between the conserved energy and the $H^1_x \times L^2_x$-norm of solutions $(A, \phi)$ to the MKG-CG system. Recall that the conserved energy is given by the expression

$$E(A, \phi) = \frac{1}{4} \sum_{\alpha, \beta} \int_{\mathbb{R}^4} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 \, dx + \frac{1}{2} \sum_\alpha \int_{\mathbb{R}^4} |\partial_\alpha \phi + iA_\alpha \phi|^2 \, dx.$$
Using the Coulomb gauge condition, this can be written as
\[
E(A, \phi) = \sum_{i,j} \int_{\mathbb{R}^4} (\partial_i A_j)^2 \, dx + \frac{1}{2} \sum_j \int_{\mathbb{R}^4} (\partial_j A_i)^2 \, dx + \frac{1}{2} \sum_a \int_{\mathbb{R}^4} |\partial_a \phi + iA_a \phi|^2 \, dx,
\]
which immediately implies
\[
E(A, \phi) \leq \sum_i \|\nabla_{t,x} A_i\|^2_{L^2_x} + \|\nabla_x A_0\|^2_{L^2_x} + \sum_a \|\nabla_{t,x} A_a\|^2_{L^2_x} \|\nabla_x \phi\|^2_{L^2_x}.
\]
Conversely, in order to exploit the conserved energy, we also need to show that the expression
\[
\sum_a \|\nabla_{t,x} A_a\|^2_{L^2_x} + \|\nabla_x \phi\|^2_{L^2_x}
\]
is bounded in terms of \(E(A, \phi)\). Here the only issue comes from bounding the terms \(\|\nabla_{t,x} \phi\|_{L^2_x}\) and \(\|\partial_j A_0\|_{L^2}\). However, the diamagnetic inequality gives the pointwise estimate
\[
|\partial_a \phi| \leq (|\partial_a + iA_a| \phi|
\]
and Sobolev’s inequality then yields
\[
|\phi|_{L^4_x} \leq \|\nabla_x \phi\|_{L^2_x} \leq \sum_j |(\partial_j + iA_j) \phi|_{L^2_x}.
\]
Thus, we find
\[
|\partial_a \phi|_{L^2_x} \leq \|\partial_a + iA_a\| \phi|_{L^2_x} + \|A_a \phi|_{L^2_x}
\leq \|\partial_a + iA_a\| \phi|_{L^2_x} + \|A_a \phi|_{L^2_x} + \|\phi|_{L^2_x}^2
\leq E(A, \phi)^\frac{1}{2} + E(A, \phi).
\]
In order to bound the time derivative \(\|\partial_t A_0\|_{L^2}\), we use the compatibility relation
\[
\Delta \partial_t A_0 = - \sum_j \partial_j \text{Im}(\phi D_j \phi)
\]
to obtain
\[
|\partial_t A_0|_{L^2_x} \leq \|\nabla_x \partial_t A_0\|_{L^2_x} \leq \sum_j \|\phi D_j \phi|_{L^2_x} \leq \sum_j \|\phi|_{L^2_x} \|D_j \phi|_{L^2_x} \leq E(A, \phi).
\]

We also recall that the notation \((A, \phi)[0]\) for initial data for the MKG-CG system only refers to the the prescribed data \(A_j[0], \phi[0]\), \(j = 1, \ldots, 4\), for the evolution of the spatial components of the connection form \(A\). The component \(A_0\) is determined via the compatibility relations.

7.2. Setting up the induction on frequency scales. Our final goal will be to show the following.

Let \((A, \phi)[0]\) be admissible Coulomb data. Then the corresponding MKG-CG evolution exists globally in time and denoting its energy by
\[
E = \frac{1}{4} \sum_{a,b} \int_{\mathbb{R}^4} (\partial_a \phi - \partial_b \phi)^2 \, dx + \frac{1}{2} \sum_{a} \int_{\mathbb{R}^4} |\partial_a + iA_a \phi|^2 \, dx,
\]
there exists an increasing function \(K : \mathbb{R}^+ \to \mathbb{R}^+\) such that
\[
\|(A, \phi)[0]\|_{L^2(\mathbb{R}^4)} \leq K(E).
\]
To prove this result we proceed by contradiction. By the small data global well-posedness result [12] we know that the assertion holds for sufficiently small energies. So assume that it does not hold for all energies \(E > 0\). Then the set of exceptional energies has a positive infimum, which
we denote by $E_{\text{crit}}$, and we can find a sequence of admissible data $\{(A^n, \phi^n)[0]\}_{n \in \mathbb{N}}$ with evolutions $\{(A^n, \phi^n)\}_{n \in \mathbb{N}}$ defined on $(-T_0^n, T_0^n) \times \mathbb{R}^4$ such that

$$\lim_{n \to \infty} E(A^n, \phi^n) = E_{\text{crit}},$$

$$\lim_{n \to \infty} \|\{A^n, \phi^n\}\|_{\mathcal{Y}((-T_0^n, T_0^n) \times \mathbb{R}^4)} = +\infty.$$  

We call such a sequence of initial data essentially singular.

We now implement a two step Bahouri-Gérard type procedure. The first step consists in selecting frequency atoms. Here we largely follow the setup of Subsection 9.1 and Subsection 9.2 in [10], which in turn is partially based on Section III.1 of [11]. See [11] for some terminology used below.

**Proposition 7.1.** Let $\{(A^n, \phi^n)[0]\}_{n \in \mathbb{N}}$ be a sequence of admissible data with energy bounded by $E$. Up to passing to a subsequence the following holds. Given $\delta > 0$, there exists an integer $\Lambda = \Lambda(\delta, E) > 0$ and for every $n \in \mathbb{N}$ a decomposition

$$A^n[0] = \sum_{a=1}^{\Lambda} A^a[0] + A^a_{\Lambda}[0],$$

$$\phi^n[0] = \sum_{a=1}^{\Lambda} \phi^a[0] + \phi^a_{\Lambda}[0].$$

For $a = 1, \ldots, \Lambda$, the functions $(A^a, \phi^a)[0]_n$ are $\lambda^a$-oscillatory for a family of pairwise orthogonal frequency scales $(\lambda^a)_n$. The error $(A^a_{\Lambda}, \phi^a_{\Lambda})[0]$ is $\lambda^a$-singular for every $1 \leq a \leq \Lambda$ and satisfies the smallness condition

$$\limsup_{n \to \infty} \|A^a_{\Lambda}[0]\|_{B^1_{2,\infty} \times B^2_{2,\infty}} < \delta, \quad \limsup_{n \to \infty} \|\phi^a_{\Lambda}[0]\|_{B^1_{2,\infty} \times B^2_{2,\infty}} < \delta.$$  

Finally, we have asymptotic decoupling of the energy

$$E(A^n, \phi^n) = \sum_{a=1}^{\Lambda} E(A^a, \phi^a) + E(A^a_{\Lambda}, \phi^a_{\Lambda}) + o(1) \quad \text{as } n \to \infty,$$

where the temporal components $A^a_{0,0}$ are determined via the compatibility relation

$$(\Lambda - |\phi^a|^2)A^a_{0,0} = -\text{Im}(\overline{\phi^a} \partial_t \phi^a),$$

and similarly for $A_{a,0}^n$.

**Proof.** We suppress the notation $[0]$ in the proof. As in Section III.1 of [11], we obtain a decomposition of the data $\{(A^n, \phi^n)\}_{n \in \mathbb{N}}$ into frequency atoms

$$A^n = \sum_{a=1}^{\Lambda} \tilde{A}^a + \tilde{A}_{\Lambda}, \quad \phi^n = \sum_{a=1}^{\Lambda} \tilde{\phi}^a + \tilde{\phi}_{\Lambda},$$

where $(\tilde{A}^a, \tilde{\phi}^a)$ are $\lambda^a$-oscillatory for a family of pairwise orthogonal scales $(\lambda^a)_n$ for $a = 1, \ldots, \Lambda$. The error $(\tilde{A}_{\Lambda}, \tilde{\phi}_{\Lambda})$ is $\lambda^a$-singular for $a = 1, \ldots, \Lambda$ and satisfies the smallness condition

$$\limsup_{n \to \infty} \|\tilde{A}_{\Lambda}[0]\|_{B^1_{2,\infty} \times B^2_{2,\infty}} < \delta, \quad \limsup_{n \to \infty} \|\tilde{\phi}_{\Lambda}[0]\|_{B^1_{2,\infty} \times B^2_{2,\infty}} < \delta.$$  

In order to get a clean separation of the frequency atoms in frequency space, we have to prepare them a bit more, because their decay from the scale $(\lambda^a)^{-1}$ might be arbitrarily slow. To this end let $R_n \to \infty$ be a sequence growing sufficiently slowly such that the intervals $[(\lambda^a)^{-1} R_n, (\lambda^a)^{-1} R_n]$
are mutually disjoint for \( n \) large enough and for different values of \( a \). Then we replace the error \( \tilde{A}^n_{\Lambda} \) by

\[
A^n_{\Lambda} = P_{[\mu R_n^\alpha]}[\mu \alpha R_n^\alpha + \log R_n]_{\Lambda} \tilde{A}^n_{\Lambda} + \sum_{a=1}^{\Lambda} P_{[\mu R_n^\alpha]}[\mu \alpha R_n^\alpha + \log R_n]_{\Lambda} \tilde{A}^{na}_{\Lambda},
\]

where \( \mu^{na} = -\log A^{na} \), and the frequency atoms \( \tilde{A}^{na}_{\Lambda} \) by

\[
A^{na} = P_{[\mu R_n^\alpha]}[\mu \alpha R_n^\alpha + \log R_n]_{\Lambda} \tilde{A}^n_{\Lambda} + \sum_{a'=1}^{\Lambda} A^{na'}
\]

for \( a = 1, \ldots, \Lambda \). In order to remove the dependence on \( \Lambda \) in the new profiles, we may replace \( \Lambda \) by \( \Lambda_n \) with \( \Lambda_n \to \infty \) sufficiently slowly as \( n \to \infty \). Analogously, we define \( \phi^{na} \) and \( \phi^n_{\Lambda} \). This new decomposition

(7.1)

\[
A^n = \sum_{a=1}^{\Lambda} A^{na}, \quad \phi^n = \sum_{a=1}^{\Lambda} \phi^{na} + \phi^n_{\Lambda},
\]

has the same properties as the original one, but that we have now arranged for a sharp separation of the frequency supports of the frequency atoms.

Finally, we turn to the asymptotic decoupling of the energy. Here we recall that the “elliptic components” \( A^0_{\Lambda} \) associated with a frequency atom \( (A^{na}, \phi^{na}) \) are determined via the elliptic compatibility equations. It therefore suffices to show that the decomposition (7.1) (which only refers to the spatial components of the connection form \( A^n \)) implies a similar frequency atom decomposition

(7.2)

\[
A^n_0 = \sum_{a=1}^{\Lambda} A^{na}_0 + A^n_{\Lambda,0} + o_{H_1^1}(1) \quad \text{as } n \to \infty,
\]

where \( A_0^{na} \) is \( \lambda^{na} \)-oscillatory and \( A^n_{\Lambda,0} \) is \( \lambda^{na} \)-singular for each \( a = 1, \ldots, \Lambda \). Then the decoupling of the energy is an immediate consequence of the construction of the frequency atoms. For example, we have the limiting relations

\[
\lim_{n \to \infty} \int_{\mathbb{R}^4} \partial_a \phi^{na} A^{na}_{\alpha} \phi^{na'} dx = 0,
\]

if not all of \( a, a', a'' \) are equal, as well as

\[
\lim_{n \to \infty} \int_{\mathbb{R}^4} A^{na}_{\alpha} \phi^{na} A^{na'}_{\alpha} \phi^{na''} dx = 0,
\]

if not all of \( a, a', a'', a''' \) are equal. It remains to prove the decomposition (7.2). To show this, we first observe that (at fixed time \( t = 0 \))

\[
- \text{Im} (\phi^n_{\alpha} \partial_{\alpha} \phi^n_{\alpha}) = \sum_{a=1}^{\Lambda} - \text{Im} (\phi^{na}_{\alpha} \partial_{\alpha} \phi^{na}_{\alpha}) - \text{Im} (\phi^n_{\alpha} \partial_{\alpha} \phi^n_{\Lambda}) + o_{L^2_{1_4}}(1) \quad \text{as } n \to \infty.
\]

It is then easy to show that

\[
(\Delta - |\phi^n|^2) A^0_{\Lambda} = - \text{Im} (\phi^{na}_{\alpha} \partial_{\alpha} \phi^{na}_{\alpha}) + o_{L^2_{1_4}}(1) \quad \text{as } n \to \infty,
\]

\[
(\Delta - |\phi^n|^2) A^n_{\Lambda,0} = - \text{Im} (\phi^n_{\alpha,0} \partial_{\alpha} \phi^n_{\alpha,0}) + o_{L^2_{1_4}}(1) \quad \text{as } n \to \infty.
\]
This in turn will easily follow once we have shown that each $A_{0}^{n_{a}}$ is $\lambda^{n_{a}}$-oscillatory, while $A_{\Lambda,0}^{n}$ is $\lambda^{n_{a}}$-singular for $a = 1, \ldots, \Lambda$. We demonstrate this for $a = 1$, where we may assume by scaling invariance of these assertions that $\lambda^{n_{1}} = 1$ throughout. We start from the compatibility relation

$$\Delta A_{0}^{n_{1}} - |\phi^{n_{1}}|^{2} A_{0}^{n_{1}} = -\text{Im} (\phi^{n_{1}} \overline{\partial_{t} \phi^{n_{1}}})$$

and distinguish between small and large frequencies.

We begin with the small frequencies. For $R \ll -1$ we write

$$\Delta P_{\leq R} A_{0}^{n_{1}} - P_{\leq R} (|\phi^{n_{1}}|^{2} A_{0}^{n_{1}}) = -P_{\leq R} (\text{Im} (\phi^{n_{1}} \overline{\partial_{t} \phi^{n_{1}}})),$$

where we have

$$\lim_{R \to -\infty} \left\| P_{\leq R} (\text{Im} (\phi^{n_{1}} \overline{\partial_{t} \phi^{n_{1}}})) \right\|_{L_{t}^{\frac{4}{3}}} = 0.$$

Next, we split

$$P_{\leq R} (|\phi^{n_{1}}|^{2} A_{0}^{n_{1}}) = P_{\leq R} (P_{\leq \frac{R}{2}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}}) + P_{\leq R} (P_{> \frac{R}{2}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}}).$$

Then we have

$$\lim_{R \to -\infty} \left\| P_{\leq \frac{R}{2}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}} \right\|_{L_{t}^{\frac{4}{3}}} = 0,$$

whence

$$\lim_{R \to -\infty} \left\| P_{\leq R} (P_{\leq \frac{R}{2}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}}) \right\|_{L_{t}^{\frac{4}{3}}} = 0,$$

while for the second term above, we obtain from Bernstein’s inequality that

$$\left\| P_{\leq R} (P_{> \frac{R}{2}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}}) \right\|_{L_{t}^{\frac{4}{3}}} \leq \sum_{k_{1}=k_{2}+O(1), \ k_{1} > \frac{R}{2}} \left\| P_{\leq R} (P_{k_{1}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}}) \right\|_{L_{t}^{\frac{4}{3}}} \leq 2^{R} \sum_{k_{1}=k_{2}+O(1), \ k_{1} > \frac{R}{2}} \left\| P_{k_{1}} |\phi^{n_{1}}|^{2} \right\|_{L_{t}^{2}} \left\| P_{k_{2}} A_{0}^{n_{1}} \right\|_{L_{t}^{2}} \leq 2^{\frac{R}{2}} \left\| |\phi^{n_{1}}|^{2} \right\|_{L_{t}^{2}} \left\| A_{0}^{n_{1}} \right\|_{H_{t}^{1}}^{2}.$$}

We immediately conclude that

$$\lim_{R \to -\infty} \left\| P_{\leq R} (P_{> \frac{R}{2}} (|\phi^{n_{1}}|^{2}) A_{0}^{n_{1}}) \right\|_{L_{t}^{\frac{4}{3}}} = 0.$$
and thus,
\[
\lim_{R \to \infty} \| P_{> R} (P_{< \frac{R}{2}} (|\phi_n^1|^2) A_0^n) \|_{L^4} = 0.
\]

On the other hand, we have
\[
\| P_{> R} (P_{< \frac{R}{2}} (|\phi_n^1|^2) A_0^n) \|_{L^4} \leq \sum_{k_1 = k_2 + O(1), \ k_1 > R} \| P_{k_1} (P_{< \frac{R}{2}} (|\phi_n^1|^2) P_{k_2} A_0^n) \|_{L^4} \\
\leq \sum_{k_1 = k_2 + O(1), \ k_1 > R} \| P_{\leq \frac{R}{2}} (|\phi_n^1|^2) \|_{L^4} \| P_{k_2} A_0^n \|_{L^2} \\
\leq \sum_{k_2 > R} 2^{-\frac{R}{2} - k_2} \| \phi_n^1 \|_{L^4}^2 \| P_{k_2} A_0^n \|_{H^1} \\
\leq 2^{-\frac{R}{2}} \| \phi_n^1 \|_{L^4}^2 \| A_0^n \|_{H^1}
\]
and so this also converges to zero as \( R \to +\infty \). We then conclude from Sobolev’s inequality that
\[
\lim_{R \to +\infty} \| P_{> R} A_0^n \|_{H^1} = 0.
\]

Given an essentially singular sequence of initial data, by Proposition 7.1 for any \( \delta > 0 \) we obtain another essentially singular sequence \( \{(A^n, \phi^n)[0]\}_{n \in \mathbb{N}} \) of the form
\[
A^n[0] = \sum_{\alpha = 1}^N A^{n\alpha}[0] + A^n_A[0],
\]
(7.3)
\[
\phi^n[0] = \sum_{\alpha = 1}^N \phi^{n\alpha}[0] + \phi^n_A[0]
\]
with
\[
\limsup_{n \to \infty} \| A^n_A[0] \|_{B^1_{2, \infty} \times B^2_{2, \infty}} < \delta, \quad \limsup_{n \to \infty} \| \phi^n_A[0] \|_{B^1_{\infty, \infty} \times B^2_{\infty, \infty}} < \delta.
\]

Eventually, we will prove that necessarily only one frequency atom \( (A^{n\alpha}, \phi^{n\alpha})[0] \) in the decomposition (7.3) is non-trivial and has to be asymptotically of energy \( E^{\text{crit}} \). In fact, the subsequent considerations will show that if there are at least two frequency atoms \( (A^{n1}, \phi^{n1})[0], (A^{n2}, \phi^{n2})[0] \) that both do not vanish asymptotically, or if there is only one frequency atom \( (A^{n1}, \phi^{n1})[0] \) with the error satisfying
\[
\limsup_{n \to \infty} \| (A^n_1, \phi^n_1)[0] \|_{H^1_2 \times L^2_4} > 0,
\]
then we get an a priori bound on the \( S^1 \) norm of the evolutions
\[
\liminf_{n \to \infty} \| (A^n, \phi^n) \|_{S^1((-T^0_0, T^0_1) \times \mathbb{R}^3)} < \infty,
\]
contradicting the assumption that \( \{(A^n, \phi^n)[0]\}_{n \in \mathbb{N}} \) is essentially singular.

We now introduce a smallness parameter \( \epsilon_0 > 0 \) that will eventually be chosen sufficiently small depending only on \( E^{\text{crit}} \). In particular, we assume that \( \epsilon_0 \) is less than the small energy threshold of the small energy global well-posedness result [12].

By passing to a suitable subsequence and by renumbering the frequency atoms, if necessary, we may assume that
\[
\limsup_{n \to \infty} E(A^{n\alpha}, \phi^{n\alpha}) \geq \epsilon_0
\]
only for \( a = 1, \ldots, \Lambda_0 \) for some integer \( \Lambda_0 > 0 \) and that
\[
\sum_{a \geq \Lambda_0 + 1} \limsup_{n \to \infty} E(A^{n_a}, \phi^{n_a}) < \varepsilon_0.
\]
Moreover, we may assume that the frequency atoms \( (A^{n_a}, \phi^{n_a})[0])_{n \in \mathbb{N}}, a = 1, \ldots, \Lambda_0 \), have “increasing frequency supports” in the sense that \((A^{n_a})^{-1}\) is growing in terms of \( a \) (for each fixed \( n \)).

The key idea now is as follows.

We approximate the initial data \((A^a, \phi^a)[0]\) by low frequency truncations, obtained by removing all or some of the atoms \((A^{n_a}, \phi^{n_a})[0], a = 1, \ldots, \Lambda_0\), and inductively obtain bounds on the \(S^1\) norm of the MKG-CG evolutions of the truncated data. As this induction stops after \( \Lambda_0 \) many steps, we will have obtained the desired contradiction, forcing eventually that there has to be exactly one frequency atom \((A^{n_1}, \phi^{n_1})[0]\) that is asymptotically of energy \( E_{\text{crit}} \).

### 7.3. Evolving the “non-atomic” lowest frequency approximation.

From now on we suppress the notation \([0]\) for the initial data. The errors \((A^a_{\Lambda_0}, \phi^a_{\Lambda_0})\) in the decomposition (7.3) are by construction supported away in frequency space from the frequency scales \((A^{n_a})^{-1}, a = 1, 2, \ldots, \Lambda_0\). It is then clear that the errors \((A^a_{\Lambda_0}, \phi^a_{\Lambda_0})\) can be written as the sum of \( \Lambda_0 + 1 \) pieces, which correspond to the \( \Lambda_0 + 1 \) shells that the frequency space gets split into by the frequency supports of the atoms \((A^{n_a}, \phi^{n_a})\). Thus, we can write

\[
A^n_{\Lambda_0} = \sum_{j=1}^{\Lambda_0+1} A^n_{\Lambda_0}^j, \quad \phi^n_{\Lambda_0} = \sum_{j=1}^{\Lambda_0+1} \phi^n_{\Lambda_0}^j,
\]

where the first pieces \((A^n_{\Lambda_0}^1, \phi^n_{\Lambda_0}^1)\) have Fourier support in the region closest to the origin, i.e. in

\[ |\xi| \leq (A^{n_1})^{-1}(R_n)^{-1}. \]

In other words, one essentially obtains the “lowest frequency approximations” \((A^a_{\Lambda_0}, \phi^a_{\Lambda_0})\) by removing all the atoms \((A^{n_a}, \phi^{n_a}), a = 1, \ldots, \Lambda_0, \) from the data.

We then start our grand inductive procedure by showing the following proposition.

**Proposition 7.2.** The parameter \( \varepsilon_0 > 0 \) can be chosen sufficiently small depending only on the size of \( E_{\text{crit}} \) such that the following holds. Constructing the lowest frequency approximations \((A_{\Lambda_0}^1, \phi_{\Lambda_0}^1)(n)\) as described in (7.4), then there exists a constant \( C(E_{\text{crit}}) > 0 \) such that for all sufficiently large \( n \), the data given by \((A_{\Lambda_0}^1, \phi_{\Lambda_0}^1)\) can be evolved globally in time and the corresponding solution satisfies

\[
\|(A_{\Lambda_0}^1, \phi_{\Lambda_0}^1)\|_{S^1(\mathbb{R}^3)} \leq C(E_{\text{crit}}).
\]

**Proof.** (Outline) The idea is to use a finite number of further low frequency approximations of \((A_{\Lambda_0}^1, \phi_{\Lambda_0}^1)(n)\) and to inductively obtain bounds on the \(S^1\) norms of their evolutions. Here it is essential that the number of these further approximations is bounded by \( C_1(E_{\text{crit}}) > 0 \) and thus independent of the choices already made. Picking these low frequency approximations requires a somewhat delicate construction involving a further frequency atom decomposition of \((A_{\Lambda_0}^1, \phi_{\Lambda_0}^1)(n)\). To
begin with, for some sufficiently small $\delta_1 = \delta_1(E_{\text{crit}}) > 0$, in particular $\delta_1 \ll \varepsilon_0$, we use decompositions

$$
A_{n_0}^{n_1} = \sum_{j=1}^{A_1(\delta_1)} A_{n_0}^{n_1(j)} + A_{n_0}(A_1),
$$

$$
\phi_{n_0}^{n_1} = \sum_{j=1}^{A_1(\delta_1)} \phi_{n_0}^{n_1(j)} + \phi_{n_0}(A_1),
$$

where the frequency atoms $(A_{n_0}^{n_1(j)}, \phi_{n_0}^{n_1(j)})$ have frequency support in mutually disjoint intervals

$$
[(\lambda^{n_1(j)} \cdot R_{n_1}^{(j)})^{-1}, (\lambda^{n_1(j)} \cdot R_{n_1}^{(j)})^{-1} R_{n_1}^{(j)}]$$

with $R_{n_1}^{(j)} \to \infty$ as $n \to \infty$, and furthermore, we have the bound

$$
\lim_{n \to \infty} \sup_{l} [||A_{n_0}^{n_1}||_{B_{\infty}^1} \times B_{\infty}^1 + ||\phi_{n_0}^{n_1}||_{B_{\infty}^1} \times B_{\infty}^1] < \delta_1.
$$

We may again assume that the atoms $(A_{n_0}^{n_1(j)}, \phi_{n_0}^{n_1(j)})$ have increasing frequency support as $j$ increases. The number of frequency atoms $A_1(\delta_1)$ here is potentially extremely large. It is crucial that the number of steps, i.e. the number of low frequency approximations of $(A_{n_0}^{n_1}, \phi_{n_0}^{n_1})$, required in the inductive procedure is in fact much smaller, of size $C_1 = C_1(E_{\text{crit}}) \ll A_1(\delta)$. As we shall see, $C_1$ can be chosen independently of $\delta_1$ and $A_1(\delta_1)$. We now pick the low frequency approximations of the data $(A_{n_0}^{n_1}, \phi_{n_0}^{n_1})$. For $\varepsilon_0$ fixed as before, we inductively construct $O(\frac{E_{\text{crit}}}{\varepsilon_0})$ closed frequency intervals $\tilde{J}_l$ for the variable $|g|$, disjoint up the the endpoints and increasing. The chosen intervals will also depend on $n$, but for notational ease we do not indicate this. So consider $n$ and $A_1$ fixed now. Having picked the intervals $\tilde{J}_1 = (-\infty, b_1]$, $\tilde{J}_l = [a_l, b_l]$ with $b_{l-1} = a_l$ for $l = 2, \ldots, L - 1$, we pick an interval $[a_L, b_L]$ with $a_L = b_L - 1$ as follows. First, pick $\tilde{b}_L$ in such a fashion that

$$
E(P_{[a_L, b_L]}A_{n_0}^{n_1}, P_{[a_L, b_L]}\phi_{n_0}^{n_1}) = \varepsilon_0
$$

or else, if this is impossible, then pick $\tilde{b}_L = \log (\lambda^{n_1})^{-1} - \log R_n$, i.e. pick the upper endpoint of the frequency interval containing the lowest frequency “large atom” $(A_{n_0}^{n_1}, \phi_{n_0}^{n_1})$. Now, in the former case assume that

$$
\tilde{b}_L \in [\log (\lambda^{n_1(j)} \cdot R_{n_1}^{(j)})^{-1} \log (\lambda^{n_1(j)} \cdot R_{n_1}^{(j)})^{-1} + \log R_{n_1}^{(j)}]
$$

for some $1 \leq j \leq A_1(\delta_1)$, i.e. $\tilde{b}_L$ falls within the frequency support of one of the (finite number of) “small frequency atoms” $(A_{n_0}^{n_1(j)}, \phi_{n_0}^{n_1(j)})$ constituting $(A_{n_0}^{n_1}, \phi_{n_0}^{n_1})$. Then we shift $\tilde{b}_L$ upwards to coincide with the upper limit, that is, we set

$$
b_L = \log (\lambda^{n_1(j)})^{-1} + \log R_{n_1}^{(j)}.
$$

Otherwise, we set

$$
b_L = \tilde{b}_L.
$$

Then we define the interval $\tilde{J}_L = [a_L, b_L]$. Observe that for sufficiently large $n$, we have

$$
E(P_{J_l}A_{n_0}^{n_1}, P_{J_l}\phi_{n_0}^{n_1}) \leq \varepsilon_0.
$$

In particular, this implies that for sufficiently large $n$ the total number of intervals $\tilde{J}_l$ is $C_1 = O(\frac{E_{\text{crit}}}{\varepsilon_0})$. We now define the low frequency approximations of the data $(A_{n_0}^{n_1}, \phi_{n_0}^{n_1})$ by truncating the frequency support of $(A_{n_0}^{n_1}, \phi_{n_0}^{n_1})$ to the intervals

$$
J_L := \cup_{l=1}^{J_l} \tilde{J}_l.
More precisely, for $1 \leq L \leq C_1$ we define the $L$-th low frequency approximation of the data $(A_{A_0}^{n_1}, \phi_{A_0}^{n_1})$ by the expression
\[
(P_L A_{A_0}^{n_1}, P_L \phi_{A_0}^{n_1}),
\]
where by construction $C_1 = C_1(E_{crit}) \leq \frac{E_{crit}}{z_0}$. In particular, we have
\[
(P_L c_1 A_{A_0}^{n_1}, P_L c_1 \phi_{A_0}^{n_1}) = (A_{A_0}^{n_1}, \phi_{A_0}^{n_1}).
\]
We also state the following key lemma, whose proof is a consequence of the preceding construction.

**Lemma 7.3.** For $L = 1, \ldots, C_1$ and for any $R > 0$, we have for all sufficiently large $n$ that
\[
\|P_{(a_L - R, a_L + R)} \nabla_{t,x} A_{A_0}^{n_1}\|_{L^2_z} + \|P_{(a_L - R, a_L + R)} \nabla_{t,x} \phi_{A_0}^{n_1}\|_{L^2_z} \leq R \delta_1.
\]

In order to prove Proposition 7.2 we inductively show that for $L = 1, \ldots, C_1$ and for all sufficiently large $n$, the evolutions of the data
\[
(P_L A_{A_0}^{n_1}, P_L \phi_{A_0}^{n_1})
\]
evolve globally and satisfy the desired global $S^1$ norm bounds, which of course get larger as $L$ grows. For $L = 1$ this is a direct consequence of the small energy theory. The main work now goes into proving the following proposition.

**Proposition 7.4.** Let us assume that the evolution of the data
\[
(P_{L-1} A_{A_0}^{n_1}, P_{L-1} \phi_{A_0}^{n_1})
\]
is globally defined for some $1 \leq L < C_1$. We denote this evolution by $(A_{A_0}^{n_1,(L-1)}, \phi_{A_0}^{n_1,(L-1)})$. Furthermore, assume that for all sufficiently large $n$, it holds that
\[
\|(A_{A_0}^{n_1,(L-1)}, \phi_{A_0}^{n_1,(L-1)})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq C_2 < \infty.
\]
Provided $\delta_1^{-1} \gg C_2$ with $\delta_1 > 0$ as above, there exists $C_3 = C_3(C_2) < \infty$ such that for all sufficiently large $n$, the data
\[
(P_{L} A_{A_0}^{n_1}, P_{L} \phi_{A_0}^{n_1})
\]
can be evolved globally and for the corresponding evolutions $(A_{A_0}^{n_1,(L)}, \phi_{A_0}^{n_1,(L)})$, it holds that
\[
\|(A_{A_0}^{n_1,(L)}, \phi_{A_0}^{n_1,(L)})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq C_3.
\]

Proposition 7.2 is then an immediate consequence of applying Proposition 7.4 $C_1$ many times. We note that there exists $\delta_{11} > 0$ depending only on $E_{crit}$ such that choosing $\delta_1 < \delta_{11}$ in each step, Proposition 7.4 can be applied. Since $C_1 = C_1(E_{crit})$ this results in a bound
\[
\|(A_{A_0}^{n_1}, \phi_{A_0}^{n_1})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq C(E_{crit}).
\]

**Proof of Proposition 7.4** We proceed in several steps.

**Step 1.** The assumed bound on $\|(A_{A_0}^{n_1,(L-1)}, \phi_{A_0}^{n_1,(L-1)})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}$ implies an exponential decay for large frequencies,
\[
\|(P_k A_{A_0}^{n_1,(L-1)}, P_k \phi_{A_0}^{n_1,(L-1)})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq 2^{-\sigma(k-b_{L-1})} \text{ for } k \geq b_{L-1}.
\]
This will follow once we can show that in fact
\[
\|(P_k A_{A_0}^{n_1,(L-1)}, P_k \phi_{A_0}^{n_1,(L-1)})\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq \xi_k^{(L-1)}.
\]
where \( |c_k^{(L-1)}|_{k \in \mathbb{Z}} \) is a sufficiently flat frequency envelope covering the initial data \((A^{n,k}_{n_0}, \phi^{n,k}_{n_0})[0]\) at time \( t = 0 \). This in turn is a consequence of Proposition 4.2, whose proof will be given in Subsection 7.4. 

**Step 2. Localizing \((A^{n,k}_{n_0}, \phi^{n,k}_{n_0})\) to suitable space-time slices.** In order to ensure that we can induct on perturbations of size \( \sim \varepsilon_0 \) that are not “too small” (such as the \( \delta_1 \)), we have to make sure that the \( S^1 \) norms of \((A^{n,k}_{n_0}, \phi^{n,k}_{n_0})\) are not too large. To simplify the notation, we label these components by \((A, \phi)\) for the rest of this step. The idea is to localize to suitable space-time slices \( I \times \mathbb{R}^4 \), whose number may be very large (depending on \( \|(A, \phi)\|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \) and \( E_{crit} \)), but such that we have on each slice 

\[
\|(A, \phi)\|_{S^1(I \times \mathbb{R}^4)} \leq C(E_{crit}),
\]

where the function \( C(\cdot) \) grows at most polynomially.

**Proposition 7.5.** There exist \( N = N(\|(A, \phi)\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}, E_{crit}) \) many time intervals \( I_1, \ldots, I_N \) partitioning the time axis \( \mathbb{R} \) such that we have for \( n = 1, \ldots, N \) a decomposition (referring to the spatial components of the connection form simply by \( A \))

\[
A|_{I_n} = A^{free,(I_n)} + A^{nonlin,(I_n)}, \quad \Box A^{free,(I_n)} = 0,
\]

where \( A^{free,(I_n)} \) and \( A^{nonlin,(I_n)} \) are in Coulomb gauge and satisfy

\[
\|\nabla_t A^{free,(I_n)}\|_{L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq E_{crit}^{1/2},
\]

\[
\|A^{nonlin,(I_n)}\|_{L_4^4 S^1(I_n \times \mathbb{R}^4)} \ll 1.
\]

Moreover, we have for \( n = 1, \ldots, N \) that

\[
\|\phi\|_{S^1(I_n \times \mathbb{R}^4)} \leq C(E_{crit}),
\]

where \( C(\cdot) \) grows at most polynomially.

**Proof.** We first define precisely the decompositions \( A = A^{free} + A^{nonlin} \) that we are using. The nonlinear structure inherent in \( A^{nonlin} \) will be pivotal for controlling the equation for \( \phi \). For a time interval \( I \subset \mathbb{R} \), say of the form \( I = [t_0, t_1] \) for some \( t_0 < t_1 \), we define for \( i = 1, \ldots, 4 \),

\[
A^{nonlin,(I)} = \chi_I \sum_{j,k} \Box^{-1} P_k Q_j \text{Im} \mathcal{P}_j((\chi_I \phi) \cdot \nabla_s (\chi_I \phi)) - \chi_I iA|\phi|^2,
\]

where \( \chi_I \) is a smooth cutoff to the interval \( I \) and \( \Box^{-1} \) denotes multiplication by the Fourier symbol. Then we define \( A^{free(I)} \) to be the free wave with initial data at time \( t_0 \) given by \( A(t_0) = A^{nonlin,(I)}[t_0] \).

By construction, we then have

\[
A = A^{free,(I)} + A^{nonlin,(I)} \text{ on } I \times \mathbb{R}^4.
\]

We now describe how to partition the time axis into \( N = N(\|(A, \phi)\|_{S^1}, E_{crit}) \) many suitable time intervals so that the bounds (7.6) - (7.8) hold on each such interval. For this, we first need the following technical lemma.

**Lemma 7.6.** Given \( \varepsilon > 0 \), there exist \( M = M(\|(A, \phi)\|_{S^1(\mathbb{R} \times \mathbb{R}^4)}, \varepsilon) \) many time intervals \( I_1, \ldots, I_M \) partitioning the time axis \( \mathbb{R} \) such that for \( m = 1, \ldots, M \) and \( i = 1, \ldots, 4 \),

\[
\sum_k \left\| \nabla_{tx} \sum_j \Box^{-1} P_k Q_j \mathcal{P}_j((\chi_{I_m} \phi) \cdot \nabla_s (\chi_{I_m} \phi)) - \chi_{I_m} iA|\phi|^2 \right\|_{L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq \varepsilon
\]

and

\[
\left\| \mathcal{P}_j((\chi_{I_m} \phi) \cdot \nabla_s (\chi_{I_m} \phi)) \right\|_{L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^4)} \leq \varepsilon.
\]
In particular, it then holds that

\begin{align}
(7.12) & \quad \| \nabla_{t,x} A_i^{free,(I_m)} \|_{L^p_t L^q_{t,x}(\mathbb{R})} \leq L^{1/2} \epsilon, \\
(7.13) & \quad \| \nabla_{t,x} A_{i}^{nonlin,(I_m)} \|_{L^p_t L^q_{t,x}(\mathbb{R})} \leq \epsilon, \\
(7.14) & \quad \| A_{i}^{nonlin,(I_m)} \|_{L^1_t L^q_{t,x}(\mathbb{R})} \leq \epsilon.
\end{align}

**Proof.** We begin with the quadratic interaction term in (7.10) and show that the time axis \( \mathbb{R} \) can be partitioned into \( M_1 = M_1(\| (A, \phi) \|_{S^i}, \epsilon) \) many intervals so that on each such interval \( I \), it holds that

\begin{equation}
(7.15) \quad \sum_k \left\| \nabla_{t,x} \sum_j \Box^{-1} P_k Q_j (\nabla_x \chi_I (\phi) \cdot \nabla_x \chi_I (\phi)) \right\|_{L^p_t L^q_{t,x}(\mathbb{R}^d)} \leq \epsilon.
\end{equation}

To this end we exploit that there is an inherent null form in the above expression

\[ \mathcal{P}_j(\phi \cdot \nabla_x \phi) = \Delta^{-1} \nabla^\prime \mathcal{N}_{ij}(\phi, \phi), \]

where

\[ \mathcal{N}_{ij}(\phi, \phi) = (\partial_i \phi)(\partial_j \psi) - (\partial_i \phi)(\partial_i \psi). \]

We first prove that on suitable intervals \( I \),

\begin{equation}
(7.16) \quad \sum_k \sum_{j \leq k+C} \left\| \nabla_{t,x} \Box^{-1} P_k Q_j \Delta^{-1} \nabla^\prime \mathcal{N}_{ij}(Q_{j \leq j-C}(\chi_I \phi), Q_{j \leq j-C}(\chi_I \phi)) \right\|_{L^p_t L^q_{t,x}(\mathbb{R}^d)} \leq \epsilon.
\end{equation}

By a Littlewood-Paley trichotomy we may reduce to the case where both inputs are at frequency \( \sim 2^k \). The singular operator \( \Box^{-1} \) costs \( 2^{-j-k} \), so we need to recover the factor \( 2^{-j} \). From the null form we gain \( 2^{\frac{j-k}{2}} \), while the inclusion \( Q_j L^p_t L^q_x \hookrightarrow L^S_{t,x} \) gains another \( 2^S \). Finally, we obtain a small power in \( j-k \) using the improved Bernstein estimate \( P_k Q_j L^p_t L^q_x \hookrightarrow L^S_{t,x} \) and that \( L^p_t L^q_x \hookrightarrow L^S_{t,x} \). Thus, we find

\begin{equation}
\sum_k \sum_{j \leq k+C} \left\| \nabla_{t,x} \Box^{-1} P_k Q_j \Delta^{-1} \nabla^\prime \mathcal{N}_{ij}(Q_{j \leq j-C}(\chi_I \phi), Q_{j \leq j-C}(\chi_I \phi)) \right\|_{L^p_t L^q_{t,x}} \leq \left( \sum_k \left( 2^{-\frac{j-k}{2}} \right) \| \chi_I \nabla_x \phi \|_{L^2_t L^2_x} \right)^{1/2} \| \phi \|_{S^i}
\end{equation}

and smallness follows from divisibility of the \( L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^d) \) norm. Next, we show that on suitable intervals \( I \), it holds that

\begin{equation}
(7.17) \quad \sum_k \sum_{j \leq k+C} \left\| \nabla_{t,x} \Box^{-1} P_k Q_j \Delta^{-1} \nabla^\prime \mathcal{N}_{ij}(\chi_I \phi, \chi_I \phi) \right\|_{L^p_t L^q_{t,x}(\mathbb{R}^d)} \leq \epsilon.
\end{equation}

By a Littlewood-Paley trichotomy we may again reduce to the case where both inputs are at frequency \( \sim 2^k \). Then we obtain, using the Bernstein inequality both in time and space, that

\begin{equation}
\sum_k \sum_{j \leq k+C} \left\| \nabla_{t,x} \Box^{-1} P_k Q_j \Delta^{-1} \nabla^\prime \mathcal{N}_{ij}(\chi_I \phi, \chi_I \phi) \right\|_{L^p_t L^q_{t,x}} \leq \left( \sum_k \left( 2^{-\frac{j-k}{2}} \right) \| \chi_I \nabla_x \phi \|_{L^2_t L^2_x} \right)^{1/2} \| \phi \|_{S^i}
\end{equation}

and smallness follows from the divisibility of the \( L^p_t L^q_x(\mathbb{R} \times \mathbb{R}^d) \) norm. In view of (7.16) and (7.17), in order to finish the proof of (7.15) we may assume that one of the two inputs has the leading modulation. It therefore suffices to show that on suitable intervals \( I \) we have bounds of the form

\begin{equation}
(7.18) \quad \sum_{k,j} \left\| \nabla_{t,x} \Box^{-1} P_k Q_{j \leq j-C} \Delta^{-1} \nabla^\prime \mathcal{N}_{ij}(Q_j(\chi_I \phi), Q_{j \leq j-C}(\chi_I \phi)) \right\|_{L^p_t L^q_{t,x}(\mathbb{R}^d)} \leq \epsilon.
\end{equation}
where we use the convention $\Box^{-1} P_t Q_{\leq j-C} = \sum_{l\leq j-C} \Box^{-1} P_t Q_l$. Using that

$$\Box^{-1} P_k Q_{\leq j-C} F(t, \cdot) = - \int_t^\infty \frac{\sin((t-s)\|\nabla\|)}{|\nabla|} (P_k Q_{\leq j-C} F)(s, \cdot) \, ds,$$

it is enough to show

$$\sum_{k,j} \left| P_k Q_{\leq j-C} \Box^{-1} \nabla' N_l (Q_j (\chi_l \phi), Q_{\leq j-C} (\chi_l \phi)) \right|_{L^1_t L^2_x} \lesssim \varepsilon.$$

By estimate (143) in [12] we may reduce to the case where $j = k + O(1)$ and both inputs are at frequency $\sim 2^k$. Then we find

$$\sum_k \left| P_k Q_{\leq k-C} \Box^{-1} \nabla' N_l (Q_k (\chi_k \phi_k), Q_{\leq k-C} (\chi_k \phi_k)) \right|_{L^1_t L^2_x} \lesssim \sum_k \left| \chi_k \nabla_x \phi_k \right|_{X_{6,\infty}^1} \lesssim \| \phi \|_{S^1} \left( \sum_k \| \chi_k \nabla_x \phi_k \|_{L^2_t L^6_x}^2 \right)^{1/2}$$

and smallness follows by divisibility. Next, we consider the cubic term in (7.10). Here we have to prove that on suitable intervals $I$ it holds that

$$(7.19) \quad \sum_k \left| \nabla_{tx} \sum_j \Box^{-1} P_k Q_j (\chi_l A |\phi|^2) \right|_{L^\infty_t L^2_x} \lesssim \varepsilon.$$  

By similar arguments as above, this reduces to showing

$$\sum_k \left| P_k (\chi_l A |\phi|^2) \right|_{L^1_t L^2_x} \lesssim \varepsilon,$$

which follows from estimate (64) in [12] and a divisibility argument. We note that the bound (7.10) implies that the estimates (7.12) and (7.13) hold on each such interval $I$.

It remains to choose the intervals so that the bound (7.11) also holds. The energy estimate for the $S_k$ and $N_k$ spaces together with the bounds (7.10) and (7.11) then also imply the bound (7.14). We pick $M_1(\| |A, \phi| \|_S, \varepsilon)$ many time intervals $I_m$, $m = 1, \ldots, M_1$, on which the bound (7.10) already holds. We show that, if necessary, each time interval $I_m$ can be subdivided into $M_2 = M_2(\| |A, \phi| \|_S, \varepsilon)$ many intervals $I_{ma}$, $a = 1, \ldots, M_2$, such that we have

$$\left| \mathcal{P}_i( (\chi_{I_{ma}} \phi) \cdot \nabla_{\lambda} (\chi_{I_{ma}} \phi) - \chi_{I_{ma}} |A| |\phi|^2) \right|_{L^\infty(\ell^1 N \cap \ell^1 L^2_{H^1 \frac{1}{2}})} \lesssim \varepsilon.$$  

For the rest of the proof of (7.20) we denote an interval $I_{ma}$ just by $I$ and say that it is of the form $I = [t_0, t_1]$ for some $t_0 < t_1$. We only outline how to make the left hand side of (7.20) small in $\ell^1 N$ for suitable intervals $I$, the $\ell^1 L^2_{H^1 \frac{1}{2}}$ component being easier. We first estimate the quadratic interaction term in (7.20),

$$\sum_k \left| \mathcal{P}_i( (\chi_l \phi) \cdot \nabla_{\lambda} (\chi_l \phi) ) \right|_{N_k} = \sum_k \left| P_k \Delta^{-1} \nabla' N_l (\chi_l \phi, \chi_l \phi) \right|_{N_k}.$$  

By (131) in [12], it suffices to consider the case where both inputs are at frequency $\sim 2^k$ and have angular separation $\sim 1$ in Fourier space,

$$\sum_k \left| P_k \Delta^{-1} \nabla' N_l (\chi_l \phi_k, \chi_l \phi_k) \right|_{N_k}.$$  

Here, the prime indicates the angular separation. We split into high and low modulation output.

\[
\sum_k \| P_k \Delta^{-1} \nabla' N_{ir}(\chi_I \phi_k, \chi_I \bar{\phi}_k) \|_{N_k} \leq \sum_k \| P_k Q_{> k-C} \Delta^{-1} \nabla' N_{ir}(\chi_I \phi_k, \chi_I \bar{\phi}_k) \|_{N_k} + \sum_k \| P_k Q_{\leq k-C} \Delta^{-1} \nabla' N_{ir}(\chi_I \phi_k, \chi_I \bar{\phi}_k) \|_{N_k}.
\]

The term with high modulation output is estimated by

\[
\sum_k \| P_k Q_{> k-C} \Delta^{-1} \nabla' N_{ir}(\chi_I \phi_k, \chi_I \bar{\phi}_k) \|_{N_k} \leq \sum_k \| 2^{-\frac{k}{2}} \chi_I \nabla \phi_k \|_{L^2 L^2_t L^2_{x}} \| \chi_I \nabla \phi_k \|_{L^\infty L^2_t}
\]

\[
\leq \left( \sum_k \left( 2^{-\frac{k}{2}} \| \chi_I \nabla \phi_k \|_{L^2 L^2_t} \right)^2 \right)^{1/2} \| \phi \|_{S^1}
\]

and can be made small on suitable intervals \( I \) using the divisibility of the quantity

\[
\sum_k \left( 2^{-\frac{k}{2}} \| \nabla \phi_k \|_{L^2 L^2_t (\mathbb{R} \times \mathbb{R}^4)} \right)^2 \leq \| \phi \|_{S^1_{(\mathbb{R} \times \mathbb{R}^4)}}^2.
\]

For the term with low modulation output we note that the angular separation of the inputs allows us to write schematically

\[
P_k Q_{\leq k-C} \Delta^{-1} \nabla' N_{ir}(\chi_I \phi_k, \chi_I \bar{\phi}_k) = P_k Q_{\leq k-C} \Delta^{-1} \nabla' N_{ir}(Q_{> k-C} (\chi_I \phi_k), \chi_I \bar{\phi}_k) + P_k Q_{\leq k-C} \Delta^{-1} \nabla' N_{ir}(Q_{\leq k-C} (\chi_I \phi_k), Q_{> k-C} (\chi_I \bar{\phi}_k)).
\]

Then we estimate

\[
\sum_k \| P_k Q_{\leq k-C} \Delta^{-1} \nabla' N_{ir}(Q_{> k-C} (\chi_I \phi_k), \chi_I \bar{\phi}_k) \|_{N_k} \leq \sum_k 2^{2k} \| Q_{> k-C} (\chi_I \nabla \phi_k) \|_{L^2 L^2_t} \| \chi_I \nabla \phi_k \|_{L^2 L^2_t}
\]

\[
\leq \| \phi \|_{S^1} \left( \sum_k \left( 2^{-\frac{k}{2}} \| \chi_I \nabla \phi_k \|_{L^2 L^2_t} \right)^2 \right)^{1/2}
\]

and similarly for the other term. Smallness follows as before by divisibility. The cubic interaction term in (7.20) is much simpler to treat, it can be made small on suitable intervals \( I \) using estimate (64) from [12] and divisibility of the \( L^2 W_{k,0} \) norm.

It remains to prove that the bound (7.8) in the statement of Proposition 7.5 holds. For \( \epsilon > 0 \) to be fixed sufficiently small further below, depending only on the size of \( \| (A, \phi) \|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \) and \( E_{crit} \), there exist \( M(\| (A, \phi) \|_{S^1(\mathbb{R} \times \mathbb{R}^4)}, \epsilon) \) many intervals \( I_m, m = 1, \ldots, M \), partitioning the time axis \( \mathbb{R} \) on which the conclusion of Lemma 7.6 holds. We pick such an interval \( I_m \) and now show that, if necessary, it can be subdivided into \( M_3(\| (A, \phi) \|_{S^1(\mathbb{R} \times \mathbb{R}^4)}, E_{crit}) \) many intervals \( I_{ma}, a = 1, \ldots, M_3 \), such that

\[
\| \phi \|_{S^1(\mathbb{R} \times \mathbb{R}^4)} \leq C(E_{crit}),
\]

where \( C(\cdot) \) grows at most polynomially. Upon renumbering the intervals \( I_{ma} \), we will then have finished the proof of Proposition 7.5.

For the remainder of the proof, we denote an interval \( I_{ma} \) just by \( I \) and assume that it is of the form \( I = [t_0, t_1] \) for some \( t_0 < t_1 \). From the equation \( \Box_A \phi = 0 \) and the decomposition \( A_I = A_{free,I} + A_{nonlin,I} \) provided by Lemma 7.6, we conclude that on \( I \times \mathbb{R}^4 \) it holds that

\[
\Box_{A_{free,I}} \phi = -2i \sum_k (P_{> k-C} A_{j}^{free,I} P_{k} \partial_j^I \phi - 2i A_{j}^{nonlin,I} \partial_j^I \phi + 2i A_0 \partial_I \phi + i(\partial_I A_0) \phi + A_0 A^\sharp) \phi + M_1 + M_2,
\]

where \( M_1 \) and \( M_2 \) are the compactness terms.
where we use the notation
\[ \Box_{A_{\text{free}, I}}^P \phi = \Box \phi + 2i \sum_k (P_{k < c} A^\text{free, (I)}_j) P_k \phi, \]
\[ M_1 = -2i \sum_k (P_{k > c} A^\text{free, (I)}_j) P_k \phi - 2i A^\text{nlin, (I)}_j \phi + 2i A_0 \phi, \]
\[ M_2 = i(\partial_t A_0) \phi + A_0 A^\alpha \phi. \]

We further split the term \( M_1 \) into
\[ M_1 \equiv \sum_k N(P_{k < c} A^\text{free, (I)}, P_k \phi) + N(A^\text{nlin, (I)}, \phi) + N_0(A_0, \phi). \]

Since \( A^\text{free, (I)} \) and \( A^\text{nlin, (I)} \) are in Coulomb gauge, we observe that the terms \( N(P_{k < c} A^\text{free, (I)}, P_k \phi) \) and \( N(A^\text{nlin, (I)}, \phi) \) exhibit a null structure,
\[ N(P_{k < c} A^\text{free, (I)}, P_k \phi) = -2i \sum_{j \neq \rho} N_{j\rho}(\Delta^{-1} \nabla_j P_{k < c} A^\text{free, (I)}_\rho, P_k \phi), \]
\[ N(A^\text{nlin, (I)}, \phi) = -2i \sum_{j \neq \rho} N_{j\rho}(\Delta^{-1} \nabla_j A^\text{nlin, (I)}_\rho, \phi). \]

We emphasize that the right hand side of (7.22) is defined on the whole space-time, but which only coincides with \( \Box_{A_{\text{free}, I}}^P \phi \) on \( \mathbb{I} \times \mathbb{R}^4 \). Using the linear estimate (3.3) for the magnetic wave operator \( \Box_{A_{\text{free}, I}}^P \) and working with suitable Schwartz extensions, we obtain that
\[ \|\phi\|_{L^1(\mathbb{I} \times \mathbb{R}^4)} \leq \|\nabla_{t,x} \phi(t_0)\|_{L_2^r} + \|\chi_I(M_1 + M_2)\|_{N(\ell^1 L_2^r H^1_n, \mathbb{R}^4)} \]
\[ \leq E_{\text{crit}} + \|\chi_I(M_1 + M_2)\|_{N(\ell^1 L_2^r H^1_n, \mathbb{R}^4)}. \]

We note that by Theorem 3.1, the implicit constant in the above estimate for the magnetic wave operator depends polynomially on \( \|\nabla_{t,x} A^\text{free, (I)}\|_{L_t^\infty L_x^2} \) and we have \( \|\nabla_{t,x} A^\text{free, (I)}\|_{L_t^\infty L_x^2} \leq E^{1/2}_{\text{crit}} \) by Lemma 7.6. In order to prove the bound (7.23), it therefore suffices to show that we can choose the intervals \( I \) such that
\[ \|M_1 + M_2\|_{N(\ell^1 L_2^r H^1_n, \mathbb{R}^4)} \leq E_{\text{crit}}. \]

Our general strategy to achieve this consists in first using the off-diagonal decay in the multilinear estimates from [12] to reduce to a situation in which a suitable divisibility argument works.

We only outline how to obtain smallness of the term \( M_1 \) in \( N(\mathbb{I} \times \mathbb{R}^4) \), the estimate of \( M_1 \) in \( \ell^1 L_2^r H^1_n \) and of \( M_2 \) in \( N \cap \ell^1 L_2^r H^1_n \) being easier. We begin with the first term in the definition of \( M_1 \),
\[ \|\sum_k N(P_{k < c} A^\text{free, (I)}, P_k \phi)\|_{N(\mathbb{I} \times \mathbb{R}^4)}. \]

From the estimate (131) in [12], we conclude that it suffices to bound the expression
\[ \sum_{k_i} \|P_{k_i} N(P_{k_2} A^\text{free, (I)}, P_{k_1} \phi)\|_{N_{k_1}^i(\mathbb{I} \times \mathbb{R}^4)}^2, \]
where \( k_2 = k_3 = k_1 + O(1) \) and both inputs have angular separation \( \sim 1 \). Similarly to the estimate of (7.21), we bound this term by
\[ \left( \sum_k (2^{-\frac{k}{2}} \|\chi_I P_k \nabla_x A^\text{free, (I)}\|_{L_2^r L_2^r})^2 \right) \|\phi\|_{S^1}^2, \]
and a divisibility argument then yields smallness. To deal with the other two terms in $M_1$, we need to achieve
\[
\left\| N(A_{nonlin}(I), \phi) + N_0(A_0, \phi) \right\|_{N(I \times \mathbb{R}^+)} \lesssim E_{crit}
\]on suitable intervals $I$. To this end we will make similar reductions as in Section 4 of [12], peeling off the “good parts” of $N(A_{nonlin}(I), \phi)$ and of $N_0(A_0, \phi)$ until we are left with three quadrilinear null form bounds.

We introduce the expressions
\[
N^{lowhi}(A_{nonlin}(I), \phi) = \sum_k N(P_{\leq k} A_{nonlin}(I), P_k \phi)
\]
and
\[
\mathcal{H}^* N^{lowhi}(A_{nonlin}(I), \phi) = \sum_k \sum_{k' \leq C} \sum_{j \leq j' + C} Q_{\leq j-C} N(Q_{j} P_{k'} A_{nonlin}(I), Q_{\leq j-C} P_k \phi).
\]

By estimate (53) in [12], we have
\[
(7.24) \quad \left\| N(A_{nonlin}(I), \phi) - N^{lowhi}(A_{nonlin}(I), \phi) \right\|_N \lesssim \left\| A_{nonlin}(I) \right\|_{\leq 1} \| \phi \|_{\leq 1}
\]
and by estimate (54) in [12], it holds that
\[
(7.25) \quad \left\| N^{lowhi}(A_{nonlin}(I), \phi) - \mathcal{H}^* N^{lowhi}(A_{nonlin}(I), \phi) \right\|_N \lesssim \left\| A_{nonlin}(I) \right\|_{\leq 1} \| \phi \|_{\leq 1}.
\]

Fixing $\varepsilon > 0$ sufficiently small, depending only on the size of $\| (A, \phi) \|_{S^1}$ and $E_{crit}$, Lemma 7.6 ensures that $\| A_{nonlin}(I) \|_{\leq 1} S^1$ is small enough so that the right hand sides of (7.24) and (7.25) are bounded by $E_{crit}$. We now define
\[
\mathcal{H}_{A_1}^{nonlin}(I) := -\chi I \sum_{k \leq \min(k_1, k_2) - C} \sum_{j \geq j + C} \Box^{-1} P_k Q_j \text{Im } \mathcal{P}_i(Q_{\leq -C} (\chi I \phi_{k_1}) \cdot \nabla x Q_{\leq -C} (\chi I \phi_{k_2})).
\]

By estimate (55) in [12] it holds that
\[
\left\| \mathcal{H}^* N^{lowhi}(A_{nonlin}(I) - \mathcal{H}_{A_1}^{nonlin}(I), \phi) \right\|_N \lesssim \left\| A_{nonlin}(I) - \mathcal{H}_{A_1}^{nonlin}(I) \right\|_{\leq 1} \| \phi \|_{\leq 1},
\]
so we have to make $\| A_{nonlin}(I) - \mathcal{H}_{A_1}^{nonlin}(I) \|_{Z}$ small. We recall the definition of the $Z$ space,
\[
\| \phi \|_Z = \sum_k \| P_k \phi \|_{Z_k}, \quad \| \phi \|_{Z_k}^2 = \sup_{j \in C} \sum_{\omega_j} 2^j \| P_{\omega_j} Q_{k+2j} \phi \|_{L_x^2 L_t^\infty}^2.
\]

Using estimate (134) in [12] and that one obtains an extra gain for very negative $l$ when estimating in the $Z$ space, we are reduced to bounding
\[
\sum_k \Box^{-1} P_k Q_{k+l(1)} \Delta^{-1} \nabla x N(\chi I \phi_k, \chi I \overline{\phi_k}).
\]

We easily find that
\[
\sum_k \| \Box^{-1} P_k Q_{k+l(1)} \Delta^{-1} \nabla x N(\chi I \phi_k, \chi I \overline{\phi_k}) \|_{L_x^2 L_t^\infty} \lesssim \left( \sum_k \| \chi I \nabla x \phi_k \|_{L_x^2 L_t^\infty} \right)^2 \| \phi \|_{\leq 1},
\]
which can be made small by a divisibility argument. We are thus left with the term
\[
\mathcal{H}^* N^{lowhi}(\mathcal{H}_{A_1}^{nonlin}(I), \phi).
\]

Carrying out similar reductions as in Section 4 of [12] for the “elliptic term” $N_0(A_0, \phi)$, we arrive at the key remaining term
\[
\mathcal{H}^* N^{lowhi}(\mathcal{H}_{A_1}^{(l)}, \phi),
\]
where
\[ \mathcal{H}^* N_0^{\text{lowhi}}(\mathcal{H}A_0^{(l)}, \phi) = \sum_k \sum_{k' \leq k-C} \sum_{j \leq k+C} Q_{j-C} \mathcal{N}(Q_j P_k \mathcal{H}A_0^{(l)}, Q_{j-C} P_k \phi) \]
and
\[ \mathcal{H}A_0^{(l)} := -\chi I \sum_{k_1, k_2, k_3} \sum_{k \leq \min(k_1, k_2) - C} \sum_{j \leq k+C} \Delta^{-1} P_k Q_j \text{Im}(Q_{j-C}(\chi_I \phi_{k_1}) \cdot Q_{j-C}(\chi_I \phi_{k_2})). \]

As in [12], we combine the “hyperbolic term” \( \mathcal{H}^* N_0^{\text{lowhi}}(\mathcal{H}A^{\text{nonlin},(l)}, \phi) \) and the preceding “elliptic term” \( \mathcal{H}^* N_0^{\text{lowhi}}(\mathcal{H}A_0^{(l)}, \phi) \) and wind up with the null forms (61) – (63) in [12]. We formulate these as quadrilinear expressions as in [12] and then prove that smallness can be achieved for each of these.

First null form ((61) in [12]). By estimate (148) in [12], it suffices to consider the following two cases. First, we show that
\[
\sum_k \sum_{k_1 = k+O(1)} \left| \langle \Box^{-1} P_k Q_j (Q_{j-C}(\chi_I \phi_{k_1}) \cdot \partial_\alpha Q_{j-C}(\chi_I \phi_{k_2})), P_k Q_j (\partial^\alpha Q_{j-C} \phi_{k_3} \cdot Q_{j-C} \psi_{k_4}) \rangle \right| \ll \|\psi\|_N^*.
\]
where \( k_1 > k + C \), \( j = k + O(1) \) and \( k_3 = k_4 + O(1) = k + O(1) \). Second, we prove that
\[
\sum_k \sum_{k_1 = k_4 + O(1)} \left| \langle \Box^{-1} P_k Q_j (Q_{j-C}(\chi_I \phi_{k_1}) \cdot \partial_\alpha Q_{j-C}(\chi_I \phi_{k_2})), P_k Q_j (\partial^\alpha Q_{j-C} \phi_{k_3} \cdot Q_{j-C} \psi_{k_4}) \rangle \right| \ll \|\psi\|_N^*.
\]
where \( k_3 > k + C \), \( j = k + O(1) \) and \( k_1 = k_2 + O(1) = k + O(1) \).

We begin with the first case. Here, the inputs \( Q_{j-C}(\chi_I \phi_{k_1}) \) and \( \partial_\alpha Q_{j-C}(\chi_I \phi_{k_2}) \) have Fourier supports in identical (or opposite) angular sectors \( \omega \) of size \( \sim 2^{k-k_1} \). Then we bound
\[
\sum_k \sum_{k_1 = k+O(1)} \left| \langle \Box^{-1} P_k Q_j (Q_{j-C}(\chi_I \phi_{k_1}) \cdot \partial_\alpha Q_{j-C}(\chi_I \phi_{k_2})), P_k Q_j (\partial^\alpha Q_{j-C} \phi_{k_3} \cdot Q_{j-C} \psi_{k_4}) \rangle \right| \leq \sum_k \sum_{k_1 = k+O(1)} 2^{\frac{1}{2}(k-k_1)} \left( \sum_\omega 2^{\frac{1}{2}k_1} \left\| \omega^\alpha Q_{j-C}(\chi_I \phi_{k_1}) \right\|_{L^2_t L^6_x} \right)^{\frac{1}{2}} \left( \sum_\omega \left\| \omega^\alpha Q_{j-C} \nabla_{t,x}(\chi_I \phi_{k_2}) \right\|_{L^\infty_t L^2_x} \right)^{\frac{1}{2}} \times 2^{-\frac{1}{2}k_1} \left\| \nabla_{t,x} \phi_{k+O(1)} \right\|_{L^\infty_t L^2_x} \left\| \psi_{k+O(1)} \right\|_{L^\infty_t L^2_x} \leq \left( \sum_{k_1 = k-C} \sup_\omega \sum_\omega 2^{\frac{1}{2}k_1} \left\| \omega^\alpha Q_{j-C}(\chi_I \phi_{k_1}) \right\|_{L^2_t L^6_x} \right)^{\frac{1}{2}} \left\| \phi_{k+O(1)} \right\|_{N^*}. \]

The desired smallness comes from the divisibility of the quantity
\[
\left( \sum_{k_1 = k-C} \sup_\omega \sum_\omega 2^{\frac{1}{2}k_1} \left\| \omega^\alpha Q_{j-C}(\chi_I \phi_{k_1}) \right\|_{L^2_t L^6_x} \right)^{\frac{1}{2}}.
\]
To see the divisibility, we write
\[ (7.26) \quad P_\gamma^\alpha Q_{j-C}(\chi_I \phi_{k_1}) = P_\gamma^\alpha Q_{j-C}(\chi_I P_\gamma^\alpha Q_{j-C}(\chi_I \phi_{k_1} + i \frac{\phi_{k_1}}{2} \phi_{k_1} + i \frac{\phi_{k_1}^*}{2} \phi_{k_1} + i \frac{\phi_{k_1}^2}{2} \phi_{k_1}) + P_\gamma^\alpha Q_{j-C}(\chi_I Q_{j-C}(\chi_I \phi_{k_1} + i \frac{\phi_{k_1}}{2} \phi_{k_1} + i \frac{\phi_{k_1}^*}{2} \phi_{k_1} + i \frac{\phi_{k_1}^2}{2} \phi_{k_1}) \]
for some $M > 0$ to be chosen sufficiently large. By disposability of the operator $P_t^{\mu} Q_{\leq k_1+l-C}$, we estimate the first term on the right hand side of (7.26) by

$$
\left( \sum_{k_1 \leq -C} \sup_{\omega} \sum_{k} 2^{k_1} \left\| P_t^{\mu} Q_{\leq k_1+l-C} (\chi_t P_t^{\mu} Q_{\leq k_1+l-M \phi_k}) \right\|_{L_t^2 L_x^6} \right) \leq \left( \sum_{k_1 \leq -C} \sup_{\omega} \sum_{k} 2^{k_1} \left\| \chi_t P_t^{\mu} Q_{\leq k_1+l-M \phi_k} \right\|_{L_t^2 L_x^6} \right) ^{\frac{1}{2}}
$$

and smallness can be forced by divisibility of the quantity

$$
\left( \sum_{k_1 \leq -C} \sup_{\omega} \sum_{k} 2^{k_1} \left\| P_t^{\mu} Q_{\leq k_1+l-M \phi_k} \right\|_{L_t^2 L_x^6} \right) ^{\frac{1}{2}} \leq \| \phi \|_{S^1}.
$$

For the second term on the right hand side of (7.26), we use

$$
P_t^{\mu} Q_{\leq k_1+l-C} (\chi_t Q_{\geq k_1+l-M \phi_k}) = P_t^{\mu} Q_{\leq k_1+l-C} (Q_{\geq k_1+l-M \phi_k} + \chi_t Q_{\geq k_1+l-M \phi_k}).
$$

By Bernstein’s inequality in space and in time, we then have

$$
\left\| P_t^{\mu} Q_{\leq k_1+l-C} (Q_{\geq k_1+l-M \phi_k}) \right\|_{L_t^2 L_x^6} \leq 2^{\frac{1}{2}(k_1+l)} 2^{\frac{3}{4}k_1} 2^l \left\| Q_{\geq k_1+l-M \phi_k} \right\|_{L_t^2 L_x^6} \leq 2^{\frac{1}{2}(k_1+l)} 2^{\frac{3}{4}k_1} 2^l \left\| Q_{\geq k_1+l-M \phi_k} \right\|_{L_t^2 L_x^6}
$$

Thus, we obtain

$$
\left( \sum_{k_1 \leq -C} \sup_{\omega} \sum_{k} 2^{k_1} \left\| P_t^{\mu} Q_{\leq k_1+l-C} (Q_{\geq k_1+l-M \phi_k}) \right\|_{L_t^2 L_x^6} \right) \leq \left( \sum_{k_1 \leq -C} \sup_{\omega} 2^{l} 2^{-\frac{3}{4}M} \left\| \nabla_x \phi_k \right\|_{L_t^6 L_x^6} \right) \leq 2^{-\frac{3}{4}M} \| \phi \|_{S^1}
$$

and a smallness factor follows for sufficiently large $M > 0$.

The second case is easier to treat. Here we estimate

$$
\sum_k \sum_{k_3 > k+O(1)} \left| \partial_{x}^{-1} P\xi Q_{k+O(1)} (Q_{\leq k-C} (\chi_t \phi_k + O(1)) \cdot \partial_{x} Q_{\leq k-C} (\chi_t \phi_k + O(1))) \right|
$$

$$
\leq \sum_k \sum_{k_3 > k+O(1)} 2^{-\frac{3}{4}k} \left\| \chi_t \nabla_x \phi_k + O(1) \right\|_{L_t^2 L_x^6} 2^{-\frac{3}{4}k} \left\| \nabla_t (\chi_t \phi_k + O(1)) \right\|_{L_t^2 L_x^6} \times
$$

$$
\times \left\| Q_{\leq k-C} \nabla_t \phi_k \right\|_{L_t^6 L_x^6} \left\| Q_{\leq k-C} \psi_k + O(1) \right\|_{L_t^6 L_x^6}
$$

$$
\leq \left( \sum_k \left( 2^{-\frac{3}{4}k} \left\| \chi_t \nabla_x \phi_k + O(1) \right\|_{L_t^2 L_x^6} \right)^2 \right)^{\frac{1}{2}} \| \phi \|_{S^1}^2 \| \psi \|_{N^2}
$$

and immediately obtain smallness by divisibility.
Second null form ((62) in [12]). By the estimates (149) and (150) in [12], we only have to show that
\[
\sum_{k} \sum_{k_3+k_4+O(1)} \left| \left( (\Box)^{-1} P_k Q_j \partial_t \partial_d \left( Q_{\leq j-C}(x I \phi_k) \cdot \partial^\alpha Q_{\leq j-C}(x I \phi_k) \right) \right) \right| \leq \|\phi\|_N',
\]
where \( j = k + O(1) \), \( k_1 = k_2 + O(1) = k + O(1) \) and \( k_3 > k + C \). Then we estimate
\[
\sum_{k} \sum_{k_3>k+C} \left| \left( (\Box)^{-1} P_k Q_{k+O(1)} \partial_t \partial_d \left( Q_{\leq k-C}(x I \phi_{k+O(1)}) \cdot \partial^\alpha Q_{\leq k-C}(x I \phi_{k+O(1)}) \right) \right) \right| \leq \sum_{k} \sum_{k_3>k+C} \left\| (\Box)^{-1} P_k Q_{k+O(1)} \partial_t \partial_d \left( Q_{\leq k-C}(x I \phi_{k+O(1)}) \cdot \partial^\alpha Q_{\leq k-C}(x I \phi_{k+O(1)}) \right) \right\|_{L^1_t L^\omega_x} \times \left\| P_k Q_{k+O(1)} \partial_t Q_{\leq k-C} \left( \partial^\alpha Q_{\leq k-C}(x I \phi_{k+O(1)}) \right) \right\|_{L^1_t L^\omega_x} \leq \sum_{k} \sum_{k_3>k+C} 2^{-\frac{3k}{4}} \left\| \nabla \chi(x I \phi_{k+O(1)}) \right\|_{L^2_t L^\omega_x} \times \left\| \nabla \chi(x I \phi_{k+O(1)}) \right\|_{L^2_t L^\omega_x} \left\| \partial_t \partial_d \left( Q_{\leq k-C}(x I \phi_{k+O(1)}) \cdot \partial^\alpha Q_{\leq k-C}(x I \phi_{k+O(1)}) \right) \right\|_{L^2_t L^\omega_x} \leq \left( \sum_{k} 2^{-\frac{3k}{4}} \left\| \nabla \chi(x I \phi_{k+O(1)}) \right\|_{L^2_t L^\omega_x} \right)^{1/2} \|\phi\|_{L^2_t L^\omega_x}^{1/2} \|\psi\|_N',
\]
and smallness follows by divisibility.

Third null form ((63) in [12]). By the estimates (152) – (154) in [12], it suffices to consider the following two cases. First, we show that
\[
\sum_{k} \sum_{k_1+k_2+O(1)} \left| \left( (\Box)^{-1} P_k Q_j \partial^i \left( Q_{\leq j-C}(x I \phi_k) \cdot \partial_t Q_{\leq j-C}(x I \phi_k) \right) \right) \right| \leq \|\phi\|_N',
\]
where \( k_1 > k + C \), \( j = k + O(1) \) and \( k_3 = k_4 + O(1) = k + O(1) \). Second, we prove that
\[
\sum_{k} \sum_{k_3+k_4+O(1)} \left| \left( (\Box)^{-1} P_k Q_j \partial^i \left( Q_{\leq j-C}(x I \phi_k) \cdot \partial_t Q_{\leq j-C}(x I \phi_k) \right) \right) \right| \leq \|\phi\|_N',
\]
where \( k_3 > k + C \), \( j = k + O(1) \) and \( k_1 = k_2 + O(1) = k + O(1) \).

In the first case we note that the first two inputs have Fourier supports in identical (or opposite) angular sectors \( \omega \) of size \( \sim 2^{k-k_1} \). Using Bernstein’s inequality, we then place the first input in \( L^2_t L^\omega_x \), the second one in \( L^\omega_t L^2_x \), the third one in \( L^\omega_t L^2_x \) and the fourth one in \( L^\omega_t L^2_x \). As in the first case of the first null form we obtain the desired smallness by divisibility of the quantity
\[
\left( \sum_{k_1} \sup_{l-C} \sum_{\omega} 2^{\frac{3k_l}{4}} \left\| P_{l+k_1} Q_{\leq l-C} \left( x I \phi_k \right) \right\|_{L^2_t L^\omega_x} \right)^{1/2}.
\]
The second case is easier to deal with and we omit the details. 

Step 3. Solution of perturbative problems on suitable space-time slices. This is the crucial technical step. We write
\[
(A^{n_1(L)}, \phi^{n_1(L)}) = (A^{n_1(L-1)}, \phi^{n_1(L-1)}) + (\delta A^{(L)}, \delta \phi^{(L)}).
\]
Then we obtain the following system of equations for the perturbations $(\delta A^{(L)}, \delta \phi^{(L)})$,

\begin{equation}
\Box_{A_{A_0}^{n_1(l-1)}+\delta A^{(L)}}(\phi_{A_0}^{n_1(l-1)}+\delta \phi^{(L)}) - \Box_{A_{A_0}^{n_1(l-1)}}\phi_{A_0}^{n_1(l-1)} = 0,
\end{equation}

\begin{equation}
\Box A^{(L)} = -\text{Im} \mathcal{P}(\phi_{A_0}^{n_1(l-1)} \cdot \nabla_x \delta \phi^{(L)} - \delta \phi^{(L)} \cdot \nabla_x \phi_{A_0}^{n_1(l-1)} + \delta \phi^{(L)} \cdot \nabla_x \phi_{A_0}^{n_1(l-1)})
+ \text{Im} \mathcal{P}(A_{A_0}^{n_1(l-1)} + \delta A^{(L)})[\phi_{A_0}^{n_1(l-1)} + \delta \phi^{(L)} - A_{A_0}^{n_1(l-1)}]^{2} = 0.
\end{equation}

We have to show that if the initial data $(\delta A^{(L)}, \delta \phi^{(L)})[0]$ are less than the absolute constant $\varepsilon_0$ in the energy sense, then we can prove frequency localized $S^1$ norm bounds via bootstrap on any space-time slice on which certain “divisible” norms of $(A_{A_0}^{n_1(l-1)}, \phi_{A_0}^{n_1(l-1)})$ are small. Furthermore, the number of such space-time slices needed to fill all of space-time depends on the a priori assumed $S^1$ norm bounds for the components $(A_{A_0}^{n_1(l-1)}, \phi_{A_0}^{n_1(l-1)})$.

One technical difficulty is the formulation of the correct frequency localized $S^1$ norm bound for the propagation of $\delta \phi^{(L)}$, because there is a contribution from low frequencies of $\phi_{A_0}^{n_1(l-1)}$, and similarly for $\delta A^{(L)}$. However, this low frequency contribution can be made arbitrarily small by picking $n$ large and $\delta_1$ small enough.

We note that while $(A_{A_0}^{n_1(l-1)}, \phi_{A_0}^{n_1(l-1)})$ exists globally in time, $(\delta A^{(L)}, \delta \phi^{(L)})$ only exists locally in time and we will have to prove global existence and $S^1$ norm bounds for it. For now, any statement we make about $(\delta A^{(L)}, \delta \phi^{(L)})$ is meant locally in time on some interval $I_0$ around $t = 0$. Proposition [7.5] yields a partition of the time axis $\mathbb{R}$ into $N = N(||(\phi_{A_0}^{n_1(l-1)}, A_{A_0}^{n_1(l-1)})||_{L^1(\mathbb{R} \times \mathbb{R}^d)})$ many time intervals $\{I_j\}_{j=1}^N$, on which the smallness conclusions (in terms of $E_{\text{crit}}$) of Proposition [7.5] hold. We tacitly assume that these intervals are intersected with $I_0$ and now fix the interval $I_1$, which we assume to contain $t = 0$. All the arguments in this step can be carried out for any of the later intervals $I_2, \ldots, I_N$.

**Bootstrap assumptions** : Suppose that there exist decompositions

\[\delta A^{(L)} = \delta A^{(L)}_1 + \delta A^{(L)}_2, \quad \delta \phi^{(L)} = \delta \phi^{(L)}_1 + \delta \phi^{(L)}_2\]

satisfying the following bounds.

(i) Let $\{c^{(L)}_{\delta A,k}\}_{k \in \mathbb{Z}}$ be a frequency envelope controlling the data $P_k \delta A^{(L)}[0]$ at time $t = 0$ and let $\{d^{(L)}_{\delta A,k}\}_{k \in \mathbb{Z}}$ be a frequency envelope that decays exponentially for $k > b_L$ but is otherwise not localized and satisfies the smallness condition

\[\sum_k d^{(L)}_{\delta A,k} \leq \delta_2 = \delta_2(\delta_1).\]

Then we assume that for all $k \in \mathbb{Z},$

\[\|P_k \delta A^{(L)}_1\|_{L^1(I_1 \times \mathbb{R}^d)} \leq C_{\delta A,k}^{(L)},\]

\[\|P_k \delta A^{(L)}_2\|_{L^1(I_1 \times \mathbb{R}^d)} \leq C_{\delta A,k}^{(L)},\]

where $C \equiv C(E_{\text{crit}})$ is sufficiently large.

(ii) Let $\{c^{(L)}_{\delta \phi,k}\}_{k \in \mathbb{Z}}$ be a frequency envelope controlling the data $P_k \delta \phi^{(L)}[0]$ at time $t = 0$ and let $\{d^{(L)}_{\delta \phi,k}\}$ be a frequency envelope that decays exponentially for $k > b_L$, but is otherwise not localized and satisfies the smallness condition

\[\left(\sum_k d^{(L)}_{\delta \phi,k}\right)^2 \leq \delta_3 = \delta_3(\delta_1).\]
Then we assume that for all \( k \in \mathbb{Z} \),
\[
\|P_k \delta \phi^{(L)}_{\Delta 0}\|_{L^1(I_1 \times \mathbb{R}^d)} \leq C \epsilon_{\phi, k}^{(L)}, \\
\|P_k \delta \phi^{(L)}_{\Delta e}\|_{L^1(I_1 \times \mathbb{R}^d)} \leq C \epsilon_{\phi, k}^{(L)},
\]
where \( C \equiv C(E_{crit}) \) is sufficiently large.

We now show that we can improve this to a similar decomposition with
\[
\|P_k \delta A^{(L)}\|_{L^1(I_1 \times \mathbb{R}^d)} \leq \frac{C}{2} \epsilon_{A, k}^{(L)}, \\
\|P_k \delta A^{(L)}_{\Delta e}\|_{L^1(I_1 \times \mathbb{R}^d)} \leq \frac{C}{2} \epsilon_{A, k}^{(L)},
\]
\[
(7.29)
\]
provided me make the additional assumption
\[
\delta_2 \ll \delta_3
\]
with implied constant depending only on \( E_{crit} \).

Observe that we have
\[
\sum_k (\epsilon_{A, k}^{(L)})^2 + (\epsilon_{\phi, k}^{(L)})^2 \leq \epsilon_0
\]
and that our smallness parameters satisfy
\[
\delta_1 \ll \delta_2 \ll \delta_3 \ll \epsilon_0.
\]

For the remainder of this step we simply write \( I \equiv I_1 \) and \( \phi \equiv \phi_{\Lambda_0}^{n_1, (L-1)}, \delta \phi \equiv \delta \phi^{(L)}, \ A \equiv A_{\Lambda_0}^{n_1, (L-1)}, \ \delta A \equiv \delta A^{(L)} \).

**Step 3a. Reorganizing the key equation** \( (7.27) \). We introduce the connection form \((A + \delta A)^{\nonlin, (I)} \)
alogously to \((7.9)\) by setting for \( i = 1, \ldots, 4 \),
\[
(7.30) \ (A + \delta A)^{\nonlin, (I)}_i := -\chi I \sum_{k,j} \phi^{-1} P_k Q_j P_i (\chi \phi + \delta \phi) \cdot \nabla (\chi \phi + \delta \phi) - \chi i (A + \delta A) \phi + \delta \phi_i^2,
\]
and define \((A + \delta A)^{\free, (I)}\) as the free wave with initial data at time \( t = 0 \) given by
\[
(A + \delta A)^{\free, (I)}[0] = (A + \delta A)[0] - (A + \delta A)^{\nonlin, (I)}[0].
\]
Then we have
\[
(A + \delta A)|_t = (A + \delta A)^{\free, (I)} + (A + \delta A)^{\nonlin, (I)}.
\]
On $I \times \mathbb{R}^4$ we may rewrite the equation (7.27) for $\delta \phi$ into the following frequency localized form

\[
\partial_t^\alpha (A + \delta A)(A + \delta A) \phi = -[P_0, \partial_t^\alpha (A + \delta A)] \phi
- P_0 \left(2i \sum_k P_{>k-C}(A + \delta A) \phi\right)
- P_0 \left(2i(A + \delta A)_{\text{nonlin}} \partial^j \delta \phi - 2i(A + \delta A)_0 \partial_i \delta \phi\right)
+ P_0 \left(i(\partial_i A_0 + \partial_j \delta A_0)(\phi + \delta \phi) - i(\partial_i A_0)\phi\right)
+ P_0 \left((A + \delta A)_\alpha (A + \delta A)^\alpha - A_\alpha A^\alpha\phi\right).
\]

(7.31)

We immediately see that compared to (7.22), a qualitatively new feature in (7.31) is the interaction term

\[
P_0 \left((\delta A)_j \partial^j \phi - (\delta A)_0 \partial_i \phi\right).
\]

(7.32)

**Step 3b. Improving the bounds for $\delta \phi$ using (7.31).** In order to obtain bounds on the $S^1(I \times \mathbb{R}^4)$ norm of $P_0 \delta \phi$ by bootstrap, we work with suitable Schwartz extensions and use the linear estimate (3.3) for the magnetic wave operator $\square^\alpha_{(A+\delta A)_{\text{rec,t}}}$. We first consider the new interaction term (7.32). As usual the main difficulty comes from the low-high interactions, so we begin with this case, i.e. the term

\[
P_{<0}(\delta A)_j \partial^j \phi - P_{<0}(\delta A)_0 \partial_i \phi.
\]

For the spatial components of $\delta A$, we define

\[
(\delta A)_{\text{free}} := (A + \delta A)_{\text{free}} - A_{\text{free}}, \quad (\delta A)_{\text{nonlin}} := (A + \delta A)_{\text{nonlin}} - A_{\text{nonlin}}.
\]

and correspondingly have on $I \times \mathbb{R}^4$ that

\[
(\delta A)_I = (\delta A)_{\text{free}} + (\delta A)_{\text{nonlin}}.
\]

We can therefore split on $I \times \mathbb{R}^4$,

\[
P_{<0}(\delta A)_j \partial^j \phi = P_{<0}(\delta A)_{\text{free}} \partial^j \phi + P_{<0}(\delta A)_{\text{nonlin}} \partial^j \phi.
\]

The first term on the right hand side can in turn be split into two contributions

\[
P_{<0}(\delta A)_{\text{free}} \partial^j \phi = P_{<0}(\delta A)_{\text{free}} \partial^j \phi + P_{<0}(\delta A)_{\text{free}} \partial^j \phi
+ P_{<0}(\delta A)_{\text{free}} \partial^j \phi,
\]

where $(\delta A)_{\text{free}}$ is the free evolution of the data $(\delta A)_{\text{free}}[0]$, while $(\delta A)_{\text{free}}$ is the free wave with data

\[
\left(\sum_{k,j} \partial^{-1} P_k Q_j \mathcal{P}(\chi_1(\phi + \delta \phi)) \nabla \chi_1(i\phi + \delta \phi)) - \chi_1 i(A + \delta A)[\phi + \delta \phi^2]\right)[0],
\]

\[
- \left(\sum_{k,j} \partial^{-1} P_k Q_j \mathcal{P}(\chi_1(\phi) \nabla \chi_1(i\phi) - \chi_1 iA[\phi]^2\right)[0].
\]

In order to estimate the terms on the right hand side of (7.33), we will invoke the following estimate from [12] for a free wave $A_{\text{free}}$ in Coulomb gauge for $k_1 \leq k_2 - C$,

\[
\|P_{k_1} A_{\text{free}} \partial^j \phi\|_N \leq \|P_{k_1} A_{\text{free}}\|_{S^1} \|P_{k_2} \phi\|_{S^1} \leq \|P_{k_1} A_{\text{free}}[0]\|_{H^1 \times L^2} \|P_{k_2} \phi\|_{S^1}.
\]

(7.34)
We begin with the first term on the right hand side of (7.35),

\[ P \triangleq \sum_{j=0}^{\infty} (\delta A)^{\text{free}},(I) P_{r,j} \partial_j \phi. \]

Here we have to take advantage of the properties of the Fourier support of the data \( P \triangleq \sum_{j=0}^{\infty} (\delta A)^{\text{free}},(I) [0]. \)

It suffices to assume that

\[ \|P_k(\delta A)^{\text{free}},[0]\|_{L^2_x \times L^2_t} \leq C(c_{\delta A,k} + d_{\delta A,k}) \]

for \( C \equiv C(E_{\text{crit}}) \) sufficiently large. This is an assumption that will hold inductively at later initial times (for the intervals \( I_2, \ldots, I_N \)). We observe that the frequency envelope \( \{c_{\delta A,k}\}_{k \in \mathbb{Z}} \) is “sharply localized” to the dyadic frequency interval \([a_L, b_L] \) in the sense that it is exponentially decaying for \( k < a_L \) and \( k > b_L \). By (7.34) we then have

\[ \|P_{\leq 0}(\delta A)^{\text{free}},(I) P_0 \partial_j \phi\|_{L^2_x(\mathbb{R}^4)} \lesssim \sum_{k < 0} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} + \sum_{k < 0} d_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)}. \]

We begin to estimate the first term on the right hand side of (7.35), where we only consider the case when \( a_L < 0 \). For \( R > 0 \) to be chosen sufficiently large later on, we split

\[ \sum_{k < 0} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} = \sum_{k \leq a_L - R} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} + \sum_{a_L - R < k \leq a_L + R} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} + \sum_{a_L + R < k < 0} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)}. \]

To make the first term on the right hand side of (7.36) small, we use the exponential decay of the frequency envelope \( \{c_{\delta A,k}\}_{k \in \mathbb{Z}} \) to bound

\[ \sum_{k \leq a_L - R} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} \lesssim \sum_{k \leq a_L - R} 2^{-\sigma(a_L - k)} \|\delta A[0]\|_{H^1_x \times L^2_t} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} \lesssim E_{\text{crit}} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)}. \]

Upon replacing the output frequency 0 by general \( l \in \mathbb{Z} \), square summing over \( l \) and choosing \( R > 0 \) sufficiently large, we bound the preceding by \( \ll E_{\text{crit}} \delta_3 \), as desired. In order to make the third term on the right hand side of (7.35) small, we exploit that by Step 1 the \( S^1 \) norms of \( P_0 \phi \) are exponentially decaying beyond the scale \( l > a_L \). We have

\[ \sum_{a_L + R < k < 0} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} \lesssim E_{\text{crit}} (|a_L| - R) c_0, \]

where \( \{c_l\}_{l \in \mathbb{Z}} \) is a sufficiently flat frequency envelope covering the initial data \( (A_0^{\text{free}},L-1), \phi_n^{\text{free}},(L-1) [0] \) as in Step 1. Then replacing the frequency 0 in \( P_0 \phi \) by a general dyadic frequency \( l > a_L + R \), square summing over \( l \) and choosing \( R > 0 \) sufficiently large, we find

\[ \sum_{l > a_L + R} \sum_{a_L + R < k < l} c_{\delta A,k} \|P_l \phi\|_{L^1_x(\mathbb{R}^4)}^2 \lesssim E_{\text{crit}} \sum_{l > a_L + R} (l - a_L - R)^2 c_1^2 \lesssim E_{\text{crit}} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)}^2. \]

which is acceptable. It remains to make the second term on the right hand side of (7.36) small. To this end we exploit the frequency evacuation property of the data \( (A_0^{\text{free}}, \phi_n^{\text{free}}) \) at the edges of the frequency intervals \([a_L, b_L] \) that we established in Lemma 7.3. For sufficiently small \( \delta_1 > 0 \) and all sufficiently large \( n \), we then have

\[ \sum_{a_L - R < k \leq a_L + R} c_{\delta A,k} \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} \lesssim R \delta_1 \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)} \ll \delta_1 \|P_0 \phi\|_{L^1_x(\mathbb{R}^4)}. \]

Upon replacing the frequency 0 in \( P_0 \phi \) by an arbitrary dyadic frequency \( l \in \mathbb{Z} \) and square summing, we obtain the desired smallness \( \ll E_{\text{crit}} \delta_3 \) for the last estimate.
The contribution of the second term on the right hand side of (7.35) is acceptable, because, upon replacing the output frequency 0 by \( l \in \mathbb{Z} \), square summing and using the bootstrap assumptions on the interval \( I \), we obtain the bound
\[
\left( \sum_{l} \sum_{k} d_{A,k} \|P_{0} \phi\|_{L^1(I \times \mathbb{R}^4)} \right)^{\frac{1}{2}} \lesssim_{E_{\text{crit}}} \delta_2 \ll \delta_3,
\]
where the implied constant in \( \lesssim_{E_{\text{crit}}} \) depends at most polynomially on \( E_{\text{crit}} \).

Next, we estimate the second term on the right hand side of (7.33),
\[
P_{<0}(\delta A)_{j}^{\text{free},(l)} P_{0} \partial_j \phi.
\]
By (7.34) we have
\[
(7.37) \quad \|P_{<0}(\delta A)_{j}^{\text{free},(l)} P_{0} \partial_j \phi\|_{L^1(I \times \mathbb{R}^4)} \leq \|P_{<0}(\delta A)_{j}^{\text{free},(l)}\|_{L^1(I \times \mathbb{R}^4)} \|P_{0} \phi\|_{L^1(I \times \mathbb{R}^4)}.
\]

We illustrate how to obtain the desired smallness in this case by assuming for simplicity that \( P_{<0}(\delta A)_{j}^{\text{free},(l)} \) is just the free evolution of the data
\[
\sum_{k=0} \sum_{j} d_{k} Q_{j} P(\chi_{I} \phi \cdot \nabla_{x}(\chi_{I} \partial_j \phi))|0|.
\]
If \( \delta \phi = (\delta \phi)_{\mathbb{Z}} \), we obtain by similar estimates as in the proof of Lemma 7.6 that
\[
\|P_{<0}(\delta A)_{j}^{\text{free},(l)}\|_{L^1(I \times \mathbb{R}^4)} \lesssim \sum_{k=0} \|\chi_{I} \partial_j \phi\|_{L^1(I \times \mathbb{R}^4)} c_{\delta \phi,k}.
\]
Then we achieve the desired smallness for
\[
\|P_{<0}(\delta A)_{j}^{\text{free},(l)} P_{0} \partial_j \phi\|_{L^1(I \times \mathbb{R}^4)} \lesssim \sum_{k=0} c_{\delta \phi,k} \|P_{0} \phi\|_{L^1(I \times \mathbb{R}^4)}
\]
by proceeding exactly as with the term (7.36). If instead \( \delta \phi = (\delta \phi)_{\mathbb{Z}} \), we find
\[
\|P_{<0}(\delta A)_{j}^{\text{free},(l)} P_{0} \partial_j \phi\|_{L^1(I \times \mathbb{R}^4)} \lesssim \left( \sum_{k} d_{A,k}^2 \right)^{\frac{1}{2}} \|P_{0} \phi\|_{L^1(I \times \mathbb{R}^4)} \lesssim \delta_3 \|\chi_{I} \partial_j \phi\|_{L^1(I \times \mathbb{R}^4)} \|P_{0} \phi\|_{L^1(I \times \mathbb{R}^4)}.
\]

Upon replacing the output frequency 0 by \( l \in \mathbb{Z} \), square summing and using that \( \|\phi\|_{L^1(I \times \mathbb{R}^4)} \lesssim C(E_{\text{crit}}) \) by the choice of the interval \( I \), we obtain the bound \( \lesssim_{E_{\text{crit}}} \delta_3 \). This is unfortunately not yet enough to close the bootstrap. To gain the extra smallness we partition the interval \( I \) further and use divisibility arguments as in the proof of Lemma 7.6 However, the number of intervals needed for this partition depends only on \( E_{\text{crit}} \) (and not on the stage of the induction), which is acceptable.

This finishes the estimate of the contribution of
\[
P_{<0}(\delta A)_{j}^{\text{free},(l)} P_{0} \partial_j \phi
\]
and we now have to bound
\[
\|P_{<0}(\delta A)_{j}^{\text{nonlin},(l)} P_{0} \partial_j \phi - P_{<0}(\delta A)_{0} P_{0} \partial_j \phi\|_{L^1(I \times \mathbb{R}^4)}.
\]
At this point we can proceed by analogy to the treatment of the \( \phi \) equation in the proof of Proposition 7.5 After a further partitioning of the interval \( I \) and similar divisibility arguments, we can replace the output frequency 0 by \( l \in \mathbb{Z} \) and upon square summing, we obtain a bound of the desired form \( \ll_{E_{\text{crit}}} \delta_3 \).

The remaining frequency interactions in the estimate of the term (7.32) as well as all other terms on the right hand side of (7.31) are easier to control. We omit the details.
Step 3c. **Improving the bounds for $\delta A$ using (7.28).** In order to deduce $S^1(I \times \mathbb{R}^4)$ norm bounds on $P_0 \delta A$ from the perturbation equation (7.28) by bootstrap, we perform the same kind of divisibility arguments as in the proof of estimate (7.11) in Lemma 7.6 for the terms linear in $\delta \phi$ or $\delta A$.

Step 4. Repetition of the bootstrap on suitable space-time slices; proof that the energy of perturbation remains small. In this final step we show that the crucial assumption on the energy of the perturbation

$$E(\delta A^{(L)}, \delta \phi^{(L)})(0) < \varepsilon_0$$

remains in tact along the evolution up to a very small correction. We recall that

$$\delta A^{(L)} = A^{n_1(L),0}_\Lambda_0 - A^{n_1(L-1),0}_\Lambda_0, \quad \delta \phi^{(L)} = \phi^{n_1(L),0}_\Lambda_0 - \phi^{n_1(L-1),0}_\Lambda_0.$$ 

**Lemma 7.7.** Assuming the bounds (7.29) on the evolution of $(\delta A^{(L)}, \delta \phi^{(L)})$ on $I_1 \times \mathbb{R}^4$, we have for sufficiently small $\delta_1 > 0$ and all sufficiently large $n$ that

$$E(\delta A^{(L)}, \delta \phi^{(L)})(t) < \varepsilon_0 \quad \text{for } t \in I_1.$$ 

**Proof.** By energy conservation for the evolutions of $(A^{n_1(L),0}_\Lambda_0, \phi^{n_1(L),0}_\Lambda_0)$ and of $(A^{n_1(L-1),0}_\Lambda_0, \phi^{n_1(L-1),0}_\Lambda_0)$, it suffices to show that

$$\left| E(A^{n_1(L),0}_\Lambda_0, \phi^{n_1(L),0}_\Lambda_0)(t) - E(A^{n_1(L-1),0}_\Lambda_0, \phi^{n_1(L-1),0}_\Lambda_0)(t) - E(\delta A^{(L)}, \delta \phi^{(L)})(t) \right|$$

can be made arbitrarily small uniformly for all $t \in I_1$ by choosing $\delta_1 > 0$ sufficiently small and $n$ sufficiently large. This reduces to bounding the following expression evaluated at any time $t \in I_1$,

$$\begin{align*}
& \left| 2 \sum_{i<j} \int_{\mathbb{R}^4} (\partial_i A^{n_1(L-1),0}_{\Lambda_0,j}) (\partial_j \delta A^{(L)})(dx) + \sum_{i} \int_{\mathbb{R}^4} (\partial_i A^{n_1(L-1),0}_{\Lambda_0,i})(\partial_i \delta A^{(L)})(dx) + (\partial_i A^{n_1(L),0}_{\Lambda_0,0})(\partial_i \delta A^{(L)})(dx) 
+ \frac{1}{2} \sum_{\alpha} \int_{\mathbb{R}^4} \left| (A^{n_1(L-1),0}_{\Lambda_0,\alpha})(\partial_\alpha \delta \phi^{(L)})(dx) + (\partial_\alpha \delta A^{(L)})(\phi^{n_1(L)-1,0}_{\Lambda_0})(dx) \right|^2 dx 
+ \sum_{\alpha} \text{Re} \left( \int_{\mathbb{R}^4} (\partial_\alpha \phi^{n_1(L-1),0}_{\Lambda_0,\alpha} + i A^{n_1(L-1),0}_{\Lambda_0,\alpha})(\phi^{n_1(L)-1,0}_{\Lambda_0})(dx) \cdot (\partial_\alpha \delta A^{(L)} + i \delta A^{(L)})(\partial_\alpha \delta \phi^{(L)})(dx) + (\partial_\alpha \phi^{n_1(L-1),0}_{\Lambda_0,\alpha} + i A^{n_1(L-1),0}_{\Lambda_0,\alpha})(\phi^{n_1(L)-1,0}_{\Lambda_0})(dx) \cdot (i A^{n_1(L-1),0}_{\Lambda_0,\alpha} \delta \phi^{(L)} + i \delta A^{(L)})(\phi^{n_1(L)-1,0}_{\Lambda_0})(dx) \right) \right|.
\end{align*}$$

We note that in this expression at least one term of the form $A^{n_1(L-1),0}_{\Lambda_0}$ or $\phi^{n_1(L-1),0}_{\Lambda_0}$ is paired against at least one term of the form $\delta A^{(L)}$ or $\delta \phi^{(L)}$. By Plancherel’s theorem (and a Littlewood-Paley trichotomy to deal with the nonlinear interactions), we reduce to estimating a sum of the form

$$\left| 2 \sum_{k \in \mathbb{Z}} \sum_{i<j} \int_{\mathbb{R}^4} P_k(\partial_i A^{n_1(L-1),0}_{\Lambda_0,j})(P_k(\partial_j \delta A^{(L)}(dx) + \ldots) \right|.$$

By the bounds (7.29) and Step 1, we estimate this by

$$\leq \sum_{k \in \mathbb{Z}} \left\| P_k \nabla_x A^{n_1(L-1),0}_{\Lambda_0}(t) \right\|_{L^2} \left\| P_k \nabla_x A^{(L)}(t) \right\|_{L^2} + \ldots$$

$$\leq \sum_{k \in \mathbb{Z}} c^{(L)}_k (C_{\delta A,k} + d^{(L)}_{\delta A,k}) + \ldots$$
To see that this expression can be made arbitrarily small, we split
\[ \sum_{k \in \mathbb{Z}} c_k^{(L-1)} c_{\delta A, k} = \sum_{k \leq a_L - R} c_k^{(L-1)} c_{\delta A, k} + \sum_{a_L - R < k \leq a_L + R} c_k^{(L-1)} c_{\delta A, k} + \sum_{k > a_L + R} c_k^{(L-1)} c_{\delta A, k}. \]

The first term can be made arbitrarily small for sufficiently large \( R > 0 \) by the exponential decay of the frequency envelope \( \{c_k^{(L)}\}_{k \in \mathbb{Z}} \) beyond \([a_L, b_L]\). Similarly, we achieve smallness for the third term for sufficiently large \( R > 0 \) by the exponential decay of \( \{c_k^{(L-1)}\}_{k \in \mathbb{Z}} \) for \( k > a_L \) established in Step 1. Finally, we gain smallness for the second term for all sufficiently large \( n \) from the frequency evacuation property in Lemma 7.3. Moreover, we have by (7.29) that
\[ \sum_{k \in \mathbb{Z}} c_k^{(L-1)} d_{\delta A, k} \lesssim_{E, \epsilon_0} \delta_2(\delta_1) \ll 1. \]

Steps 1–4 complete the proof of Proposition 7.4.

7.4. Interlude: Proofs of Proposition 4.3, Proposition 5.1, Proposition 5.11 and Proposition 6.1

Proof of Proposition 4.3 (Outline) Let \( E \) denote the conserved energy of the admissible solution \((A, \phi)\). Analogously to the proof of Proposition 7.5, we can partition the time interval \((-T_0, T_1)\) into \( N = N(||(A, \phi)||_{S^1((-T_0, T_1) \times \mathbb{R}^4)}) \) many intervals \( I \) such that
\[ A|_I = A_{\text{free}, I} + A_{\text{nonlin}, I}, \quad \Box A_{\text{free}, I} = 0 \]
and
\[ \|\nabla_{\ell^j} A_{\text{free}, I} \|_{L^\infty_t L^2_{\ell^j} (\mathbb{R}^4)} \leq E^{1/2}, \]
\[ \|A_{\text{nonlin}, I}\|_{L^4_t S^1 I (\mathbb{R}^4)} \ll 1, \]
\[ \|\phi\|_{S^1 I (\mathbb{R}^4)} \leq C(E), \]
where \( C(\cdot) \) grows at most polynomially. For each such interval \( I \), say of the form \( I = [t_0, t_1] \) for some \( t_0 < t_1 \), we let \( \{c_k\}_{k \in \mathbb{Z}} \) be a sufficiently flat frequency envelope covering the data \((A, \phi)|_{t_0}\) at time \( t_0 \). Then we show that the bootstrap assumption
\[ \|P_k A\|_{S^1_{1} (\mathbb{R}^4)} + \|P_k \phi\|_{S^1_{1} (\mathbb{R}^4)} \leq D c_k \]
for \( D = D(E) \) sufficiently large, implies the improved bound
\[ \|P_k A\|_{S^1_{1} (\mathbb{R}^4)} + \|P_k \phi\|_{S^1_{1} (\mathbb{R}^4)} \leq \frac{D}{2} c_k. \]

We only discuss the equation for \( \phi \), because the equation for \( A \) is easier. It suffices to consider the case \( k = 0 \). On \( I \times \mathbb{R}^4 \) we may rewrite the equation for \( \phi \) into the following frequency localized form
\[ \Box_{A_{\text{free}, I}} |\phi| = \sum_{k > 0} P_{k < a_L} A_{\text{free}, I} |\phi| - 2iP_0 \sum_{k > 0} P_{k < a_L} A_{\text{free}, I} |\phi| = 2iP_0 (A_{\text{nonlin}, I} |\phi| - A_0 |\phi|) + P_0 (i(\partial_t A_0 |\phi| + A_0 A_0 |\phi|). \]
In order to close the bootstrap argument we now translate the estimates in the proof of Proposition 7.5 into the language of frequency envelopes. For example, to bound the high-high interactions in the term
\[ P_0 \left( \sum_k P_{\geq k} C A_j^{free,I} P_k \partial^j \phi \right), \]
we use estimate (131) from [12] to obtain
\[ \sum_{k_1 = k_2 + O(1)} \left\| P_0( P_{k_1} A_j^{free,I} P_{k_2} \partial^j \phi) \right\|_{N_0(I \times \mathbb{R}^4)} \leq \sum_{k_1 = k_2 + O(1)} 2^{-\delta k_1} \left\| P_{k_1} A_j^{free,I} \right\|_{S^1(I \times \mathbb{R}^4)} \left\| P_{k_2} \phi \right\|_{S^1(I \times \mathbb{R}^4)} \]
\[ \leq \sum_{k_1 = k_2 + O(1)} 2^{-\delta k_1} \left\| P_{k_1} A_j^{free,I} \right\|_{S^1(I \times \mathbb{R}^4)} C_{k_2}. \]

Summing over all sufficiently large \( k_1 \gg 1 \) and using the properties of frequency envelopes, we can bound this by
\[ \leq D c_0. \]

This allows us to reduce to the case \( k_1 = k_2 + O(1) = O(1) \). Here we gain the necessary smallness by further partitioning the interval \( I \) (where the total number of subintervals depends only on the size of \( E \)), using exactly the same divisibility argument as for the term (7.23) in the proof of Proposition 7.5. All other terms on the right hand side of (7.38) can be treated analogously to the above argument. □

**Proof of Proposition 7.1** (Outline) Here we are in the situation of Step 3 of the proof of Proposition 7.3. We obtain the bound
\[ \left\| (\delta A, \delta \phi) \right\|_{S^1([-T_0, T_1] \times \mathbb{R}^4)} \leq L \delta_0 \]
for sufficiently small \( \delta_0 \) by means of a bootstrap argument performed on a finite number of space-time slices, whose number depends on \( L \). We select these space-time slices as in Proposition 7.5. The main difficulty arises from the equation for \( \phi \). As in Step 3a of the proof of Proposition 7.4, we localize the equation for \( \phi \) to frequency 0 on a suitable space-time slice \( I \times \mathbb{R}^4 \) with \( 0 \in I \). Then, as in Step 3b there, the main difficulty comes from the new low-high interaction term
\[ P_{<0} (\delta A) \partial \phi - P_{<0} (\delta A) P_0 \partial^j \phi. \]

Using notation from the proof of Proposition 7.4, the worst contribution comes from
(7.39) \[ P_{<0} \delta A_j^{free,I} P_0 \partial^j \phi, \]
where we recall that \( \delta A_j^{free,I} \) is the free evolution of the data \( \delta A[0] \). We observe that for \( \delta A_j^{free,I} \) the interaction term (7.39) vanishes by assumption on the frequency support of \( \delta A[0] \) unless \( K \leq 0 \). By our assumption on the frequency supports of the data \( (A, \phi)[0] \) and by Proposition 4.3, we obtain
\[ \left\| P_0 \phi \right\|_{S^1([-T_0, T_1] \times \mathbb{R}^4)} \leq L 2^{\sigma K}. \]

More generally, replacing the frequency 0 by \( l \in \mathbb{Z} \) with \( l \geq K \), we have
\[ \left\| P_l \phi \right\|_{S^1([-T_0, T_1] \times \mathbb{R}^4)} \leq L 2^{-\sigma(l-K)}. \]

By estimate (7.54), we then find for \( l \geq K \) that
\[ \left\| P_{<0} \delta A_j^{free,I} P_l \partial^j \phi \right\|_{N_0(I \times \mathbb{R}^4)} \leq L \delta_0 |l - K| 2^{-\sigma(l-K)}, \]
where the extra factor $|l - K|$ arises due to the $l^1$ summation over the frequencies of $P_{< l} \delta A^{free, \ell}$. But then we get the bound
\[
\left( \sum_{l \geq K} \left\| P_{< l} \delta A^{free, \ell} \right\|_{L^2_x(D_\ell \times \mathbb{R}^4)}^2 \right)^{\frac{1}{2}} \lesssim L_0 \delta_0,
\]
which gives the required smallness for this term. Then the argument proceeds as for Proposition 7.4.

**Proof of Proposition 5.11** (Outline) We write for large $R_2 \geq R_1 \geq R_0$
\[
(\tilde{A}_{R_2}, \tilde{\phi}_{R_2}) = (\tilde{A}_{R_1} + \delta A, \tilde{\phi}_{R_1} + \delta \phi).
\]
Then we analyze the equations for $(\delta A, \delta \phi)$. In fact, the only new feature occurs for the $\delta \phi$ equation and so we explain this here. We obtain the equation
\[
\square \tilde{A}_{R_1} + \delta \phi + (\square \tilde{A}_{R_1} + \delta A - \square \tilde{A}_{R_1}) \tilde{\phi}_{R_1} = 0.
\]
Here we only retain the key difficult term that cannot be treated via a perturbative argument, using suitable divisibility properties as for example done in great detail in Step 3 of the proof of Proposition 7.4. This term is given by
\[
\sum_{k \in \mathbb{Z}} P_{< k}(\delta A^{free})_k \partial^j \tilde{\phi}_{R_1}.
\]
However, since we localize to a small time interval $[-T, T]$ around $t = 0$, it will be possible to obtain good $N$ norm bounds. Note that on account of the estimates in Subsection 5.1, we may assume that
\[
\limsup_{R \to \infty} \left\| (\tilde{A}_R, \tilde{\phi}_R) \right\|_{S^1([-T, T] \times \mathbb{R}^4)} < \infty
\]
for all $R$ sufficiently large, provided that $T$ is sufficiently small. We shall assume, as we may that $T < 1$. Then write
\[
\sum_{k} P_{< k}(\delta A^{free})_k \partial^j \tilde{\phi}_{R_1} = \sum_{k} P_{< k(-\frac{1}{2} \log R_1, k)}(\delta A^{free})_k \partial^j \tilde{\phi}_{R_1}
\]
(7.40)

The last term will be estimated by taking advantage of Huygens’ principle as well as our particular choice of initial data, namely that $\phi_{R_1}[0]$ is supported on the set $\{|x| \leq \frac{R_1}{10}\}$, while $\delta A[0]$ is supported on $\{|x| \geq R_1\}$ up to tails that essentially decay exponentially fast. We now estimate both terms on the right hand side of (7.40). For the first term we find
\[
\left\| \sum_{k} P_{< k(-\frac{1}{2} \log R_1, k)}(\delta A^{free})_k \partial^j \tilde{\phi}_{R_1} \right\|_{S^1([-T, T] \times \mathbb{R}^4)} \lesssim \left( \sum_{k} \left\| P_{< k(-\frac{1}{2} \log R_1, k)}(\delta A^{free})_k \partial^j \tilde{\phi}_{R_1} \right\|_{L_x^2([-T, T] \times \mathbb{R}^4)}^2 \right)^{\frac{1}{2}}
\]
\[
\lesssim \sup_l \left\| P_{< k(-\frac{1}{2} \log R_1, l)} \delta A^{free} \right\|_{L_x^2([-T, T] \times \mathbb{R}^4)} \left( \sum_{k} \left\| P_k \nabla \tilde{\phi}_{R_1} \right\|_{L_x^2([-T, T] \times \mathbb{R}^4)}^2 \right)^{\frac{1}{2}}
\]
\[
\lesssim R_1^{-\frac{1}{2}} \left\| \delta A[0] \right\|_{H_x^1 \times L_t^2} \left\| \tilde{\phi}_{R_1} \right\|_{S^1([-T, T] \times \mathbb{R}^4)}
\]
and so this converges to 0 as $R_1 \to +\infty$. For the second term we have

\[
\| \sum_k P_{[-\frac{1}{2} \log R_1, k]} (\delta A^{free}) f_k \partial \bar{\phi}_R \|_{L^2([-T, T] \times \mathbb{R}^4)} \leq \sup_{t \approx \frac{k}{R_1}} \| P_{[-\frac{1}{2} \log R_1, k]} (\delta A^{free}) f_k \partial \bar{\phi}_R \|_{L^2([-T, T] \times \mathbb{R}^4)} \leq R_1^{-M} \]

while for the last term, we get

\[
\left( \sum_k \| X_{|x| \approx \frac{k}{R}} P_{[-\frac{1}{2} \log R_1, k]} (\delta A^{free}) f_k \partial \bar{\phi}_R \|_{L^2([-T, T] \times \mathbb{R}^4)} \right)^\frac{1}{2} \leq \left( \sum_k \| P_{[-\frac{1}{2} \log R_1, k]} (\delta A^{free}) f_k \partial \bar{\phi}_R \|_{L^2([-T, T] \times \mathbb{R}^4)} \right)^\frac{1}{2} \leq R_1^{-M}.
\]

Here we use the localization properties of $\bar{\phi}_R$ to bound the second factor by

\[
\| X_{|x| \approx \frac{k}{R}} P_{[-\frac{1}{2} \log R_1, k]} \partial \bar{\phi}_R \|_{L^2([-T, T] \times \mathbb{R}^4)} \leq (2^{\max\{|k,0|\} R_1})^{-M}
\]

as long as $k > -\frac{1}{2} \log R_1$ as we may assume and also we have the crude bound

\[
\| P_{[-\frac{1}{2} \log R_1, k]} (\delta A^{free}) f_k \|_{L^2([-T, T] \times \mathbb{R}^4)} \leq 2^k,
\]

whence we finally obtain the bound

\[
\left( \sum_k \| P_{[-\frac{1}{2} \log R_1, k]} (\delta A^{free}) f_k \partial \bar{\phi}_R \|_{L^2([-T, T] \times \mathbb{R}^4)} \right)^\frac{1}{2} \leq R_1^{-M}.
\]

Letting $R_1 \to +\infty$ then again gives the required smallness. \qed

**Proof of Proposition 6.1** (Outline) In view of Lemma 5.4 it suffices to consider the case $I = \mathbb{R}$. We argue by contradiction. Assume that we have

\[
\|(A, \phi)\|_{L^1(\mathbb{R} \times \mathbb{R}^4)} < \infty.
\]
Then the idea is that using this ingredient as well as a correct perturbative ansatz for the evolutions $(A_n, \phi_n)$ for $n$ large enough, we can show that the corresponding $S^1$ norms of $(A_n, \phi_n)$ must stay finite, contradicting the assumption. We introduce the perturbative term $\delta \phi_n$ by means of

$$\phi_n = \chi_{I_1} \phi + \chi_{I_2} \tilde{\phi}_{A,\delta A_n^{\text{free}}} + \delta \phi_n,$$

where $I_1$ is a very large time interval centered around $t = 0$ and $I_2$ represents the complement. The function $\tilde{\phi}_{A,\delta A_n^{\text{free}}}$ solves the wave equation

$$\Box_{A+\delta A_n^{\text{free}}}(\tilde{\phi}_{A,\delta A_n^{\text{free}}}) = 0,$$

$$\tilde{\phi}_{A,\delta A_n^{\text{free}}}[0] = \phi[0],$$

where

$$\Box_{A+\delta A_n^{\text{free}}} = \Box + 2i(A + \delta A_n^{\text{free}}) \partial \nu$$

and in this context we let $\delta A_n^{\text{free}}$ be the actual free evolution of the data $\delta A_n[0]$ (as usual only the spatial components). We let $\chi_{I_1}, \chi_{I_2}$ be a smooth partition of unity subordinate to dilates of the intervals $I_1, I_2$. Finally, we introduce the perturbative term $\delta A_n$ for the magnetic potential by

$$A_n = A + \delta A_n.$$

We note that in this argument one has to in fact replace the energy class solution $(A, \phi)$ by the evolution of a low frequency approximation of the energy class data very close to it and then show that this implies $S^1$ norm bounds for $(A_n, \phi_n)$ uniformly for all sufficiently close low frequency approximations.

To begin with, observe that we can show similarly to the proof of Lemma 7.11, proved later independently, that given any $\gamma > 0$ and choosing $I_1$ suitably large (depending on $A, \phi, \gamma$), we can arrange that

$$\tilde{\phi}_{A,\delta A_n^{\text{free}}} = (\tilde{\phi}_{A,\delta A_n^{\text{free}}})_1 + (\tilde{\phi}_{A,\delta A_n^{\text{free}}})_2$$

with

$$\| (\tilde{\phi}_{A,\delta A_n^{\text{free}}})_1 \|_{S^1} < \gamma, \quad \| X_{I_2} (\tilde{\phi}_{A,\delta A_n^{\text{free}}})_2 \|_{L^\infty_t L^\infty_x} < \gamma.$$

Now the equation for $\delta \phi_n$ becomes the following

$$\Box_{A+\delta A_n} \delta \phi_n = -\chi_{I_1} \Box_{A+\delta A_n} \phi - \chi_{I_2} \Box_{A+\delta A_n} \tilde{\phi}_{A,\delta A_n^{\text{free}}} + (\partial_t^2 \chi_{I_1})(\phi - \tilde{\phi}_{A,\delta A_n^{\text{free}}}) + 2(\partial_t \chi_{I_1})(\partial_t \phi - \partial_t \tilde{\phi}_{A,\delta A_n^{\text{free}}}) + 2i(\partial_t \chi_{I_1})(A + \delta A_n)[0](\phi - \tilde{\phi}_{A,\delta A_n^{\text{free}}}).$$

The error term $\partial_t^2(\chi_{I_1})(\phi - \tilde{\phi}_{A,\delta A_n^{\text{free}}})$ is potentially problematic, because we cannot place the factor

$$(\phi - \tilde{\phi}_{A,\delta A_n^{\text{free}}})$$

into $L^\infty_t L^2_x$. In fact, the latter is only possible provided we have compact spatial support (precisely, in a ball of radius $R$ with $1 \ll R \leq |I_1|$) according to the Huygens principle, because then the extra factor $|I_1|^{-1}$ stemming from $\partial_t^2(\chi_{I_1})$ will counterbalance the factor $|I_1|$ in

$$\| \phi - \tilde{\phi}_{A,\delta A_n^{\text{free}} \|}_{L^\infty_t L^2_x(I_1 \times \mathbb{R}^4)} \leq |I_1| \| \nabla_x (\phi - \tilde{\phi}_{A,\delta A_n^{\text{free}}}) \|_{L^\infty_t L^2_x(I_1 \times \mathbb{R}^4)}.$$

Here it is natural to truncate the data $\phi[0]$ in physical space to force this spatial localization later in time via Huygens’ principle, but one needs to ensure that this does not destroy the good $S^1$ norm bounds for $(A, \phi)$. In fact, since we use the same $A$ data, the argument for Proposition 5.1 applies
Furthermore, by using a divisibility argument and subdividing time axis into $\mathbb{N}$, we can write

$$\Box_A (\phi - \tilde{\phi}_{A,\delta A_n^{free}}) = 2i(\delta A_n^{free})_x \partial \tilde{\phi}_{A,\delta A_n^{free}} + \ldots,$$

and further

$$\|\partial \tilde{\phi}_{A,\delta A_n^{free}}\|_N \leq C(I_1, \phi, A)\|\delta A_n[0]\|_{H^1_t L^2_x}.$$

This gives the required smallness provided we can make $\|\nabla_x (\phi - \tilde{\phi}_{A,\delta A_n^{free}})\|_{L^\infty_t L^2_x(I_1 \times \mathbb{R}^4)}$ small. For this observe that (we omit the cubic interaction terms)

$$\Box_A (\phi - \tilde{\phi}_{A,\delta A_n^{free}}) = 2i(\delta A_n^{free})_x \partial \tilde{\phi}_{A,\delta A_n^{free}} + \ldots,$$

and further

$$\|\partial \tilde{\phi}_{A,\delta A_n^{free}}\|_N \leq C(I_1, \phi, A)\|\delta A_n[0]\|_{H^1_t L^2_x}.$$

This implies

$$\|\nabla_x (\phi - \tilde{\phi}_{A,\delta A_n^{free}})\|_{L^\infty_t L^2_x(I_1 \times \mathbb{R}^4)} \leq C(I_1, \phi, A)\|\delta A_n[0]\|_{H^1_t L^2_x}.$$

and so we conclude that

$$\|\partial \tilde{\phi}_{A,\delta A_n^{free}}\|_N \leq C(I_1, \phi, A)\|\delta A_n[0]\|_{H^1_t L^2_x}.$$

Furthermore, we can write

$$\chi_{\Omega} \Box_{A+\delta A_n} \phi = -\chi_{\Omega} \Box_{A+\delta A_n^{\text{nonlin}}} \phi - \text{error},$$

where we use the decomposition

$$\delta A_n = \delta A_n^{\text{free}} + \delta A_n^{\text{nonlin}}$$

with the first term on the right hand side the free propagation of $\delta A_n[0]$. For the error term we get

$$\|\text{error}\|_N \leq C([I_1], A)\|\delta A_n[0]\|_{H^1_t L^2_x}.$$

Furthermore, by using a divisibility argument and subdividing time axis into $N((|A, \phi|), 5)$ many time intervals $I$, using the argument for Proposition 7.2, we can force (for each such $I$)

$$\|\chi_{I_1} \Box_{A+\delta A_n^{\text{nonlin}}} \phi\|_{N(I)} \ll \|\delta A_n\|_{S^1} + \|\delta A_n\|_{S^1}.$$}

Similarly, we have

$$\|\chi_{I_1} \Box_{A+\delta A_n^{\text{nonlin}}} \phi\|_{N(I)} \ll \|\delta A_n\|_{S^1} + \|\delta A_n\|_{S^1},$$

which then suffices for the bootstrap for $\delta A_n$.

Next, we consider the equation for $\delta A_n$, which is of the schematic form

$$\Box A = \phi \cdot \nabla_x \phi - (\chi_{I_1} \phi) \cdot \nabla_x (\chi_{I_1} \phi)$$

$$- x_1 (\chi_{I_1} \phi) \cdot \nabla_x (\chi_{I_2} \tilde{\phi}_{A,\delta A_n^{free}} + \delta \phi_n)$$

$$- (\chi_{I_2} \tilde{\phi}_{A,\delta A_n^{free}} + \delta \phi_n) \cdot \nabla_x (\chi_{I_1} \phi + \chi_{I_2} \tilde{\phi}_{A,\delta A_n^{free}} + \delta \phi_n)$$

$$+ (A + \delta A_n) |\chi_{I_1} \phi + \chi_{I_2} \tilde{\phi}_{A,\delta A_n^{free}} + \delta \phi_n|^2 - A|\phi|^2.$$

Then we make the following observations. The first line on the right hand side satisfies

$$\|\phi \cdot \nabla_x \phi - (\chi_{I_1} \phi) \cdot \nabla_x (\chi_{I_1} \phi)\|_N \leq \gamma_1$$

for any prescribed $\gamma_1 > 0$, provided we pick $I_1$ sufficiently large. The reason for this is that this term is supported around the endpoints of $I_1$ (which is centered around $t = 0$), and since $\Box \phi = 0$, we
obtain similarly to the proof of Lemma \[7.1\] the dispersive decay for \( \phi \) at large times, which easily gives the desired smallness for the \( N \) norm. For the second and third line on the right, we find
\[
\| (\chi_1 \phi) \cdot \nabla (\chi_1 \tilde{\phi}_{A,\delta A_n^{\text{free}} + \delta \phi_n}) \|_{N(J \times \mathbb{R}^d)} + \| (\chi_2 \tilde{\phi}_{A,\delta A_n^{\text{free}} + \delta \phi_n}) \cdot \nabla (\chi_1 \phi + \chi_2 \tilde{\phi}_{A,\delta A_n^{\text{free}} + \delta \phi_n}) \|_{N(J \times \mathbb{R}^d)} \\
\leq \nu_1 + M^{-1} \| \delta \phi_n \|_{S^1(J \times \mathbb{R}^d)} + C \| \delta \phi_n \|_{S^1(J \times \mathbb{R}^d)},
\]
where \( J \) is a member of a suitable partition of the time axis into \( N(\| (A, \phi) \|_{S^1}, M) \) many intervals and \( C \) is a universal constant. Here we exploit the uniform dispersive decay of \( \tilde{\phi}_{A,\delta A_n^{\text{free}}} \). The last line is handled similarly.

\[
\begin{align*}
\| (A + \delta A_n) \chi_1 \phi + \chi_2 \tilde{\phi}_{A,\delta A_n^{\text{free}} + \delta \phi_n} \|_{N(J \times \mathbb{R}^d)} \\
\leq \nu_1 + C_1 M^{-1} (\| \delta \phi_n \|_{S^1(J \times \mathbb{R}^d)} + \| \delta A_n \|_{S^1(J \times \mathbb{R}^d)}) + C_2 (\| \delta \phi_n \|_{S^1(J \times \mathbb{R}^d)} + \| \delta A_n \|_{S^1(J \times \mathbb{R}^d)}) \| \delta \phi_n \|_{S^1(J \times \mathbb{R}^d)},
\end{align*}
\]

which suffices to bootstrap the bound for \( \| \delta A_n \|_{S^1} \) on \( J \). The bootstrap on the remaining intervals follows by induction (and choosing \( \nu_1 \))\] and \( \| \delta A_n(0) \|_{H^1_{\text{free}}} \) sufficiently small depending on \( M \) and \( \| (A, \phi) \|_{S^1} \). Finally, we observe that the \( S^1 \) norm bounds on \( \phi, \tilde{\phi}_{A,\delta A_n^{\text{free}}} \), and \( \delta \phi_n \) are “inherited” by the expression

\[
\chi_1 \phi + \chi_2 \tilde{\phi}_{A,\delta A_n^{\text{free}} + \delta \phi_n}
\]
on account of the support properties of the functions \( \phi, \tilde{\phi}_{A,\delta A_n^{\text{free}}} \).

\[ \square \]

7.5. Selecting concentration profiles and adding the first large frequency atom. Having established control over the evolution of the data \( (A_{\Lambda_0}^{n1}, \phi_{\Lambda_0}^{n1}) \), we now add the components \( (A^{n1}, \phi^{n1}) \), i.e., we pass to the initial data

\[
(A_{\Lambda_0}^{n1} + A^{n1}, \phi_{\Lambda_0}^{n1} + \phi^{n1}).
\]

We shall show that under the assumption

\[
\liminf_{n \to \infty} E(A^{n1}, \phi^{n1}) < E_{\text{crit}}.
\]

we can extract a subsequence from the data \( \text{(7.42)} \) for which the evolution satisfies finite \( S^1 \) norm bounds. To do so we have to crucially exploit the preceding condition on the energy as well as a careful profile decomposition in physical space of the added data \( (A^{n1}, \phi^{n1}) \), following the method in \[\text{[10]}\]. In order to select the concentration profiles, we build the influence of the low frequency magnetic potential \( A_{\Lambda_0}^{n1} \) into the linear evolution and consider the operator

\[
\square_{A_{\Lambda_0}^{\text{low}}} := \square + 2iA_{\Lambda_0}^{n1} \partial^\gamma.
\]

This is to be applied to functions of frequency support essentially contained in the set \( |\xi| \sim 1 \). We have

\[ \square \]

Lemma 7.8. Assume that the Schwartz function \( \epsilon \) has frequency support at \( |\xi| \sim 1 \) and that the frequency support of \( A_{\Lambda_0}^{n1} \) is evacuated to \( |\xi| \to 0 \) as \( n \to \infty \). Assuming that \( A_{\Lambda_0}^{n1} \) satisfies a uniform \( S^1 \) norm bound

\[
\limsup_{n \to \infty} \| A_{\Lambda_0}^{n1} \|_{S^1} < +\infty
\]
and that \( \|e(0)\|_{H^1_t \times L^2_x} \leq 1 \), then the solution \( e(t, x) \) of the linear problem (with implicit dependence on \( n \) suppressed)

\[ \Box_{A^{n_0}} e = 0 \]

satisfies

\[ \lim_{n \to \infty} \sup_{t \in \mathbb{R}} \|\nabla_t x e(t, \cdot)\|_{L^2_x}^2 - \nabla_t x e(0, \cdot)\|_{L^2_x}^2 = 0. \]

Thus, we have energy conservation in an asymptotic sense.

**Proof.** We consider the natural energy functional

\[ E_{A^{n_0}}(\epsilon)(t) := \int_{\mathbb{R}^4} \left( \frac{1}{2} |\partial_t \epsilon + iA_{n_0}^{A_0} \epsilon|^2 + \frac{1}{2} \sum_{j=1}^4 |\partial_j \epsilon + iA_{n_0}^{A_0} \epsilon|^2 \right)(t, x) \, dx. \]

Here it is to be kept in mind that the potential \( A \) is in Coulomb gauge. Differentiating the energy with respect to \( t \) and using the equation for \( \epsilon \), we infer the following relation

\[ E_{A^{n_0}}(\epsilon)(t) - E_{A^{n_0}}(\epsilon)(0) = \int_{[0,t] \times \mathbb{R}^4} (\partial_t A_{n_0}^{A_0}) \epsilon (\partial_t \epsilon + iA_{n_0}^{A_0} \epsilon) \, dx \, dt \]

\[ + \sum_{j=1}^4 \int_{[0,t] \times \mathbb{R}^4} (\partial_j + iA_{n_0}^{A_0}) (\partial_j \epsilon + iA_{n_0}^{A_0} \epsilon) \, dx \, dt \]

\[ + \int_{[0,t] \times \mathbb{R}^4} \left( \sum_{j=1}^4 (A_{n_0}^{A_0} \epsilon)^2 - (A_{n_0}^{A_0} \epsilon)^2 \right) \epsilon (\partial_t \epsilon + iA_{n_0}^{A_0} \epsilon) \, dx \, dt. \]

(7.43)

We have to show that the terms on the right hand side converge to zero as \( n \to \infty \) uniformly in \( t \). Here the quartic terms (which include all terms with at least one factor \( A_{n_0}^{A_0} \) due to its inherent nonlinear structure in the Coulomb gauge) are all expected to be straightforward, and so we focus on the delicate cubic interaction term

\[ \sum_{j=1}^4 \int_{[0,t] \times \mathbb{R}^4} \partial_j \epsilon (\partial_j \epsilon + iA_{n_0}^{A_0} \epsilon) \, dx \, dt = - \sum_{j=1}^4 \int_{[0,t] \times \mathbb{R}^4} \text{Im}(\partial_j \epsilon \overline{\epsilon})(\partial_j \epsilon) \, dx \, dt. \]

Here the Coulomb condition satisfied by \( \partial_j A_{n_0}^{A_0} \) allows us to project the term \( \text{Im}(\partial_j \epsilon \overline{\epsilon}) \) onto its divergence-free part, which means that we can replace this by a null form of the schematic type

\[ \text{Im}(\partial_j \epsilon \overline{\epsilon}) \rightarrow \sum_i \Delta^{-1} \partial_i N_{ij}(\epsilon, \overline{\epsilon}). \]

Thus we reduce to bounding uniformly the following schematic integral

\[ \sum_{i,j} \int_{[0,t] \times \mathbb{R}^4} \Delta^{-1} \partial_i N_{ij}(\epsilon, \overline{\epsilon}) \partial_j A_{n_0}^{A_0} \, dx \, dt. \]

Now we claim the microlocalized bound

\[ \left| \int_{[0,t] \times \mathbb{R}^4} \Delta^{-1} \partial_i N_{ij}(P_{k_1} \epsilon, \overline{P_{k_1} \epsilon}) P_{k_1} \partial_j A_{n_0}^{A_0} \, dx \, dt \right| \leq 2^{\sigma} (\min|k_1, k_2, k_3| - \max|k_1, k_2, k_3|) \|P_{k_1} \epsilon\|_S^1 \|P_{k_1} \epsilon\|_S^1 \|A_{n_0}^{A_0}\|_S^1 \]

for suitable \( \sigma > 0 \). Since there are at least two comparable frequencies in the above, this is enough to give the desired result in view of the frequency localizations of \( \epsilon \) and \( A_{n_0}^{A_0} \). In order to prove this,
we localize the above expression further and also omit the localization to $[0, t]$, as we may get rid of it via a suitable cutoff (which is compatible with the $S^1$ norms),

$$
\int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt = \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt \\
+ \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt,
$$

where the second term on the right is the more difficult one. Then write this term as

$$
\int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt = \sum_{j < k_0} \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt.
$$

By symmetry we may assume $k_2 \leq k_1$. Then we distinguish the following cases.

**Case 1:** $k_1 = k_2 + O(1) > k_3 + O(1)$. Since the $Q_j$ transfers to the null form $N_{ij}$, we save

$$
2^{k_3-k_2+\frac{1}{2}(j-k_3)}.
$$

Then we obtain

$$
\left| \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt \right| \\
\leq 2^{k_3-k_2+\frac{1}{2}(j-k_3)} 2^{k_1} \sum_{j=1,2} \| P_{k_0} \nabla_x \|^2_{L^2_T L^2_x} \| P_{k_0} Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \|_{L^2_T L^2_x},
$$

where we observe that the exponent pair $(4, 3)$ is Strichartz admissible in four space dimensions. Then we use the improved Sobolev type bound

$$
\| P_{k_0} Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \|_{L^2_T L^2_x} \leq 2^{\frac{2}{3} k_3 \gamma(j-k_3)} \| P_{k_0} Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \|_{L^2_T L^2_x}
$$

for suitable $\gamma > 0$, to infer

$$
\left| \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(p_{k_1}, p_{k_2}, p_k) Q_{k_0} \partial_t A_{k_0, i,j}^{n_1} \, dx \, dt \right| \\
\leq 2^{k_3-k_2+\frac{1}{2}(j-k_3)} 2^{k_1} 2^{\frac{2}{3} k_3} \frac{1}{2} 2^{\frac{2}{3} k_3} 2^{\frac{1}{2} k_3} 2^{\gamma(j-k_3)} \| P_{k_0} \nabla_x \|^2_{L^2_T L^2_x} \| P_{k_0} \partial_t A_{k_0, i,j}^{n_1} \|_{L^2_T L^2_x},
$$

which in turn can be bounded by

$$
\leq 2^{\frac{2}{3} k_3 \gamma(j-k_3)} \| P_{k_0} \nabla_x \|^2_{L^2_T L^2_x} \| P_{k_0} A_{k_0, i,j}^{n_1} \|_{L^2_T L^2_x},
$$

Summing over $j < k_3$ yields the desired bound in this case.

**Case 2:** $k_1 = k_3 + O(1) > k_2$. We distinguish between the cases $j < k_2$ and $j \geq k_2$. 
Case 2a: $j < k_2$. Here we estimate

\[
\left| \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(P_{k_1} e, \overline{P_{k_2} e}) P_{k_3} Q_j \partial_t A_{\lambda_0,j}^{n_1} \, dx \, dt \right| 
\leq 2^{-k_1} 2^{\frac{1}{2}(j-k_2)} \left\| \nabla_{t,x} P_{k_3} e \right\|_{L^2_{t,L^2_x}} \left\| \nabla_{t,x} P_{k_3} Q_j \partial_t A_{\lambda_0,j}^{n_1} \right\|_{L^2_{t,L^2_x}} 
\leq 2^{-\frac{k_1}{2}j} 2^{\frac{k_2}{2}(j-k_2)} 2^{-\frac{j}{2}} \prod_{j=1,2} \left\| P_{k_j} e \right\|_{L^2} \left\| P_{k_3} A_{\lambda_0,j}^{n_1} \right\|_{L^2}.
\]

To get summability over $j$ one can replace the norm $\| \cdot \|_{L^2_{t,L^2_x}}$ by $\| \cdot \|_{L^2_{t,L^2_x}^2}$ and then use $\| \cdot \|_{L^2_{t,L^2_x}^2}$ instead for the second factor.

Case 2b: $j \geq k_2$. Here we simply get the bound

\[
\left| \int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_t N_{ij}(P_{k_1} e, \overline{P_{k_2} e}) P_{k_3} Q_j \partial_t A_{\lambda_0,j}^{n_1} \, dx \, dt \right| \leq 2^{\frac{k_1}{2}(k_2-k_1)} \prod_{j=1,2} \left\| P_{k_j} e \right\|_{L^2} \left\| P_{k_3} A_{\lambda_0,j}^{n_1} \right\|_{L^2},
\]

which can then be summed over $k_1 > j \geq k_2$ to give the desired bound. This in essence completes the proof of the lemma.

Our next key step is to extract the concentration profiles by considering the solutions of the low-frequency “covariant” linear wave equation.

**Definition 7.9.** Given initial data $\phi^{n_1}[0] := (\phi^{n_1}(0, \cdot), \partial_t \phi^{n_1}(0, \cdot))$ at time $t = 0$, we denote by

\[ S_{A^{n,\text{low}}}(\phi^{n_1}[0]) \]

the solution of the initial value problem

\[
\Box u + 2iA^{n_1}_{\lambda_0,\nu} \partial^n u = 0, 
\]

\[ u[0] = \phi^{n_1}[0]. \]

Following [10], which in turn mimics [11], we introduce the set $\mathcal{U}_{A^{n,\text{low}}}(\phi^{n_1}[0])$, which consists of all functions that can be extracted as weak limits in the following fashion

\[ \mathcal{U}_{A^{n,\text{low}}}(\phi^{n_1}[0]) = \{ V \in L^2_{t,\text{loc}} H^1_x \cap C^1 L^2_x : \exists \{(t_n, x_n)\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^4 \text{ s.t.} \]

\[ S_{A^{n,\text{low}}}(\phi^{n_1}[0])(t + t_n, x + x_n) \to V(t,x) \}. \]

These limit functions are then distributional solutions of $\Box V = 0$, the standard flat wave equation on $\mathbb{R}^{1+4}$. We emphasize that the sequences $\{(t_n, x_n)\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^4$ are completely arbitrary. In light of the a priori energy bound, the following quantity is well-defined

\[ \eta_{A^{n,\text{low}}}(\phi^{n_1}[0]) := \sup \{ E_0(V) : V \in \mathcal{U}_{A^{n,\text{low}}}(\phi^{n_1}[0]) \} < \infty, \]

where $E_0$ refers to the standard flat energy

\[ E_0(V) = \int_{\mathbb{R}^4} \left| \nabla_{t,x} V \right|^2 \, dx. \]

The key result of this subsection is then contained in the following proposition. To clarify the notation and make it adapted to the ensuing induction procedure, we replace the superscript 1 in $\phi^{n_1}$ by $a$, indicating the frequency level of the large frequency atom.
**Proposition 7.10.** There exists a collection of sequences \( \{(t_{ab}^n, x_{ab}^n)\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^4 \), \( b \geq 1 \), as well as a corresponding family of concentration profiles

\[
\phi^{ab}[0] \in \dot{H}^1_\Lambda(\mathbb{R}^4) \times L^2(\mathbb{R}^4), \quad b \geq 1,
\]

with the following properties: Introducing the space-time translated gauge potentials

\[
\tilde{A}_{y}^{ab, low}(t, x) := A_{\Lambda_0, y}^{na}(t + t_{ab}^n, x + x_{ab}^n),
\]

we have

1. For any \( B \geq 1 \), we have a decomposition

\[
S_{A^{ab, low}}(\phi^{na}[0])(t, x) = \sum_{b=1}^B S_{A^{ab, low}}(\phi^{ab}[0])(t - t_{ab}^n, x - x_{ab}^n) + \phi^{naB}(t, x),
\]

where each function

\[
S_{A^{ab, low}}(\phi^{ab}[0])(t - t_{ab}^n, x - x_{ab}^n), \quad \phi^{naB}(t, x)
\]

solves the covariant wave equation

\[
\Box_{A^{ab, low}} u = 0.
\]

Moreover, the error satisfies the crucial asymptotic vanishing condition

\[(7.44) \quad \lim_{B \to \infty} \limsup_{n \to \infty} \eta_{A^{ab, low}}(\phi^{naB}[0]) = 0.\]

2. The sequences are mutually divergent, meaning that for \( b \neq b' \) we have

\[
\lim_{n \to \infty} (|t_{ab}^n - t_{a'b'}^n| + |x_{ab}^n - x_{a'b'}^n|) = +\infty.
\]

3. There is asymptotic energy partition

\[
E_0(\phi^{na}[0]) = \sum_{b=1}^B E_0(\phi^{ab}[0]) + E_0(\phi^{naB}[0]) + o(1),
\]

where the meaning of \( o(1) \) here is \( \lim_{B \to \infty} \limsup_{n \to \infty} o(1) = 0 \).

4. All profiles \( \phi^{ab}[0] \) as well as all errors \( \phi^{naB}[0] \) are 1-oscillatory.

The proof of this proposition is essentially contained in \[10\] Lemma 9.23. In order to control the interactions of the profiles to be discussed below, we also need the following crucial uniform dispersive type bound. Note that this is an analogue of Proposition 9.20 in \[10\] and is proved in an analogous fashion.

**Lemma 7.11.** Let \( \phi[0] \in \dot{H}^1_\Lambda(\mathbb{R}^4) \times L^2(\mathbb{R}^4) \) be fixed initial data and consider the solution \( \phi(t, x) \) of the linear problem

\[
\Box_{A^{ab, low}} \phi = 0
\]

with given data \( \phi[0] \) at time \( t = 0 \). Then for any \( \gamma > 0 \), there exists a decomposition

\[
\phi = \phi_1 + \phi_2
\]

such that

\[
\|\phi_2\|_{L^1} < \gamma
\]

and there exists a time \( t_0 = t_0(\phi[0], \gamma) \) such that for any \( |t| > t_0 \),

\[
\|\phi_1(t, \cdot)\|_{L^\infty} < \gamma.
\]
The new aspect of our setting is that the propagators \( \tilde{\psi} \) data of the special form (0 inhomogeneous Duhamel parametrix by to the situation treated in [10], by using the error analysis in [12]. Thus, denote the approximate convergent sums of further terms, which need to be analyzed. Our strategy is to reduce precisely time intervals. In fact, due to the summation over \( k \in [-\frac{C}{\sigma}, 0] \), it is seen that this number of intervals needs to be proportional to \( C_1 \) (and also depends on the energy and \( \sigma \), of course). Now we formally denote the (exact) Duhamel propagator for the equation \( \Box_A u = F \) by

\[
\begin{align*}
\text{Proof. (Outline)} \quad & \text{We first consider the case of the general problem} \\
\Box_A^p u = \Box u + \sum_k 2iA_{j,k}^{\text{free}} \partial^j u_k = 0, \quad u[0] = (f, g).
\end{align*}
\]

These can only be iterated away by using divisibility, i.e. by restricting to a finite number of suitable time intervals. In fact, due to the summation over \( k \in [-\frac{C}{\sigma}, 0] \), it is seen that this number of intervals needs to be proportional to \( C_1 \) (and also depends on the energy and \( \sigma \), of course). Now we formally denote the (exact) Duhamel propagator for the equation \( \Box_A u = F \) by

\[
\int_0^t \tilde{U}(t-s)F(s) \, ds.
\]

Also, we denote the intervals on which the error terms \( \sum_l \sum_{k\leq l} 2iA_j^{\text{free}} \partial^j u_l =: N^{\text{th}}(u_l) \) as well as the remaining errors generated by the parametric \( \tilde{U} \) and which need to be handled by divisibility, by

\[
J_0, J_1, J_2, \ldots, J_N.
\]

Assume that \( J_0 \) is the time slice containing the initial time \( t = 0 \), and also assume as we may that the intervals \( J_i \) are consecutive, with \( J_N = [t_N, \infty) \). Write \( J_i = [t_i, t_{i+1}] \), \( i \leq N - 1 \). As observed before, their number depends linearly on \( C_1 \), as well as implicitly on the energy and \( \sigma \). Then, proceeding by exact analogy to the proof of Proposition 9.20 in [10], we can write for \( u^{(j)} := u_{J_i} \),

\[
\begin{align*}
\sum_{i=1}^{\infty} u^{(J_i,j)}, \quad u^{(J,0)} &= \tilde{S}(t-t_i)u^{(j-1)}(t_i), \\
u^{(j,j)} &= \int_{t_i}^t \tilde{U}(t-s)N^{\text{th}}(u^{(j,j-1)})(s) \, ds.
\end{align*}
\]

Here, \( \tilde{S} \) is the homogeneous data propagator for \( \Box_A^p \), while \( \tilde{U} \) is the homogeneous propagator for data of the special form \( (0, g) \). Then the inductive nature of the construction is revealed by the relation (see [10])

\[
\begin{align*}
u^{(j,0)} = \tilde{S}(t,f,g) + \sum_{k=0}^{i-1} 2i \sum_{l=1}^{\infty} \int_{t_k}^{t_{l+1}} \tilde{U}(t-s)N^{\text{th}}(u^{(j,j-1)})(s) \, ds.
\end{align*}
\]

The new aspect of our setting is that the propagators \( \tilde{U}, \tilde{S} \) themselves are only obtained as infinite convergent sums of further terms, which need to be analyzed. Our strategy is to reduce precisely to the situation treated in [10], by using the error analysis in [12]. Thus, denote the approximate inhomogeneous Duhamel parametrix by

\[
\int_0^t \tilde{U}^{(app)}(t-s)F(s) \, ds.
\]
Note that due to Proposition 7 in [12], the parametrix $\tilde{U}^{(app)}(t-s)$ is given by an integral kernel which satisfies the same standard decay estimates as the standard d’Alembertian propagator, independent of the precise potential $A^{free}$ used (but with implicit constants depending on its energy, of course). Then recall from the proof of Theorem 4 in [12] that we may write
\[ \int_0^\infty \tilde{U}(t-s)F(s)\,ds = \sum_{j=0}^\infty \int_0^\infty \tilde{U}^{(app)}(t-s)F_j(s)\,ds, \]
where we have $F_0 = F$, as well as inductively, writing
\[ \phi_j := \int_0^\infty \tilde{U}^{(app)}(t-s)F_j(s)\,ds, \]
we have for $j \geq 1$
\[ F_j = F_j^1 + F_j^2 + F_j^3 + F_j^4, \]
with (schematic notation following [12])
\[ P_0F_j^1 = [\square_{A<0} e^{-i\phi_x} (t, xD) - e^{-i\phi_x} (t, xD)\square]P_0\phi_{j-1}, \]
\[ P_0F_j^2 = \frac{1}{2}[e^{-i\phi_x} (t, x, D)e^{i\phi_x} (D, y, t) - 1]P_0F_{j-1}, \]
\[ P_0F_j^3 = \frac{1}{2}[e^{-i\phi_x} (t, x, D)|D|^{-1} e^{i\phi_x} (D, y, t) - |D|^{-1}]P_0\partial_t F_{j-1}, \]
\[ P_0F_j^4 = \frac{1}{2}[e^{-i\phi_x} (t, x, D)|D|^{-1} \partial_t e^{i\phi_x} (D, y, t) - |D|^{-1}]P_0F_{j-1}. \]

Here the first term, which is treated in Section 10.2 in [12], gains a smallness factor of the form $2^{-\alpha C_1}$, which of course overwhelms any losses polynomial in $C_1$ for $C_1 \gg 1$. However, the remaining three terms do not gain smallness from $C_1$, but rather by divisibility, and so we have to be more careful to force smallness for them (we cannot make the number of intervals depend on the prescribed smallness threshold $\gamma$). Here we exploit the fact that due to Proposition 6 in [12], the kernels of the operators
\[ \frac{1}{2}[e^{-i\phi_x} (t, x, D)e^{i\phi_x} (D, y, t) - 1], \quad \frac{1}{2}[e^{-i\phi_x} (t, x, D)|D|^{-1} e^{i\phi_x} (D, y, t) - |D|^{-1}], \]
\[ \frac{1}{2}[e^{-i\phi_x} (t, x, D)|D|^{-1} \partial_t e^{i\phi_x} (D, y, t) - |D|^{-1}] \]
are rapidly decaying away from the diagonal $x = y$. This means that up to small errors (which may be incorporated into the small energy part of $u$), we may think of these operators as local ones, and then the estimates in the proof of Proposition 9.20 in [10] which rely on the inductive bound (9.81) there, go through for the error terms $F_{j-1}^r, r = 2, 3, 4$, as long as $F_{j-1}^r, r = 2, 3, 4$, satisfy these bounds. This means that the inductive argument in [10] goes through here as well.

This completes the selection of the linear concentration profiles for the $\phi$-field. It remains to pick corresponding profiles for the magnetic potential components $A_j^{n_{ab}}$. In fact, for the latter, we simply use the standard Bahouri-Gérard method to extract the profiles via the free wave evolution. Correspondingly, we quote

**Proposition 7.12.** There exists a collection of sequences $\{(\eta_{ab}^n, x_{ab}^n)\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^4, b \geq 1$, as well as a corresponding family of concentration profiles
\[ A_j^{n_{ab}}[0] \in H^1_4(\mathbb{R}^4) \times L^2_4(\mathbb{R}^4), \quad b \geq 1 \]
for $j = 1, 2, 3, 4$ with the following properties:
For any $B \geq 1$, we have a decomposition
\[ S(A_j^{na}[0])(t, x) = \sum_{b=1}^{B} S(A_j^{ab}[0])(t - t_{n}^{ab}, x - x_{n}^{ab}) + A_j^{naB}(t, x). \]

Each of the functions
\[ S(A_j^{ab}[0])(t - t_{n}^{ab}, x - x_{n}^{ab}), \quad A_j^{naB}(t, x) \]
solves the linear wave equation
\[ \square u = 0. \]

Moreover, the error satisfies the crucial asymptotic vanishing condition
\[ (7.45) \lim_{B \to \infty} \limsup_{n \to \infty} \eta(A_j^{naB}[0]) = 0. \]

• The sequences are mutually divergent, meaning that for $b \neq b'$ we have
\[ \lim_{n \to \infty} (|t_{n}^{ab} - t_{n}^{ab'}| + |x_{n}^{ab} - x_{n}^{ab'}|) = +\infty. \]

• There is asymptotic energy partition
\[ E_0(A_j^{na}[0]) = \sum_{b=1}^{B} E_0(A_j^{ab}[0]) + E_0(A_j^{naB}[0]) + o(1), \]
where the meaning of $o(1)$ is $\lim_{B \to \infty} \limsup_{n \to \infty} o(1) = 0$.

• All profiles $A_j^{ab}[0]$ as well as all errors $A_j^{naB}[0]$ are 1-oscillatory. Moreover, they all satisfy the Coulomb condition.

7.5.1. Construction of the nonlinear concentration profiles: preparations. Observe that the linear concentration profiles for $\phi, A$ exhibited in the previous subsection satisfy an asymptotic orthogonality relation with respect to the standard free energy. However, for the nonlinear problem, the energy involves a nonlinear interaction between the $\phi$ and $A$-fields, and it is this energy that we have to use. Thus to begin with, we have to carefully analyze the asymptotic orthogonality relations of the linear profiles with respect to this proper energy functional. We start by making a crucial distinction between two possible types of profiles:

• Temporally bounded profiles: Those profiles for which
\[ \liminf_{n \to \infty} |t_{n}^{ab}| < +\infty. \]

By passing to a subsequence we may as well assume
\[ t_{n}^{ab} = 0 \forall n. \]

For two distinct such profiles corresponding to $b, b'$, we must have
\[ \lim_{n \to \infty} |x_{n}^{ab} - x_{n}^{ab'}| = +\infty. \]

• Temporally unbounded profiles: Those profiles for which
\[ \lim_{n \to \infty} |t_{n}^{ab}| = +\infty. \]

We observe that by Lemma 7.11 (and its simple analogue for the free evolution used for the $A_j$), the temporally unbounded profiles are in a suitable sense “small” at time $t = 0$. The approximate orthogonality we need is then expressed by the following lemma. Here as usual we determine the temporal components $A_0^{nab}(0)$ by the elliptic compatibility equation in terms of $\hat{\phi}^{nab}[0]$, and similarly for $A_0^{nab'}(0)$.
Lemma 7.13. Given any \( \delta_4 > 0 \), there exists \( B_0 = B_0(\delta_4) \) with the following property\(^1\). We have

\[
\limsup_{n \to \infty} \left| E(A^{n}, \phi^{na}) - \sum_{b=1}^{B_0} E(\tilde{A}^{nab}, \tilde{\phi}^{nab}) - E(A^{nab_0}, \phi^{nab_0}) \right| < \delta_4,
\]

where \( E \) refers to the energy functional of the MKG system and where we denote

\[
\tilde{A}^{nab} := S(A_j^{ab}[0])(0 - \phi^{ab}(x - \hat{x}^n_{ab})), \quad \tilde{\phi}^{nab} := S_{\tilde{A}^{nab},low}(\phi^{ab}[0])(0 - \phi^{ab}(x - \hat{x}^n_{ab})).
\]

In particular, if there are at least two concentration profiles \( \phi^{ab}[0] \) (corresponding to two distinct values of \( b \)), then there exists \( \delta > 0 \) such that for all \( b \),

\[
\limsup_{n \to \infty} E(\tilde{A}^{nab}, \tilde{\phi}^{nab}) < E_{\text{crit}} - \delta.
\]

Proof. We check the various interaction terms and show that they become small when choosing \( B_0 \) as well as \( n \) sufficiently large.

1. **Two temporally bounded profiles.** This is straightforward since \( \lim_{n \to \infty} |x^a_n - x^b_n| = \infty \). In fact, we immediately infer that

\[
\lim_{n \to \infty} \sum_{\text{temporally bounded profiles, } b \neq b'} \left| \int_{\mathbb{R}^4} \text{Re} \left( (\partial_\alpha \tilde{\phi}^{nab} + i \tilde{A}^{nab} \tilde{\phi}^{nab}) \cdot (\partial_\alpha \tilde{\phi}^{nab'} + i \tilde{A}^{nab'} \tilde{\phi}^{nab'}) \right) dx \right|
\]

\[
+ \lim_{n \to \infty} \sum_{\text{temporally bounded profiles, } b \neq b'} 4 \left| \int_{\mathbb{R}^4} \nabla_{t,x} \tilde{A}^{nab} \cdot \nabla_{t,x} \tilde{A}^{nab'} \right| dx
\]

\[
+ \lim_{n \to \infty} \sum_{\text{temporally bounded profiles, } b \neq b'} 4 \left| \int_{\mathbb{R}^4} \nabla_{x} A^{nab} \cdot \nabla_{x} A^{nab'} \right| dx
\]

\[
= 0.
\]

2. **One temporally bounded and one temporally unbounded profile.** Here we exploit that the amplitude of the temporally unbounded profile vanishes asymptotically (at time \( t = 0 \)) as \( n \to \infty \), while the temporally bounded profile has bounded support. We conclude

\[
\lim_{n \to \infty} \sum_{b \text{ temporally bounded}} \left| \int_{\mathbb{R}^4} \text{Re} \left( (\partial_\alpha \tilde{\phi}^{nab} + i \tilde{A}^{nab} \tilde{\phi}^{nab}) \cdot (\partial_\alpha \tilde{\phi}^{nab'} + i \tilde{A}^{nab'} \tilde{\phi}^{nab'}) \right) dx \right|
\]

\[
+ \lim_{n \to \infty} \sum_{b \text{ temporally bounded}} 4 \left| \int_{\mathbb{R}^4} \nabla_{t,x} \tilde{A}^{nab} \cdot \nabla_{t,x} \tilde{A}^{nab'} \right| dx
\]

\[
+ \lim_{n \to \infty} \sum_{b \text{ temporally bounded}} 4 \left| \int_{\mathbb{R}^4} \nabla_{x} A^{nab} \cdot \nabla_{x} A^{nab'} \right| dx
\]

\[
= 0.
\]

3. **Two temporally unbounded profiles.** Here we exploit the asymptotic energy conservation and that the functions

\[
\phi^{ab}(x), \quad S_{\tilde{A}^{nab},low}(\phi^{ab}[0])(x^{ab}_n - x^{ab}_n)
\]

are asymptotically orthogonal.

\(^1\)The \( B_0 \) also depends on the sequence of linear concentration profiles, but we omit this dependency here.
(4) Weakly small error $\phi^{naB_0}$ and temporally bounded profile. This is handled like the interaction of a temporally bounded and a temporally unbounded profile. One uses the fact that we get

$$\phi^{naB_0} = \phi_1^{naB_0} + \phi_2^{naB_0},$$

where we have the bounds

$$\|\phi_1^{naB_0}\|_{L^\infty_t L^\infty_x} < \delta_4, \|\nabla_{t,x} \phi_2^{naB_0}\|_{L^\infty_t L^2_x} < \delta_4,$$

provided $B_0$ is sufficiently large. Of course, choosing $B_0$ large means that more and more interactions have to be controlled, and we can no longer simply use the choice of extremely large $n$ to “asymptotically kill” all such interactions as in the preceding cases. Thus, one has to argue carefully as follows: Given $\delta_4 > 0$, we pick $\tilde{B}_0$ sufficiently large such that for any $B \geq \tilde{B}_0$, we have

$$\limsup_{n \to \infty} \sum_{b=\tilde{B}_0}^B E(\tilde{\phi}^{nab}) + E(A^{nab}) \ll \delta_4,$$

where $E_0$ indicates the standard free energy. Then, passing to the interaction terms in $E(A, \phi)$ corresponding to $\phi^{naB_0}$ and $A^{naB_0}$ with the sum

$$\sum_{b=\tilde{B}_0}^B \tilde{\phi}^{nab}, \sum_{b=\tilde{B}_0}^B A^{nab}$$

leads to terms bounded by $\ll \delta_4$ for any $B_0 \geq \tilde{B}_0$, provided $n$ is chosen sufficiently large (depending on $B_0$). But then picking $\tilde{B}_0$ large enough, we can also ensure that the sum of all the interactions in $E(A, \phi)$ generated by the profiles $\tilde{\phi}^{nab}, A^{nab}, 1 \leq b \leq \tilde{B}_0$ are small, since $B_0 \geq \tilde{B}_0$ can be chosen independently. □

The preceding choice of concentration profiles

$$\tilde{\phi}^{nab} = S_{A_{nab,0}}(\phi^{nab}[0])(0 - t_n^{ab}, x - x_n^{ab})$$

for the $\phi$-concentration profiles, while adequate for the purposes of the critical wave maps problem treated in [10], is unfortunately still not quite the right choice for the critical Maxwell-Klein-Gordon problem. The reason is that the interaction term

$$A_v^{n1,\nu} \phi^{n1}$$

where both factors are essentially supported at frequency $\sim 1$, cannot be bounded due to the contribution of the free term $A_v^{n1,\text{free}}$. Thus, the $\phi$-field experiences not only an “asymptotic” twisting due to the contribution of the extremely low frequency components $A_{\Lambda_0}^{n1}$ (as is the case for wave maps), but also from the frequency $\sim 1$ field $A_v^{n1,\text{free}}$. This needs to be reflected by our choice of concentration profiles.

Finally, we now introduce the correct equation for the concentration profiles as (7.47)

$$\Box_{A^{n1,\text{true}}} u = 0,$$

where

$$\Box_{A^{n1,\text{true}}} = \Box + 2i(A_v^{n1} + A_v^{n1,\text{free}})\partial^\nu$$

and the functions $A_v^{n1,\text{free}}$ are defined by

$$\Box A_v^{n1,\text{free}} = 0,$$

$$A_v^{n1,\text{free}}[0] = A_v^{n1}[0]$$
for \( j = 1, 2, 3, 4 \), while we also simply put
\[
A^{n, free}_0 := 0.
\]
As it turns out, this modified equation still allows for an analogue of the crucial Proposition 7.10 given in the following proposition. To formulate it, we introduce the notation
\[
\tilde{A}^{n, true}_{ab}(t, x) := \tilde{A}^{n, low}_{ab}(t, x) + A^{n, free}_a(t + t_n^{ab}, x + x_n^{ab}),
\]
while the quantity \( \eta_{A^{n, true}} \) is defined like the quantity \( \eta_{A^{n, low}} \), but uses the operator \( \Box_{A^{n, true}} \) instead.

**Proposition 7.14.** There exists a collection of sequences \( \{(t_n^{ab}, x_n^{ab})\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^4, b \geq 1 \), as well as a corresponding family of concentration profiles
\[
\phi_n^{ab}[0] \in H^1_0(\mathbb{R}^4) \times L^2(\mathbb{R}^4), b \geq 1,
\]
with the following properties: Introducing the space-time translated gauge potentials
\[
\tilde{A}^{n, true}_{ab}(t, x) := A^{n, ab}_{a00}(t + t_n^{ab}, x + x_n^{ab}) + \tilde{A}^{n, true}_{a1}(t + t_n^{ab}, x + x_n^{ab}),
\]
we have
- For any \( B \geq 1 \), we have a decomposition
\[
S_{A^{n, true}}(\phi^{n, ab}[0])(t, x) = \sum_{b=1}^{B} S_{\tilde{A}^{n, true}_{ab}}(\phi^{n, ab}[0])(t - t_n^{ab}, x - x_n^{ab}) + \phi^{n, ab}(t, x).
\]
Each of the functions
\[
\tilde{\phi}^{n, ab} := S_{\tilde{A}^{n, true}_{ab}}(\phi^{n, ab}[0])(t - t_n^{ab}, x - x_n^{ab}), \quad \phi^{n, ab}(t, x)
\]
solves the covariant wave equation
\[
\Box_{A^{n, true}} u = 0.
\]
Moreover, the error satisfies the crucial asymptotic vanishing condition
\[
(7.48) \quad \lim_{B \to \infty} \limsup_{n \to \infty} \eta_{A^{n, true}}(\phi^{n, ab}[0]) = 0.
\]
- The sequences are mutually divergent, by which we mean that for \( b \neq b' \),
\[
\lim_{n \to \infty} \left[ |t_n^{ab} - t_n^{ab'}| + |x_n^{ab} - x_n^{ab'}| \right] = +\infty.
\]
- There is asymptotic energy partition
\[
E_0(\phi^{n, ab}[0]) = \sum_{b=1}^{B} E_0(\tilde{\phi}^{n, ab}[0]) + E_0(\phi^{n, ab}[0]) + o(1),
\]
where the meaning of \( o(1) \) here is \( \lim_{B \to \infty} \limsup_{n \to \infty} o(1) = 0 \).
- All profiles \( \phi^{n, ab}[0] \) as well as all errors \( \phi^{n, ab}[0] \) are 1-oscillatory.

The proof follows again exactly along the lines of the one of Lemma 9.23 in [10] except that one has to use the following lemma instead of the (analogue of the) stronger Lemma 7.8 that was used in [10]. To understand the relation of the following lemma with Proposition 7.14 we observe that the crucial Lemma 7.8 is bound to fail for the more general operator \( \Box_{A^{n, true}} \). This forces us to modify the orthogonality statement at the end of the preceding proposition, so that we cannot assert a sharp energy bound for the actual profiles \( \phi^{n, ab}[0] \) of temporally unbounded character. Fortunately, this will not doom the construction of the nonlinear concentration profiles, because there is a kind of “asymptotic orthogonality statement” that will allow us to circumvent the problem, and which also immediately implies the energy partition of the preceding proposition.
Lemma 7.15. Assume that the Schwartz $\epsilon$ has frequency at $|\xi| \sim 1$ with $\|\epsilon(0)\|_{H^1_t L^2_x} \leq 1$. Let $A^{n,true}$ be defined as above. Moreover, assume that $A^{n^1}_{A^0,\nu}$ satisfies

$$\limsup_{n \to \infty} \left\| A^{n^1}_{A^0} \right\|_{S^1} < \infty$$

as well as $\sup_n \left\| A^{n^1}_n(0, \cdot) \right\|_{H^1_t} < \infty$, and that $A^{n^1}$ is 1-oscillatory. Then the solution $\epsilon(t, x)$ of the linear problem (with implicit $n$ dependence suppressed)

$$\Box A^{true} \epsilon = 0$$

satisfies

$$\lim_{R \to +\infty} \limsup_{n \to \infty} \sup_{t \in \mathbb{R}_+} \int_{\mathbb{R}^4} \left| \nabla_{t,x} \epsilon(R + t, \cdot) \right|^2_{L^2_t} - \left| \nabla_{t,x} \epsilon(R, \cdot) \right|^2_{L^2_t} = 0.$$

The same holds even when replacing $+\infty$ by $-\infty$ and $\mathbb{R}_+$ by $\mathbb{R}_-$. Thus, we again have energy conservation in an asymptotic sense. Furthermore, assume that $\tilde{\epsilon}_k$ is a sequence of functions satisfying (again suppressing the $n$-dependence of $A^{true}$)

$$\Box A^{true} \tilde{\epsilon}_k = 0,$$

supported at frequency $|\xi| \sim 1$ (in the sense of 1-oscillatory), and satisfying $S^1$ norm bounds uniform in $k$, while $\epsilon$ is as above (with fixed profile $\epsilon(0)$). Then we have

$$\lim_{R \to +\infty} \limsup_{n \to \infty} \sup_{t \in \mathbb{R}_+} \left| \int_{\mathbb{R}^4} (\nabla_{t,x} \epsilon(t + R, x) \cdot \nabla_{t,x} \tilde{\epsilon}_k(t + R, x) - \nabla_{t,x} \epsilon(R, x) \cdot \nabla_{t,x} \tilde{\epsilon}_k(R, x)) \, dx \right| = 0$$

uniformly in $k$.

Proof. This follows the proof of Lemma 7.8. The issue is how to deal with the interactions of $\epsilon$ and $A^{n,free}$, which are now both 1-oscillatory. Consider an interaction term

$$\int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_i N_{ij}(P_{k_1} \epsilon, P_{k_2} \overline{\epsilon}) P_{k_3} \partial_i A_{j}^{n,free} \, dx \, dt.$$  

Here we may assume that all frequencies $2^{k_{1,2,3}} \sim 1$, otherwise smallness follows from the proof of Lemma 7.8. Then one uses that Lemma 7.11 applies when one replaces $A^{n,low}$ by $A^{n,true}$; in fact, this exact same proof works. This means that restricting integration to a space-time slice $[R, R + t] \times \mathbb{R}^4$ with $R \gg 1$, say, we obtain that

$$\left\| P_{k_1,2} \epsilon \right\|_{L^2_t L^{\infty}_{x}([R, R + t] \times \mathbb{R}^4)} \ll 1$$

uniformly for $t \geq 0$ and $n$ by interpolation with the endpoint Strichartz estimate. On the other hand, for the other factor $P_{k_3} \partial_i A_{j}^{n,free}$, one can use $L^6_t L^{\frac{3}{2}}_x$ instead. The last statement of the lemma follows similarly, by expressing the inner product in terms of the energies of $\epsilon$ and $\epsilon_k$, and reducing to bounding expressions such as

$$\int_{\mathbb{R}^{1+4}} \Delta^{-1} \partial_i N_{ij}(P_{k_1} \epsilon, P_{k_2} \overline{\epsilon}) P_{k_3} \partial_i A_{j}^{n,free} \, dx \, dt.$$  

To take advantage of the preceding lemma, assume that $\phi^{ab}$ is a temporally unbounded profile, say with $\lim_{n \to \infty} t_{n}^{ab} = +\infty$. Then instead of working with this profile at time $t_{n}^{ab}$, we pick $R_b$ very large and instead consider the profile at time $t_{n}^{ab} - R_b$, which is given by the data

$$\tilde{\phi}^{ab}(t_{n}^{ab} - R_b) = S_{\chi^{ab,free}}(\phi^{ab}[0])(-R_b)(\chi_{n}^{ab}).$$
The preceding lemma implies the asymptotic energy conservation
\[ \lim_{R_b \to \infty} E_0(\tilde{\phi}_{nab}^n [t_n^ab - R_b]) = E_0(\tilde{\phi}_{nab}^n [0]). \]

Also, importantly, we observe that while at first sight the profile \( \tilde{\phi}_{nab}^n [t_n^ab - R_b] \) depends on \( n \), we note that we have
\[ \tilde{\phi}_{nab}^n [t_n^ab - R_b] = S_{\tilde{\phi}_{nab}^n} \cdot -x_n^ab + o(1), \]
where \( A^ab \) is the free wave evolution associated with the profile \( A^ab [0] \) in the profile decomposition for \( A^na \) as in Proposition 7.12, and the error \( o(1) \) is in the energy sense as \( n \to \infty \). Using the asymptotic orthogonality property from the preceding proposition,
\[ E_0(\phi^{na} [0]) = \sum_{b=1}^{B} E_0(\tilde{\phi}_{nab}^n [0]) + E_0(\phi^{nab}[0]) + o(1), \]
and by passing to a suitable subsequence (in terms of \( n \)) one may then assume that
\[ \lim_{b \to \infty} \lim_{n \to \infty} E_0(\tilde{\phi}_{nab}^n [0]) = 0. \]

The natural concept of “free energy” for the temporally unbounded profiles is then
\[ \lim_{R_b \to \infty} E_0(\tilde{\phi}_{nab}^n [t_n^ab - R_b]). \]

However, the free energy is not quite the right quantity anyway for our induction on energy procedure, and in fact we need the following correct analogue of Lemma 7.13.

**Lemma 7.16.** The statement of Lemma 7.13 remains correct if \( \tilde{A}^nab, low \) is replaced by \( \tilde{A}^nab, true \). Moreover, for a temporally unbounded profile \((\tilde{A}^nab, \tilde{\phi}^{nab})\) with, say, \( t_n^ab \to +\infty \) as \( n \to \infty \), we have
\[ E(\tilde{A}^nab, \tilde{\phi}^{nab})(0) = E(\tilde{A}^nab, \tilde{\phi}^{nab})(t_n^ab - R_b) + \kappa_{ab}(n, R_b), \]
with
\[ \lim_{R_b \to +\infty} \lim_{n \to \infty} \kappa_{ab}(n, R_b) = 0. \]

Here, \( E \) refers to the energy functional of the MKG system.

**7.5.2. Construction of the nonlinear concentration profiles: the profiles.** Here we assume that the linear concentration profiles from the preceding subsection, i.e. the \((A^ab [0], \phi^{ab} [0]), b \geq 1\), have been chosen, as well as the parameter sequences \( \{ (t_n^ab, x_n^ab) \}_{n \geq 1} \). We note that when the profile is temporally bounded, i.e.
\[ \lim_{n \to \infty} |t_n^ab| < +\infty, \]
we may and shall have \( t_n^ab = 0 \) identically. We also use again the notation \((\tilde{A}^nab, \tilde{\phi}^{nab})\) as in the preceding subsection. Thus, if the profile is temporally bounded, it holds that
\[ \tilde{A}^nab [0] = A^ab [0], \quad \tilde{\phi}^{nab} [0] = \phi^{ab} [0]. \]

We can now state the key result of this subsection.
**Theorem 7.17.** Let \( a = 1 \). Assume that there are at least two non-zero profiles \( \phi^{ab} \), or all such profiles are zero, or else there is only one such profile but with

\[
\liminf_{n \to \infty} E(\tilde{A}^{nab}, \tilde{\phi}^{nab})(0) < E_{\text{crit}}.
\]

Then the initial data \( (A^{n1}_{\Lambda_0} + A^{n1}, \phi^{n1}_{\Lambda_0} + \phi^{n1}) \) can be evolved globally in time, resulting in a solution with finite \( S^1 \) norm bounds uniformly for all sufficiently large \( n \).

**Proof.** We proceed in several steps.

**Step 1:** **Construction of the nonlinear concentration profiles.** We distinguish between temporally bounded and unbounded \( (\tilde{A}^{nab}, \tilde{\phi}^{nab}) \).

**Temporally bounded case:** Here we have \( (\tilde{A}^{nab}[0], \tilde{\phi}^{nab}[0]) = (A^{ab}[0], \phi^{ab}[0]) \) with \( A^{ab} \) as usual in the Coulomb gauge. Then we define the nonlinear concentration profile

\[
(\mathcal{A}^{nab}, \Phi^{nab})
\]

as follows. Pick a large time \( T_b > 0 \) (the size will be fixed later on). On \( [-T_b, T_b] \times \mathbb{R}^4 \), we define the profiles to be the solutions to the MKG-CG system with data \( (A^{ab}[0], \phi^{ab}[0]) \), which exist by the assumption (in Theorem 7.17) with a global finite \( S^1 \) norm bound

\[
\| (\mathcal{A}^{ab}, \Phi^{ab}) \|_{S^1} < \infty.
\]

Here the profiles do not depend on \( n \), but we later include this superscript since the profile on the rest of space-time will be \( n \)-dependent. On the complement \( [-T_b, T_b]^c \times \mathbb{R}^4 \), we define the profiles as follows. On \( [T_b, \infty) \times \mathbb{R}^4 \), we let

\[
\Box \mathcal{A}^{ab} = 0
\]

with data \( \mathcal{A}^{ab}[T_b] \) given by the profile constructed on \( [-T_b, T_b] \times \mathbb{R}^4 \), and we proceed analogously on \( (-\infty, -T_b] \times \mathbb{R}^4 \). As for the \( \Phi \)-field, we postulate on \( [-T_b, T_b]^c \times \mathbb{R}^4 \) the linear equation

\[
\Box A^{ab, low} + \sum_{\nu=1}^{4} \mathcal{A}^{\nu} + A^{ab} \Phi^{nab} = 0, \quad A^{ab, low}_{v} := A^{n1}_{\Lambda_0, v}
\]

with data given at time \( T_b \), respectively \(-T_b \), by the profile on \( [-T_b, T_b] \times \mathbb{R}^4 \). Note that in order for this to make sense, we also need to know the definition of the temporally unbounded \( \mathcal{A}^{nab'} \), which is, of course, accomplished below without knowing the temporally bounded \( \Phi^{nab} \) to avoid circularity.

**Temporally unbounded case:** Assume, for example, that \( \lim_{n \to \infty} t^{ab}_n = +\infty \). Using Lemma 7.15 and the assumption of the theorem, we can pick \( R_b > 0 \) sufficiently large such that

\[
\tilde{\phi}^{nab}(t^{ab}_n - R_b, \cdot) = S_{\tilde{A}^{nab, low}}(\phi^{ab}[0])(-R_b, \cdot - x^{ab}_n)
\]

satisfies

\[
E(S(A^{ab}[0])(-R_b, \cdot - x^{ab}_n), S_{\tilde{A}^{nab, low}}(\phi^{ab}[0])(-R_b, \cdot - x^{ab}_n)) < E_{\text{crit}}.
\]

Then we use the data

\[
(S(A^{ab}[0])(-R_b)(\cdot - x^{ab}_n), S_{\tilde{A}^{nab, low}}(\phi^{ab}[0])(-R_b)(\cdot - x^{ab}_n))
\]

at time \( t = t^{ab}_n - R_b \), and evolve them forward in time using the MKG-CG system up to time \( t^{ab}_n + R_b \), say, resulting in the nonlinear profiles

\[
(\mathcal{A}^{nab}, \Phi^{nab})
\]
on \([v_{n}^{ab} - R_{b}, v_{n}^{ab} + R_{b}] \times \mathbb{R}^{4}\). Observe that this construction does not require knowledge of the other profiles \((A^{ab}_{n}, \Phi^{ab}_{n})\). Finally, on the complement \([v_{n}^{ab} - R_{b}, v_{n}^{ab} + R_{b}]^{c} \times \mathbb{R}^{4}\), we evolve \(A^{nab}\) via the free equation \(\square A^{nab} = 0\), and \(\Phi^{nab}\) via the linear evolution

\[
\square_{A^{nab} + \sum_{b=1}^{B} A^{ab}_{n} + \Lambda_{n}} \Phi^{nab} = 0.
\]

**Step 2:** Making an ansatz for the evolution \((A^{n}, \Phi^{n})\) of the full data \((A^{n}_{A_{0}} + A^{n}_{1}, \Phi^{n}_{A_{0}} + \Phi^{n}_{1})\). We now assemble the pieces that we have constructed. We shall write

\[
A^{n} := A^{n,low} + \sum_{b=1}^{B} A^{nab} + A^{n} + \delta_{n}^{a},
\]

where \(A^{nab}\) is actually simply given by \(A^{nab}\) from Proposition 7.12. We immediately observe the crucial fact that

\[
\delta_{n}^{a}[0] = 0,
\]

i.e. the choice of profiles matches the data. We proceed analogously for \(\Phi^{n}\), writing

\[
\Phi^{n} := \Phi^{n,low} + \sum_{b=1}^{B} \Phi^{nab} + \Phi^{nab} + \delta_{n}^{\phi},
\]

where \(\Phi^{nab}\) is actually simply given by \(\Phi^{nab}\) from Proposition 7.14.

**Step 3:** Showing accuracy of the ansatz. Here we finally prove the following key proposition.

**Proposition 7.18.** Assuming the conditions of Theorem 7.17 and given any \(\delta_{5} > 0\), there exists \(B\) sufficiently large (depending on the bounds on \(A^{n,low}\), the actual concentration profiles and on \(\delta_{5}\)) such that for all sufficiently large \(n\),

\[
\|\delta_{n}^{\phi} + \delta_{n}^{\phi}\|_{L^{1}S^{1}} < \delta_{5}.
\]

In light of the immediately verified facts that

\[
\limsup_{n \to \infty} \sum_{b=1}^{B} \|\Phi^{nab}\|_{L^{1}S^{1}} + \limsup_{n \to \infty} \sum_{b=1}^{B} \|A^{nab}\|_{L^{1}S^{1}} < \infty
\]

and

\[
\limsup_{n \to \infty} \|\Phi^{nab}\|_{L^{1}S^{1}} + \limsup_{n \to \infty} \|A^{nab}\|_{L^{1}S^{1}} < \infty,
\]

dthis proposition then immediately implies Theorem 7.17. □

**Proof of Proposition 7.18.** For the most part, this consists in checking that the (very large number of) interactions terms sum to something negligible upon correct choice of \(B\) and \(n\). We start with the equation for \(\delta_{n}^{\phi}\). To begin with, we note that \(\delta_{n}^{\phi}[0]\) is not necessarily 0, since the asymptotic evolution of the profiles \(\Phi^{nab}\) given by

\[
\square_{A^{n,low} + \sum_{b=1}^{B} A^{ab}_{n} + \Lambda_{n}} \Phi^{nab} = 0
\]

is different than the one used to extract the concentration profiles, i.e. \(\square_{A^{n,low}U} = 0\). But we also observe that each profile \(A^{nab}\) differs from the corresponding linear component in Proposition 7.12 given by

\[
S(A^{ab}_{j}[0])(t - t_{n}^{ab}, x - x_{n}^{ab})
\]

by a possibly large term, which however lives in a better space

\[
\|A^{nab}(t, x) - S(A^{ab}_{j}[0])(t - t_{n}^{ab}, x - x_{n}^{ab})\|_{L^{1}S^{1}} < \infty.
\]
Denote this difference by $B^{abb}(t, x)$. Then it suffices to show

**Lemma 7.19.** For any temporally unbounded profile $\Phi^{abb}$ we have

$$\lim_{n \to \infty} \sum_{B \geq 2^{b} b} \|2iB_{\nu}^{abb} \partial^{y} \Phi^{abb}\|_{N} = 0.$$ 

**Proof.** (Outline) We proceed as in the proof of Proposition 7.5, expressing the difference $B^{abb}$ in the schematic form

$$\sum_{i,j} \Box^{-1} P_{k} Q_{j} (\phi \cdot \nabla \phi + A|\phi|^{2}),$$

or else as a free wave satisfying a Besov $\ell^{1}$-bound for the data instead of the weaker energy bound. Using that all factors as well as $\Phi^{abb}$ are 1-oscillatory and the bounds from [12], we reduce to a diagonal situation, where the frequency of all factors as well as $2^{j}$, are essentially restricted to $\sim 1$, and have generic position, i.e. the Fourier supports do not have angular alignment. Then, using that the profiles $\mathcal{R}^{abb}$ disperse away from $t_{0}^{abb}$ uniformly in $n$ (Lemma 7.1), we easily infer the claim.

To be more precise, we first consider the case when $A_{n}^{j}, j = 1, 2, 3, 4$, are free waves, which are 1-oscillatory, obey the Coulomb condition, and satisfy

$$\|A_{n}^{j}\|_{L^{1} S^{1}} < \infty.$$ 

Moreover, assume that $\Phi^{n}$ is 1-oscillatory and satisfies

$$\sup_{n} \|\Phi^{n}\|_{S^{1}} < \infty$$

and also

$$\lim_{n \to \infty} \|\Phi^{n}\|_{L^{\infty}} = 0.$$ 

Then we have

$$\lim_{n \to \infty} \|2iA_{n}^{j} \partial^{y} \Phi^{n}\|_{N} = 0, \quad \lim_{n \to \infty} \|A_{n}^{j} A_{n}^{j} \Phi^{n}\|_{N} = 0.$$ 

The second assertion is easy, so we focus on the first. By the 1-oscillatory character of the inputs and the $\ell^{1}$-Besov bound for $A_{n}$, one may restrict to frequencies $\sim 1$ in both factors, and assume the output to be at modulation $\sim 1$ (else the null structure gives smallness). Then one uses the Strichartz exponents $(\frac{10}{7}, \frac{10}{3})$ for the first factor, and an interpolate of $(\infty, \infty)$ with that same space for the second factor to place the output into $L^{1} L^{2}$. 

Next, consider the case where $A_{n}^{j}$ is of the schematic form

$$\sum_{j \in \mathbb{Z}} \Box^{-1} Q_{j} (\phi \cdot \nabla \phi + A|\phi|^{2}).$$

We only consider the most difficult case, where the space-time frequency localizations have been implemented and the null form structure revealed as in [12] Theorem 12.1. For example, consider an expression

$$\Box^{-1} P_{k} Q_{j} (Q_{<j-C} \phi_{k_{1}} \partial_{n} Q_{<j-C} \phi_{k_{2}}) \partial^{y} Q_{<j-C} \phi_{k},$$

where the $k_{j}$ indicate frequency localizations, all inputs are 1-oscillatory, and satisfy uniform $S^{1}$ norm bounds, and $\Phi^{n}$ satisfies the same vanishing relation as above. Also, from [12] we have the alignments $k_{1} = k_{2} + O(1), k_{3} \geq k + O(1), j \leq k + O(1)$. One may then in fact assume $j = k + O(1)$, since else one gets smallness, and the 1-oscillatory character allows us to assume $k_{1,2,3} = k + O(1) = O(1)$. Then one places the output into $L^{1} L^{2}$ by using the Strichartz exponents $(\frac{10}{7}, \frac{10}{3})$ for the first two factors, and an interpolate of $(\frac{5}{4}, \frac{30}{11} + )$ with $(\infty, \infty)$ for the last factor. The remaining null forms (see (62) and (63) in [12]) are handled similarly. \hfill $\Box$
From the preceding lemma, we infer that we can force
\[ \|\delta_{\Phi}^n(0)\|_{H^1 \times L^2} \ll \delta, \]
provided we pick \( n \) sufficiently large. The equation for \( \delta_{\Phi}^n \) is given by

\[ (7.51) \quad \Box_{A^{n,low}} + \sum_{b'=1}^{B} \mathcal{R}_{nab'} + \mathcal{R}_{nab} + \delta_{\Phi}^n \left( \Phi_{nab,low} + \sum_{b=1}^{B} \Phi_{nab} + \Phi_{nab} + \Phi_{nab} \right) = 0. \]

We re-write this in the following form

\[ (7.52) \quad \Box_{A^{n,low}} + \sum_{b'=1}^{B} \mathcal{R}_{nab'} + \mathcal{R}_{nab} + \delta_{\Phi}^n = -(I) - (II) - (III), \]

where we put

\[ (I) := \Box_{A^{n,low}} + \sum_{b'=1}^{B} \mathcal{R}_{nab'} + \mathcal{R}_{nab} (\Phi_{nab,low}). \]

\[ (II) := \Box_{A^{n,low}} + \sum_{b'=1}^{B} \mathcal{R}_{nab'} + \mathcal{R}_{nab} (\sum_{b=1}^{B} \Phi_{nab}). \]

\[ (III) := \Box_{A^{n,low}} + \sum_{b'=1}^{B} \mathcal{R}_{nab'} + \mathcal{R}_{nab} (\Phi_{nab}). \]

Now the idea is to show smallness of all these terms (in the \( N \) norm sense) provided \( B \) and then \( n \) are chosen sufficiently large. Of course, one needs to be careful with the fact that increasing \( B \) also leads to more and more terms in the sums

\[ \sum_{b'=1}^{B} \mathcal{R}_{nab'}, \quad \sum_{b=1}^{B} \Phi_{nab}. \]

To deal with this, we use

**Lemma 7.20.** Given \( \delta > 0 \), there is a \( B_1 > 0 \) such that for all \( B \geq B_1 \) and all sufficiently large \( n \) (depending on \( B \)), it holds that

\[ \left\| \sum_{b=B_1}^{B} \mathcal{R}_{nab} \right\|_{L^1} < \delta, \quad \left\| \sum_{b=B_1}^{B} \Phi_{nab} \right\|_{L^1} < \delta. \]

**Proof.** By construction we have

\[ \Box_{\mathcal{R}_{nab}} = -\chi_{b}^{P}P \text{Im} \left( \Phi_{nab} \overline{D_{\Phi} \Phi_{nab}} \right), \]

where \( I_{b}^{p} = [-T_{b}, T_{b}] \) for temporally bounded profiles and \( I_{b}^{n} = [t_{n}^{ab} - R_{b}, t_{n}^{ab} + R_{b}] \) for temporally unbounded ones. By picking \( B_1 \) sufficiently large, so that

\[ E(\mathcal{R}_{nab}, \Phi_{nab}) \ll 1, \quad b \geq B_{1}, \]

we get

\[ \sum_{b=B_1}^{B} \left\| \chi_{b}^{P} \text{Im} \left( \Phi_{nab} \overline{D_{\Phi} \Phi_{nab}} \right) \right\|_{L^1} \lesssim \sum_{b=B_1}^{B} E(\mathcal{R}_{nab}, \Phi_{nab}), \]

where we use the notation

\[ \tilde{\mathcal{R}}_{nab} := (S(A_{nab}^{0}[0])(0 - t_{n}^{ab}), x - x_{n}^{ab}), \quad \tilde{\Phi}_{nab} = (S_{A_{nab,low}}(\Phi^{ab}[0])(0 - t_{n}^{ab}), x - x_{n}^{ab}), \]
the latter notation having been used before. But then since
\[
\limsup_{n \to \infty} \left\| \sum_{b' \geq B_1} \mathcal{R}^{nab'}[0] \right\|_{H^1_t L^2_x} < \delta_6, \quad \limsup_{n \to \infty} \sum_{b = B_1}^B E(\tilde{A}^{nab}, \tilde{\phi}^{nab}) < \delta_6,
\]
upon choosing \( B_1 \) large enough, the first bound of the lemma follows. To get the second bound, one uses
\[
\limsup_{n \to \infty} \left\| \sum_{b = B_1}^B \chi_{I_b} \Phi^{nab} \right\|_N \ll \delta_6,
\]
where now \( \chi_{I_b} \) are suitable smooth time cutoffs, while again
\[
\limsup_{n \to \infty} \left\| \sum_{b = B_1}^B \chi_{I_b} \Phi^{nab} [0] \right\|_{H^1_t L^2_x} < \delta_6,
\]
provided \( B_1 \) is chosen sufficiently large. We then infer
\[
\left\| \sum_{b = B_1}^B \chi_{I_b} \Phi^{nab} \right\|_{S^1} \ll \delta_6.
\]
For the complement
\[
\sum_{b = B_1}^B (1 - \chi_{I_b}) \Phi^{nab},
\]
we use
\[
\left\| A^{nab_1} + \sum_{b_2 \geq B_1} \mathcal{R}^{nab'} + \mathcal{A}^{nab} \left( \sum_{b = B_1}^B (1 - \chi_{I_b}) \Phi^{nab} \right) \right\|_N \ll \delta_6.
\]
Note that there are small error terms due to the cutoff, which however are harmless and can be made arbitrarily small by picking the cutoff suitably, see [10]. In fact, we make the

**Remark 7.21.** To ensure smallness of the errors generated by the cutoffs \( \chi_{I_b} \) and \( 1 - \chi_{I_b} \), it suffices to localize each \( \phi^{ab}[0] \) in physical space to a ball of radius \( 10|I_b| \), and each \( A^{ab}[0] \) to a ball of radius \( 100|I_b| \), say. The errors committed thereby may be included in \( \Phi^{nab} \), respectively \( \mathcal{R}^{nab} \).

Combining this with
\[
\left\| \left( \sum_{b = B_1}^B (1 - \chi_{I_b}) \Phi^{nab} \right) [0] \right\|_{H^1_t L^2_x} \ll \delta_6
\]
for \( B_1 \) large enough, we obtain
\[
\left\| \sum_{b = B_1}^B (1 - \chi_{I_b}) \Phi^{nab} \right\|_{S^1} \ll \delta_6,
\]
and the second inequality of the lemma follows. \( \square \)

Next, we show that each of the terms (I) - (III) can be made arbitrarily small up to certain error terms by picking \( B \) and then \( n \) sufficiently large.
The contribution of (I). One writes schematically

\[ \square A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{nab} + \delta_A^n \phi^{na,\text{low}} = \square A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} + 2i(\delta_A^n) \partial^\nu \phi^{na,\text{low}} \]

+ \left( A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} + \delta_A^n \right) \delta_A^n \phi^{na,\text{low}}. \]

Then one has for any \( B \),

\[ \lim_{n \to \infty} \left\| \square A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} \phi^{na,\text{low}} \right\|_N = 0 \]

due to the frequency localizations (up to exponential tails) of the inputs \( \sum_{b'=1}^B \mathcal{A}^{ab'} \), \( \mathcal{A}^{naB} \) and \( \phi^{na,\text{low}} \) as well as due to the fact that by construction we have

\[ \square A_{na,\text{low}} \phi^{na,\text{low}} = 0. \]

More precisely, one uses an argument as in the proof of Proposition 7.4. We then still have the error terms

\[ (7.53) \quad 2i(\delta_A^n) \partial^\nu \phi^{na,\text{low}}, \quad \left( A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} + \delta_A^n \right) \delta_A^n \phi^{na,\text{low}}. \]

The second term here shall be straightforward to treat by means of a simple divisibility argument, while the first will require the equation satisfied by \( \delta_A^n \) in conjunction with a divisibility argument.

The contribution of (II). We write schematically

\[ \square A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} + \delta_A^n \left( \sum_{b=1}^B \Phi^{nab} \right) = \sum_{b=1}^B \chi_{b} \left( \square A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} - \square \mathcal{A}^{ab} \right) \Phi^{nab} \]

+ \sum_{b=1}^B 2i(\delta_A^n) \partial^\nu \phi^{nab} \]

+ \left( A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} + \delta_A^n \right) \delta_A^n \Phi^{nab}. \]

Here the time intervals \( I_b^n \) correspond to \([-T_b, T_b]\) for the temporally bounded profiles and to \([t_0^n - R_b, t_0^n + R_b]\) for the temporally unbounded ones. We shall henceforth make the following additional assumption that

\[ |I_b^n| = M \quad \forall b \]

chosen very large (eventually depending on \( \delta_5 \) and the profiles). Then we observe that given any \( \delta_5 > 0 \), we can pick \( B \) large enough such that for any sufficiently large \( n \), we have

\[ \left\| \sum_{b=1}^B \chi_{b} \left( \square A_{na,\text{low}} + \sum_{b'=1}^B \mathcal{A}^{ab'} + \mathcal{A}^{naB} - \square \mathcal{A}^{ab} \right) \Phi^{nab} \right\|_N \ll \delta_5. \]
To show this, we need
\[ \left\| \sum_{b=1}^{B} \chi_{I_b}^n \left( A^{n_{low}} + \sum_{b' \neq b} A^{nab'} + A^{nabB} \right) \partial^\nu \Phi^{nab} \right\|_N \ll \delta_5, \]
\[ \left\| \sum_{b=1}^{B} \chi_{I_b}^n \left( A^{n_{low}} + \sum_{b' \neq b} A^{nab'} + A^{nabB} \right)^2 - (A^{nab})^2 \Phi^{nab} \right\|_N \ll \delta_5. \]

For the first expression, observe that the interactions of \( A^{nab'} \), \( b' \neq b \), with \( \Phi^{nab} \) are easily seen to vanish as \( n \to \infty \), using crude bounds, due to the time localization from \( \chi_{I_b}^n \), and the diverging supports of these profiles or their dispersive decay. Similarly, the interaction of \( A^{n_{low}} \) with \( \Phi^{nab} \) is seen to vanish asymptotically as \( n \to \infty \), due to the divergent frequency supports and again taking advantage of the extra cutoff \( \chi_{I_b}^n \). Note that at this point we have not yet used the parameter \( B \).

Finally, we also need to bound
\[ \left\| \sum_{b=1}^{B} \chi_{I_b}^n A^{nabB} \partial^\nu \Phi^{nab} \right\|_N, \]
and it is here that we shall take advantage of the size of \( B \). Precisely, we divide the above term into two. First, pick \( B_1 \) very large, depending on the parameter \( M \) (which controls the \( I_b^n \) via \( |I_b^n| \leq M \)), such that we have for any \( B \geq B_1 \),
\[ \limsup_{n \to \infty} \left\| \sum_{b=1}^{B_1} \chi_{I_b}^n A^{nabB} \partial^\nu \Phi^{nab} \right\|_N \ll \delta_5. \]

That this is possible follows from Lemma 7.20. Then, with \( B_1 \) chosen, pick \( B \geq B_1 \) sufficiently large such that
\[ \limsup_{n \to \infty} \left\| \sum_{b=1}^{B_1} \chi_{I_b}^n A^{nabB} \partial^\nu \Phi^{nab} \right\|_N \ll \delta_5. \]

Here we take advantage of the fact that we essentially have
\[ \limsup_{n \to \infty} \left\| A^{nabB} \right\|_{L^\infty_\gamma L^\infty_\nu L^1_\mu H_1^1} \to 0 \]
as \( B \to \infty \). In fact, we have to be a bit careful here, because in Remark 7.21 we assume that we have incorporated some extra errors into the tail terms \( A^{nab} \) and \( \Phi^{nab} \), which do not vanish as \( B \to \infty \). However, considering the term corresponding to a fixed \( b \in [1, B_1] \), we have that the extra contribution to \( A^{nab} \) (coming from truncating \( A^{nabB}[0] \)) interacts weakly with \( \Phi^{nab} \) (in the sense that it vanishes as \( |I_b| \to \infty \)), see e.g. the proof of Proposition 5.11. The remaining contributions from truncating \( A^{nabB}[0] \) are easily seen to result in interactions vanishing as \( n \to \infty \). The cubic term
\[ \left\| \sum_{b=1}^{B} \chi_{I_b}^n \left( A^{n_{low}} + \sum_{b' \neq b} A^{nab'} + A^{nabB} \right)^2 - (A^{nab})^2 \Phi^{nab} \right\|_N \]
is actually simpler, because the temporal cutoffs \( \chi_{I_b}^n \) are not even necessary to get the desired bound. This completes the estimate for (II) except for the error terms
\[ (7.54) \quad \Box_{A^{n_{low}} + \sum_{b' \neq b} A^{nab'} + A^{nabB} \partial^\nu} \left( \sum_{b=1}^{B} \Phi^{nab} \right) - \Box_{A^{n_{low}} + \sum_{b' \neq b} A^{nab'} + A^{nabB} \partial^\nu} \left( \sum_{b=1}^{B} \Phi^{nab} \right). \]
The contribution of (III). Here we take advantage of the fact that $\Phi^{naB}$ satisfies the equation (7.47). Of course, the profiles $A^{\alpha \beta'}$ are not free waves, but they differ from free waves by terms that are negligible as far as interactions with $\Phi^{naB}$ are concerned. In fact, we recall that

\[ \|A^{\alpha \beta'}(t, x) - S(A_j^{\alpha \beta'}[0])(t - t_j^{\alpha \beta'}, x - x_j^{\alpha \beta'})\|_{L^1(S)} < \infty. \]

Using Lemma 7.20 we can refine this to a tail estimate as follows. There exists $B_1$ sufficiently large such that denoting

\[ g^{\alpha \beta'} := A^{\alpha \beta'}(t, x) - S(A_j^{\alpha \beta'}[0])(t - t_j^{\alpha \beta'}, x - x_j^{\alpha \beta'}), \]

we have for any $B \geq B_1$,

\[ \limsup_{n \to \infty} \sum_{b' = B_1}^B \| 2i g^{\alpha \beta'} \partial_\nu \Phi^{naB} \|_N \ll \delta_5. \]

On the other hand, with this $B_1$ fixed, we can use the argument for Lemma 7.19 to conclude that there exists $B \geq B_1$ such that we have

\[ \limsup_{n \to \infty} \sum_{b' = 1}^B \| 2i \partial_\nu g^{\alpha \beta'} \partial_\nu \Phi^{naB} \|_N \ll \delta_5. \]

Finally, we are still left with the error terms

\[ \Box_{A_{na, low} + \sum_{b' = 1}^B A_{\alpha \beta'} + A_{naB} + \delta_\Phi} \Phi^{naB} - \Box_{A_{na, low} + \sum_{b' = 1}^B A_{\alpha \beta'} + A_{naB} \Phi^{naB}}. \]

We have now shown smallness of the terms (I) - (III) up to errors that are at least linear in $\delta_5$ given by (7.53) - (7.55).

Having dealt with the equation for $\delta_\Phi$, we now come to the equation for $\delta_A$ given by

\[ \Box_{A_{na, low} + \sum_{b' = 1}^B A_{\alpha \beta'} + A_{naB} + \delta_\Phi} \Phi^{naB} - \Box_{A_{na, low} + \sum_{b' = 1}^B A_{\alpha \beta'} + A_{naB} \Phi^{naB}}. \]

We rewrite this in the form

\[ (7.57) \quad \Box(\delta_A) = -(I) - (II), \]
where we put schematically

$$(I) := \text{Im} \left( \sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB} + \delta^n_{\Phi} D_x (\sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB} + \delta^n_{\Phi}) \right)$$

$$- \text{Im} \left( \sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB} + \delta^n_{\Phi} D_x \left( \sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB} + \delta^n_{\Phi} \right) \right) - \text{Im} \left( \sum_{b=1}^{B} \frac{D_x \Phi^{naB}}{\Phi^{nab}} \right).$$

The term $(I)$ can be written in terms of null forms as well as cubic terms involving at least one low frequency factor $\delta^n_{\Phi}$ and as well as at least one high frequency term from

$$\sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB},$$

or else error terms involving at least one factor $\delta^n_{\Phi}$. The former type of interaction is easily seen to converge to zero with respect to $\| \cdot \|_N$ as $n \to \infty$, and so only the latter type of error term needs to be kept. As for term $(II)$, again ignoring the terms involving at least one factor $\delta^n_{\Phi}$, we reduce this to

$$\text{Im} \left( \sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB} \right) D_x \left( \sum_{b=1}^{B} \Phi^{nab} + \Phi^{naB} \right) - \sum_{b=1}^{B} \Box A^{nab}.$$

Then from the definition of the profiles $\Phi^{nab}$ etc, we can write this for some large $B_2$ and $B \geq B_2$ as

$$\sum_{b=1}^{B_2} \chi_{\{y \}} \text{Im} \left( \Phi^{nab} D_x \Phi^{nab} \right)$$

$$+ \text{Im} \left( \sum_{b=1}^{B_2} \Phi^{nab} D_x \left( \sum_{b=1}^{B_2} \Phi^{nab} \right) \right) - \sum_{b=1}^{B_2} \text{Im} \left( \Phi^{nab} D_x \Phi^{nab} \right)$$

$$+ \text{Im} \left( \sum_{b=B_2}^{B} \Phi^{nab} D_x \left( \sum_{b=1}^{B_2} \Phi^{nab} + \Phi^{naB} \right) \right) + \text{Im} \left( \Phi^{naB} D_x \left( \sum_{b=B_2}^{B} \Phi^{nab} \right) \right)$$

$$- \sum_{b=B_2}^{B} \chi_{\{y \}} \text{Im} \left( \Phi^{nab} D_x \Phi^{nab} \right) + \text{Im} \left( \Phi^{naB} D_x \Phi^{naB} \right)$$

$$=: (II)_1 + (II)_2 + (II)_3 - (II)_4 + (II)_5.$$

Then given a $\delta_5 > 0$ arbitrarily small, we first pick $B_2$ sufficiently large such that for all sufficiently large $n$ we have

$$\| (II)_3 \|_N + \| (II)_4 \|_N \ll \delta_5,$$

using Lemma 7.20. Then one picks $n$ large enough such that

$$\| (II)_2 \|_N \ll \delta_5.$$

Further, with $B_2$ fixed, pick $B \geq B_2$ sufficiently large such that

$$\| (II)_5 \|_N \ll \delta_5.$$
Finally, with $B_2$ fixed, we choose $M = |I^n_1|$ large enough (depending on the profiles $\Phi^{a,b}, b = 1, \ldots, B_2$, where these of course depend on the $n$-independent $\phi^{a,b}[0]$), such that

$$\| (\mathcal{I})_j \|_{X} \ll \delta_5.$$  

This is then the $M$ that needs to be used in the analysis of the $\delta_{\alpha}^n$ equation in the “additional assumption” there.

7.6. Conclusion of the induction on frequency process. As in [10], we can now add in the remaining frequency atoms $(A^{a}, \phi^{a})$, $a \geq 2$, to conclude that if either there are at least two frequency atoms, or else there is only one frequency atom but with

$$\liminf_{n \to \infty} E(A^{n,1}, \phi^{n,1}) < E_{\text{crit}},$$

or if we do have

$$\lim_{n \to \infty} E(A^{n,1}, \phi^{n,1}) = E_{\text{crit}},$$

but such that there are at least two concentration profiles in the sense of the preceding section, or finally if there is only one frequency atom of asymptotic energy $E_{\text{crit}}$ and only one concentration profile $(\tilde{A}^{n,b}, \tilde{\phi}^{n,b})$ with

$$\liminf_{n \to \infty} E(\tilde{A}^{n,b}, \tilde{\phi}^{n,b})(0) < E_{\text{crit}},$$

then the sequence $(A^n, \phi^n)$ cannot possibly have been essentially singular, resulting in a contradiction to our assumption. We can then formulate the following

**Corollary 7.22.** Assume that $(A^n, \phi^n)$ is an essentially singular sequence. Then by re-scaling we may assume that the sequence of data $(A^n, \phi^n)[0]$ is $1$-oscillatory, and there exist sequences

$$\{(t_n, x_n)\}_{n \geq 1} \subset \mathbb{R} \times \mathbb{R}^4$$

and fixed profiles

$$(A, \phi)[0] \in (\mathcal{H}^1_\alpha \times \mathcal{L}^2_\alpha)^4 \times (\mathcal{H}^1_\alpha \times \mathcal{L}^2_\alpha)$$

with $A$ satisfying the Coulomb condition, such that we have

$$A^n[0] = S(A[0])(-t_n, \ldots, x_n)[0] + o_{\mathcal{H}^1_\alpha \times \mathcal{L}^2_\alpha}(1).$$

Here, $S$ denotes the standard free wave evolution. Furthermore, letting

$$\tilde{A}^{n,\text{true}}(t, x) := A^{n,\text{free}}(t + t_n, x + x_n)$$

with $A^{n,\text{free}}$ the free wave evolution of the data of $A_j$ at time $t = 0$, while $A^{n,\text{free}} = 0$, and denoting by $S_{\tilde{A}^{n,\text{true}}}$ the corresponding wave evolution as in (7.47), we have

$$\phi^n[0] = S_{\tilde{A}^{n,\text{true}}}(\phi[0])(-t_n, \ldots, x_n)[0] + o_{\mathcal{H}^1_\alpha \times \mathcal{L}^2_\alpha}(1).$$

If the sequence $(t_n)_{n \geq 1}$ admits a subsequence that is bounded, then by passing to this subsequence, we may as well replace $t_n$ by $t_n = 0$ for all $n$, and correspondingly obtain

$$(A^n, \phi^n)[0] = (A, \phi)[0] + o_{\mathcal{H}^1_\alpha \times \mathcal{L}^2_\alpha}(1).$$

Then Proposition [6.1] implies that the evolution $(\mathcal{A}^n, \Phi^n)$ of $(A, \phi)[0]$ is in fact a minimal energy blowup solution. In the case that $|t_n| \to \infty$, whence by passing to a subsequence $t_n \to +\infty$, say, we need to introduce the concept of a minimum regularity MKG-CG evolution associated with scattering data, or “a solution at infinity”. Here we have the following
Proposition 7.23. Assume that for energy class data \((A, \phi)[0]\) satisfying the Coulomb condition \(\sum_{j=1}^{4} \partial_j A_j = 0\), and further \(t_n \to +\infty\), while \(\{x_n\}_{n \geq 1} \subset \mathbb{R}^4\) is arbitrary, we introduce the scattering data (here \(\mathcal{A}\) has vanishing temporal component)
\[
(7.58) \quad \mathcal{A}^{0}[0] := S(A[0])(-t_n, -x_n)[0], \quad \Phi^{0}[0] := S_{\bar{A}, n\to \infty} (\phi[0])(t - t_n, x - x_n)[0].
\]
Here we use the notation consistent with the one in the preceding paragraph, i.e. \(\bar{A} = S(A[0])(t, x)\). Also, denote the MKG-CG evolution (in the sense of Section 5) of the Coulomb data \((\mathcal{A}^0, \Phi^0)[0]\) by \((\mathcal{A}^\infty, \Phi^\infty)(t, x)\). Then there is a \(C \in \mathbb{R}_+\) sufficiently large, such that there exist an energy class solution \((\mathcal{A}^\infty, \Phi^\infty)\) to MKG-CG on \((-\infty, -C) \times \mathbb{R}^4\), which is the limit of admissible solutions as in Section 5 with
\[
\|(\mathcal{A}^\infty, \Phi^\infty)\|_{L^1((-\infty, -C) \times \mathbb{R}^4)} < \infty \quad \forall C_0 > C,
\]
and such that for any \(t \in (-\infty, -C)\) we have in the energy topology
\[
\lim_{n \to \infty} (\mathcal{A}^n, \Phi^n)(t + t_n, x + x_n) = (\mathcal{A}^\infty, \Phi^\infty)(t, x).
\]
In particular, the expressions on the left are well-defined (in the sense of Section 5) for \(n\) sufficiently large.

Proof. This is a perturbative argument, which exploits the dispersive behaviour as evidenced by amplitude decay of the functions \(\mathcal{A}^0[0]\) and \(\Phi^0[0]\). Write
\[
\mathcal{A}^n(t, x) = \mathcal{A}^{1n}(t, x) + \delta \mathcal{A}^n(t, x), \quad \Phi^n(t, x) = \Phi^{1n}(t, x) + \delta \Phi^n(t, x),
\]
where we have introduced the notation
\[
\mathcal{A}^{1n}(t, x) := S(A[0])(t - t_n, x - x_n), \quad \Phi^{1n}(t, x) := S_{\bar{A}, n\to \infty} (\phi[0])(t - t_n, x - x_n).
\]
Also, keep in mind that \((\mathcal{A}^n, \Phi^n)(t, x)\) denotes the MKG-CG evolution (in the sense of Section 5) of the data \((\mathcal{A}^0, \Phi^0)[0]\). Then we show that \((\delta \mathcal{A}^n, \delta \Phi^n)\) satisfy good \(S^1\)-bounds on \((-\infty, t_n - C) \times \mathbb{R}^4\) for some \(C > 0\) sufficiently large, and all \(n\) large enough. This means that the evolutions \((\mathcal{A}^n, \Phi^n)\) are well-defined on \((-\infty, t_n - C) \times \mathbb{R}^4\). Furthermore, assuming as we may that \(t_n\) is monotonously increasing, we will show that for \(n' > n\), we have
\[
\lim_{n, n' \to \infty} \left\| \mathcal{A}^n[t_n' - t_n] - \mathcal{A}^0[0] \right\|_{L^1_t H^1_x L^2_t} + \lim_{n, n' \to \infty} \left\| \Phi^n[t_n' - t_n] - \Phi^0[0] \right\|_{L^1_t H^1_x L^2_t} = 0,
\]
which together with standard perturbation theory then results in the fact that
\[
\lim_{n \to \infty} (\mathcal{A}^n, \Phi^n)(t + t_n, x + x_n) = (\mathcal{A}^\infty, \Phi^\infty)(t, x),
\]
provided \(t \in (-\infty, -C)\), and the right hand side is a solution to MKG-CG in the sense of Section 5.

To get the desired bounds on \((\delta \mathcal{A}^n, \delta \Phi^n)\), we record the schematic system of equations that they satisfy
\[
0 = \Box \mathcal{A}^n \Phi^{1n} + (\Box \mathcal{A}^{1n} + \delta \mathcal{A}^n \phi^{1n} + (\Box \mathcal{A}^{1n} + \delta \mathcal{A}^n \phi^{1n}), \quad (7.59)
\]
\[
\Box \mathcal{A}^n = \mathcal{P}_J (\mathcal{A}^{1n} \mathcal{D}_x \phi^{1n}) + \mathcal{P}_J (\delta \mathcal{A}^n \mathcal{D}_x \phi^{1n}) + \mathcal{P}_J (\delta \mathcal{A}^n \mathcal{D}_x \phi^{1n} + \mathcal{P}_J (\delta \mathcal{A}^n \mathcal{D}_x \phi^{1n}).
\]
(7.60)

We then show that given \(\delta > 0\), there exists \(C = C(\delta, A[0], \phi[0])\) such that we have
\[
\left\| \delta \mathcal{A}^n \right\|_{L^1((-\infty, -C) \times \mathbb{R}^4)} + \left\| \delta \Phi^n \right\|_{L^1((-\infty, -C) \times \mathbb{R}^4)} < \delta.
\]
This follows as usual via a bootstrap argument. We show here how to obtain smallness of the non-perturbative source terms on the right hand side, i.e. the terms
\[
\Box \mathcal{A}^{1n}, \quad \mathcal{P}_J (\mathcal{A}^{1n} \mathcal{D}_x \phi^{1n}),
\]
while the remaining terms are handled either via the smallness of $\delta$ (provided they are quadratic in $\delta, A^\alpha, \delta \Phi^\alpha$), or else via a standard divisibility argument, just as in the proof of Proposition 7.4. Now the first term on the right is in effect equal to

$$A_v^{1n} A_v^{1n,v} \Phi^{1n}.$$  

To treat it, we note that we may reduce all inputs as well as the output to frequency $\sim 1$, since else we gain smallness for the $L_t^1 L_x^{\infty}$-norm of the output by using standard Strichartz norms. Then we estimate the remainder by

$$\left\| P(O(1)A_v^{1n} P(O(1)A_v^{1n,v} P(O(1)\Phi^{1n}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)} \lesssim \left\| P(O(1)A_v^{1n}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)} \left\| P(O(1)\Phi^{1n}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)}.$$  

Then by exploiting the $L_x^{\infty}$ decay and interpolation, for example, we get

$$\left\| P(O(1)A_v^{1n,v}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)} \ll \delta$$

for $C$ sufficiently large, uniformly in $n$, and this suffices to get the necessary smallness on account of the fact that uniformly in $n$,

$$\left\| P(O(1)A_v^{1n}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)} + \left\| P(O(1)\Phi^{1n}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)} \lesssim \left\| (\Delta [0], \Phi [0]) \right\|_{H^1_t \times L^2_x}.$$  

As for the quadratic term

$$P_f(\Phi^{1n} D_t \Phi^{1n}),$$

its inherent null structure allows to reduce to the case of frequencies $\sim 1$ and inputs with an angular separation between their Fourier supports so that the output is at modulation $\sim 1$. In that situation we have

$$\left\| P_f(\Phi^{1n} D_t \Phi^{1n}) \right\|_{L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)} \lesssim \left\| P_f(\Phi^{1n} D_t \Phi^{1n}) \right\|_{L^2_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)},$$

which can be estimated by placing one input into $L^1_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)$ and the other one into $L^2_t L_x^{\infty}((-\infty, t_n - C) \times \mathbb{R}^4)$. The latter norm is small uniformly in $n$ for $C$ sufficiently large on account of (a variant of) Lemma 7.11. □

**Remark 7.24.** The preceding proof implies in particular that if $(A, \Phi) [0] \in (H^1_x \times L^2_x)^5$ are given with $A$ satisfying the Coulomb condition, then there exists $t_0 > 0$ sufficiently large such that the initial data

$$(S(A[0])(\cdot - t_0, \cdot), S_{\tilde{A}_{\text{true}}}(\Phi[0])(\cdot - t_0, \cdot))[0],$$

where

$$\tilde{A}_{\text{true}}(t, \cdot) = A^{\text{free}}(t + t_0, \cdot),$$

can be evolved in the sense of Section 5 on $(-\infty, 0] \times \mathbb{R}^4$ and satisfying a global $S^1$-bound there. This is the analogue of Proposition 7.15 in [10].

Extracting a minimal blowup solution in the case of one temporally unbounded profile is still not a direct consequence of the preceding proposition on account of the somewhat delicate perturbation theory, but follows by a slightly indirect argument. Here we state

**Proposition 7.25.** Assume that the essentially singular sequence $(A^\alpha, \Phi^\alpha)$ satisfies

$$A^\alpha[0] = S(A[0])(\cdot - t_n, \cdot - x_n)[0] + o_{H^1_t \times L^2_x(1)},$$

$$\Phi^\alpha[0] = S_{\tilde{A}_{\text{true}}}(\Phi[0])(\cdot - t_n, \cdot - x_n)[0] + o_{H^1_t \times L^2_x(1)},$$

where
where \( t_n \to +\infty \), say, and we use the same notation as in the preceding corollary. Then denoting the corresponding MKG-CG evolution of these data by \((A^n, \phi^n)(t, x)\), its lifespan comprises \((-\infty, t_n - C)\) for \( C \) sufficiently large, uniformly in \( n \). Also, the sequence

\[ [(A^n, \phi^n)[t_n - 2C]]_{n \geq 1} \]

forms a pre-compact set in the energy topology. Denoting a limit point (any such satisfies the Coulomb condition) by \((\mathcal{A}^\infty, \Phi^\infty)[0]\), we have \( E(\mathcal{A}^\infty, \Phi^\infty) = E_{\text{crit}} \), and moreover, denoting the lifespan of its MKG-CG evolution by \( I \), we get

\[ \sup_{J \subseteq I} \|(\mathcal{A}^\infty, \Phi^\infty)\|_{\mathcal{S}_1(J \times \mathbb{R}^4)} = +\infty. \]

**Proof.** The fact that the evolution of \((A^n, \phi^n)(t, x)\) is defined and has finite \( S^1 \)-bounds on \((-\infty, t_n - C)\) follows by exactly the same method as in the proof of the preceding proposition. We set

\[ A^n(t, x) = A^{1n}(t, x) + \delta A^n(t, x), \quad \phi^n(t, x) = S_{A^{1n} + \delta\phi^n}(\phi[0])(t - t_n, x - x_n) + \delta\phi^n(t, x), \]

where we let \( A^{1n} \) be the free wave evolution of \( A^n[0] \), i.e.

\[ A^{1n}(t, x) = S(A[0])(t - t_n, x - x_n) + o_{H^1 \times L^2}(1), \]

and \( \delta\phi^n \) as in Corollary 7.22. Also, note that \( \delta A^n[0] = 0 \). Then choosing \( C \) large enough, we infer the bounds

\[ \|(\delta A^n, \delta\phi^n)\|_{L^1 L^1 (-\infty, t_n - C) \times \mathbb{R}^4} \ll 1 \]

via bootstrap. Since \((A^n, \phi^n)\) is essentially singular, we know by the preceding results that the data

\[ (A^n, \phi^n)[t_n - 2C] \]

are concentrated at fixed frequency \( \sim 1 \) and consist of exactly one concentration profile, which is necessarily temporally bounded. But this implies that the sequence

\[ [(A^n, \phi^n)[t_n - 2C]]_{n \geq 1} \]

is compact in the energy topology. Extracting a limiting profile \((\mathcal{A}^\infty, \Phi^\infty)[0]\), the last statement of the proposition follows directly from Proposition 6.1. \( \square \)

To conclude this section, we finally state the following crucial **compactness property** of the minimal blowup solution \((\mathcal{A}^\infty, \Phi^\infty)\) extracted in the preceding.

**Theorem 7.26.** Denote the lifespan of \((\mathcal{A}^\infty, \Phi^\infty)\) by \( I \). There exist continuous functions \( \overline{x} : I \to \mathbb{R}^4 \), \( \lambda : I \to \mathbb{R}_+ \), so that the family of functions

\[ \left\{ \left( \frac{1}{A^2(t)} \nabla_{t,x} \mathcal{A}^\infty(t, \frac{\cdot - \overline{x}(t)}{\lambda(t)}), \frac{1}{A^2(t)} \nabla_{t,x} \Phi^\infty(t, \frac{\cdot - \overline{x}(t)}{\lambda(t)}) \right) \right\}_{t \in I} \]

is pre-compact in \( (L^2_x(\mathbb{R}^4))^5 \).

The proof of this follows exactly as for Corollary 9.36 in [10], using the preceding remark.
8. Towards ruling out a minimal blowup solution with the compactness property

In this section we start key computations along the Kenig-Merle strategy. The first important step is to control the movement of the “center of mass” (or rather energy). Here as for critical wave maps in [10], the Lorentz invariance of the Maxwell-Klein-Gordon system and transformational properties of the energy under Lorentz transformations are essential. We start by considering the relativistic invariance properties of our system: Assume that

\[ L : \mathbb{R}^{1+4} \to \mathbb{R}^{1+4} \]

is a Lorentz transformation, acting on column vectors via multiplication with the matrix \( L \). Then we transform \( \phi \) according to

\[ \phi \mapsto \phi^L(x) := \phi(Lx), \]

which results in

\[ \nabla_{t,x} \phi \mapsto L^t \nabla_{t,x} \phi(Lx). \]

Then the potential \( A_{\alpha} \) needs to transform accordingly, i.e. writing this as a column vector (indexed by \( \alpha \)), we transform

\[ A_{\alpha} \mapsto A^L := L^t A_{\alpha}(Lx). \]

Then the expression \( \partial_\beta F_{\alpha\beta} \), when interpreted as a column vector in \( \alpha \), also transforms according to multiplication with \( L^t \), as does the expression

\[ \text{Im}(\phi D_{\alpha} \phi). \]

Under these transformations, the Maxwell-Klein-Gordon system is then invariant. The conserved energy does not remain invariant under general Lorentz transformations, and the first step consists in quantifying this. As a result, we will obtain that a certain generalized moment has to vanish for the solution to be minimal blowup, which in turn influences the movement of the center of mass later on. To see why the non-invariance of the energy under Lorentz transformations may be of decisive use for us, we include the following technical subsection.

8.1. \( S^1 \)-bounds and Lorentz transformations. In the sequel we restrict to a special class of Lorentz transforms, namely of the form

\[ t \mapsto \frac{t}{\sqrt{1-d^2}} - \frac{dx_1}{\sqrt{1-d^2}}, \quad x_1 \mapsto \frac{x_1}{\sqrt{1-d^2}} - \frac{dt}{\sqrt{1-d^2}}, \quad x_j \mapsto x_j, \quad j = 2, 3, 4, \]

or the cases where \( x_1 \) gets replaced by \( x_j, \ j = 2, 3, 4 \). In order to ensure that the Lorentz transformed solution \( (A^L, \phi^L) \) is actually defined, we need to ensure that it is defined at time \( t = 0 \) at least, so we can evolve it on some maximal interval from there. Thus we assume that \( (A, \phi) \) is a globally existing and admissible solution. Throughout we assume that \( A \) is of Coulomb class. Under the preceding hypotheses, we can then define \( (A^L, \phi^L)[0] \). We shall assume that these data are well-defined in the sequel. Moreover, the fact that \( (A, \phi) \) is admissible implies that \( (A^L, \phi^L)[0] \) is of energy class. To see this, note that if \( (t, x) \) are restricted to a spacelike plane containing the origin, then we have

\[ |F_{jk}(t,x)| \lesssim (|t| + |x|)^{-N} \]

and from the equation satisfied by \( F_{\alpha\beta} \) we get upon simple integration in time that \( |F_{\alpha\beta}(t,x)| \lesssim (|t| + |x|)^{-3} \). Then the curvature components of \( (A^L, \phi^L) \) at time \( t = 0 \) also satisfy the \( |x|^{-3} \)-decay property, which ensures \( L^2 \)-integrability, and the components \( \nabla_{t,x} \phi^L \) decay rapidly with respect to \( x \). In particular, it is meaningful to consider the \( S^1 \)-norm of the evolution of the data \( (A^L, \phi^L)[0] \). Then we prove the following important technical
Proposition 8.1. Let \((A, \phi)[0]\) be a admissible Coulomb class data at time \(t = 0\) as above. Assume that the Lorentz transformed \((A^L, \phi^L)[0]\), when transformed into the Coulomb gauge, result in a global in time smooth solution \((\tilde{A}^L, \tilde{\phi}^L)\) with
\[
\|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1} < +\infty.
\]
Then the original data \((A, \phi)[0]\) admit a global admissible evolution that satisfies
\[
\| (A, \phi) \|_{S^1} \leq C \|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1, L}.
\]

Proof. Let \(\gamma\) be chosen such that \(\tilde{A}^L = A^L - \nabla \gamma\) is in the Coulomb gauge. Then we have \(\tilde{\phi}^L = e^{i\gamma} \phi\).

Note that we have
\[
(\tilde{A}^L)^{-1} = A - (\nabla \gamma)^{-1} = A - \nabla (\gamma(L^{-1}(t, x)))
\]
and so we get
\[
\gamma(L^{-1}(t, x)) = \Delta^{-1} \partial_i ((\tilde{A}^L)^{-1})_i.
\]

The difficulty in controlling the \(S^1\)-norm of \((A, \phi)\) here is that this norm is far from invariant under the operation of Lorentz transforms. Nonetheless, one can establish control over a certain set of norms of \((A, \phi)\) that are essentially invariant under Lorentz transforms, and which in turn imply control over the full set of \(S^1\)-norms. We do this in the following observations.

**Observation 1:** For a \(C = C(L)\) with \(C(L) \to \infty\) as \(L \to Id\), we have the bounds
\[
\left( \sum_{k \in \mathbb{Z}} \left\| \nabla_x P_k \mathcal{Q}_{\{k+\xi, C\}} \phi^L \right\|_{L^2_x L^2_t}^2 \right)^{1/2} \lesssim \|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1},
\]
\[
\left( \sum_{k \in \mathbb{Z}} 2^{-\nu k} \left\| \nabla_x P_k \mathcal{Q}_{\{k+\xi, C\}} \phi^L \right\|_{L^2_x L^2_t}^2 \right)^{1/2} \lesssim \|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1}.
\]

Similarly for \(A^L\), we have the bounds
\[
\left( \sum_{k \in \mathbb{Z}} \left\| \nabla_x P_k \mathcal{Q}_{\{k+\xi, C\}} A^L \right\|_{L^2_x L^2_t}^2 \right)^{1/2} \lesssim \|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1},
\]
\[
\left( \sum_{k \in \mathbb{Z}} 2^{-\nu k} \left\| \nabla_x P_k \mathcal{Q}_{\{k+\xi, C\}} A^L \right\|_{L^2_x L^2_t}^2 \right)^{1/2} \lesssim \|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1}.
\]

Finally, we also have the bound
\[
\left( \sum_{k \in \mathbb{Z}} \left\| \nabla_x P_k \mathcal{Q}_{\{k+\xi, C\}} \phi^L \right\|_{L^2_x L^4_t}^2 \right)^{1/2} \lesssim \|(\tilde{A}^L, \tilde{\phi}^L)\|_{S^1}.
\]

**Proof.** (Observation 1) We first compute \(\gamma\) in terms of \(\tilde{A}^L\), for which we have good bounds by assumption. Note that
\[
\gamma = (\Delta^{-1} \partial_i ((\tilde{A}^L)^{-1}))^L.
\]

Now for any fixed dyadic frequency \(k \in \mathbb{Z}\), we can write
\[
\Delta^{-1} \partial_t P_k f^{L^{-1}}(t, x) = \int_{\mathbb{R}^4} m_k(a) f^{L^{-1}}(t, x - a) \, da
\]
for a suitable \(L^1\)-function \(m_k(a)\) with \(L^1\)-mass \(\sim 2^{-k}\), and further
\[
\left( \int_{\mathbb{R}^4} m_k(a) f^{L^{-1}}(t, x - a) \, da \right)^L = \int_{\mathbb{R}^4} m_k(a) f((t, x) - L^{-1}(0, a)) \, da.
\]
Also, if \( l < k + \frac{C}{2} \) for suitable \( C = C(L) \), then Fourier localization to dyadic modulation \( 2^l \) and spatial frequency \( 2^k \) essentially commutes with the Lorentz transform, provided \( C \) is not too large depending on \( d \), and so we also have

\[
\Delta^{-1} \partial_t P_k Q l^{L^{-1}}(t, x) = \int_{\mathbb{R}^4} m_k(a)(P_{k+O(1)} Q l+O(1), f)^{L^{-1}}(t, x - a) \, da.
\]

From the preceding, we have that for \( C \) suitably large (depending on \( d \) in the definition of \( L \)) the schematic relation (in the last line one needs linear combinations of components \( \tilde{A}_i^L \))

\[
P_k Q l(\Delta^{-1} \partial_t ((\tilde{A}_i^L)^{L^{-1}}))^L = P_k Q l(\Delta^{-1} \partial_t P_{k+O(1)} Q l+O(1)((\tilde{A}_i^L)^{L^{-1}}))^L
\]

\[
= P_k Q l \int_{\mathbb{R}^4} m_k(a)(P_{k+O(1)} Q l+O(1) \tilde{A}_i^L)((t, x) - L^{-1}(0, a)) \, da.
\]

This immediately implies \( (l < k + \frac{C}{2}) \)

\[
2 \frac{1}{2} \left\| \nabla_x P_k Q l \gamma \right\|_{L^2_{t,x}}^2 = 2 \frac{1}{2} \left\| \nabla_x P_k Q l(\Delta^{-1} \partial_t ((\tilde{A}_i^L)^{L^{-1}}))^L \right\|_{L^2_{t,x}}^2 \leq \left\| P_k \tilde{A}_i^L \right\|_{\dot{X}^0_{\infty, x}}^4.
\]

One similarly shows that for \( l > k + C \) we have

\[
2 \frac{1}{2} \left\| P_k Q l \nabla_x^2 \gamma \right\|_{L^2_{t,x}}^2 \leq 2^{k-l} \left\| \nabla_{t,x} P_l \tilde{A}_i^L \right\|_{\dot{X}^0_{\infty, x}}^4.
\]

Finally, for the expression \( P_k Q l(k + \frac{C}{2}, k + C) \gamma \), the Lorentz transformation may lead to small frequencies \( \leq 2^k \), which is why we can only place the expression into \( L^2_{2} L^2_{t,x} \) then via Bernstein, i.e.

\[
(\sum_{k \in \mathbb{Z}} 2^{-1(k)} \left\| \nabla_{t,x} P_k Q l(k + \frac{C}{2}, k + C) \gamma \right\|_{L^2_{2} L^2_{t,x}}^2)^\frac{1}{2} \leq \left\| \tilde{A}_i^L \right\|_{S^1}.
\]

One also easily infers by similar reasoning that

\[
(\sum_{k \in \mathbb{Z}} \left\| \nabla_{t,x} P_k Q l(k + \frac{C}{2}, k + C) \gamma \right\|_{L^2_{2} L^2_{t,x}}^2)^\frac{1}{2} \leq \left\| \tilde{A}_i^L \right\|_{S^1}
\]

as well as

\[
\left\| \nabla_{t,x} P_k Q l(k + \frac{C}{2}, k + C) \gamma \right\|_{L^\infty_{2} L^2_{t,x}} \lesssim \left\| P_{k+O(1)} \tilde{A}_i^L \right\|_{L^\infty_{t} L^2_{x}}.
\]

With these bounds on \( \gamma \) in hand, we can now approach the bounds for \( \phi^L = e^{-i\gamma} \tilde{\phi}^L \). We write for \( l < k + \frac{C}{2} \)

\[
P_k Q l(e^{-i\gamma} \tilde{\phi}^L) = P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) \tilde{\phi}^L) + P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) \tilde{\phi}^L)
\]

\[
+ P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) \tilde{\phi}^L).
\]

For the first term on the right, we have

\[
P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) \tilde{\phi}^L) = P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) P_{k+O(1)} Q l=O(1) \tilde{\phi}^L),
\]

and so we infer

\[
(8.1) \quad 2 \frac{1}{2} \left\| \nabla_{t,x} P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) \tilde{\phi}^L) \right\|_{L^2_{t,x}}^2 \leq \left\| \nabla_{t,x} P_k \tilde{\phi}^L \right\|_{\dot{X}^0_{\infty, x}}^4.
\]

For the second term on the right, write schematically

\[
P_k Q l(P_{cl+10} Q l=10(e^{-i\gamma}) \tilde{\phi}^L) = 2^{-l} P_k Q l(P_{cl+10} Q l=10(\partial_t \gamma e^{-i\gamma}) \tilde{\phi}^L)
\]
and so we get from the preceding
\[
2\tilde{t}\|P_k Q_l\|_{L^2_t L^\infty_x} \lesssim 2\tilde{t} \cdot 2^{-\frac{\tilde{t}}{2}} \|P_{<\tilde{t} - 20} Q_{\tilde{t} - 10}(e^{-iy})\|_{L^2_x} \|\nabla_x P_k \bar{\phi}^L\|_{L^\infty_t L^2_x} \\
\lesssim \|A^L\|_{S^1} \|\nabla_x P_k \bar{\phi}^L\|_{L^\infty_t L^2_x}.
\] (8.2)

For the last term on the right, write it as
\[
P_k Q_l[P_{\geq \tilde{t} - 20}(e^{-iy})\bar{\phi}^L] = P_k Q_l[P_{[\tilde{t} - 20, k - 10]}(e^{-iy})\bar{\phi}^L] + P_k Q_l[P_{[k - 10, k + 10]}(e^{-iy})\bar{\phi}^L] \\
+ P_k Q_l[P_{> k + 10}(e^{-iy})\bar{\phi}^L].
\]
The first term on the right is bounded by
\[
2\tilde{t}\|\nabla_x P_k Q_l[P_{[\tilde{t} - 20, k - 10]}(e^{-iy})\bar{\phi}^L]\|_{L^2_x} \lesssim 2\tilde{t} \cdot 2^{-\frac{\tilde{t}}{2}} \|P_{[\tilde{t} - 20, k - 10]}(\nabla_x \gamma e^{-iy})\|_{L^2_t L^\infty_x} \|\nabla_x P_k \bar{\phi}^L\|_{L^\infty_t L^2_x} \\
\lesssim \|A^L\|_{S^1} \|\nabla_x P_k \bar{\phi}^L\|_{L^\infty_t L^2_x}.
\] (8.3)

For the second term on the right, write it as
\[
P_k Q_l[P_{[k - 10, k + 10]}(e^{-iy})\bar{\phi}^L] \\
= 2^{-k} P_k Q_l[P_{[k - 10, k + 10]}(\nabla \gamma P_{< k - 20} Q_{< k - 20}(e^{-iy}))P_{< k + 20}\bar{\phi}^L] \\
+ 2^{-k} P_k Q_l[P_{[k - 10, k + 10]}(\nabla \gamma P_{< k - 20} (e^{-iy}))P_{< k + 20}\bar{\phi}^L] \\
+ 2^{-k} P_k Q_l[P_{[k - 10, k + 10]}(\nabla \gamma P_{\geq k - 20}(e^{-iy}))P_{< k + 20}\bar{\phi}^L].
\]

Then estimate
\[
2\tilde{t}\|\nabla_x 2^{-k} P_k Q_l[P_{[k - 10, k + 10]}(\nabla \gamma P_{< k - 20} Q_{< k - 20}(e^{-iy}))P_{< k + 20}\bar{\phi}^L]\|_{L^2_x} \\
\lesssim 2\tilde{t}\|P_{[k - 10, k + 10]} Q_{< k + 20} \nabla \gamma\|_{L^\infty_t L^2_x} \|P_{< k + 20}\bar{\phi}^L\|_{L^\infty_t L^2_x} \\
+ 2\tilde{t}\|P_{[k - 10, k + 10]} \nabla \gamma\|_{L^\infty_t L^2_x} \|P_{< k + 20} Q_{\geq k - 20}\bar{\phi}^L\|_{L^\infty_t L^2_x} \\
\lesssim 2\tilde{t}^{-\frac{k}{2}} \|P_{k + O(1)} \sqrt{A^L}\|_{L^\infty_t L^2_x} \|\bar{\phi}^L\|_{S^1}.
\] (8.4)

Further, we get
\[
2\tilde{t}\|\nabla_x 2^{-k} P_k Q_l[P_{[k - 10, k + 10]}(\nabla \gamma P_{< k - 20} Q_{\geq k - 20}(e^{-iy}))P_{< k + 20}\bar{\phi}^L]\|_{L^2_x} \\
\lesssim 2\tilde{t}^{-k} \|P_{k + O(1)} \nabla \gamma\|_{L^\infty_t L^2_x} \|\nabla \gamma\|_{L^\infty_t L^2_x} \|P_{< k + 20} \bar{\phi}^L\|_{L^\infty_t L^2_x} \\
\lesssim 2\tilde{t}^{-k} 2\tilde{t}^{-\frac{k}{2}} \|P_{k + O(1)} \nabla \gamma\|_{L^\infty_t L^2_x} \|\bar{\phi}^L\|_{S^1} \|\bar{\phi}^L\|_{S^1}.
\] (8.5)

The term
\[
2^{-k} P_k Q_l[P_{[k - 10, k + 10]}(\nabla \gamma P_{\geq k - 20}(e^{-iy}))P_{< k + 20}\bar{\phi}^L]
\]
is handled similarly, which concludes treating the contribution of
\[
P_k Q_l[P_{[k - 10, k + 10]}(e^{-iy})\bar{\phi}^L].
\]
To treat the high-high interaction term
\[
P_k Q_l[P_{> k + 10}(e^{-iy})\bar{\phi}^L],
\]
write it as
\[ P_k Q [P_{>k} e^{-iy} \bar{\phi}^L] = \sum_{k_1 \geq k+10} \sum_{k_1 = k+O(1)} P_{k_1} Q [P_{k_1} (e^{-iy}) P_{>k} \bar{\phi}^L] = \sum_{k_1 \geq k+10} 2^{-k_1} P_k Q [P_{k_1} (\nabla y e^{-iy}) P_{>k} \bar{\phi}^L] \]
and so we can estimate this by
\[
2^k \left\| \nabla_y P_k Q [P_{>k} e^{-iy} \bar{\phi}^L] \right\|_{L^2_x L^4_t} \leq \sum_{k_1 \geq k+10} 2^{\frac{k_1}{k+1}} \left\| \nabla y \right\|_{L^2_x L^4_t} \left\| P_{k_1} \bar{\phi}^L \right\|_{L^{10}_x L^4_t}.
\]
Combining the bounds (8.1) - (8.6), applying \( \nabla_y \) and square-summing over \( k \), the bound
\[
\left( \sum_{k \in \mathbb{Z}} \left\| \nabla_y P_k Q \bar{\phi}^L \right\|_{L^2_x L^4_t}^2 \right)^{\frac{1}{2}} \leq \left\| (\bar{\phi}^L) \right\|_{S^1}
\]
with implied constant also depending on \( \left\| (\bar{\phi}^L) \right\|_{S^1} \) easily follows. We omit the estimate for
\[
\left( \sum_{k \in \mathbb{Z}} \left\| \nabla_y P_k Q \bar{\phi}^L \right\|_{L^2_x L^4_t}^2 \right)^{\frac{1}{2}}
\]
as it is similar. Next, consider
\[ P_k Q [Q_{k+\xi} (e^{-iy} \bar{\phi}^L)] = P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L)) + P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L)) + P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L)) + P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L)) + P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L)) + P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L))
\]
Each of these terms is straightforward to estimate. Thus for the first term on the right, we obtain
\[
\left\| P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L))] \right\|_{L^2_x L^4_t} \lesssim \left\| P_k Q [O(1) Q_{<k-20} (e^{-iy} \bar{\phi}^L)] \right\|_{L^2_x L^4_t} \lesssim 2^{(v-1)k} \left\| P_k \bar{\phi}^L \right\|_{S^1}.
\]
Also, we get
\[
\left\| P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L))] \right\|_{L^2_x L^4_t} \lesssim 2^{-k} \left\| \nabla_y y \right\|_{L^2_x L^4_t} \left\| P_k Q [O(1) \bar{\phi}^L] \right\|_{L^{10}_x L^4_t} \lesssim 2^{(v-1)k} \left\| P_k \bar{\phi}^L \right\|_{S^1},
\]
and
\[
\left\| P_k Q [Q_{k+\xi} (P_{<k-20} Q_{<k-20} (e^{-iy} \bar{\phi}^L))] \right\|_{L^2_x L^4_t} \lesssim 2^{-k} \left\| \nabla_y y \right\|_{L^2_x L^4_t} \left\| P_k Q [O(1) \bar{\phi}^L] \right\|_{L^{10}_x L^4_t} \lesssim 2^{(v-1)k} \sup_i 2^{-i} \left\| P_k \bar{\phi}^L \right\|_{S^1}.
\]
These bounds are then enough to furnish the second inequality of Observation 1. We also observe that the estimates on \( y \) established earlier give the required bounds for \( A^L = \bar{A}_L + \nabla y \).

Now we prove the last bound of Observation 1: Write
\[ P_k Q_{<k, \xi} \bar{\phi}^L = P_k Q_{<k, \xi} (P_{<k-10} (e^{-iy} \bar{\phi}^L) + P_k Q_{<k, \xi} (P_{<k+10} (e^{-iy} \bar{\phi}^L)) + P_k Q_{<k+10} (e^{-iy} \bar{\phi}^L)).
\]
The first term is directly bounded by
\[
\left\| \nabla_x P_k Q [P_{<k-10} (e^{-iy} \bar{\phi}^L)] \right\|_{L^{10}_x L^4_t} \lesssim \left\| \nabla_x P_k Q [O(1) \bar{\phi}^L] \right\|_{L^{10}_x L^4_t}.
\]
The second term on the right is a bit more complicated: Write schematically

\[ P_k Q_{ck+\frac{\lambda}{2}} (P_{[k-10,k+10]}(e^{-iy})\tilde{\phi}^L) \]

\[ = P_k Q_{ck-2c}(2^{-k} P_{[k-10,k+10]}(\nabla \gamma e^{-iy})P_{ck+O(1)}\tilde{\phi}^L) \]

\[ = P_k Q_{ck} (2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \]

\[ + P_k Q_{ck} (2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \]

\[ + P_k Q_{ck} (2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \]

Then we get for the first term of the last list of four terms

\[ \| \nabla P_k Q_{ck+\frac{\lambda}{2}} (2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \|_{L^\infty_t L^2_x} \]

(8.8)

\[ \leq \| P_{[k-15,k+15]} Q_{ck+\frac{\lambda}{2}} 5 \nabla \gamma \|_{L^\infty_t L^2_x} \| P_{ck-30} \tilde{\phi}^L \|_{L^\infty_t L^2_x} \leq \sum_l 2^{-\sigma l - k} \| (P_l \tilde{A}^L, P_l \tilde{\phi}^L) \|_{S^1}, \]

where we have taken advantage of our previous considerations on the structure of \( \gamma \). For the second term on the above list, we get

\[ \| \nabla P_k Q_{ck+\frac{\lambda}{2}} (2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \|_{L^\infty_t L^2_x} \]

(8.9)

\[ \leq \| P_{ck+O(1)} \nabla \gamma \|_{L^\infty_t L^2_x} \| P_{ck-30} Q_{ck-30}(e^{-iy}) \|_{L^\infty_t L^2_x} \| P_{ck-30} \tilde{\phi}^L \|_{L^\infty_t L^2_x} \]

\[ \leq \sum_l 2^{-\sigma l - k} \| (P_l \tilde{A}^L, P_l \tilde{\phi}^L) \|_{S^1}. \]

The term

\[ P_k Q_{ck-2c}(2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \]

is handled similarly. Finally, we have

\[ \| P_k Q_{ck} (2^{-k} P_{[k-10,k+10]}(\nabla \gamma P_{ck-30} Q_{ck-30}(e^{-iy}))P_{ck-30} \tilde{\phi}^L) \|_{L^\infty_t L^2_x} \]

(8.10)

\[ \leq \| P_{ck+O(1)} \nabla \gamma \|_{L^\infty_t L^2_x} \| P_{ck+O(1)} \tilde{\phi}^L \|_{L^\infty_t L^2_x}. \]

The bounds (8.7) - (8.10) suffice to perform the square summation over \( k \) in the last inequality of Observation 1. The term

\[ P_k Q_{ck+\frac{\lambda}{2}} (P_{>k+10}(e^{-iy})\tilde{\phi}^L) \]

treated similarly and hence omitted here. \( \Box \)

**Observation 2:** We have the bound

\[ \sum_{|k_1| \leq k_2} 2^{-k_1} \| Q_{k_1}^{1/2} \phi_{k_1} \nabla L, x Q_{k_1}^{1/2} \phi_{k_1} \|_{L^2_x} \leq \| (A^L, \phi^L) \|_{S^1}^{3/4}. \]

Moreover, for any \( L^1 \)-space-time integrable weight function \( m(a), a \in \mathbb{R}^{+1} \), we have

\[ \sum_{|k_1| \leq k_2} 2^{-k_1} \| \int_{\mathbb{R}^{+1}} m(a) Q_{k_1}^{1/2} \phi_{k_1} (-a) \nabla L, x Q_{k_1}^{1/2} \phi_{k_2} (-a) da \|_{L^2_x} \leq \| (A^L, \phi^L) \|_{S^1}^{3/4}. \]

Similar bounds can be obtained upon replacing one or more factors by \( A^L \). We note that these bounds are essentially invariant under mild Lorentz transforms. Thus we infer similar bounds for \( \phi, A \).
Proof. (Observation 2) Here one places the low frequency input
\[ Q_{<k_1+\frac{\epsilon}{2}k_1} \]
into \( L^2_{\infty}L^\infty_x \), and the high-frequency input
\[ Q_{<k_2+\frac{\epsilon}{2}k_2} \nabla_{t,x}\phi_{k_2} \]
into \( L^\infty_xL^2_t \), using Observation 1. The fact that \( P_k\phi^L \) can be placed into \( L^2_{\infty}L^\infty_x \) and that we have
\[ \sum_k 2^{-k}\|P_k\phi^L\|^2 \leq \|A^L,\tilde{\phi}^L\|^2_{S^1} \]
follows easily from the proof of Observation 1. Also, note that the operators \( P_{k_2}Q_{<k_1+\frac{\epsilon}{2}k_1} \) act boundedly on spaces of the type \( L^p_tL^q_x \). 

Using Observations 1 and 2, we can now move toward controlling the norm \( \|(A,\phi)\|_{S^1} \). From above, we know a priori that we control
\[ \left( \sum_{k \in \mathbb{Z}} \|\nabla_{t,x}P_kQ_{\left\lfloor k+\frac{\epsilon}{4}k \right\rfloor} \phi^L\|^2 \right)^{\frac{1}{2}} \]
as well as norms of the form
\[ \left( \sum_{k_1 \leq k_2} 2^{-k_1}\int_{\mathbb{R}^{n+1}} m(a)Q_{<k_1+\frac{\epsilon}{2}k_1} \phi_{k_1}^L \cdot \nabla_{t,x}Q_{<k_1+\frac{\epsilon}{2}k_2} \phi_{k_2}^L \right) \, da \right)^{\frac{1}{2}} \]
The latter has the crucial divisibility property. Thus for any \( \delta > 0 \) we can divide the time axis \( \mathbb{R} \) into intervals \( J_1, J_2, \ldots, J_N \), with \( N \) depending on the size of the norm as well as \( \delta \), such that we have for each \( j \)
\[ \left( \sum_{k_1 \leq k_2} 2^{-k_1}\int_{\mathbb{R}^{n+1}} m(a)Q_{<k_1+\frac{\epsilon}{2}k_1} \phi_{k_1}^L \cdot \nabla_{t,x}Q_{<k_1+\frac{\epsilon}{2}k_2} \phi_{k_2}^L \right) \, da \right)^{\frac{1}{2}} \leq \delta. \]

Of course we get a similar statement for weakened versions of the former norm, such as
\[ \left( \sum_{k \in \mathbb{Z}} \|\nabla_{t,x}P_kQ_{\left\lfloor k-\frac{\epsilon}{4}k \right\rfloor} \phi^L\|^2 \right)^{\frac{1}{2}}. \]

In order to infer the desired bound on \( (A,\phi) \), we shall partition \( \mathbb{R} \) into finitely many intervals \( I_1, \ldots, I_N \), whose number depends on
\[ \|(A,\phi)\|_{S^1(I \times \mathbb{R}^n)} \]
and such that on each of these \( I_j \), we can infer via a direct bootstrap argument a bound on
\[ \|(A,\phi)\|_{S^1(I_j \times \mathbb{R}^n)} \]
This will then suffice to obtain the desired bound on \( \|(A,\phi)\|_{S^1(I \times \mathbb{R}^{n+1})} \). We do this in two steps, which we outline below.

Step 1: Given \( \delta_1, \delta_2 > 0 \), using the a priori bounds known and choosing the \( I_j \) suitably as above (whose number will depend on \( \delta_1, \delta_2 \), as well as the assumed bound on \( \|(A^L,\tilde{\phi}^L)\|_{S^1} \)), upon writing
\[ A|_{I \times \mathbb{R}^4} = A^\text{free}_{I \times \mathbb{R}^4} + A^\text{nonlin}_{I \times \mathbb{R}^4}, \]
we can infer from the A equation that there is a decomposition
\[ A^\text{nonlin}_{I \times \mathbb{R}^4} = A^\text{nonlin}^1_{I \times \mathbb{R}^4} + A^\text{nonlin}^2_{I \times \mathbb{R}^4}, \]
where we have schematically
\[
A^{2\text{nonlin}} = \sum_k \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)} Q_{\ll -C_1} \phi \nabla_x P_{k+O(1)} Q_{\ll -C_1} \phi),
\]
while we also have the bound
\[
\left\| A^{1\text{nonlin}} \right\|_{L^2(I \times \mathbb{R}^4)} \leq \delta_1 + \delta_2 \left\| (A, \phi) \right\|_{L^2(I \times \mathbb{R}^4)}^3 + \left\| (A, \phi) \right\|_{L^2(I \times \mathbb{R}^4)}^3, \forall j = 1, 2, \ldots, N.
\]
We also have the bound
\[
\left\| A^{2\text{nonlin}} \right\|_{L^2(I \times \mathbb{R}^4)} \leq \delta_1 \forall l.
\]
The idea behind this bound is to insert this into the \( \phi \) equation, and pick \( \delta_1 \) and \( \delta_2 \) small depending on \( E_{\text{crit}} \).

To accomplish Step 1, proceeding as in the proof of Lemma 7.6 we write the source term in the \( A \)-equation in the schematic form
\[
\Box A_i = \Delta^{-1} \nabla_j N_{ij}(\phi, \phi) + A\phi^2.
\]
Localizing this to frequency \( k = 0 \), we write the right hand side in the form
\[
P_0(\Delta^{-1} \nabla_j N_{ij}(\phi, \phi) + A\phi^2) = P_0(\sum_{k_1 \neq k_2} \Delta^{-1} \nabla_j N_{ij}(P_{k_1} \phi, P_{k_2} \phi) + \sum_{k_1, k_2, \lambda} P_{k_1} P_{k_2} \phi P_{k_3} \phi).
\]

We first deal with the quadratic null form term. We first reduce this to moderate frequencies by observing that for \( C = C(\delta_2) \) large enough, we get (for a suitable absolute constant \( \sigma \) independent of all other constants)
\[
\left\| \sum_{|k_1| > C, k_2} P_0 \Delta^{-1} \nabla_j N_{ij}(P_{k_1} \phi, P_{k_2} \phi) \right\|_{H^0} \leq \delta_2 \sum_{k_1 > C} 2^{-|\sigma|} \left\| P_{k_1} \phi \right\|_{L^2}^2.
\]
Generalizing to arbitrary output frequencies, one easily gets from here the bound
\[
\sum_k \left\| \sum_{|k_1| > C, |k_2|} P_k \Delta^{-1} \nabla_j N_{ij}(P_{k_1} \phi, P_{k_2} \phi) \right\|_{H^0} \leq \delta_1 \left\| \phi \right\|_{L^2}^2.
\]

Next, we pick \( C_1 = C_1(E_{\text{crit}}) \) such that
\[
\left\| P_0 \left( \sum_{|k_1, k_2| < C} \Delta^{-1} \nabla_j N_{ij}(P_{k_1} Q_{\ll -C_1} \phi, P_{k_2} \phi) \right) \right\|_{H^0} \leq \delta_2 \sum_{|k_1, k_2| < C} \left\| P_{k_1} \phi \right\|_{S[|k_1|]} \left\| P_{k_2} \phi \right\|_{S[|k_2|]},
\]
and generalizing to general output frequencies, we then reduce to
\[
\sum_k P_k \left( \sum_{|k_1|, |k_2| < C} \Delta^{-1} \nabla_j N_{ij}(P_k Q_{\ll -C_1} \phi, P_{k_2} Q_{\ll -C_1} \phi) \right)
\]
Depending on our choice \( C_1 \) we may assume the Lorentz transform \( L \) to be chosen sufficiently close to the identity (i.e. \( |d| \) sufficiently small) such that according to Observation 1 we have
\[
\left( \sum_k \right\| P_{k_1} Q_{k_1} \Delta^{-1} \nabla_j N_{ij}(P_{k_1} Q_{\ll -C_1} \phi, P_{k_2} Q_{\ll -C_1} \phi) \right\|_{L^2}^2 \leq \left\| (\hat{A}^L, \hat{\phi}^L) \right\|_{S^1}^2.
\]
As observed before, this norm has the divisibility property, so that restricting to suitable time intervals \( I_l, l = 1, 2, \ldots, N \), which form a partition of the time axis \( \mathbb{R} \), we may assume
\[
\left( \sum_k \right\| P_{k_1} Q_{k_1} \Delta^{-1} \nabla_j N_{ij}(P_{k_1} Q_{\ll -C_1} \phi, P_{k_2} Q_{\ll -C_1} \phi) \right\|_{L^2}^2 \leq \delta_2 \forall l.
\]
But then we easily infer the bound
\[ P_0\left( \sum_{|k_1,2|<C} \Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1+k_1+C_1|}\phi, P_{k_2}\phi) \right)_{N[0][I\times\mathbb{R}^4]} \]
\[ \leq \|P_k Q_{|k_1-C_1+k_1+C_1|}\phi\|_{L^2(I\times\mathbb{R}^4)} \|P_{k_2}\phi\|_{L^2(I\times\mathbb{R}^4)} \leq \delta_2 \|P_{k_2}\phi\|_{S[k_1]} \]
and this suffices again, after generalizing this to arbitrary output frequency. In fact, we get
\[ \sum_k \|P_k(\sum_{|k_1,2-k|<C} \Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1+k_1+C_1|}\phi, P_{k_2}\phi))\|_{N[k](I\times\mathbb{R}^4)} \]
\[ \leq (\sum_{k_1} \|P_k Q_{|k_1-C_1+k_1+C_1|} \nabla_j \phi \|_{L^2(I\times\mathbb{R}^4)}^{1/2}) \|\phi\|_{S^1(I\times\mathbb{R}^4)} \]
\[ \leq \delta_1 \|\phi\|_{S^1(I\times\mathbb{R}^4)} \]
We have now reduced to the expression
\[ \sum_k P_k(\sum_{|k_1,2-k|<C} \Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1}\phi, P_{k_2} Q_{|k_2-C_1|}\phi)) \]
The last reduction here consists in removing extremely small angular separation between the inputs
\[ P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi. \]
Thus, there is a \( C_2 = C_2(\delta_2) \), such that we have
\[ \|P_0(\sum_{|k_1,2-k|<C} \Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi))\|_{N[0][I\times\mathbb{R}^4]} \leq \delta_2 \sum_{|k_1,2-k|<C} \|P_k \phi\|_{S[k_1]} \|P_{k_2}\phi\|_{S[k_2]} \]
where the prime indicates that the inputs are reduced to have closely aligned Fourier supports of angular separation \( C_2^{-1} \). Finally, we write
\[ \sum_{|k_1,2|<C} P_0(\Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi)) \]
\[ = \sum_{|k_1,2|<C} P_0(\Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi))' \]
\[ + \sum_{|k_1,2|<C} P_0(\Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi))'' , \]
where the second term is of the form \( A^{nonlin} \) as required for Step 1. In fact, the angular separation of the inputs and small modulation forces the output to have modulation \( \sim 1 \). Moreover, replacing the output frequency by \( k \) and \textit{square-summing} over \( k \) results in a small norm due to the fact that
\[ \|\sum_{|k_1,2|<C} P_0(\Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi))''\|_{N[0][I\times\mathbb{R}^4]} \]
\[ \leq \|\sum_{|k_1,2|<C} P_0(\Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi))''\|_{L^2(I\times\mathbb{R}^4)} \]
and we can take advantage of Observation 2 to obtain
\[ (\sum_{k} \|\sum_{|k_1,2-k|<C} P_k(\Delta^{-1}\nabla_j N_{ij}(P_k Q_{|k_1-C_1|}\phi, P_{k_2} Q_{|k_2-C_1|}\phi))''\|_{N[k](I\times\mathbb{R}^4)} \|^{2} \leq \delta_1 \]
by choosing the intervals \( I_I \) suitably. The cubic term \( \sum_{k_1,2,3} P_{k_1} A P_{k_2} \phi P_{k_3} \phi \) is handled similarly.
Step 2: Choosing the time intervals $I_l$ suitably as in Step 1, we obtain the equation

$$ \sum_{k \in \mathbb{Z}} \Box_{A_{ck}} P_k \phi = F, $$

where we have

$$ \|F\|_{L^2(I \times \mathbb{R}^4)} \lesssim \delta_1 + \delta_2 \|\langle A, \phi \rangle\|_{L^1(I \times \mathbb{R}^4)} + \|\langle A, \phi \rangle\|_{L^1(I \times \mathbb{R}^4)}, $$

where $A^{\text{free}}$ is the free wave evolution of the data for $A$ at the beginning endpoint of $I_l$.

This to a large extent mimics the argument for the proof of Proposition 7.5. In fact, we recall from there that we can write $F = \sum_{k \in \mathbb{Z}} F_k$ with

$$ F_k = -2iP_k(A^{\text{free}}_{\geq k} \phi^v) - [P_k, \Box_{A^{\text{free}}_{ck}}] \phi - P_k([\Box A - \Box A^{\text{free}}] \phi) + P_k(\Box_{A^{\text{free}}_{ck}} \phi + 2i(A^{\text{free}}_{\geq k} \phi^v) - \Box A^{\text{free}} \phi). $$

As usual, we treat each term separately:

First term. Similar to the proof of Proposition 7.5, reduce it to

$$ -2iP_k(P_{k+O(1)}A^{\text{free}} \phi^v P_{k+O(1)} \phi) $$

up to terms satisfying the conclusion of Step 2. Then using divisibility for the norm

$$ \sum_{k \in \mathbb{Z}} 2^{-k} \|P_k A^{\text{free}} \|_{L^2 L^{\infty}} $$

as well as the inequality

$$ \| \sum_{k \in \mathbb{Z}} -2iP_k(P_{k+O(1)}A^{\text{free}} \phi^v P_{k+O(1)} \phi) \|_{L^1(I \times \mathbb{R}^4)} \lesssim (\sum_{k \in \mathbb{Z}} 2^{-k} \|P_k A^{\text{free}} \|_{L^2 L^{\infty}})^{1/2} \|\phi\|_{L^1(I \times \mathbb{R}^4)}, $$

we get the conclusion of Step 2 by choosing the $I_l$ suitably and by subdividing the intervals obtained from Step 1, if necessary.

Second term. This is handled like the first term, since it can be written in the form

$$ \sum_k \tilde{P}_k(2i\nabla_s A^{\text{free}}_{< k} \phi^v) + \tilde{P}_k(\nabla_s ((A^{\text{free}}_{< k})^2) \phi). $$

Third term. This is as usual the most difficult term, since it contains

$$ \sum_{k \in \mathbb{Z}} 2iP_k A_{\text{nonlin}}^v \phi^v \phi. $$

We then essentially follow the reductions performed in the proof of Proposition 7.5, whence we shall be correspondingly brief.

Reduction to $\mathcal{H}^s N^{\text{lowhi}}$. Using the same notation as in that proof and restricting to frequency $k = 0$, and also keeping in mind Step 1, we get

$$ \|N^{\text{lowhi}}(A_{1,\infty}^{\text{nonlin}},0,\phi_0) - \mathcal{H}^s N^{\text{lowhi}}(A_{1,\infty}^{\text{nonlin}},0,\phi_0)\|_{L^1(I \times \mathbb{R}^4)} \lesssim \|A_{1,\infty}^{\text{nonlin}}\|_{L^1(I \times \mathbb{R}^4)} \|\phi_0\|_{L^1(I \times \mathbb{R}^4)} $$

and hence replacing the output frequency by general $k \in \mathbb{Z}$ and square-summing gives the bound

$$ \lesssim \|\langle A, \phi \rangle\|_{L^1(I \times \mathbb{R}^4)} \|\delta_1 + \delta_2 \|\langle A, \phi \rangle\|_{L^1(I \times \mathbb{R}^4)} + \|\langle A, \phi \rangle\|_{L^1(I \times \mathbb{R}^4)}^3, $$

which is of the desired form. This then reduces the estimate of

$$ N^{\text{lowhi}}(A_{1,\infty}^{\infty},0,\phi_0) - \mathcal{H}^s N^{\text{lowhi}}(A_{1,\infty}^{\infty},0,\phi_0) $$
to the contribution of $A_{2\text{nonlin}}^{k<0}$, whose explicit form we recall from Step 1. This means we have to estimate the expression

$$(1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C}) \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) \nabla_{t,x} \phi_0).$$

The idea here is to use the a priori bounds from Observations 1 and 2 to arrive at the required estimate. For this, we split the above expression into the following:

$$(1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C} \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) Q_{\xi,C} \nabla_{t,x} \phi_0)
= (1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C} \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) Q_{\xi,C} \nabla_{t,x} \phi_0)
+ (1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C} \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) Q_{\xi,C} \nabla_{t,x} \phi_0)
+ (1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C} \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) Q_{<C} \nabla_{t,x} \phi_0).$$

We now estimate each of the terms on the right in turn. Call them $A, B, C, D$.

**Estimate for term $A$.** We distinguish between very small $k$ and $k = O(1)$. In the latter case, we schematically estimate the term in the following fashion: We shall suppress the distinction between space-time translates of $\phi$ and $\phi$, as our norms are invariant under these, and also keep in mind that the operator

$\Box^{-1} P_k Q_{k+O(1)}$

is given by (space-time) convolution with a kernel of $L^1$-mass $\sim 2^{-2k}$. Then we get in case $k = O(1)$,

$$\|(1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C} \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) Q_{\xi,C} \nabla_{t,x} \phi_0)\|_{N(I,\mathbb{R}^2)}$$

$$\leq \|(1 - \mathcal{H}^*)(\sum_{k < 0} \Box^{-1} P_k Q_{k+O(1)}(P_{k+O(1)}Q_{<k-C} \phi \nabla_x P_{k+O(1)}Q_{<k-C} \phi) Q_{\xi,C} \nabla_{t,x} \phi_0)\|_{L^1_t L^2_x(I,\mathbb{R}^2)}$$

$$\leq 2^{-2k}\|P_{k+O(1)}Q_{<k-C} \phi Q_{\xi,C} \nabla_{t,x} \phi_0\|_{L^2_t L^2_x}.$$

Here the second factor is essentially invariant under mild Lorentz transformations, and so we get (up to changing the meaning of the constants slightly)

$$\|\nabla_x P_{k+O(1)}Q_{<k-C} \phi Q_{\xi,C} \nabla_{t,x} \phi_0\|_{L^2_{t,x}} \leq \|\nabla_x P_{k+O(1)}Q_{<k-C} \phi^L Q_{\xi,C} \nabla_{t,x} \phi^L_{<O(1)}\|_{L^2_{t,x}}.$$

We estimate the last norm using Observation 1, resulting in the bound

$$\|\nabla_x P_{k+O(1)}Q_{<k-C} \phi Q_{\xi,C} \nabla_{t,x} \phi_0\|_{L^2_{t,x}} \leq \|\nabla_x P_{k+O(1)}Q_{<k-C} \phi^L\|_{L^2_{t}L^4_{x}} \|Q_{\xi,C} \nabla_{t,x} \phi^L_{<O(1)}\|_{L^2_{t}L^4_{x}}$$

$$\leq C([\tilde{A}^L, \tilde{\phi}^L]_{S^1}).$$

Then by divisibility of the norm

$$(\sum_{r \in \mathbb{Z}} 2^{-r}[\sum_{k - r = O(1)} \|\nabla_x P_{k+O(1)}Q_{<k-C} \phi Q_{\xi,C} \nabla_{t,x} \phi_0\|_{L^2_{t,x}}]^2)^{1/2},$$
we arrive upon suitable choice of the $I_j$ at the conclusion that
\[
\| \sum_{r \in \mathbb{Z}} (1 - \mathcal{H}^*) (\sum_{|k-r| = O(1)} \Box^{-1} P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi) Q_{[\frac{k}{2}, k+C]_r} \nabla_{t,x} \phi_r) \|_{L^1_t L^4_x(x \in \mathbb{R}^4)} \leq \delta_2 \| \phi \|_{S^1(I_j \times \mathbb{R}^4)}.
\]

This completes the contribution of $A$ when $k = O(1)$. On the other hand, when $k \ll -1$, the smallness gain comes directly from $k$. Indeed, we can then estimate
\[
\| (1 - \mathcal{H}^*) (\sum_{k \ll -1} \Box^{-1} P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi) Q_{[k-2C, k+C]} \nabla_{t,x} \phi_0) \|_{L^1_t L^4_x(x \in \mathbb{R}^4)} \leq \sum_{k \ll -1} 2^{-2k} \| P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi) Q_{[k-2C, k+C]} \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \sum_{k \ll -1} 2^{-2k} \| \phi_0 \|_{S^1(I_j \times \mathbb{R}^4)} \| \phi \|_{S^1(I_j \times \mathbb{R}^4)} \leq \delta_2 \| \phi \|_{S^1(I_j \times \mathbb{R}^4)}.
\]

Replacing $\phi_0$ by $\phi_r$, $r \in \mathbb{Z}$, and square-summing over $r$ results in the desired bound. This completes the estimate for term $A$.

**Estimate for term B.** Here we use the bound
\[
\| (1 - \mathcal{H}^*) (\sum_{k \ll 0} \Box^{-1} P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi) Q_{[k-2C, k+C]} \nabla_{t,x} \phi_0) \|_{L^1_t L^4_x(x \in \mathbb{R}^4)} \leq \sum_{k \ll 0} 2^{-2k} \| P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi) Q_{[k-2C, k+C]} \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \sum_{k \ll 0} 2^{-2k} \| \phi_0 \|_{S^1(I_j \times \mathbb{R}^4)} \| \phi \|_{S^1(I_j \times \mathbb{R}^4)} \leq \delta_2 \| \phi \|_{S^1(I_j \times \mathbb{R}^4)}.
\]

Now if we further restrict the above term to $|k - r| \gg 1$, we easily bound it by
\[
\| \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \| \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^4_t L^2_x(x \in \mathbb{R}^4)} \leq \| \Box^{-1} P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \delta_2 \| \phi \|_{S^1(I_j \times \mathbb{R}^4)}.
\]

which is as desired. On the other hand, when restricting the modulation of $Q_{[k-2C, k+C]} \nabla_{t,x} \phi_0$ to $r = k + O(1)$, we use the fact that for $k \ll -1$,
\[
\| \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \| \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^4_t L^2_x(x \in \mathbb{R}^4)} \leq \| P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \| (\Box^{-1} P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq \delta_2 \| \phi \|_{S^1(I_j \times \mathbb{R}^4)}.
\]

Then changing the frequency 0 to general $p \in \mathbb{Z}$ and using Observation 1, we infer
\[
\sum_{k < p - C} 2^{-3k} \| \nabla_x P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)} \leq C(\| \Box^{-1} P_k Q_{k+O(1)}(P_k+O(1)Q_{<k-C_1} - C_1) \phi \nabla_{t,x} \phi_0) \|_{L^2_t L^4_x(x \in \mathbb{R}^4)}).
\]
Also, the square-sum norm on the left has the divisibility property, whence by restricting to suitable
time intervals $I_j$, we may arrange it to be $\ll \delta_2$. Finally, we infer the bound
\[ \sum_p \left\| (1 - \mathcal{H}) (\sum_{k < p - C} \Box^{-1} P_k Q_{k_0+O(1)}(P_k + O(1)Q_{ck-C}e^{-\phi} \nabla_x P_k + O(1)Q_{ck-C}e^{-\phi})Q_{k+O(1)}\nabla_{t,x} \phi) p) \right\|_{L^2(I_j \times \mathbb{R}^4)}^2 \leq (\sum_k 2^{-k} \left\| P_k Q_{ck-C}e^{-\phi} \right\|_{L^2_t L^6_x})^2 (\sum_{k < p - C} 2^{-3k} \left\| \nabla_x P_k Q_{ck-C}e^{-\phi} Q_{k+O(1)}\nabla_{t,x} \phi \right\|_{L^2_t L^6_x})^{1/2} \leq \delta_2 \left\| \phi \right\|_{S^1(I_j \times \mathbb{R}^4)}^2.

**Estimate for term C.** This follows the same pattern as term $B$, by placing the product
\[ \nabla_x P_k Q_{k_0+O(1)}(P_k + O(1)Q_{ck-C}e^{-\phi}) \] into $L^2_t L^\infty_x$ and using Observation 2.

**Estimate for term D.** Here one places
\[ \Box^{-1} P_k Q_{k_0+O(1)}(P_k + O(1)Q_{ck-C}e^{-\phi}) \nabla_x P_k + O(1)Q_{ck-C}e^{-\phi} \nabla_{t,x} \phi_0 \] into $L^2_t L^\infty_x$ and
\[ Q_{ck} \nabla_{t,x} \phi_0 \] into $L^2_t L^\infty_x$, keeping in mind that $C \gg C_1 = C_1(E_{crit})$ is very large.

**Reduction to $H^*N^{lowhi}(H^A_{t,x \times \mathbb{R}^4}, \phi_0)$.** To begin with, recall the notation from the proof of Proposition 7.5 for the definition of the symbol $\mathcal{H}$ applied to bilinear expressions. To reduce to this term, we need to estimate the difference
\[ \left\| H^*N^{lowhi}(A^A_{t,x \times \mathbb{R}^4}, \phi_0) - H^*N^{lowhi}(H^A_{t,x \times \mathbb{R}^4}, \phi_0) \right\|_{L^2_t L^\infty_x}. \]

Here we recall that
\[ H^\phi M(\phi, \psi) = \sum_{j < k + C} Q_j M(Q_{<j-C}e^{-\phi}, Q_{<j-C}e^{-\phi}) \] as well as
\[ H^\phi M(\phi, \psi) = \sum_{k < k_{1,2} - C} H^k M(P_{k_1} e^{-\phi} P_{k_2} e^{-\phi}). \]

Then write for the spatial components of $(I - H)^{nonlin}_{t,x \times \mathbb{R}^4, <0}$
\[ (I - H)^{nonlin}_{t,x \times \mathbb{R}^4, <0} = \sum_{k < 0} \Box^{-1} P_k P_x (P_{k_1} e^{-\phi} P_{k_2} e^{-\phi}) + \sum_{k < k_{1,2} - C} \Box^{-1} P_k Q_{j+k+C} (P_{k_1} e^{-\phi} P_{k_2} e^{-\phi}) + \sum_{k < k_{1,2} - C} \Box^{-1} P_k Q_{j+k+C} (P_{k_1} e^{-\phi} P_{k_2} e^{-\phi}) + \sum_{k < k_{1,2} - C} \Box^{-1} P_k Q_{j+k+C} (P_{k_1} e^{-\phi} P_{k_2} e^{-\phi}). \]
For the first term on the right, employing notation introduced in [12] and also used in the proof of Proposition 7.5, we get upon further restricting to

\[ |\max\{k_1, k_2\} - |\min\{k_1, k_2\}| \gg 1, \]

the smallness gain

\[ \| \sum_{k < 0} |k| \| P_k \mathcal{P}_x (P_{k_1} \phi \nabla_x P_{k_2} \phi) \|_Z \ll \delta_2 \| \phi_0 \|_{S^1}, \]

and the corresponding contribution to \( \mathcal{H}^s N^\text{lowhi} ((I - \mathcal{H}) A^\text{nonlin}_{I_j \times \mathbb{R}^4}, \phi_0) \) can then be bounded with respect to \( \| \cdot \|_{N(I_j \times \mathbb{R}^4)} \) by

\[ \leq \delta_2 \| \phi \|_{S^1} \| \phi_0 \|_{S^1}, \]

which upon replacing 0 by general frequencies and square summing gives the desired bound. Similarly, for the remaining terms on the right above, one may reduce to \( k_{1,2} = k + O(1) \), see estimate (134) in [12]. Finally, in each of these terms, we may reduce the output to modulation \( \sim 2^k \), since else one gains smallness due to the null form structure for

\[ \| \mathcal{H}^s N^\text{lowhi} (A^\text{nonlin}_{I_j \times \mathbb{R}^4}, \phi_0) \|_{N(I_j \times \mathbb{R}^4)} \]

Thus we have now reduced to estimating (and gaining a smallness factor) for the schematic expression

\[ \sum_{k < 0, k_{1,2} = k + O(1)} |k| \| P_k \mathcal{Q}_{k_1 + O(1)} (P_{k_1} \phi \nabla_x P_{k_2} \phi) \partial^\nu Q_{< k - C} \phi_0 \|_{N(I_j \times \mathbb{R}^4)}, \]

Here we can suppress the operator \( \square^{-1} P_k Q_{k + O(1)} \), which is given by convolution with a space-time kernel of \( L^1 \)-norm \( \sim 2^{-2k} \), and then schematically estimate the preceding via

\[ \| \sum_{k < 0, k_{1,2} = k + O(1)} |k| \| P_k \mathcal{Q}_{k_1 + O(1)} (P_{k_1} \phi \nabla_x P_{k_2} \phi) \partial^\nu Q_{< k - C} \phi_0 \|_{N(I_j \times \mathbb{R}^4)} \]

\[ \leq \sum_{k_1 = k_{1,2} + O(1)} 2^{-2k_1} \| P_{k_1} \phi \|_{L^2_{x,t} L^\infty_y (I_j \times \mathbb{R}^4)} \| \nabla_x P_{k_2} \phi Q_{< k - C} \phi_0 \|_{L^2_{x,t}(I_j \times \mathbb{R}^4)}. \]

Here we exploit Lorentz invariance of the norm of the right factor to obtain

\[ \sum_{k_2 < 0} 2^{-3k_2} \| \nabla_x P_{k_2} \phi Q_{< k - C} \phi_0 \|_{L^2_{x,t}}^2 \leq \sum_{k_2 < 0} 2^{-3k_2} \| \nabla_x (P_{k_2} \phi L^\nu Q_{< k - C} \phi_0) \|_{L^2_{x,t}}^2 \leq \| (A^L, \tilde{\phi}^L) \|_{S^1}^4. \]

In fact, distinguishing as usual between different frequency/modulation configurations for either of the factors, one estimates the \( L^2_{x,t} \)-norm of the input by placing the first input into \( L^2_{x,t} L^\infty_y \) and the second into \( L^\nu_y L^2_{x,t} \), both of which are controlled (by the proof of ) Observation 1. Using divisibility of the norm \( \| \cdot \|_{L^2_{x,t}} \), it now follows that upon proper choice of the intervals \( I_j \) (whose number of
course only depends on \( \| ( \mathcal{A}, \phi) \|_{S^1} \), we get the estimate
\[
\| \sum_{k<0, k_{1,2}=k+O(1)} \Box^{-1} P_k Q_{k+O(1)}(P_{k,1} \phi \nabla_x P_{k,2} \phi) \partial^\nu Q_{<k-C} \phi_0 \|_{L^2(I \times \mathbb{R}^4)} \\
\leq \left( \sum_{k_1<0} 2^{-k_1} \| P_{k,1} \phi \|_{L^2(I \times \mathbb{R}^4)}^2 \right)^{1/2} \left( \sum_{k_2<0} 2^{-2k_2} \left\| \nabla_x P_{k,2} \phi Q_{<k-C} \phi_0 \right\|_{L^2_r(I \times \mathbb{R}^4)}^2 \right)^{1/2} \\
\ll \delta_2�(\| \phi \|_{S^1(I \times \mathbb{R}^4)}).
\]

Of course, one gets the same bound upon replacing the frequency 0 by general \( p \in \mathbb{Z} \) and square summing.

As usual similar reductions can be applied to the elliptic interaction term \( A_{0,0} \partial_x \phi_0 \).

**Dealing with \( \mathcal{H}^s \mathcal{N}^{\text{lowhi}}(\mathcal{H}^s \mathcal{N}^{\text{nonlin}} \phi_0) \).** Here we exploit the null structure arising from combining the elliptic as well as hyperbolic terms, just as in the proof of Proposition 7.5 or as in [12]. Correspondingly, we have to analyze three null forms, each in turn:

The first null form. We can write it as
\[
\sum_{j<k, k_{1,2}=k+O(1)} \Box^{-1} P_k Q_j(Q_{<j-C} \phi_{k_1} \partial_\alpha Q_{<j-C} \phi_{k_2}) \partial^\nu Q_{<k-C} \phi_0.
\]
From (148) in [12], it follows that we may restrict to \( j = k+O(1) \), as otherwise the desired smallness follows (even without restriction to smaller time intervals). Furthermore, if \( k_{1,2} < 0 \), then we gain exponentially in the difference \( k - k_1 \), while if \( k_{1,2} \geq 0 \), we gain exponentially in \( k \). So we may further restrict to
\[
\sum_{k<0, k_{1,2}=k+O(1)} \Box^{-1} P_k Q_{k+O(1)}(Q_{<k-C} \phi_{k_1} \partial_\alpha Q_{<k-C} \phi_{k_2}) \partial^\nu Q_{<k-C} \phi_0
\]
and from here the argument proceeds exactly as before by suppressing the operator \( \Box^{-1} P_k Q_{k+O(1)} \) and schematically using
\[
\| \partial_\alpha Q_{<k-C} \phi_{k_1} \partial^\nu Q_{<k-C} \phi_0 \|_{L^2_r L^\infty_x}
\]
while placing \( Q_{<k-C} \phi_{k_1} \) into \( L^2_t L^\infty_x \).

The second and third null forms. These are treated identically and hence omitted here.

This completes Step 2. Together with Step 1, the linear theory for the operator \( \Box \mathcal{A} P_k \) and a standard bootstrap argument, this easily yields the bounds claimed in Proposition 8.1 for the localized norms \( \| (\mathcal{A}, \phi) \|_{S^1(I \times \mathbb{R}^4)} \). From there one can glue the localized components together to get the global bounds. \( \square \)

8.2. **The effect of Lorentz transformations on the energy.** In this subsection we assume that \( (A, \phi) \) is an admissible global solution to MKG-CG on \( \mathbb{R}^{1+4} \). For our purposes, it shall suffice to consider very special Lorentz transforms of the following form

\[
L = \begin{pmatrix}
\frac{1}{\sqrt{1-d^2}} & -\frac{d}{\sqrt{1-d^2}} & 0 & 0 & 0 \\
\frac{d}{\sqrt{1-d^2}} & \frac{1}{\sqrt{1-d^2}} & 0 & 0 & 0 \\
\frac{1}{\sqrt{1-d^2}} & \frac{d}{\sqrt{1-d^2}} & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0
\end{pmatrix}.
\]
Accordingly, the magnetic potential $A_a$ is transformed into $\tilde{A}_a$ as follows:

$$
\tilde{A}_0 = \frac{1}{\sqrt{1 - d^2}} A_0 - \frac{d}{\sqrt{1 - d^2}} A_1, \quad \tilde{A}_1 = \frac{1}{\sqrt{1 - d^2}} A_1 - \frac{d}{\sqrt{1 - d^2}} A_0, \quad \tilde{A}_j = A_j,
$$

where $j = 2, 3, 4$ (of course these expressions are to be evaluated at $Lx$). Then we compute the corresponding curvature components. We obtain

$$
\tilde{F}_{01} = \partial_t \tilde{A}_1 - \partial_1 \tilde{A}_0 = -\frac{d}{\sqrt{1 - d^2}} \left[ \frac{1}{\sqrt{1 - d^2}} \partial_t A_0 - \frac{d}{\sqrt{1 - d^2}} \partial_1 A_0 \right]
+ \frac{1}{\sqrt{1 - d^2}} \left[ \frac{1}{\sqrt{1 - d^2}} \partial_t A_1 - \frac{d}{\sqrt{1 - d^2}} \partial_1 A_1 \right]
- \frac{1}{\sqrt{1 - d^2}} \left[ -\frac{d}{\sqrt{1 - d^2}} \partial_t A_0 + \frac{1}{\sqrt{1 - d^2}} \partial_1 A_0 \right]
+ \frac{d}{\sqrt{1 - d^2}} \left[ -\frac{d}{\sqrt{1 - d^2}} \partial_t A_1 + \frac{1}{\sqrt{1 - d^2}} \partial_1 A_1 \right]
= \partial_t A_1 - \partial_1 A_0.
$$

Further, we get

$$
\partial_t \tilde{A}_2 - \partial_2 \tilde{A}_0 = \frac{1}{\sqrt{1 - d^2}} \partial_t A_2 - \frac{d}{\sqrt{1 - d^2}} \partial_1 A_2 - \partial_2 \left[ \frac{1}{\sqrt{1 - d^2}} A_0 - \frac{d}{\sqrt{1 - d^2}} A_1 \right]
= \frac{1}{\sqrt{1 - d^2}} [\partial_t A_2 - \partial_2 A_0] + \frac{d}{\sqrt{1 - d^2}} [\partial_2 A_0 - \partial_1 A_2]
$$

and similarly

$$
\partial_1 \tilde{A}_2 - \partial_2 \tilde{A}_1 = \frac{1}{\sqrt{1 - d^2}} \partial_1 A_2 - \frac{d}{\sqrt{1 - d^2}} \partial_2 A_2 - \partial_2 \left[ \frac{1}{\sqrt{1 - d^2}} A_1 - \frac{d}{\sqrt{1 - d^2}} A_0 \right]
= \frac{1}{\sqrt{1 - d^2}} [\partial_1 A_2 - \partial_2 A_1] + \frac{d}{\sqrt{1 - d^2}} [\partial_2 A_1 - \partial_1 A_2].
$$

We can immediately deduce analogous results with $2$ replaced by $j = 3, 4$. Finally, for $i, j = 2, 3, 4$, we find

$$
\partial_i \tilde{A}_j - \partial_j \tilde{A}_i = \partial_i A_j - \partial_j A_i.
$$

Now we investigate how the energy functional changes under this type of Lorentz transform. Of course, the plane $t = 0$ gets transformed into a spacelike hyperplane, and we shall consider this carefully in the proof of Proposition 9.5. We compute

$$
(\partial_t \tilde{A}_2 - \partial_2 \tilde{A}_0)^2 + (\partial_1 \tilde{A}_2 - \partial_2 \tilde{A}_1)^2 = \frac{1 + d^2}{1 - d^2} (\partial_t A_2 - \partial_2 A_0)^2 + \frac{1 + d^2}{1 - d^2} (\partial_1 A_2 - \partial_2 A_1)^2
- \frac{4d}{1 - d^2} (\partial_2 A_0)(\partial_1 A_2 - \partial_2 A_1).
$$

Doing the same for $j = 3, 4$ instead of $j = 2$, we find

$$
\sum_{\alpha, \beta} (\partial_\alpha \tilde{A}_\beta - \partial_\beta \tilde{A}_\alpha)^2 = \sum_{\alpha, \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 - \frac{8d}{1 - d^2} \sum_{j=1}^4 (\partial_i A_j - \partial_j A_i)(\partial_1 A_j - \partial_j A_1) + O(d^2).
$$
The reason for the factor 8 instead of 4 is that the terms \((\partial_t A_2 - \partial_2 A_0)^2\) etc. get counted twice in the above sum. Now the same needs to be done for the scalar potential energy, which is given by

\[
\sum_{\alpha} |\nabla_{\alpha} \phi + i \tilde{A}_{\alpha} \phi|^2 = |\nabla_1 \phi + i \tilde{A}_0 \phi|^2 + |\nabla_1 \phi + i \tilde{A}_1 \phi|^2 + \sum_{j=2}^{4} |\nabla_j \phi + i \tilde{A}_j \phi|^2,
\]

and we have for the first two terms on the right the relation

\[
|\nabla_1 \phi + i \tilde{A}_0 \phi|^2 + |\nabla_1 \phi + i \tilde{A}_1 \phi|^2 = \left(1 - \frac{d}{1 - d^2}\right) |\nabla_1 \phi + i \tilde{A}_1 \phi|^2
\]

\[
\left(1 - \frac{d}{1 - d^2}\right) |\nabla_1 \phi + i \tilde{A}_1 \phi|^2 + \sum_{j=2}^{4} |\nabla_j \phi + i \tilde{A}_j \phi|^2 = \frac{1 + d^2}{1 - d^2} |\nabla_1 \phi + i \tilde{A}_1 \phi|^2 - \frac{4d}{1 - d^2} \text{Re}((\partial_t \phi + i A_0 \phi)(\partial_t \phi + i A_1 \phi)) + O(d^2).
\]

and so we obtain that

\[
\sum_{\alpha} |\nabla_{\alpha} \phi + i \tilde{A}_{\alpha} \phi|^2 = \sum_{\alpha} |\nabla_{\alpha} \phi + i \tilde{A}_{\alpha} \phi|^2 - \frac{4d}{1 - d^2} \text{Re}((\partial_t \phi + i A_0 \phi)(\partial_t \phi + i A_1 \phi)) + O(d^2).
\]

To summarize the preceding computations, we state

\[
\frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha \tilde{A}_\beta - \partial_\beta \tilde{A}_\alpha)^2 + \frac{1}{2} \sum_{\alpha} |\nabla_\alpha \phi + i \tilde{A}_\alpha \phi|^2
\]

\[
= \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 + \frac{1}{2} \sum_{\alpha} |\nabla_\alpha \phi + i A_\alpha \phi|^2
\]

\[
- \frac{2d}{1 - d^2} \sum_{j=1}^{4} (\partial_t A_j - \partial_j A_0)(\partial_t A_j - \partial_j A_1) + \text{Re}((\partial_t \phi + i A_0 \phi)(\partial_t \phi + i A_1 \phi)) + O(d^2),
\]

where the right hand side has to be evaluated at \(L(t, x)\). The preceding computation strongly suggests that if it is impossible to lower the energy by means of such a Lorentz transform for very small \(d\) with a suitable sign, this implies that the following momentum type expressions vanish

\[
\int_{\mathbb{R}^4} \left[ \sum_{j=1}^{4} (\partial_t A_j - \partial_j A_0)(\partial_t A_j - \partial_j A_1) + \text{Re}((\partial_t \phi + i A_0 \phi)(\partial_t \phi + i A_1 \phi)) \right] dx = 0,
\]

where \(k = 1, \ldots, 4\). This will be made rigorous below, in special situations.

8.3. Virial identities. Again suppose that \((A, \phi)\) is an admissible solution to MKG-CG, not necessarily global. The goal of this subsection is to prove Proposition 8.3 below. To derive it, we always split the functional under consideration into an \(A\)-part and a \(\phi\)-part, and check a suitable cancellation between their time derivatives. In the identities below, any terms that are bounded by

\[
\int_{|x| \geq R} \left[ \sum_{\alpha} |D_\alpha \phi|^2 + \sum_{\alpha, \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 \right] dx
\]

for some \(R > 0\) are denoted by “error”.

\[
\sum_{\alpha} |\nabla_{\alpha} \phi + i \tilde{A}_{\alpha} \phi|^2 = \sum_{\alpha} |\nabla_{\alpha} \phi + i \tilde{A}_{\alpha} \phi|^2 - \frac{4d}{1 - d^2} \text{Re}((\partial_t \phi + i A_0 \phi)(\partial_t \phi + i A_1 \phi)) + O(d^2).
\]
Weighted energy. Consider the expression

$$\int_{\mathbb{R}^4} \psi_R(x) \left[ \frac{1}{2} (\partial_i A_j - \partial_j A_i)^2 + \frac{1}{4} \sum_{i,j} (\partial_i A_j - \partial_j A_i)^2 \right] dx,$$

where we define the vector-valued function $\psi_R(x) = x \varphi(\frac{x}{R})$ for $R > 0$ and a smooth cutoff $\varphi \in C_0^\infty(\mathbb{R}^4)$ with $\varphi \equiv 1$ for $|x| \leq 1$ and $\varphi \equiv 0$ for $|x| > 2$, as in [5]. Differentiating with respect to time $t$, we obtain the expression

$$\int_{\mathbb{R}^4} \psi_R(x) \left[ \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) + \frac{1}{2} \sum_{i,j} \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) \right] dx.$$

Then for the second term on the right, we get

$$\frac{1}{2} \sum_{i,j} \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) = \frac{1}{2} \sum_{i,j} \left[ \partial_t (\partial_i A_j - \partial_j A_i) - \partial_j (\partial_i A_i - \partial_i A_j) \right] (\partial_i A_j - \partial_j A_i)$$

$$= \frac{1}{2} \sum_{i,j} \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) - \frac{1}{2} \sum_{i,j} \partial_j (\partial_i A_i - \partial_i A_j)(\partial_i A_j - \partial_j A_i)$$

$$= \sum_{i,j} \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i).$$

Looking at the $k$-th component, $k = 1, \ldots, 4$, it follows by integration by parts that

$$\int_{\mathbb{R}^4} \psi_R^k(x) \frac{1}{2} \sum_{i,j} \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) dx$$

$$= \int_{\mathbb{R}^4} \psi_R^k(x) \sum_{i,j} \partial_t (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) dx$$

$$= -\int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \sum_j (\partial_i A_j - \partial_j A_i)(\partial_k A_j - \partial_j A_k) dx$$

$$- \int_{\mathbb{R}^4} \psi_R^k(x) \sum_{i,j} (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_i) dx$$

$$+ \text{error}.$$
We reformulate the second expression on the right as
\[ \int_{\mathbb{R}^4} \psi_R(x) \sum_{j=1}^{4} \text{Re} \left[ \partial_j (\partial_j \phi + iA_j \phi) (\partial_j \phi + iA_j \phi) \right] dx \]
\[ = \int_{\mathbb{R}^4} \psi_R(x) \sum_{j=1}^{4} \text{Re} \left[ (\partial_j + iA_0)(\partial_j \phi + iA_j \phi) (\partial_j \phi + iA_j \phi) \right] dx \]
\[ = \int_{\mathbb{R}^4} \psi_R(x) \sum_{j=1}^{4} \text{Re} \left[ (\partial_j + iA_j)(\partial_j \phi + iA_0 \phi) (\partial_j \phi + iA_j \phi) \right] dx \]
\[ + \int_{\mathbb{R}^4} \psi_R(x) \sum_{j=1}^{4} \text{Re} \left[ i(\partial_j A_j - \partial_j A_0)(\partial_j \phi + iA_j \phi) \right] dx, \]

where the last term is used to compensate for the term generated by the integration by parts in the \( A \) expression via the equation for \( A \). For the first term on the right, we integrate by parts with respect to \( j \), which generates an additional term when \( \partial_j \) falls on \( \psi_R(x) \). If we look at the \( k \)-th component, we obtain
\[ \int_{\mathbb{R}^4} \psi_R^k(x) \sum_{j=1}^{4} \text{Re} \left[ (\partial_j + iA_j)(\partial_j \phi + iA_0 \phi) (\partial_j \phi + iA_j \phi) \right] dx \]
\[ = -\int_{\mathbb{R}^4} \psi_R^k(x) \sum_{j=1}^{4} \text{Re} \left[ (\partial_j \phi + iA_0 \phi)(\partial_j + iA_j)^2 \phi \right] dx - \int_{\mathbb{R}^4} \varphi(x) R \text{Re} \left[ (\partial_k \phi + iA_0 \phi)(\partial_k + iA_k) \phi \right] dx + \text{error}. \]

Combining with the result for the \( A \)-integral and inserting the Maxwell-Klein-Gordon equations, we obtain for the \( k \)-th component
\[ -\int_{\mathbb{R}^4} \varphi(x) R \text{Re} \left[ (\partial_k \phi + iA_0 \phi)(\partial_k + iA_k) \phi \right] + \sum_{j} (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) dx + \text{error}, \]

which, up to the additional factor \( \varphi(x) \), is exactly the expression appearing in (8.12).

**Weighted momentum.** Next, we consider a weighted generalized momentum operator, given by the expression
\[ \int_{\mathbb{R}^4} \psi_R(x) \cdot \left[ \text{Re} \left[ (\partial_l \phi + iA_0 \phi)(\partial_k + iA_k) \phi \right] + \sum_{j} (\partial_l A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \right] dx, \]

where \( \psi_R(x) \) is the vector valued function with components \( x_k \varphi_R^k(x) \), so the integrand is to be interpreted as an inner product. We shall compute the time derivative of this expression (without the weight \( \psi_R(x) \), this would be a time-independent function). Start with the product involving the magnetic potential. We have
\[ \frac{d}{dt} \int_{\mathbb{R}^4} \psi_R(x) \sum_{j} (\partial_l A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) dx \]
\[ = \int_{\mathbb{R}^4} \psi_R(x) \sum_{j} \partial_l (\partial_l A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) dx + \int_{\mathbb{R}^4} \psi_R(x) \sum_{j} (\partial_l A_j - \partial_j A_0) \partial_l (\partial_k A_j - \partial_j A_k) dx. \]
For the first term on the right, use that

\[ \partial_t (\partial_j A_j - \partial_j A_0) = \partial_t F_{0j} = -\sum_l \partial_l F_{jl} + \text{Im} (\phi D_j \phi). \]

We then conclude that

\[
\int_{\mathbb{R}^4} \psi_R(x) \sum_j \partial_t (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \, dx
= \sum_l \int_{\mathbb{R}^4} \psi_R(x) \sum_j \partial_l (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \, dx
+ \int_{\mathbb{R}^4} \psi_R(x) \sum_j \text{Im} (\phi D_j \phi)(\partial_k A_j - \partial_j A_k) \, dx.
\]

To simplify things further, observe that the original expression

\[
\int_{\mathbb{R}^4} \psi_R(x) \cdot \left[ \text{Re} \left( (\partial_t \phi + iA_0 \phi)(\partial_k + iA_k) \phi \right) \right] + \sum_j (\partial_t A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \, dx
\]

is gauge invariant, so we may in fact suppose that we are already in the Coulomb gauge. Then we have

\[
\sum_{l,j} \int_{\mathbb{R}^4} \psi_R(x) \partial_l (\partial_t A_j - \partial_j A_i)(\partial_k A_j - \partial_j A_k) \, dx
= \int_{\mathbb{R}^4} \psi_R(x) \left( \sum_j \Delta A_j (\partial_k A_j - \partial_j A_k) \right) \, dx
= \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) \Delta A_k \cdot A_k \, dx
- \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) \sum_j (\partial_k A_j)^2 \, dx
- \sum_{l,j} \int_{\mathbb{R}^4} \psi_R(x) \partial_l A_j \partial_k (\partial_l A_j) \, dx + \text{error}.
\]

This can be further simplified to

\[
- \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) \sum_l (\partial_l A_k)^2 \, dx
- \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) \sum_j (\partial_k A_j)^2 \, dx
+ \frac{1}{2} \sum_{l,j} \varphi \left( \frac{x}{R} \right) (\partial_l A_j)^2 \, dx + \text{error}.
\]

Observe that if one sums here over \( k = 1, 2, 3, 4 \), then the result is simply error, due to the dimension \( n = 4! \). In short, we have shown that

\[
\int_{\mathbb{R}^4} \psi_R(x) \sum_j \partial_t (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \, dx
= \int_{\mathbb{R}^4} \psi_R(x) \sum_j \text{Im} (\phi D_j \phi)(\partial_k A_j - \partial_j A_k) \, dx
+ \text{error}.
\]
Next, we compute

\[
\int_{\mathbb{R}^4} \psi_R(x) \sum_j (\partial_t A_j - \partial_j A_0) \partial_t (\partial_k A_j - \partial_j A_k) \, dx
\]

\[
= \int_{\mathbb{R}^4} \psi_R(x) \sum_j (\partial_t A_j - \partial_j A_0) (\partial_k (\partial_t A_j - \partial_j A_0) - \partial_j (\partial_t A_k - \partial_k A_0)) \, dx
\]

\[
= -\frac{1}{2} \sum_j \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) (\partial_t A_j - \partial_j A_0)^2 \, dx - \int_{\mathbb{R}^4} \psi_R(x) \sum_j (\partial_t A_j - \partial_j A_0) \partial_j (\partial_t A_k - \partial_k A_0) \, dx
\]

\[
+ \text{error}
\]

\[
= -\frac{1}{2} \sum_j \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) (\partial_t A_j - \partial_j A_0)^2 \, dx + \int_{\mathbb{R}^4} \psi_R(x) \sum_j (\partial_t A_k - \partial_k A_0)^2 \, dx
\]

\[
- \int_{\mathbb{R}^4} \psi_R(x) \Delta A_0 (\partial_t A_k - \partial_k A_0) \, dx + \text{error},
\]

where we have again invoked the Coulomb condition, as we may. Then from the equation for \( A_0 \) we can equate the preceding with

\[
\int_{\mathbb{R}^4} \psi_R(x) \sum_j (\partial_t A_j - \partial_j A_0) \partial_t (\partial_k A_j - \partial_j A_k) \, dx
\]

\[
= -\sum_k \frac{1}{2} \sum_j \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) (\partial_t A_j - \partial_j A_0)^2 \, dx + \sum_k \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) (\partial_t A_k - \partial_k A_0)^2 \, dx
\]

\[
+ \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Im}(\phi D_t \bar{\phi})(\partial_t A_k - \partial_k A_0) \, dx + \text{error}
\]

\[
= -\sum_k \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) (\partial_t A_k - \partial_k A_0)^2 \, dx + \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Im}(\phi D_t \bar{\phi})(\partial_t A_k - \partial_k A_0) \, dx + \text{error}.
\]

Now we repeat the same exercise for the contribution to the momentum from \( \phi \), given by

\[
\int_{\mathbb{R}^4} \psi_R(x) \cdot \left[ \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_k + iA_k) \phi \right] \right].
\]
We get

\[
\frac{d}{dt} \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
= \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
+ \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
= \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
+ \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
= \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ \sum_l (\partial_l + iA_l)^2 \phi(\partial_k + iA_k \phi) \right] \, dx
\]

\[
+ \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx.
\]

We treat the last two expressions separately. For the last one, we find

\[
\int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
= \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t A_k + \partial_k A_0) \right] \, dx
\]

\[
+ \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\partial_t + iA_k \phi) \right] \, dx
\]

\[
= - \sum_k \frac{1}{2} \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) |\partial_k \phi + iA_k \phi|^2 \, dx + \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Im} (D_t \phi \overline{\phi})(\partial_t A_k - \partial_k A_0) \, dx + \text{error}.
\]

Carefully note how the last expression is exactly negative the analogous one coming from the calculation for the $A$-terms. Next, for the term involving the $l$ spatial derivatives, we get

\[
\int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ \sum_l (\partial_l + iA_l)^2 \phi(\partial_k + iA_k \phi) \right] \, dx
\]

\[
= - \sum_k \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) |\partial_k \phi + iA_k \phi|^2 \, dx
\]

\[
- \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ \sum_l (\partial_l + iA_l)\phi(\partial_l + iA_l)(\partial_k + iA_k \phi) \right] \, dx + \text{error}
\]
and for the last term, we can also write

\[- \int_{\mathbb{R}^n} \psi_R(x) \cdot \text{Re} \left[ \sum_l (\partial_l + iA_l)\phi\overline{(\partial_l + iA_l)(\partial_l + iA_k)\phi}) \right] \, dx \]

\[= - \int_{\mathbb{R}^n} \psi_R(x) \cdot \text{Re} \left[ \sum_l (\partial_l + iA_l)\phi\overline{(\partial_l + iA_k)(\partial_l + iA_l)\phi}) \right] \, dx \]

\[\quad + \int_{\mathbb{R}^n} \psi_R(x) \cdot \text{Re} \left[ \sum_l (\partial_l + iA_l)\phi(\partial_l A_k - \partial_k A_l)\phi \right] \, dx \]

\[= \sum_l \frac{1}{2} \int_{\mathbb{R}^n} \psi_R(x) \phi_l^2 \, dx - \int_{\mathbb{R}^n} \psi_R(x) \cdot \text{Im} (D_l\phi\overline{\phi})(\partial_l A_k - \partial_k A_l) \, dx + \text{error.} \]

Note that the last term is again exactly the negative of the corresponding term arising from the $A$-expression, as it has to be.

Finally, we summarize the preceding computations in the following identity

\[\frac{d}{dt} \int_{\mathbb{R}^4} \psi_R(x) \cdot \text{Re} \left[ (\partial_l \phi + iA_l \phi)(\overline{\partial_l + iA_k \phi}) \right] + \sum_j (\partial_j A_j - \partial_j A_0)(\partial_j A_j - \partial_j A_0) \, dx \]

\[= \sum_k \int_{\mathbb{R}^4} \psi_R(x) |\partial_k \phi + iA_k \phi|^2 \, dx - 2 \int_{\mathbb{R}^4} \psi_R(x) |\partial_t \phi + iA_0 \phi|^2 \, dx - \sum_k \int_{\mathbb{R}^4} \psi_R(x)(\partial_t A_k - \partial_k A_0)^2 \, dx \]

\[+ \text{error.} \]

This is of the right sign, except for the positive term

\[\sum_k \int_{\mathbb{R}^4} \psi_R(x) |\partial_k \phi + iA_k \phi|^2 \, dx. \]

To deal with it, we need to introduce a second functional

\[\int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ \phi(\partial_t \phi + iA_0 \phi) \right] \, dx. \]

Then observe that we have

\[\frac{d}{dt} \int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ \phi(\partial_t \phi + iA_0 \phi) \right] \, dx \]

\[= \int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ \partial_t \phi(\overline{\partial_t \phi + iA_0 \phi}) \right] \, dx + \int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ \phi(\overline{\partial_t \phi + iA_0 \phi}) \right] \, dx \]

\[= \int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ (\partial_t \phi + iA_0 \phi)(\overline{\partial_t \phi + iA_0 \phi}) \right] \, dx + \int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ \phi(\overline{\partial_t + iA_0})(\partial_t \phi + iA_0 \phi) \right] \, dx \]

\[= \int_{\mathbb{R}^4} \psi_R(x) |\partial_t \phi + iA_0 \phi|^2 \, dx + \sum_l \int_{\mathbb{R}^4} \psi_R(x) \text{Re} \left[ \phi(\partial_t + iA_l)^2 \phi \right] \, dx \]

\[= \int_{\mathbb{R}^4} \psi_R(x) |\partial_t \phi + iA_0 \phi|^2 - \sum_l |\partial_l \phi + iA_l|^2 \, dx + \text{error.} \]
We conclude that if we introduce the functional
\[
M_R(t) := \int_{\mathbb{R}^4} \psi_R(x) \cdot \left[ \Re \left( (\partial_t \phi + iA_0 \phi)(\partial_k + iA_k)\phi \right) + \sum_j (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \right] dx + \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \Re \left( \phi(\partial_t \phi + iA_0 \phi) \right) dx,
\]
then we have the relation
\[
\frac{d}{dt} M_R(t) = -\int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \left| \partial_t \phi + iA_0 \phi \right|^2 dx - \sum_k \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) (\partial_t A_k - \partial_k A_0)^2 dx + \text{error}
\]
Thus, \( M_R(t) \) is the desired Lyapunov functional.

Finally, we make the following simple remark, which follows directly from the preceding proof by replacing \( \psi(\frac{x}{R}) \) by 1.

**Remark 8.2.** The momentum is independent of time. For \( k = 1, \ldots, 4 \), we have
\[(8.13) \quad \frac{d}{dt} \int_{\mathbb{R}^4} \left[ \Re \left( (\partial_t \phi + iA_0 \phi)(\partial_k + iA_k)\phi \right) + \sum_j (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \right] dx = 0.\]

To summarize the preceding computations, we formulate

**Proposition 8.3.** For \( R > 0 \), set
\[
r(R) := \int_{|x| \geq R} \left[ \sum_\alpha |D_\alpha \phi|^2 + \sum_{\alpha, \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 \right] dx.
\]
For a smooth cutoff \( \varphi \in C_0^\infty(\mathbb{R}^4) \) with \( \varphi(x) \equiv 1 \) for \( |x| \leq 1 \) and \( \varphi(x) \equiv 0 \) for \( |x| > 2 \), we define \( \psi_R(x) := x\varphi(\frac{x}{R}) \) for \( R > 0 \). Then we have the following relations for an admissible solution \((A, \phi)\) to MKG-CG.

- **Energy conservation:**
  \[
  (8.14) \quad \frac{d}{dt} \int_{\mathbb{R}^4} \left[ \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 + \frac{1}{2} \sum_\alpha |D_\alpha \phi|^2 \right] dx = 0.
  \]

- **Momentum conservation:** For \( k = 1, \ldots, 4 \), it holds that
  \[
  (8.15) \quad \frac{d}{dt} \int_{\mathbb{R}^4} \left[ \Re \left( (\partial_t \phi + iA_0 \phi)(\partial_k + iA_k)\phi \right) + \sum_j (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \right] dx = 0.
  \]

- **Weighted energy:** For \( k = 1, \ldots, 4 \), we have
  \[
  (8.16) \quad \frac{d}{dt} N_{R,k}(t) = -\int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \left[ \Re \left( (\partial_t \phi + iA_0 \phi)(\partial_k + iA_k)\phi \right) + \sum_j (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \right] dx + O(r(R)),
  \]
  where
  \[
  N_{R,k}(t) := \int_{\mathbb{R}^4} \psi_R^k(x) \left[ \frac{1}{2} \sum_j (\partial_j A_j - \partial_j A_0)^2 + \frac{1}{4} \sum_{i,j} (\partial_i A_j - \partial_j A_i)^2 \right] dx + \int_{\mathbb{R}^4} \psi_R^k(x) \left[ |\partial_t \phi + iA_0 \phi|^2 + \sum_j |\partial_j \phi + iA_j \phi|^2 \right] dx.
  \]
• Weighted momentum monotonicity:

$$
\frac{d}{dt} M_R(t) = - \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) |\partial_t \phi + iA_0 \phi|^2 \, dx - \sum_k \int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) (\partial_k A_k - \partial_k A_0)^2 \, dx + O(R),
$$

where

$$
M_R(t) := \sum_k \int_{\mathbb{R}^4} \psi_R^k(x) \left[ \text{Re} \left( (\partial_t \phi + iA_0 \phi)(\partial_k \phi + iA_k \phi) \right) \right] + \sum_j (\partial_j A_j - \partial_j A_0)(\partial_k A_j - \partial_j A_k) \, dx
$$

and

$$
\int_{\mathbb{R}^4} \varphi \left( \frac{x}{R} \right) \text{Re} \left( \phi(\partial_t \phi + iA_0 \phi) \right) \, dx.
$$

To conclude this section, we also mention the following local energy conservation law and an important consequence.

**Lemma 8.4.** Let \((A, \phi)\) be an admissible solution to MKG-CG on \(\mathbb{R}^{1+4}\) or a suitable space-time slab. Then we have

$$
\frac{1}{2} \sum_j (\partial_t A_j - \partial_j A_0)^2 + \frac{1}{4} \sum_{i,j} (\partial_i A_j - \partial_j A_i)^2 \geq \sum_{i,j} (\partial_i A_j - \partial_j A_i)(\partial_i A_j - \partial_j A_0) + \frac{1}{2} \sum_j |\partial_t \phi + iA_0 \phi|^2 - \sum_j \text{Re} \left( (\partial_t \phi + iA_0 \phi)(\partial_j \phi + iA_j \phi) \right) = 0.
$$

(8.18)

This entails the following consequence. Let \(K\) be a solid forward light cone, and \(t_0 < t_1\). Then denoting

$$
E_K(t) := \int_{(t) \times \mathbb{R}^4 \setminus K} \left[ \sum_{\alpha \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 + \sum_\alpha |\partial_\alpha \phi + iA_\alpha \phi|^2 \right] \, dx,
$$

we have

$$
E_K(t_1) \leq E_K(t_0).
$$

In fact, these two quantities differ by a suitable energy flux across the part of the mantle of the light cone enclosed by the time slices \(t_0 \times \mathbb{R}^4\) and \([t_1] \times \mathbb{R}^4\).

**Proof.** The identity (8.18) follows by direct computation. To get the inequality, we use the divergence theorem. Assuming as we may that \(K\) is centered around \(x = 0\) and denoting the outward unit normal vector to \(S^\alpha\) by \(\sigma = \frac{\delta}{|\delta|}\), we obtain

$$
E(t_0) = E(t_1) + \int_{\mathcal{C}_{t_0}^{t_1}} [\mathcal{A}_\omega + \Phi_\omega] \, d\sigma,
$$

where \(\mathcal{C}_{t_0}^{t_1}\) denotes the portion of the mantle of \(K\) between times \(t_0, t_1\), and \(d\sigma\) denotes the standard surface measure. Further, we use the notation

$$
\sqrt{2} \mathcal{A}_\omega = \frac{1}{2} \sum_j (\partial_t A_j - \partial_j A_0)^2 + \frac{1}{4} \sum_{i,j} (\partial_i A_j - \partial_j A_i)^2 + \sum_{i,j} \omega_j (\partial_i A_j - \partial_j A_i)(\partial_t A_j - \partial_j A_0),
$$

$$
\sqrt{2} \Phi_\omega = \frac{1}{2} \sum_j |\partial_t \phi + iA_0 \phi|^2 + \sum_j |\partial_j \phi + iA_j \phi|^2 + \sum_j \omega_j \text{Re} \left( (\partial_t \phi + iA_0 \phi)(\partial_j \phi + iA_j \phi) \right).
$$

Then it suffices to show that \(\mathcal{A}_\omega \geq 0, \Phi_\omega \geq 0\). Observe that

$$
\sum_{i,j} \omega_j (\partial_i A_j - \partial_j A_i)(\partial_t A_j - \partial_j A_0) = \frac{1}{2} \sum_{i,j} (\partial_i A_j - \partial_j A_0)(\omega_j (\partial_i A_j - \partial_j A_0) - \omega_j (\partial_i A_i - \partial_i A_0)).
$$
Also, for any \( r \in \mathbb{R}^4 \), we have the general inequality
\[
\sum_{i,j} (\omega_i r_j - \omega_j r_i)^2 = \sum_{i,j} (\omega_i r_j^\perp - \omega_j r_i^\perp)^2 = 2|r^\perp|^2 \leq 2|r|^2,
\]
where \( r^\perp \) denotes projection orthogonal to \( \omega \), and so we have
\[
\left| \sum_{i,j} \omega_i (\partial_i A_j - \partial_j A_i) (\partial_i A_j - \partial_j A_0) \right| \\
\leq \frac{1}{4} \sum_{i,j} (\partial_i A_j - \partial_j A_i)^2 + \frac{1}{4} \sum_{i,j} (\omega_i (\partial_i A_j - \partial_j A_0) - \omega_j (\partial_i A_i - \partial_j A_0))^2 \\
\leq \frac{1}{4} \sum_{i,j} (\partial_i A_j - \partial_j A_i)^2 + \frac{1}{2} \sum_j (\partial_t A_j - \partial_j A_0)^2,
\]
and so \( \mathcal{A}_\omega \geq 0 \). We also have
\[
\left| \sum_j \omega_j \text{Re} \left( (\partial_t \phi + i A_0 \phi) (\partial_t \phi + i A_1 \phi) \right) \right| \leq \frac{1}{2} |\partial_t \phi + i A_0 \phi|^2 + \frac{1}{2} |\sum_j \omega_j (\partial_t \phi + i A_j \phi)|^2 \\
\leq \frac{1}{2} |\partial_t \phi + i A_0 \phi|^2 + |\sum_j |\partial_t \phi + i A_j \phi|^2|,
\]
which entails \( \Phi_\omega \geq 0 \).

8.4. **A virial identity in self-similar coordinates.** Here we introduce the tool that will allow us later to rule out self-similar solutions. This virial identity is of a somewhat delicate character, since it is of singular nature. To start the computations below, we make the following assumptions: Let \( (\mathcal{A}^\infty, \Phi^\infty) \) be an energy class solution to the MKG-CG system in the sense of Section 5 with lifespan \( I = [0, 1) \). Assume that restricting to the slice \( [0, t_0] \times \mathbb{R}^4 \) for any \( t_0 \in [0, 1) \), \( \nabla_{t,A} \Phi^\infty \) and all of the curvature components are supported in a compact subset of the solid cone \( |x| < 1 - t \). In order to justify the calculations below we shall have to assume smoothness of the components; of course, smoothing the components (as in Section 5) will destroy the support properties of the curvature components. In that sense, certain of the expressions below which involve weights of the form \((1 - |y|^2)^{-\frac{1}{2}}, (1 - |y|^2)^{-\frac{3}{4}}, \) become singular at \( |y| = 1 \). Here, \( y = \frac{x}{|x|} \) is the self-similar coordinate introduced below. To deal with this, one needs to introduce an additional smooth cutoff \( \chi_\varepsilon(y) := \chi \left( \frac{1 - |y|}{\varepsilon} \right) \) that smoothly localizes away from the boundary, but such that \( \lim_{\varepsilon \to 0} \chi_\varepsilon(y) = \chi_{[0,1]}(|y|) \), the sharp characteristic cutoff. Thus, for the calculations below to be rigorous, one really needs to consider the weight
\[
\chi_\varepsilon(y)(1 - |y|^2)^{-\frac{1}{2}},
\]
which will lead to additional error terms localized near the boundary. But then letting the frequency cutoff converge toward \( |\xi| = +\infty \) in the regularization (for fixed but sufficiently small \( \varepsilon > 0 \)), it will be easy to convince oneself that the additional errors can be killed in the limit (due to the support properties of the underlying \((\mathcal{A}^\infty, \Phi^\infty)\)). In that sense, to simplify the calculations, we shall formally omit this additional cutoff in the computations below.

Further, in order to simplify the computations below, we shall assume that we have transformed things into the **Cronstrom gauge**, i.e. we impose the condition
\[
\sum_{k=1}^{4} \chi_k A_k = 0
\]
or also \( \sum_{k=1}^{4} y_k A_k = 0 \) in terms of the self-similar coordinates employed below. This again leads to a technical complication, in that the \( C^\infty \) smoothness of the regularized \((A^\infty, \Phi^\infty)\) will be lost. This can again be dealt with via smooth truncation of the functional, this time away from the origin by including

\[
\chi\left(\frac{|y|}{\epsilon}\right).
\]

Since all the integrations by parts to be performed below involve an operator \( y \cdot \nabla_y \), the error terms are seen to be controllable in terms of the energy on smaller and smaller balls, and hence negligible in the limit as \( \epsilon \to 0 \). Again, we shall gloss over this technicality in the formulas below. Of course, there is also a gauge invariant version of the main identity (8.24), as we shall explain after proving it.

Introduce the standard change of variables

\[
s = -\log(1 - t), \quad y = \frac{x}{1 - t},
\]

and then consider the new functions

\[
\tilde{\phi}(s, y) := e^{-s}\phi(1 - e^{-s}, ye^{-s}), \quad \tilde{A}_\nu = e^{-s}A_\nu(1 - e^{-s}, ye^{-s}).
\]

We first analyze the \( \tilde{\phi} \)-equation

\[
(\partial_s + y \cdot \nabla_y + i\tilde{A}_0)(\partial_s + y \cdot \nabla_y + i\tilde{A}_0 + 1)\tilde{\phi} = \sum_{k=1}^{4} (\partial_k + i\tilde{A}_k)^2 \tilde{\phi},
\]

where \( \partial_k := \frac{\partial}{\partial y_k} \). We develop this into

\[
(\partial_s + i\tilde{A}_0)^2 \tilde{\phi} + (\partial_s + i\tilde{A}_0)(y \cdot \nabla_y + 1)\tilde{\phi} + (y \cdot \nabla_y + 2)(\partial_s + i\tilde{A}_0)\tilde{\phi} + (y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1)\tilde{\phi}
\]

\[
= \sum_{k=1}^{4} (\partial_k + i\tilde{A}_k)^2 \tilde{\phi}.
\]

Alternatively, we can also write this as

\[
(\partial_s + i\tilde{A}_0)^2 \tilde{\phi} + (3 + 2y \cdot \nabla_y)(\partial_s + i\tilde{A}_0)\tilde{\phi} - i(y \cdot \nabla\tilde{A}_0)\tilde{\phi} + (y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1)\tilde{\phi}
\]

\[
= \sum_{k=1}^{4} (\partial_k + i\tilde{A}_k)^2 \tilde{\phi}.
\]

Now we start by analyzing the following energy functional

\[
\frac{d}{ds} \frac{1}{2} \int_{\mathbb{R}^4} \left( |(\partial_s + i\tilde{A}_0)\tilde{\phi}|^2 + \sum_{k=1}^{4} |(\partial_k + i\tilde{A}_k)\tilde{\phi}|^2 \right) \rho(y) dy,
\]
where $\rho(y) = (1 - |y|^2)^{-\frac{1}{2}}$. We obtain

\[
\frac{d}{ds} \frac{1}{2} \int_{\mathbb{R}^4} \left[ |(\partial_s + i\tilde{A}_0)\tilde{\phi}|^2 + \sum_{k=1}^4 |(\partial_k + i\tilde{A}_k)\tilde{\phi}|^2 \right] \rho(y) dy
= \int \Re \left[ \partial_s (\partial_s + i\tilde{A}_0)\tilde{\phi} (\partial_s + i\tilde{A}_0)\tilde{\phi} \right] \rho(y) dy
+ \sum_{k=1}^4 \int \Re \left[ (\partial_k + i\tilde{A}_k)\tilde{\phi} (\partial_k + i\tilde{A}_k)\tilde{\phi} \right] \rho(y) dy
= \int \Re \left[ (\partial_s + i\tilde{A}_0)^2 \tilde{\phi} (\partial_s + i\tilde{A}_0)\tilde{\phi} \right] \rho(y) dy
+ \sum_{k=1}^4 \int \Re \left[ (\partial_k + i\tilde{A}_k)\tilde{\phi} (\partial_k + i\tilde{A}_k)\tilde{\phi} \right] \rho(y) dy.
\]

For the last term but one, we insert the equation for $\tilde{\phi}$, replacing it by

\[
\int \Re \left[ (\partial_s + i\tilde{A}_0)^2 \tilde{\phi} (\partial_s + i\tilde{A}_0)\tilde{\phi} \right] \rho(y) dy
= \int \Re \left[ \sum_{k=1}^4 (\partial_k + i\tilde{A}_k)^2 \tilde{\phi} (\partial_k + i\tilde{A}_k)\tilde{\phi} \right] \rho(y) dy
+ \int \Re \left[ i(y \cdot \nabla \tilde{A}_0)\tilde{\phi} (\partial_s + i\tilde{A}_0)\tilde{\phi} \right] \rho(y) dy
- \int \Re \left[ (3 + 2y \cdot \nabla y)(\partial_s + i\tilde{A}_0)\tilde{\phi} (\partial_s + i\tilde{A}_0)\tilde{\phi} \right] \rho(y) dy
- \int \Re \left[ (y \cdot \nabla y + 2)(\partial_s + i\tilde{A}_0)\tilde{\phi} (\partial_s + i\tilde{A}_0)\tilde{\phi} \right] \rho(y) dy.
\]
For the first term on the right, one performs integration by parts, re-writing it as

\[
\int \text{Re} \left[ \sum_{k=1}^{4} (\partial_{k} + i\tilde{A}_{k})^{2} \phi(\partial_{s} + i\tilde{A}_{0})\phi \right] \rho(y) \, dy
\]

\[
= - \int \text{Re} \left[ \sum_{k=1}^{4} (\partial_{k} + i\tilde{A}_{k})\phi(\partial_{s} + i\tilde{A}_{0})(\partial_{s} + i\tilde{A}_{0})\phi \right] \rho(y) \, dy
\]

\[
- \int \text{Re} \left[ \sum_{k=1}^{4} (\partial_{k} + i\tilde{A}_{k})\phi(\partial_{s} + i\tilde{A}_{0})\phi \right] \partial_{k}\rho(y) \, dy
\]

\[
= - \int \text{Re} \left[ \sum_{k=1}^{4} (\partial_{k} + i\tilde{A}_{k})\phi(\partial_{s} + i\tilde{A}_{0})(\partial_{k} + i\tilde{A}_{k})\phi \right] \rho(y) \, dy
\]

\[
+ \int \text{Re} \left[ \sum_{k=1}^{4} (\partial_{k} + i\tilde{A}_{k})\phi(\partial_{s} + i\tilde{A}_{k} - \partial_{s}\tilde{A}_{0})\phi \right] \rho(y) \, dy
\]

\[
- \int \text{Re} \left[ \sum_{k=1}^{4} (\partial_{k} + i\tilde{A}_{k})\phi(\partial_{s} + i\tilde{A}_{0})\phi \right] \partial_{k}\rho(y) \, dy
\]

\[
= - \frac{d}{ds} \frac{1}{2} \int \left| (\partial_{k} + i\tilde{A}_{k})\phi \right|^2 \rho(y) \, dy - \int \text{Im} \left( \bar{\phi} D_{k}\phi \right) (\partial_{s}\tilde{A}_{k} - \partial_{s}\tilde{A}_{0}) \rho(y) \, dy
\]

\[
- \int \text{Re} \left[ \sum_{k=1}^{4} \partial_{k}\phi(\partial_{s} + i\tilde{A}_{0})\phi \right] \partial_{k}\rho(y) \, dy,
\]

where we use the notation \( \tilde{D}_{k} := \partial_{k} + i\tilde{A}_{k} \) and in the last step we have taken advantage of the Cronstrom condition \( \sum_{k} y_{k}\tilde{A}_{k} = 0 \). Here we expect the last term but one to cancel against a corresponding term from the \( A \)-energy. On the other hand, the last term is expected to cancel against other terms from the \( \phi \)-equation. Specifically, we have

\[
- \int \text{Re} \left[ (y \cdot \nabla_{y} + 2)(y \cdot \nabla_{y} + 1)\phi(\partial_{s} + i\tilde{A}_{0})\phi \right] \rho(y) \, dy
\]

\[
= \int \text{Re} \left[ (y \cdot \nabla_{y} + 1)\phi(\partial_{s} + i\tilde{A}_{0})\phi \right] \rho(y) \, dy
\]

\[
+ \int \text{Re} \left[ (y \cdot \nabla_{y} + 1)\phi(\partial_{s} + i\tilde{A}_{0})(y \cdot \nabla_{y} + 1)\phi \right] \rho(y) \, dy
\]

\[
+ \int \text{Re} \left[ (y \cdot \nabla_{y})\phi(y \cdot \nabla_{y}\tilde{A}_{0})\phi \right] \rho(y) \, dy
\]

\[
+ \int \text{Re} \left[ (y \cdot \nabla_{y})\tilde{\phi}(\partial_{s} + i\tilde{A}_{0})\phi \right] (y \cdot \nabla_{y})\rho(y) \, dy,
\]
which we write more succinctly as
\[
\int \operatorname{Re} [(y \cdot \nabla_y + 1) \bar{\phi}(\partial_s + i\bar{A}_0)(y \cdot \nabla_y + 1) \bar{\phi}] \rho(y) \, dy \\
+ \int \operatorname{Re} [(y \cdot \nabla_y + 1) \bar{i} i(y \cdot \nabla_y \bar{A}_0) \bar{\phi}] \rho(y) \, dy \\
+ \int \operatorname{Re} [(y \cdot \nabla_y) \bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] (y \cdot \nabla_y + 1) \rho(y) \, dy \\
+ \int \operatorname{Re} [\bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] \rho(y) \, dy.
\]

Here the last term is an exact \( s \)-derivative,
\[
\int \operatorname{Re} [\bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] \rho(y) \, dy = \frac{d}{ds} \int |\bar{\phi}|^2 \rho(y) \, dy,
\]
which we will incorporate into the energy functional. Moreover, the last term but one cancels against a term obtained further above: Since
\[
(y \cdot \nabla_y + 1) \rho(y) = (1 - |y|^2)^{-\frac{3}{2}},
\]
we get
\[
\int \operatorname{Re} [(y \cdot \nabla_y) \bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] (y \cdot \nabla_y + 1) \rho(y) \, dy = \int \operatorname{Re} \left[ \sum_{k=1}^{4} \partial_k \bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi} \right] \partial_k \rho(y) \, dy
\]
and so these two terms cancel out, as announced earlier. The first term above is again an exact \( s \)-derivative:
\[
\int \operatorname{Re} [(y \cdot \nabla_y + 1) \bar{\phi}(\partial_s + i\bar{A}_0)(y \cdot \nabla_y + 1) \bar{\phi}] \rho(y) \, dy = \frac{d}{ds} \int |(y \cdot \nabla_y + 1) \bar{\phi}|^2 \rho(y) \, dy.
\]
Finally, we need to account for the term
\[
\int \operatorname{Re} [i(y \cdot \nabla \bar{A}_0) \bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] \rho(y) \, dy
\]
generated by the integration by parts above. In fact, we shall combine this term with the expression
\[
\int \operatorname{Re} [i(y \cdot \nabla \bar{A}_0) \bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] \rho(y) \, dy
\]
which arose when inserting the \( \dot{\phi} \) equation. Observe that we get
\[
\int \operatorname{Re} [(y \cdot \nabla_y + 1) \bar{\phi}(y \cdot \nabla_y \bar{A}_0) \bar{\phi}] \rho(y) \, dy + \int \operatorname{Re} [\bar{i} i(y \cdot \nabla \bar{A}_0) \bar{\phi}(\partial_s + i\bar{A}_0) \bar{\phi}] \rho(y) \, dy
\]
\[
= \int \operatorname{Re} [\bar{i} i(y \cdot \nabla \bar{A}_0) \bar{\phi}(\partial_s + i\bar{A}_0 + y \cdot \nabla_y + 1) \bar{\phi}] \rho(y) \, dy
\]
\[
= - \int (y \cdot \nabla \bar{A}_0) \operatorname{Im} (\bar{\phi} D_i \phi) \rho(y) \, dy.
\]
Here we link with the \( A \)-equation, i.e. we observe that
\[
- \operatorname{Im} (\bar{\phi} D_i \phi) = - \sum_{k=1}^{4} \partial_k [\partial_i A_k - \partial_k A_0].
\]
We can now summarize the preceding computations by the following identity: We have

\[
\frac{d}{ds} \frac{1}{2} \int \left( |(\partial_s + i\tilde{A}_0)\tilde{\phi}|^2 + \sum_{k=1}^{4} |(\partial_k + i\tilde{A}_k)\tilde{\phi}|^2 - |(y \cdot \nabla_y + 1)\tilde{\phi}|^2 - |\tilde{\phi}|^2 \right) \rho(y) \, dy
\]

\[\tag{8.20}\]

\[= - \int \text{Im}(\overline{\phi D_k \phi})(\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) \rho(y) \, dy
\]

\[\quad - \int \text{Re} \left[ (3 + 2y \cdot \nabla_y)(\partial_s + i\tilde{A}_0)\overline{\phi}(\partial_s + i\tilde{A}_0)\phi \right] \rho(y) \, dy
\]

\[\quad - \int (y \cdot \nabla \tilde{A}_0) \sum_{k=1}^{4} \partial_k[\partial_j \tilde{A}_k - \partial_k \tilde{A}_j] \rho(y) \, dy,
\]

where we use the notation

\[\tilde{A}(s, y) = e^{-2s}A(1 - e^{-s}, e^{-s}y)\]

Here we expect the first and the last term on the right to cancel against corresponding terms generated by differentiating a suitable A-energy, while the middle term furnishes the key monotonicity: We get

\[- \int \text{Re} \left[ (3 + 2y \cdot \nabla_y)(\partial_s + i\tilde{A}_0)\overline{\phi}(\partial_s + i\tilde{A}_0)\phi \right] \rho(y) \, dy = \int |(\partial_s + i\tilde{A}_0)\tilde{\phi}|^2 (1 + y \cdot \nabla_y) \rho(y) \, dy,
\]

where the weight \((1 + y \cdot \nabla_y)\rho(y)\) is positive.

At this point we have to pass to the corresponding \(\tilde{A}\)-equation. It is given by

\[
(\partial_s[\partial_s \tilde{A}_j - \partial_j \tilde{A}_0] + (3 + 2y \cdot \nabla_y)\partial_s \tilde{A}_j + (2 + y \cdot \nabla_y)(1 + y \cdot \nabla_y)\tilde{A}_j \\
- (2 + y \cdot \nabla_y)\partial_j \tilde{A}_0 - \sum_{k=1}^{4} \partial_k[\partial_k \tilde{A}_j - \partial_j \tilde{A}_k] = \text{Im}(\phi D_j \phi).
\]

\[\tag{8.21}\]

Now we start with a tentative ansatz for the correct \(\tilde{A}\)-energy (to leading order), which we differentiate with respect to \(s\):

\[\frac{d}{ds} \int \left[ \frac{1}{2} \sum_{j=1}^{4} |(\partial_j \tilde{A}_j - \partial_j \tilde{A}_0)|^2 + \frac{1}{4} \sum_{k,j} |(\partial_k \tilde{A}_k - \partial_k \tilde{A}_j)|^2 \right] \rho(y) \, dy
\]

\[= \int \sum_{j} \partial_s(\partial_s \tilde{A}_j - \partial_j \tilde{A}_0)(\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy + \int \frac{1}{2} \sum_{j,k} \partial_s(\partial_j \tilde{A}_k - \partial_k \tilde{A}_j)(\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) \rho(y) \, dy.
\]
Then substitute the $\tilde{A}$-equation for the first term on the right: We obtain

$$\int \sum_j \partial_s (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy$$

$$= -\sum_j \int (3 + 2y \cdot \nabla_y) \partial_s \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy$$

$$- \sum_j \int (2 + y \cdot \nabla_y) (1 + y \cdot \nabla_y) \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy$$

$$+ \sum_j \int (2 + y \cdot \nabla_y) \partial_j \tilde{A}_0 (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy$$

$$+ \int \sum_{j,k} \partial_k [\partial_k \tilde{A}_j - \partial_j \tilde{A}_k] (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy$$

$$+ \sum_j \int \text{Im} (\phi D_j \tilde{\phi}) (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy.$$ 

Here we immediately see that the last term cancels against the term

$$- \int \sum_k \text{Im} (\phi D_k \tilde{\phi}) (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) \rho(y) \, dy$$

in (8.20). We also expect the last term but one to essentially cancel against the term

$$\int \frac{1}{2} \sum_{j,k} \partial_s (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) \rho(y) \, dy.$$ 

In fact, we compute

$$\int \frac{1}{2} \sum_{j,k} \partial_s (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) \rho(y) \, dy$$

$$= \int \frac{1}{2} \sum_{j,k} (\partial_j (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) - \partial_k (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0)) (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) \rho(y) \, dy$$

$$= -\int \sum_{j,k} \partial_j (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) \rho(y) \, dy - \int \sum_{j,k} (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) \partial_j \rho(y) \, dy$$

and we have

$$\int \sum_{j,k} \partial_j (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) \rho(y) \, dy = \int \sum_{j,k} \partial_k [\partial_k \tilde{A}_j - \partial_j \tilde{A}_k] (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy.$$ 

So these two terms cancel and we are left with

$$- \int \sum_{j,k} (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) \partial_j \rho(y) \, dy$$

(8.22)

$$= -\int \sum_{j,k} y_j (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) (\partial_s \tilde{A}_k - \partial_k \tilde{A}_0) (1 - |y|^2)^{-\frac{3}{2}} \, dy.$$
We now manipulate these terms to simplify the resulting sum. Call these terms \( I, II, III \).

(I) For the first one, we get
\[
I = - \sum_j \int (3 + 2y \cdot \nabla y) \partial_j \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy
\]
\[
= - \sum_j \int [(3 + 2y \cdot \nabla y) \partial_j \tilde{A}_j] \, \partial_s \tilde{A}_j \rho(y) \, dy + \sum_j \int [(3 + 2y \cdot \nabla y) \partial_j \tilde{A}_j] \, \partial_j \tilde{A}_0 \rho(y) \, dy.
\]
Here the first term on the right will again lead to the crucial monotone term, this time involving \( \partial_s \tilde{A}_j \), while the second term is of ambiguous sign and expected to cancel later on.

(II) For the second term, we have
\[
II = - \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy
\]
\[
= - \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy
\]
\[
+ \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_j (3 + y \cdot \nabla y) (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy
\]
\[
+ \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) (y \cdot \nabla y) \rho(y) \, dy.
\]
This can be further simplified to
\[
II = \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_j (1 + y \cdot \nabla y) (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) \rho(y) \, dy
\]
\[
+ \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_j (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0) (y \cdot \nabla y + 1) \rho(y) \, dy.
\]
At this stage, we recall the Cronstrom gauge condition, which implies that
\[
\sum_j y_j (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j) = (y \cdot \nabla y + 1) \tilde{A}_k
\]
and so we can formulate the error term (8.22) left over from earlier as

\[ (8.22) = - \int \sum_k (y \cdot \nabla y + 1) \hat{A}_k (\partial_s \hat{A}_k - \partial_k \hat{A}_0)(1 - |y|^2)^{-\frac{1}{2}} \ dy, \]

which is exactly negative the second term in the last expression for \( II \). Thus, taking this cancellation into account, we have reduced the contribution from \( II \) to the term

\[ \sum_j \int (1 + y \cdot \nabla y) \hat{A}_j(1 + y \cdot \nabla y) \partial_s \hat{A}_j \rho(y) \ dy \]

\[ - \sum_j \int (1 + y \cdot \nabla y) \hat{A}_j(1 + y \cdot \nabla y) \partial_j \hat{A}_0 \rho(y) \ dy. \]

(III) For the last of these terms, write it as

\[ III = \sum_j \int (2 + y \cdot \nabla y) \partial_j \hat{A}_0 \partial_s \hat{A}_j \rho(y) \ dy - \sum_j \int (2 + \frac{1}{2}y \cdot \nabla y)(\partial_j \hat{A}_0)^2 \rho(y) \ dy. \]

Finally, we compute the sum \( I + II + III \), where we omit terms cancelling as already noted, indicated by a \((...)'\). Then we get

\[ (I + II + III)' = - \sum_j \int [(3 + 2y \cdot \nabla y) \partial_s \hat{A}_j] \partial_s \hat{A}_j \rho(y) \ dy \]

\[ + \sum_j \int (1 + y \cdot \nabla y) \hat{A}_j(1 + y \cdot \nabla y) \partial_s \hat{A}_j \rho(y) \ dy \]

\[ - \sum_j \int (2 + \frac{1}{2}y \cdot \nabla y)(\partial_j \hat{A}_0)^2 \rho(y) \ dy \]

\[ + \sum_j \int [(3 + 2y \cdot \nabla y) \partial_s \hat{A}_j] \partial_j \hat{A}_0 \rho(y) \ dy \]

\[ - \sum_j \int (1 + y \cdot \nabla y) \hat{A}_j(1 + y \cdot \nabla y) \partial_j \hat{A}_0 \rho(y) \ dy \]

\[ + \sum_j \int (2 + y \cdot \nabla y) \partial_j \hat{A}_0 \partial_s \hat{A}_j \rho(y) \ dy. \]

Here the first three terms either have the right sign or are a pure \( s \)-derivative. The last three terms, call them \( A + B + C \), are of ambiguous sign and need to cancel. First, we have

\[ A = \sum_j \int [(3 + 2y \cdot \nabla y) \partial_s \hat{A}_j] \partial_j \hat{A}_0 \rho(y) \ dy \]

\[ = \sum_j \int \partial_s \hat{A}_j(- 5 - 2y \cdot \nabla y) \partial_j \hat{A}_0 \rho(y) \ dy - \sum_j \int \hat{A}_j \partial_j \hat{A}_0 2(y \cdot \nabla y) \rho(y) \ dy. \]
It follows that

$$A + B + C = \sum_j \int \partial_s \tilde{A}_j (-3 - y \cdot \nabla) \partial_j \tilde{A}_0 \rho(y) \, dy$$

$$- \sum_j \int (1 + y \cdot \nabla) \tilde{A}_j (1 + y \cdot \nabla) \partial_j \tilde{A}_0 \rho(y) \, dy$$

$$- \sum_j \int \partial_s \tilde{A}_j \partial_j \tilde{A}_0 (y \cdot \nabla) \rho(y) \, dy$$

$$= - \sum_j \int (1 + y \cdot \nabla) \partial_j \tilde{A}_0 (\partial_s + 1 + y \cdot \nabla) \tilde{A}_j \rho(y) \, dy$$

$$- \sum_j \int \partial_s \tilde{A}_j \partial_j \tilde{A}_0 (1 + y \cdot \nabla) \rho(y) \, dy.$$ 

We reformulate this further as

$$A + B + C = \sum_j \int (1 + y \cdot \nabla) \tilde{A}_0 (\partial_s + 2 + y \cdot \nabla) \partial_j \tilde{A}_j \rho(y) \, dy$$

$$+ \sum_j \int \partial_j \tilde{A}_0 (\partial_s + 1 + y \cdot \nabla) \tilde{A}_j \rho(y) \, dy$$

$$+ \sum_j \int (1 + y \cdot \nabla) \tilde{A}_0 (\partial_s + 2 + y \cdot \nabla) \tilde{A}_j \partial_j \rho(y) \, dy$$

$$- \sum_j \int \partial_s \tilde{A}_j \partial_j \tilde{A}_0 (1 + y \cdot \nabla) \rho(y) \, dy.$$ 

To simplify this further, we write the second term on the right as

$$\sum_j \int \partial_j \tilde{A}_0 (\partial_s + 1 + y \cdot \nabla) \tilde{A}_j \rho(y) \, dy = - \sum_j \int \tilde{A}_0 (\partial_s + 1 + y \cdot \nabla) \partial_j \tilde{A}_j \rho(y) \, dy$$

$$- \sum_j \int \tilde{A}_0 \partial_j \tilde{A}_j \rho(y) \, dy$$

$$- \sum_j \int \tilde{A}_0 (\partial_s + 1 + y \cdot \nabla) \tilde{A}_j \partial_j \rho(y) \, dy.$$ 

Note that due to the Cronstrom gauge condition, we have

$$\sum_j (\partial_s + 1 + y \cdot \nabla) \tilde{A}_j \partial_j \rho(y) = \sum_j (\partial_s + 2 + y \cdot \nabla) \tilde{A}_j \partial_j \rho(y) = 0.$$
We conclude that

\[
A + B + C = \sum_j \int (1 + y \cdot \nabla y) \tilde{A}_0(\partial_s + 2 + y \cdot \nabla y) \partial_j \tilde{A}_0 \rho(y) \, dy - \sum_j \int \tilde{A}_0(\partial_s + 1 + y \cdot \nabla y) \partial_j \tilde{A}_0 \rho(y) \, dy - \sum_j \int \tilde{A}_0 \partial_j \tilde{A}_0 \rho(y) \, dy - \sum_j \int \partial_s \tilde{A}_0 \partial_j \tilde{A}_0 2(1 + y \cdot \nabla y) \rho(y) \, dy.
\]

Finally, this can be simplified to

\[
\sum_j \int (y \cdot \nabla y) \tilde{A}_0(\partial_s + 2 + y \cdot \nabla y) \partial_j \tilde{A}_0 \rho(y) \, dy - \sum_j \int \partial_s \tilde{A}_0 \partial_j \tilde{A}_0 2(1 + y \cdot \nabla y) \rho(y) \, dy.
\]

Then note that

\[
\sum_j \int (y \cdot \nabla y) \tilde{A}_0(\partial_s + 2 + y \cdot \nabla y) \partial_j \tilde{A}_0 \rho(y) \, dy = \sum_j \int (y \cdot \nabla y) \tilde{A}_0 \partial_j \tilde{x}_j \tilde{A}_0 \rho(y) \, dy,
\]

which partly cancels the last term in (8.20). We can now finally summarize all the preceding computations by the following conclusion: We get

\[
\frac{d}{ds} \int \left[ \frac{1}{2} \sum_{j=1}^4 (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0)^2 + \frac{1}{4} \sum_{j,k} (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j)^2 - \frac{1}{2} \sum_{j=1}^4 [(1 + y \cdot \nabla y) \tilde{A}_j]^2 \right] \rho(y) \, dy \\
+ \frac{d}{ds} \int \left[ \frac{1}{2} \sum_{j=1}^4 [(\partial_s + i \tilde{A}_0) \phi]^2 + \sum_{k=1}^4 |(\partial_k + i \tilde{A}_k) \phi|^2 - |(y \cdot \nabla y + 1) \phi|^2 - |\phi|^2 \right] \rho(y) \, dy \\
= - \int \text{Re} \left[ (3 + 2y \cdot \nabla y)(\partial_s + i \tilde{A}_0) \phi(\partial_s + i \tilde{A}_0) \phi \right] \rho(y) \, dy \\
+ \sum_{k=1}^4 \int (y \cdot \nabla y) \tilde{A}_0 \partial_k^2 \tilde{A}_0 \rho(y) \, dy \\
- \int (3 + 2y \cdot \nabla y) \partial_s \tilde{A}_0 \partial_j \tilde{A}_0 \rho(y) \, dy \\
- \int (2 + y \cdot \nabla y) \partial_j \tilde{A}_0 \partial_j \tilde{A}_0 \rho(y) \, dy \\
- \sum_j \int \partial_s \tilde{A}_j \partial_j \tilde{A}_0 2(1 + y \cdot \nabla y) \rho(y) \, dy.
\]
Performing integrations by parts in the last three terms but one, as well as the first term, we can write this equation in its final form as

\[
\frac{d}{ds} \int \left[ \frac{1}{2} \sum_{j=1}^{4} (\partial_s \tilde{A}_j - \partial_j \tilde{A}_0)^2 + \frac{1}{4} \sum_{j,k} (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j)^2 - \frac{1}{2} \sum_{j=1}^{4} [(1 + y \cdot \nabla_y)\tilde{A}_j]^2 \right] \rho(y) \, dy \\
+ \frac{d}{ds} \frac{1}{2} \int \left[ [\phi] + \sum_{k=1}^{4} [(\partial_k \tilde{A}_0)^2 + (\partial_s \tilde{A}_k)^2 - 2 \partial_s \tilde{A}_k \partial_j \tilde{A}_0] (1 + y \cdot \nabla_y) \rho(y) \, dy \right. \\
- \frac{1}{2} \int \left. \left[ [\phi] + \sum_{k=1}^{4} [(\partial_k \tilde{A}_0)^2 - (\partial_j \tilde{A}_k)^2] (1 - |y|^2)^{-\frac{1}{2}} \right] \rho(y) \, dy. \right)
\]

(8.24)

The left hand side of this relation appears to be non gauge invariant, but recalling the assumed Cronstrom condition, we can immediately replace it by a gauge invariant one: Recall that

\[
\sum_k y_j (\partial_k \tilde{A}_j - \partial_j \tilde{A}_0) = (y \cdot \nabla_y + 1) \tilde{A}_j
\]

for \( k = 1, \ldots, 4 \). Hence, we can write

\[
\frac{1}{2} \int \left[ \sum_{j=1}^{4} (\partial_j \tilde{A}_j - \partial_j \tilde{A}_0)^2 + \frac{1}{4} \sum_{j,k} (\partial_j \tilde{A}_k - \partial_k \tilde{A}_j)^2 - \frac{1}{2} \sum_{j=1}^{4} [(1 + y \cdot \nabla_y)\tilde{A}_j]^2 \right] \rho(y) \, dy \\
+ \frac{1}{2} \int \left[ [\phi] + \sum_{k=1}^{4} [(\partial_k \tilde{A}_0)^2 + (\partial_s \tilde{A}_k)^2 - 2 \partial_s \tilde{A}_k \partial_j \tilde{A}_0] (1 + y \cdot \nabla_y) \rho(y) \, dy \right. \\
- \frac{1}{2} \int \left. \left[ [\phi] + \sum_{k=1}^{4} [(\partial_k \tilde{A}_0)^2 - (\partial_j \tilde{A}_k)^2] (1 - |y|^2)^{-\frac{1}{2}} \right] \rho(y) \, dy. \right)
\]

Note that all expressions here are gauge invariant with the exception of \( \partial_s \tilde{A}_j - \partial_j \tilde{A}_0, (\partial_s + i\tilde{A}_0)\tilde{\phi} \).

Recalling the definition of \( \tilde{A}_v \) and again exploiting the Cronstrom assumption, we can write this last expression as

\[
\partial_s \tilde{A}_j - \partial_j \tilde{A}_0 = e^{-2s}(\partial_s A_j - \partial_j A_0)(1 - e^{-s}, ye^{-s}) - \sum_k y_k (\partial_k \tilde{A}_j - \partial_j \tilde{A}_k),
\]

which is now gauge independent. The latter identity also needs to be used for the expression \( \sum_{k=1}^{4} (\partial_k \tilde{A}_0 - \partial_s \tilde{A}_k)^2 \) on the right hand side of (8.24) to make it in turn gauge invariant. Furthermore, we have

\[
(\partial_s + i\tilde{A}_0)\tilde{\phi} = e^{-2s}(\phi_t + iA_0\phi)(1 - e^{-s}, ye^{-s}) - e^{-s}[y \cdot (\nabla_y + i\tilde{A}) + 1]\phi
\]

which is manifestly gauge invariant.
9. Technical details for ruling out a minimal blowup solution

In this section we outline how the same steps as in Section 10 in [10] can be carried out for the Maxwell-Klein-Gordon equation. To begin with, we note the analogue of Corollary 10.2 in [10].

**Corollary 9.1.** The conclusions of Proposition 8.3 hold provided that \((A, \phi)\) are a weak solution to MKG-CG in the sense of Section 5.

In fact such solutions are locally uniform limits of admissible solutions, and hence the corresponding conservation laws follow by passing to the limit in an integrated formulation. In a similar vein, we have the analogue of Lemma 10.3 in [10], for which we actually make the following stronger statement.

**Lemma 9.2.** Let \((A, \phi)\) a weak solution to MKG-CG in the sense of Section 5. For given \(\epsilon > 0\), let \(M > 0\) be such that

\[
\int_{|x| > M} \left[ \frac{1}{2} \sum_{\alpha} |D_{\alpha} \phi(0, \cdot)|^2 + \frac{1}{4} \sum_{\alpha, \beta} (\partial_{\alpha} A_{\beta}(0, \cdot) - \partial_{\beta} A_{\alpha}(0, \cdot))^2 \right] dx < \epsilon.
\]

Then if \(I\) denotes the lifespan of the solution and \(I^+ := I \cap [0, \infty)\), then for any \(t \in I^+\), we have

\[
\int_{|x| > M+t} \left[ \frac{1}{2} \sum_{\alpha} |D_{\alpha} \phi(t, \cdot)|^2 + \frac{1}{4} \sum_{\alpha, \beta} (\partial_{\alpha} A_{\beta}(t, \cdot) - \partial_{\beta} A_{\alpha}(t, \cdot))^2 \right] dx < \epsilon.
\]

This is proved by approximating \((A, \phi)\) by a sequence of admissible solutions and passing to the limit in Lemma 8.4.

Following the procedure in [10], which in turn is following the one in [5], one distinguishes between the case of finite and infinite lifespan \(I\) for the minimal blowup solution \((A^\infty, \Phi^\infty)\). As above we denote \(I^+ = I \cap [0, \infty)\). Then in case of finite lifespan we have the analogue of Lemma 10.4 in [10].

**Lemma 9.3.** Assume \(I^+\) is finite and after re-scaling that \(I^+ = [0, 1)\). Let \(\lambda(t) : I^+ \rightarrow \mathbb{R}_+\) be as in Theorem 7.26. Then there exists a constant \(C_0 > 0\) (depending on the solution) such that

\[0 < \frac{C_0}{1 - t} \leq \lambda(t)\]

for all \(0 \leq t < 1\).

The proof of this follows exactly as in [10] by combining Corollary 6.3 and Theorem 7.26.

The preceding lemma, when combined with Lemma 9.2, gives immediately the following analogue of Lemma 10.5 in [10].

**Lemma 9.4.** Make the same assumptions as in the preceding lemma. Then there is \(x_0 \in \mathbb{R}^4\) such that

\[\text{supp} (\partial_{\alpha} A^\infty_{\beta} - \partial_{\beta} A^\infty_{\alpha}, \nabla_{t,x} \Phi^\infty) \subset B(x_0, 1 - t)\]

for all \(0 \leq t < 1\) and \(\alpha, \beta = 0, \ldots, 4\).

**Proof.** We follow the proof of Lemma 4.8 in [5]. First, we show that for each \(t \in [0, 1)\), there exists a ball \(B_{1-t}\) of radius \(1 - t\) such that all components

\[\partial_{\alpha} A^\infty_{\beta} - \partial_{\beta} A^\infty_{\alpha}, \nabla_{t,x} \Phi^\infty\]
are supported in $B_{1-t}$. If not, then for some $t \in [0, 1)$, there exist $\epsilon_0 > 0, \eta_0 > 0$ such that for all $x_0 \in \mathbb{R}^4$ we have

$$
(9.1) \quad \int_{|x-x_0| \geq (1+\eta_0)(1-t)} \left[ \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha \mathcal{A}_\alpha^0 - \partial_\beta \mathcal{A}_\alpha^0)^2 + \frac{1}{2} \sum_\alpha |D_\alpha \Phi^0|^2 \right] dx \geq \epsilon_0.
$$

Next, pick a sequence $t_n \rightarrow 1$. We know from the preceding lemma that there exists a constant $C_0$ with

$$
\lambda(t_n) \geq \frac{C_0}{1-t_n}.
$$

By the compactness property expressed in Theorem 7.26, we know that for any $R_0 > 0$ and $n$ sufficiently large, we have

$$
\int_{|x+\frac{tn}{n-1}| \geq R_0} \left[ \sum_{\alpha, \beta} |\partial_\beta \mathcal{A}_\alpha^0(t, \cdot)|^2 \right] dx < \frac{\epsilon_0}{10}.
$$

In particular, using Lemma 9.2, we then find that

$$
\int_{|x+\frac{tn}{n-1}| \geq R_0, t_n-t} \left[ \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha \mathcal{A}_\beta^0(t, \cdot) - \partial_\beta \mathcal{A}_\alpha^0(t, \cdot))^2 + \frac{1}{2} \sum_\alpha |D_\alpha \Phi^0(t, \cdot)|^2 \right] dx < \frac{\epsilon_0}{10}.
$$

This contradicts (9.1) with $x_0 = \frac{tn}{n-1}$, provided

$$
(1+\eta_0)(1-t) \geq R_0 + t_n - t,
$$

and this can always be forced by choosing $n$ large enough and $R_0$ small enough, depending on $\eta_0$ and $t \in [0, 1)$. Again, arguing as in [5], one shows that $\frac{\alpha(t_n)}{\lambda(t_n)}$ is a bounded function, and so by passing to a suitable sequence $t_n \rightarrow 1$ one may assume

$$
\frac{\alpha(t_n)}{\lambda(t_n)} \rightarrow x_* \in \mathbb{R}^4.
$$

As before, for any fixed $R_0 > 0, \epsilon_0 > 0$ and $n$ sufficiently large we get

$$
\int_{|x+\frac{tn}{n-1}| \geq R_0, t_n-t} \left[ \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha \mathcal{A}_\beta^0(t, \cdot) - \partial_\beta \mathcal{A}_\alpha^0(t, \cdot))^2 + \frac{1}{2} \sum_\alpha |D_\alpha \Phi^0(t, \cdot)|^2 \right] dx < \frac{\epsilon_0}{10}
$$

and then passing to the limit $n \rightarrow \infty$, we infer

$$
\int_{|x+\frac{tn}{n-1}| \geq R_0, t_n-t} \left[ \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha \mathcal{A}_\beta^0(t, \cdot) - \partial_\beta \mathcal{A}_\alpha^0(t, \cdot))^2 + \frac{1}{2} \sum_\alpha |D_\alpha \Phi^0(t, \cdot)|^2 \right] dx < \frac{\epsilon_0}{10}.
$$

Since $\epsilon_0 > 0, R_0 > 0$ were arbitrary, we conclude that for all $\alpha, \beta$ and any $t \in [0, 1)$, we have

$$
\text{supp} (\partial_\alpha \mathcal{A}_\beta^0(t, \cdot) - \partial_\beta \mathcal{A}_\alpha^0(t, \cdot), D_\alpha \Phi^0) \subset B_{1-t}(-x_*).
$$

We next render the heuristic considerations about the transformation behavior of the energy under Lorentz transformations at the beginning of Section 8 rigorous. For technical reasons, we need to distinguish between the case of infinite lifespan $I$ and finite lifespan.

**Proposition 9.5.** Let $(\mathcal{A}_\alpha^0, \Phi^0)$ be as in the preceding discussion. Assume that $I^+$ is finite, whence by re-scaling, we may assume the conclusion of the preceding lemma applies. Then we have

$$
(9.2) \quad \int_{\mathbb{R}^+} \left[ \sum_{j=1}^4 (\partial_j \mathcal{A}_j^0 - \partial_j \mathcal{A}_j^0)(\partial_j \mathcal{A}_j^0 - \partial_j \mathcal{A}_j^0) + \text{Re} ((\Phi_i^0 + i\mathcal{A}_0^0 \Phi^0)(\Phi_i^0 + i\mathcal{A}_0 \Phi^0)) \right] dx = 0
$$
for $k = 1, \ldots, 4$, where the integration is over an arbitrary time slice $\{t\} \times \mathbb{R}^4$ with $t \in I^+$.

**Proof.** This follows essentially by combining Proposition 8.1 with the calculation on the change of energy under Lorentz transformations. In order to apply the latter proposition, we have to use smooth solutions, which are globally defined, as otherwise we cannot meaningfully apply a Lorentz transformation. In fact, we may exploit that by the preceding lemma, the function $\Phi^n(0)$ is compactly supported, which means its Fourier transform cannot also be compactly supported (we may of course assume $\Phi^n(0)$ to be non-vanishing, since otherwise, the solution extends trivially in global fashion and cannot be singular). But then, truncating the data in Fourier space as in Proposition 5.1 and the discussion following it, we may construct a sequence of smooth Coulomb data $(A^n, \phi^n)[0]$ converging to $(\mathcal{R}^\infty, \Phi^\infty)[0]$, and if necessary, multiplying the $\phi^n[0]$ in the resulting $(A^n, \phi^n)[0]$ by a small scalar $\lambda_n \in [0, 1]$ with $\lambda_n \to 1$ as $n \to \infty$, we may force

$$E(A^n, \phi^n) < E_{\text{crit}}, \forall n.$$  

Note that then the perturbation theory developed in Proposition 5.1 still applies in relation to $(\mathcal{R}^\infty, \Phi^\infty)$, since we have not changed the data for $A^n$. This means that the data $(A^n, \phi^n)[0]$ do admit a global MKG-CG evolution by definition of $E_{\text{crit}}$, and can be Lorentz transformed. In order to justify various conservation laws for the Lorentz transformed $(A^n, \phi^n)$, we observe that we may also localize the data $(A^n, \phi^n)[0]$ in physical space to a sufficiently large ball, using the argument in Subsection 5.2 as well as the result [4], such that the Lorentz transformed solution also has compact support on bounded time slices, and we still have the above inequality for the energy.

We shall then make the hypothesis that the above momentum does not vanish, whence there is some $k$ and $\gamma > 0$ such that for all sufficiently large $n$ we have, say,

$$\int_{\mathbb{R}^4} \left[ \sum_{j=1}^{4} \left( \partial_i A^n_j - \partial_j A^n_i \right) \left( \partial_i \tilde{A}^n_j - \partial_j \tilde{A}^n_i \right) + \text{Re} \left( \left( \phi^n_i + iA^n_i \phi^n \right) \left( \tilde{\phi}^n_k + i\tilde{A}^n_k \phi^n \right) \right) \right] \, dx > \gamma. \tag{9.3}$$

Then we shall show that a suitable Lorentz transformation $L$ exists such that the transformed

$$(A^n, \phi^n) =: (\tilde{A}^n, \tilde{\phi}^n)$$

does have energy $< E_{\text{crit}} - \kappa(\gamma, \mathcal{R}^\infty, \Phi^\infty)$, where $\kappa(\gamma, \mathcal{R}^\infty, \Phi^\infty) > 0$, and so a global $S^1$-bound applies to this solution when transformed into the Coulomb gauge. But then, using Proposition 8.1 we can infer a global $S^1$-bound uniformly in $n$ for

$$(A^n, \phi^n)[0],$$

which then contradicts the fact that $(\mathcal{R}^\infty, \Phi^\infty)$ is a singular solution. To implement this strategy, we shall combine the argument for Proposition 4.10 in [5] with the computation (8.11). We shall and may assume that $k = 1$. We already note that in the sequel, there is only a mild technical complication due to the fact that the $(A^n, \phi^n)$ will not satisfy the same compact support property as the limiting object, but since they approximate the limiting object, this does not really cause problems.

Using the same notation as in Section 8.2 we use the same Lorentz transform $L$ introduced there, and denote the transformed components of $(A^n, \phi^n)$ by $(\tilde{A}^n, \tilde{\phi}^n)$. We follow the argument for the proof of Proposition 4.10 in [5] and shift the backward light cone in the conclusion of the preceding lemma to have vertex at the origin, and so we shall now restrict to $t \in [-1, 0)$. Then, using (8.14) for $(\tilde{A}^n, \tilde{\phi}^n)$, we have the energy relation

$$\frac{1}{4} E(\tilde{A}^n(-\frac{1}{2}, \cdot), \tilde{\phi}^n(-\frac{1}{2}, \cdot)) = \int_{-\frac{1}{2}}^{\frac{1}{2}} E(\tilde{A}^n(t, \cdot), \tilde{\phi}^n(t, \cdot)) \, dt,$$
where we denote (on a fixed time slice)

\[ E(A, \phi) = \int_{\mathbb{R}^4} \left[ \frac{1}{4} \sum_{\alpha, \beta} (\partial_\alpha A_\beta - \partial_\beta A_\alpha)^2 + \frac{1}{2} \sum_\alpha |D_\alpha \phi|^2 \right] dx. \]

We now follow [5] quite closely: According to (8.11) and the exact computations preceding it, we can write the above as \( I_1 + I_2 \), where

\[
I_1 = \int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} \left[ \frac{1}{4} \sum_{j=1}^4 \left( (\partial_j A_n^n - \partial_j A_0^n)^2 + \frac{1}{2} \left( \partial_j A_j^n - \partial_j A_i^n \right)^2 \right) \right] dx \, dt
\]

\[
+ \int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} \left[ \frac{1}{4} \left( \partial_j A_0^n - \partial_j A_1^n \right)^2 + \frac{1}{4} \sum_{i,j \neq 0,1} (\partial_i A_j - \partial_j A_i)^2 \right] dx \, dt
\]

\[
+ \int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} \left[ \frac{1}{2} \left( |D_\alpha \phi|^2 + |D_1 \phi|^2 \right) + \sum_{j=1}^4 |D_j \phi|^2 \right] dx \, dt
\]

\[ = \int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} I_1 \, dx \, dt \]

and

\[
-I_2 = \frac{2d}{1-d^2} \int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} \left[ \sum_{j=1}^4 \left( \partial_j A_j^n - \partial_j A_0^n \right) \left( \partial_1 A_j^n - \partial_j A_0^n \right) + \text{Re} \left( \left( \phi_i^n + iA_0^n \phi^n \right) \left( \phi_i^n + iA_1^n \phi^n \right) \right) \right] dx \, dt
\]

\[ = \int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} I_2 \, dx \, dt. \]

Throughout it is to be kept in mind that we are abusing notation in the above, in that the factors in each integrand need to be evaluated at

\( \left( \frac{t-dx_1}{\sqrt{1-d^2}}, \frac{x_1-dt}{\sqrt{1-d^2}}, x_2, x_3, x_4 \right) =: \Phi(t, x) = (s, y_1, y_2, y_3, y_4). \)

Following [5], we now analyze the limit of \( I_2 \) as \( d \to 0 \). Writing \( [-\frac{1}{d}, \frac{1}{d}] \times \mathbb{R}^4 =: D_{-\frac{1}{d}} \) and denoting the distorted time slice after coordinate change by \( D_{-\frac{1}{d}} := \Phi(D_{-\frac{1}{d}}) \), we have

\[
\int_{-\frac{1}{d}}^{\frac{1}{d}} \int_{\mathbb{R}^4} I_2 \, dx \, dt = \int_{D_{-\frac{1}{d}}} I_2 \, dy_1 \, d\gamma \, ds,
\]

where we of course denote \( y_1 = \frac{x_1-dt}{\sqrt{1-d^2}} \), \( s = \frac{-dx_1}{\sqrt{1-d^2}}, \gamma = (y_2, y_3, y_4) = (x_2, x_3, x_4) \). Also, in the second integral, the inputs \( A^n, \phi^n \) are now evaluated in \( s, y, i.e. A^n(s, y) \) etc.

Now fix a small \( d > 0 \). Then we shall pick \( n = n(d) \) large enough such that the energy of \( (A^n, \phi^n) \) is sufficiently concentrated in the set \( |y| \leq 2 \) such that we have

\[
\int_{D_{-\frac{1}{d}} \cap \{ |y| > 2 \}} I_2 \, dy_1 \, d\gamma \, ds \ll d^2.
\]
Then it follows that

\[
\int_{D_{1/2}} I_2 \, dy_1 \, d\vec{y} \, ds = \int_{|y| \leq 2} \int_{-\frac{1}{2} \sqrt{1 - d^2} - dy_1}^{\frac{1}{2} \sqrt{1 - d^2} - dy_1} I_2 \, ds \, dy_1 \, d\vec{y} + O(d^2)
\]

\[
= \int_{|y| \leq 2} \int_{-\frac{1}{2} \sqrt{1 - d^2} - dy_1}^{\frac{1}{2} \sqrt{1 - d^2} - dy_1} I_2 \, ds \, dy_1 \, d\vec{y} + O(d^2) + O(E_{\text{crit}}) + O(d^2).
\]

Further, again by the localization of \((\mathcal{R}^\infty, \Phi^\infty)\), picking \(n\) large enough (in relation to \(d\)), we may assume

\[
\int_{|y| \leq 2} \int_{-\frac{1}{2} \sqrt{1 - d^2} - dy_1}^{\frac{1}{2} \sqrt{1 - d^2} - dy_1} I_2 \, ds \, dy_1 \, d\vec{y} = \int_{\mathbb{R}^4} \int_{-\frac{1}{2} \sqrt{1 - d^2} - dy_1}^{\frac{1}{2} \sqrt{1 - d^2} - dy_1} I_2 \, ds \, dy_1 \, d\vec{y} + O(d^2)
\]

\[
= \frac{1}{4} \int_{\mathbb{R}^4} \left[ \sum_{j=1}^{4} (\partial_j A^n_j - \partial_j A^n_0) (\partial_1 A^n_1 - \partial_j A^n_1) + \text{Re} ((\phi^n_1 + iA^n_0\phi^n_0)(\phi^n_1 + iA^n_1\phi^n_0)) \right] \, dx + O(d^2).
\]

In total, we have shown that for \(n\) sufficiently large (depending on \(d > 0\) fixed), we have

\[-I_2 = \frac{d}{2} \int_{\mathbb{R}^4} \left[ \sum_{j=1}^{4} (\partial_j A^n_j - \partial_j A^n_0) (\partial_1 A^n_1 - \partial_j A^n_1) + \text{Re} ((\phi^n_1 + iA^n_0\phi^n_0)(\phi^n_1 + iA^n_1\phi^n_0)) \right] \, dx + O(d^2).
\]

For the first term \(I_1\), we use that for \(n\) again large enough

\[
I_1 = \int_{D_{1/4} \cap \{|y| \leq 2\}} I_1 \, dy_1 \, d\vec{y} \, ds + O(d^2) = \int_{|y| \leq 2} \int_{-\frac{1}{2} \sqrt{1 - d^2} - dy_1}^{\frac{1}{2} \sqrt{1 - d^2} - dy_1} I_1 \, ds \, dy_1 \, d\vec{y} + O(d^2)
\]

and this we further carefully write as

\[
I_1 = d \int_{|y| \leq 2} \left[ y_1 I_1 \left( \frac{1}{2} \cdot, \cdot \right) - y_1 I_1 \left( -\frac{1}{4} \cdot, \cdot \right) \right] \, dy_1 \, d\vec{y} + d \cdot c(d) + O(d^2) + \int_{|y| \leq 2} \int_{-\frac{1}{2} \sqrt{1 - d^2}}^{\frac{1}{2} \sqrt{1 - d^2}} I_1 \, ds \, dy_1 \, d\vec{y},
\]

where the function \(c(d)\) satisfies \(\lim_{d \to 0} c(d) = 0\) with a rate of convergence depending on the profile \((\mathcal{R}^\infty, \Phi^\infty)\), but independent of \(n\), provided the latter is sufficiently large. Then again choosing \(n\) large enough we can force that

\[
\int_{|y| \leq 2} \int_{-\frac{1}{2} \sqrt{1 - d^2}}^{\frac{1}{2} \sqrt{1 - d^2}} I_1 \, ds \, dy_1 \, d\vec{y} = \int_{\mathbb{R}^4} \int_{-\frac{1}{2} \sqrt{1 - d^2}}^{\frac{1}{2} \sqrt{1 - d^2}} I_1 \, ds \, dy_1 \, d\vec{y} + O(d^2) = \frac{1}{4} E_{\text{crit}} + O(d^2).
\]

On the other hand, for the integral involving the difference expression above, using \((8.16)\) where we let \(R \to \infty\) (it is straightforward to justify this limit by using our assumption that the spatial components of \(A^n\) as well as \(\phi^n\) are compactly supported on fixed time slices), we obtain

\[
\int_{|y| \leq 2} \left[ y_1 I_1 \left( \frac{1}{2} \cdot, \cdot \right) - y_1 I_1 \left( -\frac{1}{4} \cdot, \cdot \right) \right] \, dy_1 \, d\vec{y}
\]

\[
= \frac{1}{4} \int_{\mathbb{R}^4} \left[ \sum_{j=1}^{4} (\partial_j A^n_j - \partial_j A^n_0) (\partial_1 A^n_1 - \partial_j A^n_1) + \text{Re} ((\phi^n_1 + iA^n_0\phi^n_0)(\phi^n_1 + iA^n_1\phi^n_0)) \right] \, dx + O(d^2).
\]
To summarize the preceding discussion, we have now shown that
\[
\frac{1}{4} E(\tilde{A}^n(-\frac{1}{2}, \cdot), \tilde{\phi}^n(-\frac{1}{2}, \cdot)) = \frac{1}{4} E_{crit} - \frac{d}{4} \int_{\mathbb{R}^4} \left[ \sum_{j=1}^4 (\partial_t A_j^n - \partial_j A_0^n)(\partial_t A_j^n - \partial_j A_0^n) + \text{Re} \left( (\phi_i^n + iA_0^n\phi^n)(\phi_i^n + iA_0^n\phi^n) \right) \right] \, dx \\
+ O(d^2) + d \cdot c(d),
\]
provided that \( n = n(d) \) is sufficiently large. On the other hand, the term \( c(d) \) is \( \ll 1 \), provided \( d \) is sufficiently small (depending on the fixed profile \((\mathcal{A}^\infty, \Phi^\infty)\)), and \( n \) is sufficiently large. If we further shrink \( d \), if necessary, depending on the constant \( \gamma \) in (9.3), we can force that
\[
E(\tilde{A}^n(-\frac{1}{2}, \cdot), \tilde{\phi}^n(-\frac{1}{2}, \cdot)) < E_{crit} - \kappa
\]
for some \( \kappa > 0 \), which is uniform for all sufficiently large \( n \), and this furnishes the desired a priori \( S^1 \)-bound on the components obtained by Coulomb gauging the \((\tilde{A}^n, \tilde{\phi}^n)\).

We next mention the corresponding result provided \( I^+ \) is infinite. Its proof follows essentially along the lines of the proof of Proposition 4.11 in [5], using the same modifications as in the preceding proof.

**Proposition 9.6.** Let \((\mathcal{A}^\infty, \Phi^\infty)\) be as before, with \( I^+ = [0, \infty) \cap I = [0, \infty) \), and assume the scaling parameter \( \lambda(t) \) from Theorem 7.26 is uniformly bounded from below by \( \lambda_0 > 0 \). Then the same conclusion as in the preceding proposition applies.

With these preparations in place, we can now follow essentially the same argument as the one given in Section 10.2 in [10]. We state the following fundamental

**Proposition 9.7.** Let \((\mathcal{A}^\infty, \Phi^\infty)\) be as before, and write \( I = (-T_0, T_1) \) for its lifespan. Then one cannot have \( T_1 \) or \( T_0 \) finite. Moreover, if \( \lambda(t) \geq \lambda_0 > 0 \) for all \( t \in \mathbb{R} \), then necessarily \((\mathcal{A}^\infty, \Phi^\infty) = (0, 0) \), whence there cannot be a minimal energy blowup solution under the given assumptions.

We follow the method of proof in [10], which in turn follows the strategy in [5], but also adds a crucial Vitali type covering argument. We start with the case when \( T_1 = +\infty \), and \( \lambda(t) \geq \lambda_0 > 0 \) for all \( t \in I^+ \). From Theorem 7.26 we know that for any \( \epsilon > 0 \), there exists \( R_0(\epsilon) > 0 \) with the property that (using the terminology from that theorem) we have
\[
\int_{|x| \geq \|x\|_{A_0}} \left[ \sum_{j} |\nabla_x \mathcal{A}^\infty_{\alpha}(t, \cdot)|^2 + |\nabla_x \Phi|^2 \right] \, dx \leq \epsilon.
\]

Then in perfect analogy with [10] and [5], we have the following

**Lemma 9.8.** Under the preceding assumptions, there exist \( \epsilon_1 > 0 \), \( C > 0 \) such that if \( \epsilon \in (0, \epsilon_1) \), there exists \( R_0(\epsilon) \) so that if \( R > 2R_0(\epsilon) \), then there exists \( t_0 = t_0(R, \epsilon) \), \( 0 \leq t_0 \leq CR \), with the property that for all \( 0 < t < t_0 \) one has
\[
\left| \frac{\overline{\lambda}(t)}{\lambda(t)} \right| < R - R_0(\epsilon), \quad \left| \frac{\overline{\lambda}(t_0)}{\lambda(t_0)} \right| = R - R_0(\epsilon).
\]

**Proof.** We adjust the proof of [10] by using the conservation laws from Proposition 8.3. To begin with, we show that there exists \( \alpha > 0 \) with
\[
\int_I \int_{\mathbb{R}^4} \left[ |\Phi^\infty + iA_0^\infty\Phi^\infty|^2 + \sum_k (\partial_t \mathcal{A}^\infty_k - \partial_k A_0^\infty)^2 \right] \, dx \, dt \geq \alpha > 0
\]
for all intervals $I \subset I^+$ of length one. We argue by contradiction. If not, then there exists a sequence of intervals $J_n = [t_n, t_n + 1]$ with $t_n \to +\infty$ and such that

\begin{equation}
\int_{J_n} \int_{\mathbb{R}^4} \left[ |\Phi_t^\infty + iA_0^\infty \Phi^\infty|^2 + \sum_{k} (\partial_k A_k^\infty - \partial_k A_0^\infty)^2 \right] \, dx \, dt \leq \frac{1}{n}.
\end{equation}

Then pick $s_n \in \frac{I}{4}$, i.e. the concentric interval of half the length. By the compactness property in Theorem 7.26, there exists a non-empty open interval $I^*$ around $t = 0$, and such that (passing to a subsequence and not making a distinction)

\[ \lambda(s_n)^{-2}(\nabla_{I^*} A_0^\infty, \nabla_{I^*} \Phi^\infty)(s_n + t, \lambda(s_n)^{-1}, (-\nabla(s_n))\lambda(s_n)^{-1}) \]

converges toward a limiting function

\[ (\nabla_{I^*} A^*, \nabla_{I^*} \Phi^*) \in C^0(I^*, (L^2_\lambda)^7) \]

in the given topology, such that $(A^*, \Phi^*)$ represents a weak solution of MKG on $I^* \times \mathbb{R}^4$ in the $L^2_\lambda H^1_\lambda$-sense. Again this solution will satisfy the Coulomb condition.

First, assume that $\lambda(s_n)$ remains bounded. Then one may replace $I^*$ by a smaller non-empty open $I^*$ with

\[ s_n + \lambda(s_n)^{-1}I^* \subset J_n, \ \forall n \geq 1, \]

and then by (9.6) we infer that

\[ \int_{I^*} \int_{\mathbb{R}^4} \left[ |\Phi_t^* + iA_0^* \Phi^*|^2 \right] \, dx \, dt = 0, \]

whence $\Phi_t^* + iA_0^* \Phi^* = 0$ on $I^* \times \mathbb{R}^4$. But then

\[ \sum_{k=1}^4 (\partial_k + i\mathcal{A}_k^\infty)^2 \Phi^* = 0 \]

on $I^* \times \mathbb{R}^4$, of course in the weak sense. But this equation then implies that

\[ \Phi^*|_{I^* \times \mathbb{R}^4} = 0 \]

and from there also $\Phi_t^*|_{I^*} = 0$. This then implies that $(A^*, \Phi^*)$ is a “trivial” solution in that the spatial components of $A^*$ are finite energy free waves, while its temporal component vanishes, and we have $\Phi^* = 0$. But this solution has finite $S^1$-bounds, which is a contradiction upon applying Proposition 7.23.

Next, assume that $\lambda(s_n)$ is not bounded. Then we essentially replicate the preceding argument, but need to also add a Vitali type covering trick. Thus write

\[ J_n = \bigcup_{s \in J_n} [s - \lambda^{-1}(s), s + \lambda^{-1}(s)] \cap J_n. \]

Then applying Vitali’s covering lemma we may pick a set $A_n \subset J_n$ such that the corresponding intervals are disjoint, and such that we have

\[ \sum_{s \in A_n} \left[ [s - \lambda^{-1}(s), s + \lambda^{-1}(s)] \cap J_n \right] \geq \frac{1}{5}. \]

It follows that we may pick a sequence $s_n, n \geq 1$, with $s_n \in A_n$, and such that with

\[ I_s := [s - \lambda^{-1}(s), s + \lambda^{-1}(s)], \]
we have
\[
\int_{I_0 \cap J_n} \int_{\mathbb{R}^4} \left[ |\Phi_1^\infty + iA_0^\infty \Phi^\infty|^2 + \sum_k (\partial_x A_k^\infty - \partial_t A_0^\infty)^2 \right] dx dt = o(\lambda(s_n)^{-1}).
\]
In particular, we get
\[
\int_{-1}^1 \int_{\mathbb{R}^4} \left[ |\Phi_1^\infty + iA_0^\infty \Phi^\infty|^2 + \sum_k (\partial_x A_k^\infty - \partial_t A_0^\infty)^2 \right] (s_n + t\lambda(s_n)^{-1}, \cdot) dx dt = o(1).
\]
But then, passing to a subsequence, we may again extract a limit from
\[
\lambda(s_n)^{-2}\left( \nabla_{t,x} A^\infty, \nabla_{t,x} \Phi^\infty \right) (s_n + t\lambda(s_n)^{-1}, \cdot - \bar{\pi}(s_n))\lambda(s_n)^{-1},
\]
which is a time independent solution, which leads to a contradiction as before.

This shows that (9.5) is indeed valid for suitable \( \alpha > 0 \). Next, use (8.17). From (9.4), and assuming
\[
\left| \frac{\bar{\pi}(t)}{\lambda(t)} \right| < R - R_0(\epsilon), \ 0 \leq t < CR
\]
for sufficiently large \( C \) to be picked later on, we have
\[
r(R) \leq \epsilon.
\]
But then for any interval \( I \) of length one, we have
\[
\int_I \left( \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \left[ |\Phi_1^\infty + iA_0^\infty \Phi^\infty|^2 + \sum_k (\partial_x A_k^\infty - \partial_t A_0^\infty)^2 \right] dx + r(R) \right) dt \geq \frac{\alpha}{2}
\]
provided \( R \) is sufficiently large. In particular, this implies that for \( CR > t_1 \geq t_0 + 1 \) and \( t_0 \geq 0 \), we have
\[
\int_{t_0}^{t_1} \left( \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \left[ |\Phi_1^\infty + iA_0^\infty \Phi^\infty|^2 + \sum_k (\partial_x A_k^\infty - \partial_t A_0^\infty)^2 \right] dx + r(R) \right) dt \geq \frac{\alpha}{2}(t_1 - t_0 - 1).
\]
On the other hand, using (8.17) and bounding the last term in \( M_R(t) \) via Hardy’s inequality, we get
\[
\int_{t_0}^{t_1} \left( \int_{\mathbb{R}^4} \varphi(\frac{x}{R}) \left[ |\Phi_1^\infty + iA_0^\infty \Phi^\infty|^2 + \sum_k (\partial_x A_k^\infty - \partial_t A_0^\infty)^2 \right] dx + r(R) \right) dt \leq RE_{\text{crit}}
\]
with a universal implied constant. The preceding bound contradicts the one before it, provided \( C \gg \alpha^{-1}E_{\text{crit}} \).

To finish off the case in Proposition 9.6 when \( I^+ = [0, \infty) \), we state the next lemma, which contradicts the preceding one.

**Lemma 9.9.** There exists \( \epsilon_2 > 0 \), \( R_1(\epsilon) > 0 \), \( C_0 > 0 \) such that if \( R > R_1(\epsilon) \), \( t_0 = t_0(R, \epsilon) \) as in the preceding lemma, and \( 0 < \epsilon < \epsilon_2 \), then
\[
t_0(R, \epsilon) > \frac{C_0R}{\epsilon}.
\]

The proof is identical to the one in 5, using (8.16) as well as the vanishing momentum relation in Proposition 9.6

In order to complete the proof of Proposition 9.7, we still need to handle the case of \( T_1 \) finite, whence by scaling \( T_1 = 1 \), say. Then we have the conclusion of Lemma 9.4 where by translation invariance we may of course assume \( x_0 = 0 \). Then we state
Lemma 9.10. Under the preceding assumptions, there exists a positive constant $C_1$ such that

$$\frac{C_1}{1 - t} \geq \lambda(t), \ t \in [0, 1).$$

Proof. As for the analogous Lemma 10.11 in [10], this follows basically from the argument in [5], by considering the functional

$$z(t) := \int_{\mathbb{R}^4} \sum_k \alpha_k \cdot \left[ \text{Re} \left[ (\Phi_t^\infty + iA_0^\infty \Phi_t^\infty)(\partial_k + iA_k^\infty)\Phi_t^\infty \right] + \sum_j (\partial_j A_j^\infty - \partial_j A_0^\infty) (\partial_k A_j^\infty) \right] dx$$

$$+ \int_{\mathbb{R}^4} \text{Re} \left[ \Phi_t^\infty (\Phi_t^\infty + iA_0^\infty \Phi_t^\infty) \right] dx,$$

compare with (8.17). Then from Proposition 8.3, we obtain

$$z'(t) = -\int_{\mathbb{R}^4} \left| \Phi_t^\infty + iA_0^\infty \Phi_t^\infty \right|^2 \ dx - \sum_k \int_{\mathbb{R}^4} (\partial_k A_k^\infty - \partial_k A_0^\infty)^2 \ dx.$$

As we clearly have \( \lim_{t \to 1} z(t) = 0 \), we can write

$$z(t) = \int_t^1 \int_{\mathbb{R}^4} \left| \Phi_t^\infty + iA_0^\infty \Phi_t^\infty \right|^2 \ dx ds + \sum_k \int_{\mathbb{R}^4} (\partial_k A_k^\infty - \partial_k A_0^\infty)^2 \ dx ds.$$

Then we distinguish between two possibilities: Either there exists \( \alpha > 0 \) such that for all \( t \in [0, 1) \) we have

$$\int_t^1 \int_{\mathbb{R}^4} \left| \Phi_t^\infty + iA_0^\infty \Phi_t^\infty \right|^2 \ dx ds + \sum_k (\partial_k A_k^\infty - \partial_k A_0^\infty)^2 \ dx ds \geq \alpha(1 - t),$$

or else there exists a sequence \( \{t_n\}_{n \geq 1} \subset [0, 1) \) with \( t_n \to 1 \) and such that denoting \( J_n := [t_n, 1) \), we have

$$|J_n|^{-1} \int_{J_n} \int_{\mathbb{R}^4} \left| \Phi_t^\infty + iA_0^\infty \Phi_t^\infty \right|^2 \ dx ds + \sum_k (\partial_k A_k^\infty - \partial_k A_0^\infty)^2 \ dx ds \to 0.$$

In the former case, one argues exactly as in [5], to obtain the conclusion of the lemma. In the latter case, a contradiction ensues as follows: Using the same Vitali’s covering argument as in the proof of Lemma 9.8 we can select a sequence of intervals \( J_n' = [s_n - \lambda(s_n)^{-1}, s_n + \lambda(s_n)^{-1}] \), \( s_n \in J_n \), and such that

$$|J_n'|^{-1} \int_{J_n'} \int_{\mathbb{R}^4} \left| \Phi_t^\infty + iA_0^\infty \Phi_t^\infty \right|^2 \ dx ds + \sum_k (\partial_k A_k^\infty - \partial_k A_0^\infty)^2 \ dx ds \to 0.$$

But then using compactness one again extracts a trivial solution, and obtains a contradiction as in the proof of Lemma 9.8 \( \Box \)

An immediate consequence of the preceding is the following

Corollary 9.11. Under the hypotheses of the preceding lemma, the family of functions indexed by \( t \in [0, 1) \) and given by

$$\left\{ (1 - t)^2 \nabla_{t,x} A_0^\infty(t, (1 - t)x), (1 - t)^2 \nabla_{t,x} \Phi^\infty(t, (1 - t)x) \right\}$$

is compact in \( (L^2(\mathbb{R}^4))^2 \).
We have finally reduced the case of a finite time minimal blowup solution to an “essentially self-similar” scenario, which will be treated via the computations performed in Section 8.4. Since the limiting object \((\mathcal{A}^\infty, \Phi^\infty)\) is only of minimal regularity, however, we have to be careful to ensure that all quantities are well-defined; this is non trivial on account of the singular weight \(\rho(y) = (1 - |y|^2)^{-\frac{1}{2}}\), see the comments at the beginning of Section 8.4. Correspondingly, we introduce the slightly shifted coordinate change

\[ y = \frac{x}{1 - t + \delta}, \quad s = -\log(1 - t + \delta) \]

and correspondingly we have the transformed variables

\[ \tilde{\mathcal{A}}^\infty_v := e^{-s} \mathcal{A}^\infty_v(1 + \delta - e^{-s}, ye^{-s}), \quad \tilde{\Phi}^\infty := e^{-s} \Phi^\infty(1 + \delta - e^{-s}, ye^{-s}). \]

Then the usual limiting argument leads to the following conclusion

**Lemma 9.12.** The gauge invariant version of (8.24) is valid for \((\tilde{\mathcal{A}}^\infty, \tilde{\Phi}^\infty)\).

Proceeding in analogy to [10], we state the

**Proposition 9.13.** Denote the gauge invariant version of the differentiated expression on the left of (8.24) by \(E(\tilde{\mathcal{A}}, \tilde{\Phi})(s)\). Also, denote the gauge invariant version of the right hand side \(\Xi(\tilde{\mathcal{A}}, \tilde{\Phi})(s)\). Then we have the identity

\[ E(\mathcal{A}^\infty, \Phi^\infty)(s_2) - E(\mathcal{A}^\infty, \Phi^\infty)(s_1) = \int_{s_1}^{s_2} \Xi(\mathcal{A}^\infty, \Phi^\infty)(\tilde{s}) \, d\tilde{s}, \quad 0 < s_1 < s_2 < \log(\delta^{-1}). \]

Moreover, we have

\[ \limsup_{s \to \log(\delta^{-1})} E(\tilde{\mathcal{A}}^\infty, \tilde{\Phi}^\infty)(s) \leq E_{\text{crit}}. \]

The first relation is simply the conclusion of the preceding lemma, and the last limiting relation follows from straightforward computation. Keeping in mind the implicit \(\delta\)-dependence of \((\tilde{\mathcal{A}}^\infty, \tilde{\Phi}^\infty)\), we can then infer the following

**Lemma 9.14.** There exists \(\tilde{s}_0 \in \left(\frac{\log \delta}{2}, \log \delta\right)\), such that

\[ \int_{\tilde{s}_0}^{\tilde{s}_0 + \log \delta^2} (\tilde{\mathcal{A}}^\infty, \tilde{\Phi}^\infty)(s) \, ds \leq \frac{E_{\text{crit}}}{|\log \delta|^2}. \]

At this point, we can follow the reasoning in [10] after Lemma 10.15 there. Using the compactness property from Theorem 7.26 with \(\lambda(t) = (1 - t)^{-1}\), and passing to the Cronstrom gauge for \((\mathcal{A}^\infty, \Phi^\infty)\) (which does not change the energy regularity, and also preserves compactness), we can select a limiting solution

\[(\mathcal{A}^*, \Phi^*)\]

which, when translated into the \((s, y)\) coordinates, results in a \((\tilde{\mathcal{A}}^*, \tilde{\Phi}^*)\) with

\[ (\nabla_{s,y} \tilde{\mathcal{A}}^*, \nabla_{s,y} \tilde{\Phi}^*) \in C^0([0, S), L^2_2(\mathbb{R}^4)) \]

for some \(S > 0\) and such that \((\tilde{\mathcal{A}}^*, \tilde{\Phi}^*)\) weakly solves (8.21), (8.19), and furthermore satisfies the crucial self-similarity property

\[ \partial_s \tilde{\Phi}^* + i\tilde{\mathcal{A}}_0^* \tilde{\Phi}^* = 0, \quad \partial_k \tilde{\mathcal{A}}_0^* - \partial_s \tilde{\mathcal{A}}_k^* = 0 \]

Also, we recall that \((\tilde{\mathcal{A}}^*, \tilde{\Phi}^*)\) is in the Cronstrom gauge, that \((\partial_a \tilde{\mathcal{A}}_0^* - \partial_b \tilde{\mathcal{A}}_a^*, \tilde{\Phi}^*)\) are all supported in \(|y| \leq 1\). Then we have
Proposition 9.15. We necessarily have $\Phi^* = 0$, whence going back to the $(t, x)$ coordinates and passing into the Coulomb gauge, we have that $\mathcal{A}'_t$ are free waves, while $\mathcal{A}'_0 = 0$.

The preceding proposition then contradicts Proposition 6.1, which completes the proof of Proposition 9.7.

Proof of Proposition 9.15 To simplify the notation, we write $(\tilde{A}, \tilde{\phi})$ instead of $(\tilde{A}^*, \tilde{\Phi}^*)$. Thus we assume

$$\partial_t \phi + i \tilde{A}_0 \phi = 0, \quad \partial_t \tilde{A}_0 - \partial_j \tilde{A}_k = 0.$$ 

Recall that this solution is supported in $|y| \leq 1$. Since we are in the Cronstrom gauge, the latter relation implies that

$$\partial_r \tilde{A}_0 = 0.$$ 

We now need to rule out a nontrivial energy-class solution of this type. Note that the $\tilde{\phi}$-equation in our present situation can be written in the form

$$(y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1) \tilde{\phi} = \sum_{k=1}^{4} (\partial_k + i \tilde{A}_k)^2 \tilde{\phi},$$

since, as just observed, we have $(y \cdot \nabla_y) \tilde{A}_0 = 0$. We split this into real and imaginary parts:

$$(y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1)(\text{Re} \, \tilde{\phi}) = \sum_{k=1}^{4} (\partial_k^2 - \tilde{A}_k^2) \text{Re} \, \tilde{\phi} - \partial_k (\tilde{A}_k \text{Im} \, \tilde{\phi}) - \tilde{A}_k \partial_k \text{Im} \, \tilde{\phi},$$

$$(y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1)(\text{Im} \, \tilde{\phi}) = \sum_{k=1}^{4} (\partial_k^2 - \tilde{A}_k^2) \text{Im} \, \tilde{\phi} + \tilde{A}_k (\tilde{A}_k \text{Re} \, \tilde{\phi}) + \tilde{A}_k \partial_k \text{Re} \, \tilde{\phi}.$$ 

Integrating against $\text{Re} \, \tilde{\phi}$, respectively $\text{Im} \, \tilde{\phi}$, we obtain

$$\int (y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1)(\text{Re} \, \tilde{\phi}) \text{Re} \, \tilde{\phi} \, dy + \int (y \cdot \nabla_y + 2)(y \cdot \nabla_y + 1)(\text{Im} \, \tilde{\phi}) \text{Im} \, \tilde{\phi} \, dy$$

$$= - \sum_{k=1}^{4} \int [\partial_k \text{Re} \, \tilde{\phi}]^2 + (\partial_k \text{Im} \, \tilde{\phi})^2 + \tilde{A}_k^2 [\text{Re} \, \tilde{\phi}]^2 + \text{Im} \, \tilde{\phi}^2] \, dy$$

$$+ \sum_{k=1}^{4} \int [2 \tilde{A}_k \partial_k \text{Re} \, \tilde{\phi} \text{Im} \, \tilde{\phi} - 2 \tilde{A}_k \text{Re} \, \tilde{\phi} \partial_k \text{Im} \, \tilde{\phi}] \, dy.$$ 

Now observe that in the last integral, due to the Cronstrom condition $\sum_{k=1}^{4} \gamma_k \tilde{A}_k = 0$, we may project the vectors $(\partial_k \text{Re} \, \tilde{\phi})_{k=1}^{4}, (\partial_k \text{Im} \, \tilde{\phi})_{k=1}^{4}$ onto their “angular part”. In particular, it follows that

$$- \sum_{k=1}^{4} \int [\partial_k \text{Re} \, \tilde{\phi}]^2 + (\partial_k \text{Im} \, \tilde{\phi})^2 + \tilde{A}_k^2 [\text{Re} \, \tilde{\phi}]^2 + \text{Im} \, \tilde{\phi}^2] \, dy$$

$$+ \sum_{k=1}^{4} \int [2 \tilde{A}_k \partial_k \text{Re} \, \tilde{\phi} \text{Im} \, \tilde{\phi} - 2 \tilde{A}_k \text{Re} \, \tilde{\phi} \partial_k \text{Im} \, \tilde{\phi}] \, dy$$

$$\leq - \int |\partial_r \tilde{\phi}|^2 \, dy.$$
Finally, consider the left hand side of the above equation. We get
\[
\int (y \cdot \nabla y + 2) (y \cdot \nabla y + 1) (\text{Re} \, \tilde{\phi}) \text{Re} \, \tilde{\phi} \, dy
\]
\[
= - \int (y \cdot \nabla y + 1) (\text{Re} \, \tilde{\phi}) (y \cdot \nabla y + 2) \text{Re} \, \tilde{\phi} \, dy
\]
\[
= - \int ((y \cdot \nabla y) (\text{Re} \, \tilde{\phi}))^2 \, dy - \int ((2 + 3 y \cdot \nabla y) \text{Re} \, \tilde{\phi}) \text{Re} \, \tilde{\phi} \, dy
\]
\[
= 4 \int \text{Re} \, \tilde{\phi}^2 \, dy - \int ((y \cdot \nabla y) (\text{Re} \, \tilde{\phi}))^2 \, dy
\]
with a similar result for the integral involving \( \text{Im} \, \tilde{\phi} \). Combining this with the preceding inequality, we infer the inequality
\[
4 \int |\tilde{\phi}|^2 \, dy \leq -\left( \int |\partial_r \tilde{\phi}|^2 \, dy \right. - \left. \int |(y \cdot \nabla y) \tilde{\phi}|^2 \, dy \right),
\]
which in light of the fact that the solution is supported on the set \(|y| \leq 1\) immediately implies \( \tilde{\phi} = 0 \).

To conclude the program of this paper, we finally need to reduce to the assumption \( \lambda(t) > \lambda_0 > 0 \) made in the statement of Proposition 9.7. However, this follows by exactly the same argument as the one for Lemma 10.18 in [10].

**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.
2. 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.
3. ______, *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.
4. Tosio Kato, Marius Mitrea, Gustavo Ponce, and Michael Taylor, *Extension and representation of divergence-free vector fields on bounded domains*, Math. Res. Lett. 7 (2000), no. 5-6, 643–650.
5. Carlos E. Kenig and Frank Merle, *Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation*, Acta Math. 201 (2008), no. 2, 147–212.
6. S. Klainerman and M. Machedon, *On the Maxwell-Klein-Gordon equation with finite energy*, Duke Math. J. 74 (1994), no. 1, 19–44.
7. ______, *Finite energy solutions of the Yang-Mills equations in \( \mathbb{R}^{3+1} \)*, Ann. of Math. (2) 142 (1995), no. 1, 39–119.
8. Sergiu Klainerman and Daniel Tataru, *On the optimal local regularity for Yang-Mills equations in \( \mathbb{R}^{4+1} \)*, J. Amer. Math. Soc. 12 (1999), no. 1, 93–116.
9. Joachim Krieger, *Global regularity of wave maps from \( \mathbb{R}^{3+1} \) to \( H^2 \)*. Small energy, Comm. Math. Phys. 250 (2004), no. 3, 507–580.
10. Joachim Krieger and Wilhelm Schlag, *Concentration compactness for critical wave maps*, EMS Monographs in Mathematics, European Mathematical Society (EMS), Zürich, 2012.
11. Joachim Krieger and Jacob Sterbenz, *Global regularity for the Yang-Mills equations on high dimensional Minkowski space*, Mem. Amer. Math. Soc. 223 (2013), no. 1047, vi+99.
12. Joachim Krieger, Jacob Sterbenz, and Daniel Tataru, *Global well-posedness for the Maxwell-Klein-Gordon equation in 4+1 dimensions. Small energy*, arXiv:1211.3527, to appear Duke Math. J.
13. Matei Machedon and Jacob Sterbenz, *Almost optimal local well-posedness for the (3 + 1)-dimensional Maxwell-Klein-Gordon equations*, J. Amer. Math. Soc. 17 (2004), no. 2, 297–359 (electronic).
14. Sung-Jin Oh, *Finite Energy Global Well-posedness of the Yang-Mills equations on \( \mathbb{R}^{1+3} \): An Approach Using the Yang-Mills Heat Flow*, arXiv:1210.1557, to appear Duke Math. J.
15. ______, *Gauge choice for the Yang-Mills equations using the Yang-Mills heat flow and local well-posedness in \( H^1 \)*, J. Hyperbolic Differ. Equ. 11 (2014), no. 1, 1–108.
16. Sung-Jin Oh and Daniel Tataru, *Energy dispersed solutions for the (4 + 1)-dimensional Maxwell-Klein-Gordon equation*, arXiv:1503.01561.

17. Sung-Jin Oh, *Global well-posedness and scattering of the (4 + 1)-dimensional Maxwell-Klein-Gordon equation*, arXiv:1503.01562.

18. Sung-Jin Oh, *Local well-posedness of the (4 + 1)-dimensional Maxwell-Klein-Gordon equation at energy regularity*, arXiv:1503.01560.

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.

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.

21. Jacob Sterbenz and Daniel Tataru, *Energy dispersed large data wave maps in 2 + 1 dimensions*, Comm. Math. Phys. 298 (2010), no. 1, 1, 139–230.

22. Terence Tao, *Regularity of wave-maps in dimension 2 + 1*, Comm. Math. Phys. 298 (2010), no. 1, 231–264.

23. Terence Tao, *Global regularity of wave maps. III – VII*, Preprints.

24. Terence Tao, *Global regularity of wave maps. I. Small critical Sobolev norm in high dimension*, Internat. Math. Res. Notices (2001), no. 6, 299–328.

25. Terence Tao, *Global regularity of wave maps. II. Small energy in two dimensions*, Comm. Math. Phys. 224 (2001), no. 2, 443–544.

26. Daniel Tataru, *On global existence and scattering for the wave maps equation*, Amer. J. Math. 123 (2001), no. 1, 37–77.

27. Daniel Tataru, *Rough solutions for the wave maps equation*, Amer. J. Math. 127 (2005), no. 2, 293–377.