



Quasilinear Schrödinger Equations III: Large Data and Short Time

JEREMY L. MARZUOLA , JASON METCALFE & DANIEL TATARU

Communicated by J. BEDROSSIAN

Abstract

In this article we prove short time local well-posedness in low-regularity Sobolev spaces for large data general quasilinear Schrödinger equations with a nontrapping assumption. These results represent improvements over the small data regime considered by the authors in MARZUOLA et al. (Adv Math 231(2):1151–1172, 2012; Kyoto J Math 54(3):529–546, 2014), as well as the pioneering works by KENIG et al. (Invent Math 158:343–388, 2004; Adv Math 196(2), 402–433, 2005; Adv Math 206(2):373–486, 2006), where viscosity methods were used to prove the existence of solutions for localized data in high regularity spaces. Our arguments here are purely dispersive. The function spaces in which we show existence are constructed in ways motivated by the results of MIZOHATA, ICHINOSE, DOI, and others, including the authors.

1. Recap of History and Discussion of Main Results

In this article we consider the large data local well-posedness for quasilinear Schrödinger equations, extending the earlier small data results of the authors in [26, 27]. Specifically, we will study equations of the form

$$\begin{cases} iu_t + g^{jk}(u, \bar{u}, \nabla u, \nabla \bar{u}) \partial_j \partial_k u = F(u, \bar{u}, \nabla u, \nabla \bar{u}), & u: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{C}^m \\ u(0, x) = u_0(x). \end{cases} \quad (1.1)$$

Here, g and F are assumed to be smooth functions of their variables, with g real and positive definite. In particular this allows for them to depend on both u and $\bar{u}$.

In the small initial data case these equations have already been considered by the authors [26, 27] in spaces of relatively low Sobolev regularity. The latter paper considers problems which have only cubic and higher nonlinearities, where the initial data is in Sobolev spaces H^s . The former paper allows for general nonlinearities,

i.e. including quadratic terms, but with smaller Sobolev spaces $l^1 H^s$, which are still translation invariant but have some stronger summability assumptions. These will be made precise in the discussion below.

The aim of this paper is to prove instead a local well-posedness result for the large data problem. Compared with the small data case, here we need to contend with an additional obstacle, namely trapping. To prevent this, we impose a nontrapping condition on the initial data. Then, as a part of our results, we prove that nontrapping persists for some small time. In this context, the lifespan of the solutions no longer depends only on the data size. Instead, our lower bound on the lifespan will also depend on a quantitative form of the nontrapping assumption.

Here we will work primarily with quadratic nonlinear interactions, as in [26]. We will also state the counterpart of the result in the cubic case, as considered for small data in [27]; however, as the proofs only differ slightly we will focus on the quadratic setting and only remark where the proofs need adjustment for the cubic interactions. Specifically, we will study the equations (1.1) assuming that

$$\begin{aligned} g: \mathbb{C}^m \times \mathbb{C}^m \times (\mathbb{C}^m)^d \times (\mathbb{C}^m)^d &\rightarrow \mathbb{R}^{d \times d}, \\ F: \mathbb{C}^m \times \mathbb{C}^m \times (\mathbb{C}^m)^d \times (\mathbb{C}^m)^d &\rightarrow \mathbb{C}^m \end{aligned}$$

are smooth functions which, for $y, z \in \mathbb{C}^m \times (\mathbb{C}^m)^d$, satisfy

$$g(0) = I_d, \quad |F(y, z)| \sim O(|y|^2 + |z|^2) \text{ near } (y, z) = (0, 0) \quad (1.2)$$

in the quadratic interaction problem and

$$g(y, z) = I_d + O(|y|^2 + |z|^2), \quad F(y, z) = O(|y|^3 + |z|^3) \text{ near } (y, z) = (0, 0), \quad (1.3)$$

in the cubic interaction problem. We will also assume uniform ellipticity of g in both cases. Namely, we will assume that

$$c_0 |\xi|^2 \leq g^{jk} \xi_j \xi_k \leq c_0^{-1} |\xi|^2$$

for a fixed $c_0 > 0$.

As in [26, 27], we also consider a second class of quasilinear Schrödinger equations

$$\begin{cases} iu_t + \partial_j g^{jk}(u, \bar{u}) \partial_k u = F(u, \bar{u}, \nabla u, \nabla \bar{u}), & u: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{C}^m \\ u(0, x) = u_0(x), \end{cases} \quad (1.4)$$

with g and F as in (1.2), but where the metric g depends on u but not on ∇u . Such an equation is obtained for instance by differentiating the first equation (1.1). Precisely, if u solves (1.1) then the vector $(u, \nabla u)$ solves an equation of the form (1.4), with a nonlinearity F which depends at most quadratically on ∇u .

We note that the second order operator in (1.4) is written in divergence form. This is easily achieved by commuting the first derivative with g and moving the output to the right hand side. However, the second order operator in (1.1) cannot be written in divergence form without possibly changing the type of the equation.

The proof of the large data result presented here follows the same strategy as in the works [26,27]. The main novelty in this paper is in the proof of the local energy decay estimates for the linearized equation. The difficulty is not only that we need to work with a large nontrapping metric, but also that, in order to prevent a nonlinear energy cascade to high frequencies, we have to produce a very accurate bound for the (exponentially large) constant in the local energy bounds in terms of our quantitative nontrapping parameters. Our proof requires a new multiplier construction for the estimates since here we assume no quantitative decay of the solution in the physical space. Also, a careful set-up and ordering of large constants is essential in order to avoid a circular argument.

Even at the linear level, an obstruction to well-posedness comes from the infinite speed of propagation phenomena. We recall some of the issues here as it is even more relevant for the large data problem. From [13,24,29–31,36] it is known that even in the case of linear problems of the form

$$(i\partial_t + \Delta_g)v = A_i(x)\partial_i v, \quad (1.5)$$

a necessary condition for L^2 well-posedness is an integrability condition for the (imaginary part of) the magnetic potential A along the Hamilton flow of the leading order differential operator. In the case of (1.1), we would have to look instead at the corresponding linearized problem, which would exhibit a magnetic potential of the form $A = A(u, \nabla u)$. If one considers equations with quadratic terms and with H^s initial data, then such a potential does not generically satisfy Mizohata's integrability condition. Thus, some further decay condition on the initial data is necessary. A further motivation for decay conditions comes from the large data problem, where one seeks to confine the trapping analysis to a compact set. Together, these two observations show that in our context nontrapping is a compact phenomena, and also that it is stable with respect to small perturbations of u_0 .

Indeed, such a decay condition was manifest in the seminal papers [17–19], where the first local well-posedness results for this problem were obtained. There, local well-posedness results for this problem were proved for data (and solutions) in $H^s \cap L^2(\langle x \rangle^N)$, where $\langle x \rangle = (1 + |x|^2)^{\frac{1}{2}}$, for some large s and N .

By contrast, our previous results in [26,27] apply for initial data in spaces which are not only low regularity, but also translation invariant. Precisely, for the quadratic problem (1.2) we use the $l^1 H^s$ spaces (see a precise definition below), while for the cubic problem (1.3) we revert to the classical H^s spaces. Maintaining this natural setting is one of our objectives for the large data problem.

We now turn our attention to the nontrapping condition for the initial data. This is defined in a qualitative manner in terms of the Hamilton flow associated with $g(u_0)$ (or equivalently, the geodesic flow associated to the Riemannian metric $g(u_0)$):

Definition 1.1. We say that the metric $g(u_0)$ is nontrapping if all nontrivial bicharacteristics for $\Delta_{g(u_0)}$ escape to spatial infinity at both ends.

To motivate the fact that this definition is meaningful for u_0 in our initial data spaces we make several observations. Firstly, our choice for the initial data space

guarantees that $g(u_0) \in C^2$, therefore its Hamilton flow is well defined locally. Secondly, while u_0 may be large, the fact that we are using L^2 based spaces implies that u_0 is small in our function spaces outside a large compact set. Thus trapping is necessarily confined to bicharacteristics which intersect this compact set.

The above qualitative definition of trapping suffices in order to state our main results. However, in order to prove the results, as well as to provide a lifespan bound, we will have to use a more precise quantitative characterization of nontrapping.

Before stating our main results, we briefly recall the definition of the $l^1 H^s$ spaces, following [26]. These are defined using a standard spatial Littlewood–Paley decomposition

$$1 = \sum_{k \in \mathbb{N}} S_k,$$

where S_0 selects all frequencies of size $\lesssim 1$. Corresponding to each dyadic frequency scale $2^j \geq 1$ we consider an associated partition $\mathcal{Q}_j$ of $\mathbb{R}^d$ into cubes of side length 2^j and an associated smooth partition of unity

$$1 = \sum_{Q \in \mathcal{Q}_j} \chi_Q.$$

Then we can define the $l_j^1 L^2$ norm by

$$\|u\|_{l_j^1 L^2} = \sum_{Q \in \mathcal{Q}_j} \|\chi_Q u\|_{L^2},$$

and the space $l^1 H^s$ with norm given by

$$\|u\|_{l^1 H^s}^2 = \sum_{j \geq 0} 2^{2sj} \|S_j u\|_{l_j^1 L^2}^2.$$

With our spaces in hand, we can now state our main result concerning the quasilinear problem (1.1) with data $u_0(x) \in l^1 H^s$ and quadratic interactions as in (1.2).

Theorem 1.2. (a) *Let $s > \frac{d}{2} + 3$. Let $u_0 \in l^1 H^s$ be a nontrapping initial datum for the Eq. (1.1) with quadratic interactions (1.2). Then, there exists $T = T(u_0) > 0$ sufficiently small such that the Eq. (1.1) is locally well-posed in $l^1 H^s(\mathbb{R}^d)$ on the time interval $I = [0, T]$.*
 (b) *The same result holds for the Eq. (1.4) with $s > \frac{d}{2} + 2$.*

The counterpart of this result for the cubic problem (1.3) is as follows using standard Sobolev spaces, which extends [28] to the low regularity regime:

Theorem 1.3. (a) *Let $s > \frac{d+5}{2}$. Let $u_0 \in H^s$ be a nontrapping initial datum for the Eq. (1.1) with cubic interactions (1.3). Then, there exists $T = T(u_0) > 0$ sufficiently small such that the Eq. (1.1) is locally well-posed in $l^2 H^s(\mathbb{R}^d)$ on the time interval $I = [0, T]$.*
 (b) *The same result holds for the Eq. (1.4) with cubic nonlinear interactions with $s > \frac{d+3}{2}$.*

Remark 1.3.1. The well-posedness result in the statement of the theorems above has to be interpreted in the classical quasilinear fashion. Precisely, in the setting of Theorem 1.2, it asserts that

- (i) (Regular solutions) For large σ and regular nontrapping data $u_0 \in l^1 H^\sigma$ there exists a unique local nontrapping solution $u \in C(l^1 H^\sigma)$ on a nonempty maximal time interval $I = [0, T_{\max}(u_0))$.
- (ii) (Rough solutions) For s as in the theorem and nontrapping data $u_0 \in l^1 H^s$ there exists a unique local nontrapping solution $u \in C(l^1 H^s) \cap l^1 X^s$ on a nonempty maximal time interval $I = [0, T_{\max}(u_0))$. The spaces $l^1 X^s$, which capture the space-time local energy decay structure, will be described in the next section.
- (iii) (Continuous dependence) The maximal time $T_{\max}(u_0)$ is a lower semicontinuous function of u_0 in the $l^1 H^s$ topology, and the data to solution map $v_0 \rightarrow v$ is continuous at u_0 from $l^1 H^s$ into $C([0, T]; l^1 H^s) \cap l^1 X^s([0, T])$ for all $T < T_{\max}(u_0)$.

The latter property allows one to alternately uniquely identify rough solutions as limits of regular solutions.

Remark 1.3.2. The existence time T in the theorem is allowed to depend on the full profile of the initial data u_0 , and not only on its size. This dependence will be made more clear later on. The difficulty is that the well-posedness depends not only on the size of the data, but also on the nontrapping property of the metric. Later we will introduce a more explicit parameter $L = L(u_0)$, which quantitatively measures nontrapping, and is stable with respect to small perturbations of u_0 . Then we will show that it suffices to choose $T \ll 1$ so that

$$|\log T| \gg \|u_0\|_{l^1 H^s} L, \quad (1.6)$$

where the implicit dependence on the data size is polynomial.

Remark 1.3.3. In order to define nontrapping, we require a well-defined Hamilton flow. A sufficient condition which guarantees this is to have a $C^{1,1}$ metric, and that is satisfied in the context of Theorem 1.2. However, this is not guaranteed in the case when $s < \frac{d}{2} + 2$, in Theorem 1.3. Nevertheless, we can instead guarantee $\nabla^2 g \in L^2(\mathbb{R}_{x_1}, L^\infty(\mathbb{R}_{x'}^{d-1}))$, as well as in any other rotated frame; this still suffices in order to define the flow as a bi-Lipschitz map. In turn, implementing this would require appropriate changes in Section 6; we omit these, and instead refer the interested reader to [4] where a similar analysis was conducted.

We further remark that the short time large data result cannot be obtained by scaling from the small data result. This is due to the fact that the spaces used are inhomogeneous Sobolev spaces, and spatial localization is not allowed due to the infinite speed of propagation. A reflection of this is the fact that the small data case is nontrapping, while in the large data regime, one must also take into account the existence of trapping.

The paper is organized as follows. In Section 2 we describe the space-time function spaces in which we will solve (1.1) and (1.4); these are identical to those

in the small data setting in [26, 27]. In Section 3 we introduce some key notations, including the main size parameters which govern our lifespan bound, and give an overview of the proof. Section 4 contains the necessary multilinear and nonlinear estimates in order to close the eventual bootstrap estimates; some of these are from [26, 27], but the main bound for the paradifferential error term is new. The stability of the nontrapping assumption will be discussed in Section 5; this is critical in order to propagate nontrapping to positive times. In Section 6 we establish local energy decay for a linear, nontrapping, inhomogeneous paradifferential version of the Schrödinger equation. Finally, in Section 7, we combine the above estimates with the proper paradifferential decomposition of the equation in order to conclude the proof of Theorem 1.2.

2. Recap of Function Spaces and Notations

In this section we recall the definition of the main function spaces as well as some of their key properties. For this we follow our previous works [26, 27].

We will use an inhomogeneous Littlewood–Paley decomposition $\sum_{k \in \mathbb{N}} S_k = 1$. We set $u_j = S_j u$ and

$$S_{\leq N} f = \sum_{i=0}^N S_i f, \quad S_{\geq N} f = \sum_{i=N}^{\infty} S_i f, \quad S_{[N_1, N_2]} f = \sum_{i=N_1}^{N_2} S_i f.$$

When it is clear from the context, we may abuse notation and use $u_0 = S_0 u$.

Given a translation invariant Sobolev-type space U , we define $l_j^p U$ via

$$\|u\|_{l_j^p U}^p = \sum_{Q \in \mathcal{Q}_j} \|X_Q u\|_U^p,$$

which generalizes the notion of $l_j^1 L^2$ defined in the introduction. We make the natural modification when $p = \infty$. Upon replacing the sum over the cubes with an integral, the existence of translation invariant norms that are equivalent can easily be checked. As noted previously in [26], the smooth partition of compactly supported cutoffs in the $l_j^1 U$ spaces can be replaced by cutoffs which are frequency localized when it is convenient.

We recall briefly the local energy type space X of functions on $[0, T] \times \mathbb{R}^d$, with norm

$$\|u\|_X = \sup_l \sup_{Q \in \mathcal{Q}_l} 2^{-\frac{l}{2}} \|u\|_{L_{t,x}^2([0, T] \times Q)}.$$

Note, these spaces are dependent upon T , and hence the dependence upon the time interval in the estimates below will be treated with some care.

We define $Y \subset L_{t,x}^2([0, T] \times \mathbb{R}^d)$ to satisfy $X = Y^*$. This will be the space in which we measure the forcing terms for the Schrödinger equation. See, e.g., [5], [26] for more details on the construction of this atomic space.

We set

$$X_j = 2^{-\frac{j}{2}} X \cap L^\infty L^2,$$

which will incorporate the half-degree of smoothing into our local energy spaces. We then add the l^p spatial summation on the 2^j scale to obtain the space $l_j^p X_j$ with norm

$$\|u\|_{l_j^p X_j}^p = \sum_{Q \in \mathcal{Q}_j} \|\chi_Q u\|_{X_j}^p.$$

We finally define the spaces $l^p X^s$ by

$$\|u\|_{l^p X^s}^2 = \sum_j 2^{2js} \|S_j u\|_{l_j^p X_j}^2. \quad (2.1)$$

For quadratic interactions, we shall use $l^1 X^s$ to bound the high frequencies exterior to a large ball in our solutions to (1.1), (1.4) with nontrapping $l^1 H^s$ data. For cubic interactions, we shall use $p = 2$ and data in $l^2 H^s \approx H^s$.

For the inhomogeneous terms at frequency 2^j , we shall use

$$Y_j = 2^{\frac{j}{2}} Y + L^1 L^2$$

which has norm

$$\|f\|_{Y_j} = \inf_{f=2^{\frac{j}{2}} f_1 + f_2} \|f_1\|_Y + \|f_2\|_{L^1 L^2}.$$

We then similarly consider

$$\|f\|_{l^p Y^s}^2 = \sum_j 2^{2js} \|S_j f\|_{l_j^p Y_j}^2, \quad (2.2)$$

where $p = 1$ will be utilized for the case of quadratic interactions and $p = 2$ for the cubic case.

We also record the spaces X^s without the summability. These are given by the norm

$$\|u\|_{X^s}^2 = \sum_j 2^{2js} \|S_j u\|_{X_j}^2. \quad (2.3)$$

Similarly, we define Y^s via

$$\|f\|_{Y^s}^2 = \sum_j 2^{2js} \|S_j f\|_{Y_j}^2. \quad (2.4)$$

In the regimes where we can apply paradifferential analysis, it is convenient to present our bilinear and nonlinear estimates using the method of frequency envelopes, which we recall below. For a Sobolev-type space U so that

$$\|u\|_U^2 \sim \sum_{k=0}^{\infty} \|S_k u\|_U^2$$

a positive sequence c_j is called an admissible frequency envelope for u in U provided that it

(1) controls the dyadic U size,

$$\|S_k u\|_U \lesssim c_k,$$

(2) is controlled by the U norm,

$$\sum_{k \in \mathbb{N}} c_k^2 \lesssim \|u\|_U^2$$

(3) is slowly varying to the left,

$$c_j \geq 2^{\delta(j-k)} c_k, \quad j < k,$$

and

(4) is uniformly varying to the right,

$$c_j \geq 2^{\sigma(k-j)} c_k, \quad j > k$$

for a fixed (large) σ .

These properties are easily adapted to the case when the l^2 dyadic summability is replaced by l^p with $1 \leq p < \infty$.

An admissible frequency envelope always exists, say by

$$c_j = \max_{k > j} 2^{-\delta|j-k|} \|S_k u\|_U + \max_{k < j} 2^{-\sigma|j-k|} \|S_k u\|_U. \quad (2.5)$$

In the sequel we will use frequency envelopes for the spaces $l^p H^s$, $l^p X^s$ and $l^p Y^s$ for $p = 1, 2$.

3. Outline of the Proof

Let us briefly outline the ideas we will pursue below for the case of quadratic interactions and the Eq. (1.4). The cubic case will follow similarly. We seek to solve the equation

$$\begin{cases} iu_t + \partial_j g^{jk}(u, \bar{u}) \partial_k u = F(u, \bar{u}, \nabla u, \nabla \bar{u}), & u: \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{C}^m, \\ u(0, x) = u_0(x), \end{cases} \quad (3.1)$$

in $l^1 H^s$ for $s > s_0 > \frac{d}{2} + 2$. Here we have explicitly included the dependence upon u and $\bar{u}$ separately since the large data dynamics will depend upon them both more delicately than in the small data case.

1. *The linearized and paradiifferential equation.* An important role in the analysis will be played by the linearized equation, which has the form

$$\begin{cases} i v_t + \partial_j g^{jk} \partial_k v + b^j \partial_j v + \tilde{b}^j \partial_j \bar{v} + c v + \tilde{c} \bar{v} = 0, \\ v(0, x) = v_0(x), \end{cases} \quad (3.2)$$

where the coefficients g , respectively b^j , $\tilde{b}^j$, c , $\tilde{c}$, are smooth nonlinear expressions in u , respectively u , ∇u , which can be explicitly calculated in terms of g and F . In particular, we have

$$\begin{aligned} b^j &= \partial_u g^{jk} \partial_k u - \partial_{(\nabla u)_j} F, & \tilde{b}^j &= \partial_{\bar{u}} g^{jk} \partial_k u - \partial_{(\nabla \bar{u})_j} F, \\ c &= \partial_j (\partial_u g^{jk}) \partial_k u - \partial_u F, & \tilde{c} &= \partial_j (\partial_{\bar{u}} g^{jk}) \partial_k u - \partial_{\bar{u}} F \end{aligned} \quad (3.3)$$

for a fixed $u \in l^1 X^s$.

From the linearized equation we extract its associated linear paradifferential flow

$$\begin{cases} i \partial_t w + \partial_j T_{g^{jk}} \partial_k w + T_{bj} \partial_j w + T_{\tilde{b}^j} \partial_j \bar{w} = \tilde{f}, \\ w(0, x) = w_0(x), \end{cases} \quad (3.4)$$

where we take the paraproduct operator to be

$$T_a b = \sum_{N \geq 4} S_{\leq N-4} a S_N b. \quad (3.5)$$

For some of the analysis it will be more convenient to use the Weyl quantization for the paraproduct, which is denoted as follows:

$$T_a^w b = OP^w \left(\sum_{N \geq 4} S_{\leq N-4} a(x) s_N(\xi) \right).$$

If v solves (3.2) then it also solves (3.4) with

$$\begin{aligned} \tilde{f} &= (\partial_j T_{g^{jk}}(u, \bar{u}) \partial_k w - \partial_j g^{jk} \partial_k w) + (T_{bj} \partial_j w - b^j \partial_j w) \\ &\quad + (T_{\tilde{b}^j} \partial_j \bar{w} - \tilde{b}^j \partial_j \bar{w}) - cw - \tilde{c} \bar{w}. \end{aligned}$$

Here heuristically $\tilde{f}$ contains only high-high frequency interactions; e.g. in the multilinear case, the two highest frequencies must always¹ be balanced. Because of this, its contribution will always be treated perturbatively.

2. Rewriting the equation. Based on the expressions above for the linearized equation and its paradifferential truncation, we write the full nonlinear equation (3.1) in a paradifferential form, namely

$$\begin{cases} i \partial_t u + \partial_j T_{g^{jk}}^w \partial_k u + T_{bj}^w \partial_j u + T_{\tilde{b}^j}^w \partial_j \bar{u} = G, \\ u(0, x) = u_0(x). \end{cases} \quad (3.6)$$

¹ This is not entirely accurate in what we do. Instead, $\tilde{f}$ is also allowed to contain low-high interactions as long as the high frequency factors are undifferentiated; this makes such terms perturbative.

Here the nonlinearity $G = G(u, \bar{u}, \nabla u, \nabla \bar{u})$ is no longer purely algebraic, as it involves frequency localizations. The key idea in our proof of the local well-posedness result is that G plays a perturbative role. In order that we have it, we record that

$$G(u, \bar{u}, \nabla u, \nabla \bar{u}) = F(u, \bar{u}, \nabla u, \nabla \bar{u}) - \partial_j (g^{jk} - T_{g^{jk}}^w) \partial_k u + T_{b_j}^w \partial_j u + T_{\tilde{b}_j}^w \partial_j \bar{u}$$

where b and $\tilde{b}$ contain both contributions arising from g and from F ; see (3.3).

The solutions will be constructed via an iterative scheme, where we set $u^{(0)} = 0$, and successively define $u^{(n+1)}$ as the solution to the linear equation

$$\begin{cases} \left(i \partial_t + \partial_j T_{g^{jk}, (n)}^w \partial_k + T_{b(n)}^w \cdot \nabla \right) u^{(n+1)} + T_{\tilde{b}(n)}^w \cdot \nabla \bar{u}^{(n+1)} = G(u^{(n)}, \nabla u^{(n)}), \\ u^{(n+1)}(0) = u_0 \end{cases} \quad (3.7)$$

where

$$g^{(n)} = g(u^{(n)}), \quad b^{(n)} = b(u^{(n)}, \nabla u^{(n)}), \quad \tilde{b}^{(n)} = \tilde{b}(u^{(n)}, \nabla u^{(n)}).$$

Each of $G, g^{(n)}, b^{(n)}, \tilde{b}^{(n)}$ also depends on the conjugates of the solution, but this is suppressed here. In order to guarantee the convergence of this scheme we will carefully choose time T small enough, depending on the initial data profile.

3. *Quantifying nontrapping: the parameters M, R, L .* For the purpose of characterizing the nontrapping properties of the metric $g(u_0)$ we do not need the full H^s regularity. Instead we will use a smaller exponent s_0 so that

$$\frac{d}{2} + 2 < s_0 < s; \quad (3.8)$$

its choice within these bounds is not important, but we fix it once and for all. The gap between s_0 and s will be critical in order to propagate the nontrapping property.

The first parameter M we use to describe nontrapping measures the size of the data,

$$M = \|u_0\|_{l^1 H^{s_0}}. \quad (3.9)$$

Outside a compact spatial region $B = B(x_0, R)$ where both x_0 and R depend on u_0 , the metric $g_{ij}(u_0)$ will have a small $l^1 H^{s_0}$ norm, and thus be nontrapping. The nontrapping assumption guarantees that all geodesics intersecting B will eventually leave $2B$ at both ends, and never return to B once leaving $2B$. We denote by $L \gtrsim R$ the maximum Euclidean length of any such geodesic within $2B$. The nontrapping condition is then shown to be stable with respect to perturbations, $g \rightarrow g + \delta g$, of the metric which satisfy an exponential smallness condition

$$\|\delta g\|_{l^1 X^{s_0}} \lesssim e^{-C_0(M)L}. \quad (3.10)$$

This is proved in Section 5.

4. *Nontrapping and norm inflation.* In order to carry out the above iteration, we need to consider energy estimates and local energy decay for the linear paradifferential flow (3.4). Precisely, we would like to have bounds of the form

$$\|w\|_{l^1 X^\sigma} \leq C(\|w_0\|_{l^1 H^\sigma} + \|f\|_{l^1 Y^\sigma}), \quad 0 \leq \sigma \quad (3.11)$$

in a time interval $[0, T]$ where T depends only on the initial data w_0 .

However, even if $w_0 = 0$, $\sigma = 0$ and T is arbitrarily small, the energy bounds for this system will exhibit L^2 growth, due both to the large metric in the compact set $B(x_0, R)$ and to the large coefficient b for the first order term. Hence, the best we could hope for is a bound of the type

$$\|w\|_{X^0} \lesssim e^{C(M)L}(\|w_0\|_{L^2} + \|f\|_{Y^0}), \quad (3.12)$$

with the only redeeming feature that only a lower regularity $l^1 X^{s_0}$ bound for u , which occurs in the coefficients, is needed. Naturally, the constant C in (3.11) would have to be at least as large as the exponential in (3.12). Indeed, in Section 6 we establish that the bound (3.12) holds.

The key to handle this exponential growth is to restrict to a very short time interval $[0, T]$, with T satisfying

$$T \lesssim e^{-C(M)L}. \quad (3.13)$$

To balance the choice of T and the exponential growth we divide and conquer. We first prove a high frequency energy estimate via positive commutator methods that control the high frequencies while allowing low frequency errors; this part is independent of the length of the time interval. Then, we are able to use the short time in a more direct fashion to control the contribution from the low frequencies.

5. *Nontrapping and high frequency energy estimates.* The first step in the proof of (3.12) is to use a positive commutator method in order to establish local energy bounds with low frequency errors,

$$\|w\|_{X^0} \lesssim e^{C(M)L}(\|w_0\|_{L^2} + \|f\|_{Y^0} + \|w\|_{L^2 L^2}) \quad (3.14)$$

with no restriction on the time T . This is done in three stages:

- (i) Bounds for incoming rays. Here we estimate the energy along geodesics which approach the compact region $2B$ without any norm inflation by introducing a suitable incoming multiplier Q_{in} .
- (ii) Bounds in a compact set $2B$. This is where we bound the local energy norm of the solution in $2B$ in terms of the incoming part, using the nontrapping condition to construct a suitable multiplier Q_{comp} . This is where the norm inflation occurs.
- (iii) Global bounds. Here we use the local energy estimate in $2B$ in order to produce a global exterior bound, which follows very similarly to the small data metric perturbation theoretic arguments of [26, 27].

6. *Low frequency estimates for short time.* The last step in the proof of (3.12) is to complement the above high frequency bound with an estimate for the low frequencies. This is quite trivial and is obtained by Hölder's inequality in time, which gives

$$\|w\|_{L^2L^2} \lesssim T^{\frac{1}{2}} \|w\|_{L^\infty L^2} \lesssim T^{\frac{1}{2}} \|w\|_{X^0}. \quad (3.15)$$

7. *Uniform bounds for the iteration scheme.* The difficulty we face here is that, in view of the bound (3.12), the best we can expect of the sequence $u^{(n)}$ in $l^1 X^s$ is a bound with exponential growth of the form

$$\|u^{(n)}\|_{l^1 X^s} \lesssim e^{C(M)L} \|u_0\|_{l^1 H^s}. \quad (3.16)$$

Such a bound cannot be directly obtained in a self-contained inductive argument and is predicated on the additional assumption that

$$\|u^{(n)}\|_{l^1 X^{s_0}} \lesssim M.$$

To avoid a circular argument, we will obtain this last bound not directly from (3.16), but rather by interpolating with a lower regularity bound but which has a T factor,

$$\|u^{(n)} - u_0\|_{l^1 L^\infty L^2} \lesssim TM.$$

After interpolation, we will be able to leverage the remaining (small) power of T against the exponential provided that

$$T \ll_{M_s} e^{-C(M)L}, \quad M_s = \|u_0\|_{l^1 H^s}. \quad (3.17)$$

8. *Lipschitz bounds and weak convergence for the iteration scheme.* Here we consider the difference equations for $u^{(n+1)} - u^{(n)}$ and use the bound (3.12) directly to prove the convergence of the iteration scheme in $l^1 X^0$. Given the uniform $l^1 X^s$ bound, we also obtain convergence in all intermediate topologies. The same type of argument also yields Lipschitz dependence of the solutions on the initial data in the weaker topology, and in particular uniqueness.

9. *Frequency envelopes and continuous dependence.* Frequency envelope bounds are only needed at high frequency. They are derived from similar frequency envelope bounds for the paradifferential equation and allow us to (i) propagate higher regularity and (ii) prove continuous dependence on the initial data in the strong topology.

4. Multilinear and Nonlinear Estimates in $l^1 X$ Type Spaces

In this section we recall the main bilinear and nonlinear estimates from [26], and add several related bounds that can be derived from them.

4.1. Bilinear and Moser Estimates

Our aim here is to recall some of the estimates in [26], as well as to provide some improvements adapted to the context of this paper. We begin with bounds in $l^1 X^s$ spaces, where we first recall the dyadic bilinear bounds from [26] (these are contained within the proof of Proposition 3.2 there).

Lemma 4.1. *The following bilinear estimates hold in $l^1_j X_j$ spaces:*

(a) *High-low interactions $j < k - 4$:*

$$\|S_j u S_k v\|_{l^1_k X_k} \lesssim 2^{\frac{jd}{2}} \|u\|_{l^1_j X_j} \|v\|_{l^1_k X_k}, \quad (4.1)$$

(b) *Balanced interactions, $|j - k| \leq 4$:*

$$\|S_i(S_j u S_k v)\|_{l^1_i X_i} \lesssim 2^{dk} 2^{-\frac{di}{2}} \|u\|_{l^1_j X_j} \|v\|_{l^1_k X_k}. \quad (4.2)$$

We note that in case (a) it suffices in effect to bound the low frequency factor in L^∞ , as the $l^1_k X_k$ norms depend only on the pointwise size of functions. For the L^∞ norm, on the other hand, we have the Bernstein-type inequality

$$\|S_j u\|_{L^\infty L^\infty} \lesssim 2^{\frac{jd}{2}} \|u\|_{l^1_j X_j}. \quad (4.3)$$

We continue with a refinement of [26, Prop. 3.1(a)].

Proposition 4.2. (a) *Let $s > \frac{d}{2}$. Then the $l^1 X^s$ spaces satisfy the bilinear estimates*

$$\|uv\|_{l^1 X^\sigma} \lesssim \|u\|_{l^1 X^\sigma} \|v\|_{l^1 X^s} + \|u\|_{l^1 X^s} \|v\|_{l^1 X^\sigma}, \quad d - s \leq \sigma, \quad (4.4)$$

respectively

$$\|uv\|_{l^1 X^\sigma} \lesssim \|u\|_{l^1 X^\sigma} \|v\|_{l^1 X^s}, \quad d - s \leq \sigma \leq s. \quad (4.5)$$

(b) *For all smooth F with $F(0) = F'(0) = F''(0) = 0$ we have the Moser-type estimate*

$$\|F(u)\|_{l^1 X^\sigma} \lesssim \|u\|_{l^1 X^\sigma} \|u\|_{l^1 X^s}^2 c(\|u\|_{L^\infty}), \quad d - s \leq \sigma, \quad (4.6)$$

as well as the difference estimates

$$\|F(u) - F(v)\|_{l^1 X^\sigma} \lesssim \|u - v\|_{l^1 X^\sigma} (\|u\|_{l^1 X^s} + \|v\|_{l^1 X^s})^2 c(\|u\|_{L^\infty}, \|v\|_{L^\infty}), \quad d - s \leq \sigma \leq s. \quad (4.7)$$

Here in (b) we are assuming that $F(u)$ is cubic in u near $u = 0$ just for convenience, as the linear part is uninteresting and the quadratic part is dealt with in (4.4) and (4.5).

Proof. The bilinear estimates (4.4) and (4.5) follow almost directly from [26, (3.1)]. For the high-high frequency interactions, using (4.2) we get

$$\left\| \sum_{\substack{i,j \geq k-4 \\ |i-j| \leq 4}} S_k(S_i u S_j v) \right\|_{l^1 X^\sigma} \lesssim \sum_{i \geq k-4} 2^{(d-s-\sigma)i+k(\sigma-\frac{d}{2})} \|u\|_{l^1 X^\sigma} \|v\|_{l^1 X^s}. \quad (4.8)$$

For these interactions, the bound (4.5) follows provided $\sigma \geq d-s$. And hence (4.4) follows trivially. For the low-high frequency interactions, by (4.1) we have

$$\left\| \sum_{\substack{j < k-4 \\ |i-k| \leq 4}} S_k(S_i u S_j v) \right\|_{l^1 X^\sigma} \lesssim \sum_{j < k-4} 2^{(\frac{d}{2}-s)j} \sum_{|i-k| < 4} \|S_i u\|_{l^1 X^\sigma} \|v\|_{l^1 X^s}.$$

In order to obtain (4.5), and hence (4.4), for these interactions, we only require $s > \frac{d}{2}$. In the remaining case, from (4.1) we instead obtain

$$\left\| \sum_{\substack{i < k-4 \\ |j-k| \leq 4}} S_k(S_i u S_j v) \right\|_{l^1 X^\sigma} \lesssim \sum_{i < k-4} 2^{(\frac{d}{2}-\sigma)i} 2^{k(\sigma-s)} \|u\|_{l^1 X^\sigma} \sum_{|j-k| < 4} \|S_j v\|_{l^1 X^s}.$$

Here, when $\sigma \leq \frac{d}{2}$, we again only require $s > \frac{d}{2}$. When, however, $\sigma > \frac{d}{2}$, in order to obtain (4.5), we additionally require that $\sigma \leq s$. This in turn justifies the need for the symmetric term in (4.4) when such an upper bound on the range of σ is not assumed. In [26], the corresponding estimates were phrased in a more precise way using the concept of frequency envelopes, which one could also do here.

The Moser-type estimate (4.7) is a refinement of [26, (3.2)], which showed that if $s > \frac{d}{2}$ and F is smooth with $F(0) = 0$, then

$$\|F(u)\|_{l^1 X^s} \lesssim \|u\|_{l^1 X^s} (1 + \|u\|_{l^1 X^s}) c(\|u\|_{L^\infty}). \quad (4.9)$$

A trivial improvement of this is obtained in an identical manner if we eliminate the linear part of F , and assume instead that $F(0) = F'(0) = 0$:

$$\|F(u)\|_{l^1 X^s} \lesssim \|u\|_{l^1 X^s}^2 c(\|u\|_{L^\infty}). \quad (4.10)$$

To prove (4.7), we write

$$F(u) - F(v) = (v - u) \int_0^1 F'((1-t)u + tv) dt,$$

so that (4.5) gives

$$\|F(u) - F(v)\|_{l^1 X^\sigma} \lesssim \|u - v\|_{l^1 X^\sigma} \sup_t \|F'((1-t)u + tv)\|_{l^1 X^s}.$$

Using (4.10), this is in turn bounded by

$$\|u - v\|_{l^1 X^\sigma} \left(\|u\|_{l^1 X^s} + \|v\|_{l^1 X^s} \right)^2 c(\|u\|_{L^\infty}, \|v\|_{L^\infty}),$$

as desired.

For (4.6), this follows directly from (4.7) with $v = 0$ when $d - s \leq \sigma \leq s$. However, we require this estimate for all $d - s \leq \sigma$; this will be critical later on in order to achieve higher regularity bounds for our nonlinear evolution. In order to prove (4.6) for $\sigma > s$, we must revisit the proof of (4.9) from Proposition 3.1 in [26]. We use the expansion

$$S_k F(u) = S_k F(u_0) + \int_0^\infty S_k(u_j F'(u_{<j})) dj.$$

Here we have temporarily replaced the discrete Littlewood-Paley decomposition by a continuous one with $\text{Id} = S_0 + \int_0^\infty S_j dj$.

For the first term we can simply use (4.6) in the already studied case $s = \sigma$ since the norms are equivalent on u_0 . Indeed, we have

$$\|S_k F(u_0)\|_{l^1 X^\sigma} \lesssim 2^{k(\sigma-s)} \|S_k F(u_0)\|_{l^1 X^s} \lesssim 2^{k(\sigma-s)} 2^{-Nk} \|S_k \partial^N (F(u_0))\|_{l^1 X^s}.$$

Upon choosing $N \geq \sigma - s$ and computing the derivative, the bound for $S_k F(u_0)$ follows immediately from (4.5) (with $\sigma = s$) except for the terms

$$\|S_k((\partial^N u_0) F'(u_0))\|_{l^1 X^s} + \|S_k((\partial^{\leq N-1} u_0^2) F''(u_0))\|_{l^1 X^s}.$$

The bounds for these terms instead follow from (4.5) followed by an application of (4.10) and (4.9) respectively.

For the integrand, as an intermediate step we need to estimate the expression $f_j = F'(u_{<j})$. A direct application of (4.10) yields the bound

$$\|f_j\|_{l^1 X^s} \lesssim \|u\|_{l^1 X^s}^2 C(\|u\|_{L^\infty}). \quad (4.11)$$

Differentiating any number of times and then applying (4.10) yields the better high frequency bound

$$\|f_j\|_{l^1 X^{s+N}} \lesssim \|\partial^N f_j\|_{l^1 X^s} \lesssim 2^{Nj} \|u\|_{l^1 X^s}^2 C(\|u\|_{L^\infty}), \quad N \geq 0, \quad (4.12)$$

where we have argued as in the $F(u_0)$ case above to obtain the last inequality.

Then the desired estimate (4.7) is obtained using the Littlewood-Paley trichotomy as follows:

- (a) For high-low interactions we must have $|j - k| \leq 4$, and we bound

$$\|S_k(u_j S_{<j-4} f_j)\|_{l^1 X^\sigma} \lesssim \|u_j\|_{l^1 X^\sigma} \|f_j\|_{L^\infty}.$$

A subsequent application of Bernstein's inequality and (4.11) yields the desired estimate.

- (b) For high-high interactions we must have $j > k - 4$, and we use (4.2) to bound

$$\|S_k(u_j S_{[j-4, j+4]} f_j)\|_{l^1 X^\sigma} \lesssim 2^{(\sigma+s-d)(k-j)} 2^{(\frac{d}{2}-s)k} \|u_j\|_{l^1 X^\sigma} \|f_j\|_{l^1 X^s}.$$

The rapid decay with respect to both k and j provides the summability, and an application of (4.11) completes the estimate.

(c) For low-high interactions we must have $j < k - 4$, and we use (4.1) to bound

$$\|u_j S_k f_j\|_{l^1 X^\sigma} \lesssim 2^{(\sigma-s-N)(k-j)} 2^{\left(\frac{d}{2}-s\right)j} \|u_j\|_{l^1 X^\sigma} 2^{-Nj} \|S_k f_j\|_{l^1 X^{s+N}}.$$

For $N > \sigma - s$, we obtain sufficient decay with respect to both j and k to obtain summability so that (4.12) completes the estimate.

This concludes the proof of the proposition. $\square$

We next turn our attention to bilinear estimates from $X \times X$ into Y -type spaces; we first recall the two main dyadic estimates from there.

Lemma 4.3. *The following bilinear estimates hold in $l_j^1 Y_j$ spaces:*

(a) *High-low interactions $j < k - 4$:*

$$\|S_j u S_k v\|_{l_k^1 Y_k} \lesssim 2^{\frac{jd}{2}} 2^{j-k} \|S_j u\|_{l_j^1 X_j} \|S_k v\|_{l_k^1 X_k}, \quad (4.13)$$

(b) *Balanced interactions, $|j - k| \leq 4$:*

$$\|S_i(S_j u S_k v)\|_{l_i^1 Y_i} \lesssim 2^{\frac{kd}{2}} \|S_j u\|_{l_j^1 X_j} \|S_k v\|_{l_k^1 X_k}. \quad (4.14)$$

These estimates are the main building blocks for the proof of [26, Proposition 3.1 (b)]. The results there are no longer sufficient in the present paper, where, in order to deal with the large data problem, we also need bilinear bounds where a gain is obtained when $T \ll 1$. Our result is as follows:

Proposition 4.4. *Let $s > \frac{d}{2} + 2$ and $T \leq 1$. Then for any $\delta \geq 0$ sufficiently small, the following bilinear bounds hold on the time interval $[0, T]$:*

$$\|uv\|_{l^1 Y^{\sigma-\delta}} \lesssim T^\delta \|u\|_{l^1 X^{\sigma-1}} \|v\|_{l^1 X^{s-1}}, \quad 0 \leq \sigma \leq s, \quad (4.15)$$

$$\|uv\|_{l^1 Y^{\sigma-\delta}} \lesssim T^\delta \|u\|_{l^1 X^\sigma} \|v\|_{l^1 X^{s-2}}, \quad 0 \leq \sigma \leq s-1. \quad (4.16)$$

We remark that setting $\delta = 0$ one recovers [26, Proposition 3.1 (b)].

Proof of Proposition 4.4. The main step of the proof is to establish the extensions of the bounds (4.13) and (4.14) where we gain a power of T . In the case of (4.13) we will trade the balance of derivatives for the T^δ gain and will prove:

$$\|S_j u S_k v\|_{l_k^1 Y_k} \lesssim T^\delta 2^{\frac{jd}{2}} 2^{(1-\delta)(j-k)} \|S_j u\|_{l_j^1 X_j} \|S_k v\|_{l_k^1 X_k}, \quad j < k - 4, \quad \delta \in [0, 1]. \quad (4.17)$$

On the other hand, as it turns out, the bound (4.14) already has the T gain built in, so we will prove that

$$\|S_i(S_j u S_k v)\|_{l_i^1 Y_i} \lesssim T 2^{\frac{kd}{2}} \|S_j u\|_{l_j^1 X_j} \|S_k v\|_{l_k^1 X_k}. \quad (4.18)$$

Once these dyadic bounds are proved, the conclusion of the proposition easily follows after applying the Littlewood-Paley trichotomy and appropriate dyadic summation.

We first prove (4.17). Due to (4.13), we already know it for $\delta = 0$, so by interpolation it suffices to prove it when $\delta = 1$. Using Hölder's inequality in time, we have

$$\begin{aligned} \|S_j u S_k v\|_{l_k^1 Y_k} &\lesssim \|S_j u S_k v\|_{l_k^1 L^1 L^2} \\ &\lesssim \|S_j u\|_{L^\infty L^\infty} T \|S_k v\|_{l_k^1 L^\infty L^2} \lesssim T 2^{\frac{id}{2}} \|S_j u\|_{l_j^1 X_j} \|S_k v\|_{l_k^1 X_k}, \end{aligned}$$

as needed.

The bound (4.18) follows by a similar repeated application of Hölder's inequality and an application of Bernstein's inequality:

$$\begin{aligned} \|S_i(S_j u S_k v)\|_{l_i^1 Y_i} &\lesssim \|S_i(S_j u S_k v)\|_{l_i^1 L^1 L^2} \lesssim T^{\frac{1}{2}} \|S_i(S_j u S_k v)\|_{l_i^1 L^2 L^2} \\ &\lesssim T^{\frac{1}{2}} 2^{\frac{d}{2}(k-i)} \|S_i(S_j u S_k v)\|_{l_k^1 L^2 L^2} \lesssim T 2^{\frac{d}{2}(k-i)} \|S_i(S_j u S_k v)\|_{l_k^1 L^\infty L^2} \\ &\lesssim T 2^{\frac{d}{2}k} \|S_j u S_k v\|_{l_k^1 L^\infty L^1} \lesssim T 2^{\frac{d}{2}k} \|S_j u\|_{l_j^1 X_j} \|S_k v\|_{l_k^1 X_k}. \end{aligned}$$

□

4.2. The Paradifferential Source Term

Our goal here is to obtain estimates for the paradifferential remainder term $G(u)$. We recall here that

$$G(u, \bar{u}, \nabla u, \nabla \bar{u}) = F(u, \bar{u}, \nabla u, \nabla \bar{u}) - \partial_j (g^{jk} - T_{g^{jk}}^w) \partial_k u + T_{b_j}^w \partial_j u + T_{\bar{b}_j}^w \partial_j \bar{u}.$$

Proposition 4.5. *Assume that $s_0 > \frac{d}{2} + 2$. Then the nonlinearity G satisfies uniform bounds,*

$$\|G(u)\|_{l^1 Y^\sigma} \lesssim T^\delta C(\|u\|_{l^1 X^{s_0}}) \|u\|_{l^1 X^\sigma}, \quad 0 \leq \sigma. \quad (4.19)$$

as well as Lipschitz bounds,

$$\|G(u_1) - G(u_2)\|_{l^1 Y^\sigma} \lesssim T^\delta C(\|u_{1,2}\|_{l^1 X^{s_0}}) \|u_1 - u_2\|_{l^1 X^\sigma} \quad 0 \leq \sigma \leq s_0. \quad (4.20)$$

Remark 4.5.1. We note that in addition to the T^δ gain, here one could also obtain a slight gain in regularity for G , which is common in quasilinear problems for the paradifferential remainders. We do not pursue this here. By the same token, the Lipschitz bound extends to a slightly larger range for σ .

Remark 4.5.2. Another refinement of the bounds (4.19) and (4.20) is in the description of the constant C . Above, C is allowed to fully depend on the $l^1 X^{s_0}$ norm. One could refine this and restrict the dependence on the full $l^1 X^{s_0}$ norm to at most quadratic, but with a full dependence on the L^∞ norm, akin to Proposition 4.2. Such an improvement is implicit in the proof below, but not needed for our results.

Proof. Rewriting the Eq. (1.4) in nondivergence form,

$$iu_t + g^{jk}(u, \bar{u}) \partial_j \partial_k u = \tilde{F}(u, \bar{u}, \nabla u, \nabla \bar{u}),$$

where

$$\tilde{F}(u, \bar{u}, \nabla u, \nabla \bar{u}) = F(u, \bar{u}, \nabla u, \nabla \bar{u}) - \partial_u g^{jk}(u, \bar{u}) \partial_j u \partial_k \bar{u} - \partial_{\bar{u}} g^{jk}(u, \bar{u}) \partial_j \bar{u} \partial_k u,$$

one can verify by a direct computation that we can split $G = G_1 + G_2$ where

$$\begin{aligned} G_1 &= (g^{jk} - T_{g^{jk}}^w) \partial_j \partial_k u, \\ G_2 &= \tilde{F} - T_{b_j}^w \partial_j u - T_{\tilde{b}_j}^w \partial_j \bar{u}, \end{aligned}$$

where by a slight abuse of notation we have redefined

$$b_j = \partial_{\partial_j u} \tilde{F}, \quad \tilde{b}_j = \partial_{\partial_j \bar{u}} \tilde{F}.$$

We will work with G_1 directly in the form above, but G_2 needs some further processing. Precisely, using a continuous Littlewood–Paley truncation indexed by the dyadic parameter k , we write G_2 in the form

$$\begin{aligned} G_2 &= \tilde{F}(u, u_{<0}) + \int_0^\infty (b(u, \nabla u_{<k}) - T_{b(u, \nabla u)}^w) \nabla u_k \\ &\quad + (\tilde{b}(u, \nabla u_{<k}) - T_{\tilde{b}(u, \nabla u)}^w) \nabla \bar{u}_k \, dk. \end{aligned} \quad (4.21)$$

The two integrated terms are similar and can be estimated separately so we consider the first one. There we re-expand to rewrite the integrand as

$$S_{>k-4} b(u, \nabla u_{<k}) \nabla u_k + \int_k^\infty S_{<k-4} (b(u, \nabla u_{<j}) \nabla u_j) \nabla u_k \, dj. \quad (4.22)$$

At this point we divide the argument into two cases, depending on how σ compares with s_0 .

Case 1, $\sigma \geq s_0$. Here we only need to prove (4.19). For G_1 we apply the Moser estimate (4.6) to get

$$\|g(u) - I\|_{l^1 X^\sigma} \lesssim_M \|u\|_{l^1 X^\sigma}$$

and then apply Proposition 4.4 (specifically (4.18)).

For the first term of G_2 in (4.21) we simply apply the Moser estimates (4.6) to place it in $l^1 X^\sigma \subset Tl^1 Y^\sigma$.

Next we consider the two terms in (4.22). For the expression $b_k := b(u, \nabla u_{<k})$ we use the Moser estimate (4.6) to bound

$$\|b_k\|_{l^1 X^{\sigma-1+h}} \lesssim 2^{hk} C(\|u\|_{L^\infty}) \|u\|_{l^1 X^\sigma} (1 + \|u\|_{l^1 X^{s_0}}), \quad h \in [0, 1], \quad s_0 \leq \sigma. \quad (4.23)$$

Here the case $h = 0$ follows by applying (4.6) directly to $b(u, \nabla u_{<k})$, whereas the case $h = 1$ is obtained by applying (4.6) to its gradient. In the first term of (4.22)

we can conclude now via (4.15) (or, more precisely, the dyadic bounds (4.17) and (4.18)).

In the integrand in the second term of (4.22) we can freely insert a projector and rewrite it as

$$S_{<k-4}(S_j b_j \nabla u_j) \nabla u_k,$$

after which we estimate applying (4.17) and (4.5) sequentially

$$\begin{aligned} \|S_{<k-4}(S_j b_j \nabla u_j) \nabla u_k\|_{l^1 Y^\sigma} &\lesssim T^\delta 2^{\delta k} \|S_j b_j \nabla u_j\|_{l^1 X^{s_0-1}} \|u_k\|_{l^1 X^\sigma} \\ &\lesssim T^\delta 2^{\delta(k-j)} \|b_j\|_{l^1 X^{s_0-1}} \|\nabla u_j\|_{l^1 X^{s_0-1}} \|u_k\|_{l^1 X^\sigma}. \end{aligned}$$

Now we can conclude by (4.23) with $\sigma = s_0$ and $h = 0$.

Case 2, $0 \leq \sigma \leq s_0$. Here we only need to prove (4.20), as (4.19) is then a straightforward consequence. We denote $v = u_2 - u_1$ and use u for either u_1 or u_2 .

For G_1 we represent

$$g(u_1) - g(u_2) = vh(u)$$

for a smooth function h . Then

$$G_1(u_1) - G_1(u_2) = \sum_k S_{>k-4}(vh(u)) \partial^2 u_k + S_{>k-4}(g(u)) \partial^2 v_k.$$

The terms $h(u) - h(0)$ and $g(u) - g(0)$ are estimated in $l^1 X^{s_0}$ by the Moser estimate (4.6). The second term as well as the contribution of constants in h are estimated by (4.17), (4.18). One might want to also use bilinear estimates to fully include $w = h(u) - h(0)$ in v , but this is not allowed by the limited range of σ in (4.5). Nevertheless, we can estimate the high-low interactions in vw in this manner. For the remaining terms we need to treat this as a trilinear bound,

$$\|S_j v S_j w S_k u\|_{l^1 Y^\sigma} \lesssim T^\delta 2^{-2j-\delta j} \|v\|_{l^1 X^\sigma} \|w\|_{l^1 X^{s_0}} \|S_k u\|_{l^1 X^{s_0}}, \quad j \geq k - 5. \quad (4.24)$$

For this we separate into two cases. If $j > k + 4$ we use the algebra property to absorb $S_k u$ into w , and then apply (4.15). Else, $|j - k| \leq 4$ so we can place the product $S_j w S_k u$ into $l^1 X^{s_0-1}$ with a $2^{-\delta k}$ gain, and then use (4.15).

We now turn our attention to G_2 . The contributions arising from the first term in (4.21) are easily estimated using the Moser estimates in (4.6) and (4.15). We next consider the differences corresponding to the first term in (4.22). Using again the notation b_k for expressions of the form $b(u, \nabla u_{<k})$, the differences will be linear combinations of expressions of the form

$$S_{>k-4}(vb_k) \nabla u_k, \quad S_{>k-4}(\nabla v_{<k} b_k) \nabla u_k, \quad S_{>k-4} b_k \nabla v_k.$$

Given the bound (4.23) for b_k , the first term is like in the case of G_1 but better. The second term is worst in the case $\sigma = s_0$, which was already covered in Case 1. Finally the last term is as in Case 1, and can be handled using (4.23) and (4.15).

Next we consider the integrand in (4.22). For any expression $b_j = b(u, \nabla u_{\leq j})$ we use (4.23), which allows us to reduce the problem to estimating quadrilinear terms as follows:

$$\begin{aligned} S_{<k-4}(vb_j\nabla u_j)\nabla u_k & \quad S_{<k-4}(b_j\nabla v_{<j}\nabla u_j)\nabla u_k, \\ S_{<k-4}(b_j\nabla v_j)\nabla u_k, & \quad S_{<k-4}(b_j\nabla u_j)\nabla v_k. \end{aligned}$$

If the frequency of v is $\leq k+4$ then this is no different from Case 1. In particular, we can fully discard the last term. Else, the worst case is $\sigma = 0$, which we assume from here on. For the first two terms we can use the algebra property to bound $b_j\nabla u_j$ in $2^{-\delta j}l^1X^{s_0-1}$, where the gain is only needed to insure the j summation. Then in all three cases we are left with a trilinear bound

$$\|S_{<k-4}(v_j w_j)u_k\|_{l^1Y^0} \lesssim T^\delta \|v\|_{l^1X^{-1}} \|w\|_{l^1X^{s_0-1}} \|u\|_{l^1X^{s_0-1}}, \quad j > k+5. \quad (4.25)$$

We discard the multiplier, use the algebra property for $w_j u_k$ and then apply (4.15). $\square$

As a corollary of this, we also obtain frequency envelope bounds for G .

Corollary 4.6. *Assume that $s_0 > \frac{d}{2} + 2$. Let c_k be a frequency envelope for u in $l^1X^{s_0}$. Then, provided $\delta \geq 0$ is sufficiently small,*

$$\|S_k G(u)\|_{l^1X^{s_0}} \lesssim_M T^\delta c_k. \quad (4.26)$$

Proof. Expand

$$S_k G(u) = S_k G(u_{<0}) + \sum_{j=0}^{\infty} S_k \left[G(u_{<j+1}) - G(u_{<j}) \right].$$

When $j \geq k$, we estimate the differences in l^1Y^0 using (4.20). When $j < k$, we instead estimate the terms separately in l^1Y^N where $N > s_0 + \sigma$, with σ as in part (4) of the definition of a frequency envelope. $\square$

4.3. Estimates for the Cubic Problem

In the same way the proof of Theorem 1.2 relies primarily on quadratic estimates covered in the previous subsection, for the proof of Theorem 1.3 we need instead trilinear estimates. The statements for these trilinear bounds are provided in this section; the proofs are similar and largely omitted. See [27] for some related background. We note that here we only require $s > \frac{d+3}{2}$, which accounts for the improvement in regularity as compared to the quadratic case. We start with the replacement of Proposition 4.2 for l^2X type spaces.

Proposition 4.7. (a) *Let $s > \frac{d}{2}$. Then the l^2X^s spaces satisfy the bilinear estimate*

$$\|uv\|_{l^2X^\sigma} \lesssim \|u\|_{l^2X^\sigma} \|v\|_{l^2X^s}, \quad 0 \leq \sigma \leq s. \quad (4.27)$$

(b) For all smooth F with $F(0) = F'(0) = F''(0) = 0$ we have the Moser estimate

$$\|F(u)\|_{l^2 X^\sigma} \lesssim \|u\|_{l^2 X^\sigma} \|u\|_{l^2 X^s}^2 c(\|u\|_{L^\infty}), \quad 0 \leq \sigma \quad (4.28)$$

as well as the difference estimates

$$\begin{aligned} \|F(u) - F(v)\|_{l^2 X^\sigma} &\lesssim \|u - v\|_{l^2 X^\sigma} (\|u\|_{l^2 X^s} + \|v\|_{l^2 X^s})^2 c(\|u\|_{L^\infty}, \|v\|_{L^\infty}) \\ 0 &\leq \sigma \leq s. \end{aligned} \quad (4.29)$$

Compared to Proposition 4.2, we note that here we have a larger range for σ . This is due to a corresponding improvement in the balanced bilinear interactions. Precisely, the bound (4.2) is replaced by

$$\|S_i(S_j u S_k v)\|_{l_i^2 X_i} \lesssim 2^{\frac{dk}{2}} \|u\|_{l_j^2 X_j} \|v\|_{l_k^2 X_k}. \quad (4.30)$$

The bilinear $l^1 X \times l^1 X \rightarrow l^1 Y$ bounds in Proposition 4.4 are now replaced by trilinear $l^2 X \times l^2 X \times l^2 X \rightarrow l^2 Y$ type bounds. Again we will also need the slight improvement where we trade regularity for a slight gain in time.

Proposition 4.8. *Let $s > \frac{d+3}{2}$ and $T \leq 1$. Then for any $\delta \geq 0$ sufficiently small, the following trilinear bounds hold on the time interval $[0, T]$:*

$$\|uvw\|_{l^2 Y^{\sigma-\delta}} \lesssim T^\delta \|u\|_{l^2 X^{\sigma-1}} \|v\|_{l^2 X^{s-1}} \|w\|_{l^2 X^{s-1}}, \quad 0 \leq \sigma \leq s, \quad (4.31)$$

and

$$\|uvw\|_{l^2 Y^{\sigma-\delta}} \lesssim T^\delta \|u\|_{l^2 X^\sigma} \|v\|_{l^2 X^{s-2}} \|w\|_{l^2 X^{s-1}}, \quad 0 \leq \sigma \leq s-1. \quad (4.32)$$

Just as in the quadratic case, the next proposition is the main tool in the proof of the cubic bound for the paradifferential error term.

Proposition 4.9. *Assume that $s_0 > \frac{d+3}{2}$. Then the nonlinearity G satisfies uniform bounds,*

$$\|G(u)\|_{l^2 Y^\sigma} \lesssim T^\delta C(\|u\|_{\ell^2 X^{s_0}}^2) \|u\|_{l^1 X^\sigma}, \quad \sigma \geq 0, \quad (4.33)$$

as well as Lipschitz bounds,

$$\|G(u) - G(v)\|_{\ell^2 Y^\sigma} \lesssim T^\delta C(\|u\|_{\ell^2 X^{s_0}}^2, \|v\|_{\ell^2 X^{s_0}}^2) \|u - v\|_{l^2 X^\sigma} \quad 0 \leq \sigma \leq s_0. \quad (4.34)$$

The proofs of Propositions 4.8 and 4.9 follow directly by modifying the proofs of Propositions 4.4 and 4.5 in the quadratic case and are left to the reader.

5. Nontrapping Metrics

Here we begin with $u_0 \in l^1 H^{s_0}$, $s_0 > \frac{d}{2} + 2$, and fix M so that

$$\|u_0\|_{l^1 H^{s_0}} \leq M. \quad (5.1)$$

Then the associated metric $g(u_0)$ satisfies the Moser type bound [cf. (4.6)]

$$\|g(u_0) - I\|_{l^1 H^{s_0}} \lesssim M^2, \quad (5.2)$$

with an implicit constant depending on the L^∞ norm of u_0 . Here I denotes the flat background metric. In particular, by Sobolev embeddings we have

$$\|g(u_0)\|_{C^2} \lesssim 1 + M^2.$$

This guarantees that the Hamilton flow $(x, \xi) \rightarrow (x^t, \xi^t)$ given by

$$(\dot{x}^t, \dot{\xi}^t) = (a_\xi(x^t, \xi^t), -a_x(x^t, \xi^t)), \quad (x^0, \xi^0) = (x, \xi), \quad (5.3)$$

with $a(x, \xi) = g^{ij}(u_0(x))\xi_i\xi_j$ is well-defined.

In order to guarantee that the qualitative assumption that the metric $g(u_0)$ is nontrapping is meaningful we will show that this is in effect a condition about the bicharacteristic flow within a compact set. Furthermore, in order to prove the local well-posedness result with a bound from below for the lifespan of the solution we need to turn this assumption into a quantitative statement.

We begin by selecting a ball $B := B_R$ of radius $R \gg 1$ so that outside B we have the smallness condition

$$\|\chi_{>R/2}(g(u_0) - I)\|_{l^1 H^{s_0}(\mathbb{R}^d)} \leq \varepsilon, \quad (5.4)$$

with a universal small constant ε . Here $\chi_{>R/2}$ is a smooth cutoff which equals 1 outside $B_{R/2}$ and zero inside $B_{R/4}$. We observe that for large frequencies such a bound holds globally, so only small frequencies contribute to this.

However, R could still be arbitrarily large, independently of M , so we retain it as one of the main parameters in our problem. Without any restriction in generality we will make the assumption

$$\log R \gg \log M \quad (5.5)$$

with a universal implicit constant. The smallness outside B ensures that no trapping can happen there. More precisely, we have the following:

Lemma 5.1. *Let B be as in (5.4) where $s_0 > \frac{d}{2} + 2$ and $\varepsilon > 0$ is sufficiently small. Then any geodesic for the metric $g = g(u_0)$ that exits the ball B will escape to spatial infinity without reentering $B/2$.*

Proof. The result will follow by proving uniform bounds similar to those in [24]. We begin with the Hamilton flow for $g = g(u_0)$,

$$\dot{x}_i^t = g^{ij}(x^t)\xi_j^t, \quad \dot{\xi}_i^t = -(g^{jk})_i(x^t)\xi_j^t\xi_k^t, \quad (5.6)$$

which we will compare with that driven by the flat flow,

$$\dot{\tilde{x}}_i^t = \tilde{\xi}_i^t, \quad \dot{\tilde{\xi}}_i^t = 0. \quad (5.7)$$

We will show that all outward pointing rays are forced to live in a small angular sector. Precisely, consider a geodesic which exits B_R , i.e. a solution (x^t, ξ^t) to (5.6) starting say at $t = 0$ at the point (x, ξ) so that $x \in \partial B_R$ and $\dot{x}(0) \cdot x \geq 0$. This will be compared to the corresponding trajectory $(x + \xi t, \xi)$ for the flat flow (5.7). We will show that the two trajectories stay close to each other,

$$\|\xi^t - \xi\|_{L^\infty} < c|\xi|, \quad \|x^t - x - t\xi\|_{L^\infty} < ct|\xi| \quad (5.8)$$

for $\varepsilon \ll c \ll 1$ sufficiently small.

At the starting point we must have

$$|\dot{x}^0 - \xi| \lesssim \varepsilon|\xi|.$$

By the continuity of $\dot{x}^s$, for a sufficiently small t , we have

$$|\dot{x}^s - \xi| < c|\xi|, \quad s \in [0, t].$$

It, thus, follows that (5.8) holds for sufficiently small $t \geq 0$. We then use a bootstrap argument to show that (5.8) holds globally. We assume the bounds (5.8) for $t \in [0, T]$ and shall show that the same hold with, say, c replaced by $c/2$. We note two immediate consequences of our bootstrap assumptions. On the time scale $[0, T]$, for which (5.8) is assumed,

- (i) The bicharacteristic (x^t, ξ^t) cannot re-enter $B/2$;
- (ii) The bicharacteristic (x^t, ξ^t) is nearly straight, $|\dot{x}^t - \xi| \lesssim c|\xi|$.

It remains to complete the bootstrap. By property (i), we will freely assume that

$$\|g - I\|_{L^1 H^{s_0}} \leq \varepsilon$$

in a unit neighbourhood of the bicharacteristic (x^t, ξ^t) , $t \in [0, T]$. By property (ii), the bicharacteristic remains in a cube of sidelength 2^j for an $\mathcal{O}(2^j/|\xi|)$ amount of time. Thus, Bernstein's inequality yields the uniform bound

$$\int_0^t |(g - I)(x^s)| + |\nabla_x g(x^s)| ds \lesssim |\xi|^{-1} \|\chi_B^{ext}(g - I)\|_{L^1 H^{s_0}} \lesssim \varepsilon|\xi|^{-1}, \quad (5.9)$$

provided that $t \in [0, T]$.

We first close the bootstrap for $\xi^t - \xi$. We have

$$\frac{d}{dt}(\xi^t - \xi)_i = \partial_i(g^{kl})\xi_k^t\xi_l^t.$$

By our bootstrap assumption we have $|\dot{\xi}^t| \approx |\xi|$. Hence integrating in the last relation and using (5.9), we obtain

$$|\xi^t - \xi| \lesssim \int_0^t |\partial_i(g^{kl})\xi_k^s \xi_l^s| ds \leq C\varepsilon|\xi|, \quad t \in [0, T],$$

which suffices provided $\varepsilon < \frac{c}{2C}$.

Next we consider the difference $x^t - x - t\xi$, for which we have

$$\frac{d}{dt}(x^t - x - t\xi)_i = g^{ij}\xi_j^t - \xi_i.$$

Hence, using that $\|\chi_{>R/2}(g - I)\|_{L^\infty} \lesssim \varepsilon$, we have

$$|x^t - x - t\xi| \lesssim \int_0^t |g - I||\dot{\xi}^\tau| + |\xi - \dot{\xi}^\tau| d\tau \lesssim \varepsilon t|\xi|,$$

which gives the desired result with ε sufficiently small as above. $\square$

Lemma 5.1 ensures that any trapping must be confined to a compact set, namely the ball B . The nontrapping assumption guarantees that all geodesics exit B at both ends. We now seek to quantify that.

In order to measure the length of a geodesic within the ball B we remark that for the geodesic flow (5.3) we have the scaling symmetry

$$\xi \rightarrow \lambda\xi, \quad t \rightarrow \lambda t,$$

where t is used to denote the parameter in (5.3) along geodesics. This one dimensional degree of freedom needs to be removed in order to uniquely define the length of geodesics. Our strategy will be to project the geodesic flow on the cosphere bundle $\{|\xi| = 1\}$, i.e. replace (5.3) with

$$(\dot{x}^t, \dot{\xi}^t) = (a_\xi(x^t, \xi^t), -a_x(x^t, \xi^t) + (a_x(x^t, \xi^t) \cdot \xi^t)\xi^t), \quad |\xi^t| = 1. \quad (5.10)$$

Then we define the length of a geodesic γ between two points as

$$\ell(\gamma) = \int_{t_0}^{t_1} |\dot{x}^t| dt.$$

We remark that for a single metric g it would be more natural to measure its length using the g metric. However, here we also need to allow for changes in the metric, so it is better to have a common reference frame.

Now we are ready to define our last parameter L which measures the maximum length of a geodesic within B . Indeed, a straightforward compactness argument applied to the projection of the flow to the cosphere bundle shows that the quantity

$$L = \sup\{\ell(\gamma \cap 2B_R); \quad \gamma \text{ geodesic for } g(u_0)\} \quad (5.11)$$

is finite, where $\ell(\gamma)$ stands for the Euclidean length as in (5.10). We have an obvious lower bound $L \gtrsim R$, but no upper bound for L in terms of R . We will use L as the second parameter in the quantitative description of nontrapping.

Our next task is to see that our nontrapping assumption is stable with respect to a class of small perturbations.

Proposition 5.2. Assume that $u_0 \in l^1 H^{s_0}$ is so that $g(u_0)$ is nontrapping. Let M, R, L be as above. Then, there exists $C_0(M) > 0$ such that for $w \in l^1 X^{s_0}$ satisfying

$$\|w\|_{l^1 X^{s_0}} \leq e^{-C_0(M)L} \quad (5.12)$$

the metrics $g(u_0 + w)$ are uniformly nontrapping, with comparable parameters R, L .

Proof. We denote the two metrics by

$$g_0 = g(u_0), \quad g_1 = g(u_0 + w).$$

By hypothesis, we have

$$\|u_0 - (u_0 + w)\|_{l^1 X^{s_0}} \lesssim e^{-C_0(M)L}.$$

Hence, for R chosen as in (5.4), outside $B_R/2$ both metrics are close to the Euclidean metric,

$$\|\chi_{B/2}^{ext}(g_0 - I)\|_{l^1 X^{s_0}} + \|\chi_{B/2}^{ext}(g_1 - I)\|_{l^1 X^{s_0}} \lesssim \varepsilon + e^{-C_0(M)L} \lesssim \varepsilon$$

for some $\varepsilon > 0$ and the analysis in Lemma 5.1 equally applies. It remains to compare their Hamilton flows in B_R , where the two metrics are close,

$$\|g_0 - g_1\|_{C^2(B_R)} \lesssim e^{-C_0(M)L}.$$

We begin with the Hamilton flows of both problems and show that the trajectories of the two are close on B_R . Namely, we take

$$\begin{aligned} \dot{x}_i &= g_0^{ij}(x) \xi_j, \\ \dot{\xi}_i &= -\partial_i g_0^{jk}(x) \xi_j \xi_k + (\xi \cdot \partial g_0^{jk}(x) \xi_j \xi_k) \xi_i \end{aligned} \quad (5.13)$$

and

$$\begin{aligned} \dot{\tilde{x}}_i &= g_1^{ij}(t, \tilde{x}) \tilde{\xi}_j, \\ \dot{\tilde{\xi}}_i &= -\partial_i g_1^{jk}(t, \tilde{x}) \tilde{\xi}_j \tilde{\xi}_k + (\tilde{\xi} \cdot \partial g_1^{jk}(t, \tilde{x}) \tilde{\xi}_j \tilde{\xi}_k) \tilde{\xi}_i, \end{aligned} \quad (5.14)$$

with the same initial data (x_0, ξ_0) at $t = 0$, where $|\xi_0| = 1$, and $x_0 \in B_R$. Our goal will be to prove

$$|x(t) - \tilde{x}(t)| + |\xi(t) - \tilde{\xi}(t)| \lesssim e^{-C(M)L} \ll 1. \quad (5.15)$$

Rather than estimating the difference directly, it is perhaps easiest to consider a one parameter family of metrics

$$g(\cdot; h) = (1 - h)g_0 + hg_1, \quad h \in [0, 1],$$

define the flow $(y(t; h), \eta(t; h))$ using

$$\begin{aligned} \dot{y}_i &= g^{ij}(t, y; h) \eta_j, \\ \dot{\eta}_i &= -\partial_i g^{jk}(t, y; h) \eta_j \eta_k + (\eta \cdot \partial g^{jk}(t, y; h) \eta_j \eta_k) \eta_i \end{aligned} \quad (5.16)$$

and then differentiate in h . Note that using this notation we have $(y, \eta)(t; 0) = (x(t), \xi(t))$ and $(y, \eta)(t; 1) = (\tilde{x}(t), \tilde{\xi}(t))$. For this family we seek to prove

$$|y(t) - x(t)| + |\eta(t) - \xi(t)| \lesssim e^{-C(M)L} \ll 1 \quad (5.17)$$

uniformly in $h \in [0, 1]$, which will yield (5.15) upon choosing $h = 1$. We shall prove (5.17) using a bootstrapping argument. We assume (5.17), and we shall prove the same with an improved constant.

The h derivatives (y_h, η_h) solve the differentiated system

$$\begin{cases} \dot{y}_h = (g_1 - g_0)(t, y)\eta + y_h \cdot \partial(g(t, y; h))\eta + g(t, y; h)\eta_h, \\ \dot{\eta}_h = -\eta \partial(g_1 - g_0)(t, y)\eta - \eta y_h \cdot \partial \partial g(t, y; h)\eta - 2\eta \partial g(t, y; h)\eta_h \\ \quad + (\eta \eta_h \cdot \partial g(t, y; h)\eta) + (\eta \eta \cdot \partial(g_1 - g_0)(t, y)\eta) \\ \quad + (\eta(y_h \cdot \partial)\eta \cdot \partial g(t, y; h)\eta) + 2(\eta \eta \cdot \partial g(t, y; h)\eta_h) + (\eta \eta \cdot \partial g(t, y; h)\eta)\eta_h. \end{cases} \quad (5.18)$$

By the Mean Value Theorem, in order to establish (5.17), it suffices to establish the bound

$$|y_h(t)| + |\eta_h(t)| \lesssim e^{-2C(M)L} \ll 1 \quad (5.19)$$

uniformly for $h \in [0, 1]$.

Due to the bootstrapping hypothesis (5.17), it suffices to consider (5.18) on I , which denotes the maximal time interval that a bicharacteristic for g_0 spends in B_R . We note that $|I| \lesssim L$.

Since the flow (5.16) is projected onto the cosphere bundle, we observe that

$$\frac{d}{dt}(y_h^2 + \eta_h^2) \lesssim M^{-1}e^{-2C_0(M)L} + M(y_h^2 + \eta_h^2).$$

Using Grönwall's inequality and that $|I| \lesssim L$, we obtain

$$|y_h| + |\eta_h| \lesssim M^{-1/2}L^{1/2}e^{-C_0(M)L}e^{CML} \lesssim e^{-C_0(M)L}e^{C(M)L},$$

which for a choice of $1 \ll C(M) < \frac{1}{3}C_0(M)$ proves (5.19) with the improved constant as desired.

We now have the uniformity of the nontrapping assumption as all trajectories exiting the ball are now sufficiently close. Now it is clear why the C^2 difference between the metrics needs to be small compared to $e^{-C_0(M)L}$. $\square$

6. The Linear Flow

In this section we consider the L^2 well-posedness question for the linear Schrödinger flow

$$(i\partial_t + \partial_k g^{kl}\partial_l + b \cdot \nabla)w + \tilde{b} \cdot \nabla \bar{w} = f, \quad w(0, x) = w_0, \quad (6.1)$$

for large but nontrapping metrics g . We also consider the corresponding linear paradifferential flow

$$(i\partial_t + \partial_k T_{g^{kl}}\partial_l + T_b \cdot \nabla)w + T_{\tilde{b}} \cdot \nabla \bar{w} = f, \quad w(0, x) = w_0. \quad (6.2)$$

To frame the question in the context of the previous section where nontrapping is discussed, we consider metrics g and lower order perturbations b satisfying the following properties with respect to parameters M , R and L , which themselves are subject to the constraints

$$\log M \ll \log R \leq \log L :$$

- Large size:

$$\|g - I_d\|_{L^1 X^{s_0}} + \|g_t\|_{L^1 X^{s_0-2}} + \|(b, \tilde{b})\|_{L^1 X^{s_0-1}} + \|(b_t, \tilde{b}_t)\|_{L^1 X^{s_0-3}} \leq M. \quad (6.3)$$

- Smallness outside a ball B_R :

$$\|g - I_d\|_{L^1 X^{s_0}(B_R^c)} + \|(b, \tilde{b})\|_{L^1 X^{s_0-1}(B_R^c)} \leq \varepsilon \ll 1 \quad (6.4)$$

for a fixed universal constant ε .

- Uniform nontrapping: For each geodesic [projected onto the cosphere bundle as in (5.10)] γ at fixed time, we have

$$\ell(\gamma \cap 2B_R) \leq L. \quad (6.5)$$

Our goal is to understand the energy and local energy bounds for such a flow. Since we will want to apply these bounds in order to solve a nonlinear equation, it is crucial to have good control in a way that does not grow rapidly in time. This is a delicate matter, since the nontrapping property allows energy growth by a $e^{C(M)L}$ factor within B_R . Further, the coefficients at high frequencies may repeatedly redirect inwards some of the outgoing energy, for additional potential exponential growth (as we cannot assume ε is exponentially small with respect to L).

We can prevent such iterated growth by restricting the time to a short enough interval.

Theorem 6.1. *Suppose that $g, b, \tilde{b}$ are as above, with $s_0 > \frac{d}{2} + 2$ and associated parameters M, R, L . Then the Eqs. (6.1) and (6.2) are well-posed in L^2 . Further, for all small enough times T ,*

$$T \leq e^{-C(M)L} \quad (6.6)$$

we have a uniform bound

$$\|w\|_{X^0[0,T]} \lesssim e^{C(M)L} (\|w_0\|_{L^2} + \|f\|_{Y^0[0,T]}). \quad (6.7)$$

For simplicity the above result is stated in the L^2 setting, but a similar result also holds in higher norms.

Corollary 6.2. *Suppose that $g, b, \tilde{b}$ are as above, with $s_0 > \frac{d}{2} + 2$ and associated parameters M, R, L . Then the Eq. (6.1) is well-posed in H^σ and $l^1 H^\sigma$ for $0 \leq \sigma \leq s_0$. Further, for all small enough times T ,*

$$T \leq e^{-C(M)L}, \quad (6.8)$$

we have a uniform bound

$$\|w\|_{l^1 X^\sigma[0,T]} \lesssim e^{C(M)L} (\|w_0\|_{l^1 H^\sigma} + \|f\|_{l^1 Y^\sigma}) \quad (6.9)$$

$$\|w\|_{\ell^2 X^\sigma[0,T]} \lesssim e^{C(M)L} (\|w_0\|_{H^\sigma} + \|f\|_{\ell^2 Y^\sigma}). \quad (6.10)$$

The same result holds for (6.2) for all $\sigma \geq 0$.

This is not so much a corollary of the previous theorem, but rather of its proof. Precisely, we will be able to reuse the key part of the proof of the theorem, and then replicate the (short) remaining parts of the proof.

Since the last result holds for all σ in the case of the paradifferential equation, it easily implies the corresponding frequency envelope version, as a direct application of the techniques in [26, Propositions 5.1 and 5.3].

Corollary 6.3. *Let $s_0 > \frac{d}{2} + 2$, and suppose $w_0 \in l^1 H^{s_0}$, $f \in l^1 Y^{s_0}$ with admissible frequency envelopes $\{a_k\}$, $\{b_k\}$ respectively. Then the solution w to the paradifferential flow (6.2) satisfies*

$$\|S_k w\|_{l^1 X^{s_0}} \lesssim e^{C(M)L} (a_k + b_k). \quad (6.11)$$

The essential part of Theorem 6.1 is the energy estimate (6.7). The L^2 well-posedness follows in a standard fashion from a similar energy estimate for the (backward) adjoint equation. Since the adjoint equation has a similar form, with similar bounds on the coefficients, such an estimate follows directly from (6.7). Thus, in what follows we focus on the proof of the bound (6.7).

An important part of the theorem is to keep good track of the dependence of the constants on our parameters M, L and T . In order to avoid circular arguments, it is useful to apply a divide and conquer strategy. One step in this direction is to take T out of the equation, at the expense of allowing a lower order term in the estimate. We state this intermediate result as follows:

Proposition 6.4. *Suppose that g, b , and $\tilde{b}$ are as above, with $s_0 > \frac{d}{2} + 2$ and associated parameters M, R, L . Then for any solution w to (6.1) or (6.2) and all $T \leq 1$ we have a uniform bound*

$$\|w\|_{X^0[0,T]} \lesssim e^{C(M)L} (\|w_0\|_{L^2} + \|f\|_{Y^0[0,T]} + \|w\|_{L^2 L^2[0,T]}). \quad (6.12)$$

This trivially implies Theorem 6.1 by applying Hölder's inequality in time to bound the last L^2 norm, and then taking T small enough.

The main bound above, (6.12), admits an exact counterpart in the context of Corollary 6.2, which is as follows:

$$\|w\|_{l^1 X^\sigma[0,T]} \lesssim e^{C(M)L} (\|w_0\|_{l^1 H^\sigma} + \|f\|_{l^1 Y^\sigma[0,T]} + \|w\|_{l^1 L^2 H^\sigma[0,T]}), \quad (6.13)$$

$$\|w\|_{\ell^2 X^\sigma[0,T]} \lesssim e^{C(M)L} (\|w_0\|_{H^\sigma} + \|f\|_{\ell^2 Y^\sigma[0,T]} + \|w\|_{\ell^2 L^2 H^\sigma[0,T]}). \quad (6.14)$$

Similarly, this trivially implies Corollary 6.2 by applying Hölder's inequality in time.

The bulk of this section is devoted to the proof of this last proposition for the paradifferential equation (6.2). The advantage to working with the paradifferential formulation is that it transfers easily to all higher regularities, whereas the paradifferential errors will only play a perturbative role. After that, it is much easier to obtain the rest of the result going in reverse order.

Before presenting the proof of Proposition 6.4 in full detail, we first outline our strategy. We seek to decompose the estimate into three pieces, roughly corresponding to a decomposition of w into three components,

$$w = w_{in} + w_R + w_{out}$$

which represent the portions of w which are microlocalized near the incoming rays outside B_R , near B_R , respectively near the outgoing rays. Heuristically the energy travels along bicharacteristics, and may go through all three stages or only the incoming and outgoing part, depending on the bicharacteristic. The main steps of the argument can be described as follows:

- (1) Prove an exterior incoming high frequency estimate of the form

$$\|\chi_{<100R} w_{in}\|_{X^0} \leq C (\|w_0\|_{L^2} + \|f\|_{Y^0}) + C(M) \|w\|_{L^2 L^2}$$

with a universal constant C . This amounts to constructing a multiplier which selects the incoming region and positive commutator estimates. Crucially, the metric is only used in the exterior region (outside $4B_R$), where it is a small perturbation of a flat metric. Here the time is taken in any interval $[0, T]$ with $T \leq 1$; the smallness of T is not used.

- (2) Using the fact that the metric g satisfies the nontrapping condition, we estimate the local energy inside the compact set B_{2R} in terms of the incoming part,

$$\|w_R\|_{X^0} \lesssim e^{C(M)L} (\|w_0\|_{L^2} + \|f\|_{Y^0} + \|\chi_{<100R} w_{in}\|_{X^0} + \|w\|_{L^2 L^2}),$$

in a way that quantifies the potential exponential growth.

- (3) Rather than controlling the remaining component w_{out} on the outgoing rays, we estimate the full exterior part $w_{ext} = w_{in} + w_{out}$

$$\|w_{ext}\|_{X^0} \lesssim \|w_R\|_{X^0} + \|w_0\|_{L^2} + \|f\|_{Y^0},$$

simply by truncating to the exterior region and applying the small data estimate directly.

- (4) Combining the three estimates above we obtain

$$\|w\|_{X^0} \lesssim e^{C(M)L} (\|w_0\|_{L^2} + \|f\|_{Y^0} + \|w\|_{L^2 L^2}),$$

independently of the time interval size $0 < T \leq 1$.

6.1. A Review of the Small Data Results

The small data problem, studied in [26], provides us with a baseline for the study of the current problem. The main assumption for the linear result there is

$$\|g - I_d\|_{l^1 X^s} + \|(b, \tilde{b})\|_{l^1 X^{s-1}} \leq \varepsilon \ll 1, \quad (6.15)$$

where $g, b, \tilde{b}$ are the coefficients in (6.1). Here the smallness of ε suffices to guarantee that the metric g is nontrapping. Under this assumption, we have the following counterpart of Theorem 6.1:

Theorem 6.5. [26, Proposition 5.2] *Suppose that g, b , and $\tilde{b}$ are as in (6.15). Then the Eq. (6.1) is well-posed in L^2 . Further more, we have the uniform bounds*

$$\|w\|_{X^0[0,1]} \lesssim \|w_0\|_{L^2} + \|f\|_{Y^0[0,1]}, \quad (6.16)$$

$$\|w\|_{l^1 X^\sigma[0,1]} \lesssim \|w_0\|_{l^1 H^\sigma} + \|f\|_{l^1 Y^\sigma[0,1]}, \quad 0 \leq \sigma \leq s. \quad (6.17)$$

The same result holds for the corresponding paradifferential problem (6.2).

Remark 6.5.1. The original result in [26] also allows for zero order terms in (6.1), as in (3.2), with coefficients $c, \tilde{c}$ with regularity $c, \tilde{c} \in l^1 X^{s-2}$; in that case one has to limit the upper limit for σ in (6.17) to $s - 1$. The X^0 bound (6.16) follows from the proof of Proposition 4.1. Note also that in this result the contributions of the b and $\tilde{b}$ terms are also perturbative.

6.2. The $\bar{w}$ Correction

One difficulty in the large data problem, which is absent in the small data problem, is due to the $\tilde{b} \nabla \bar{w}$ term. This is perturbative in the small data case but not in the large data case. Our goal in this section is to show that we can eliminate this term from the paradifferential equation (6.2) at the expense of more perturbative terms.

Our correction will be of the form

$$\tilde{w} = Sw := w + \mathcal{R}\bar{w}, \quad (6.18)$$

where $\mathcal{R}$ is a paradifferential operator of order -1 . A similar conjugation was used in [6] to remove the complex conjugate leading order terms. Assuming that w solves (6.2), we obtain the following as the equation for $\tilde{w}$:

$$\begin{aligned} (i\partial_t + \partial_j T_{g^{jk}}^w \partial_k + T_b^w \cdot \nabla) \tilde{w} &= f - \mathcal{R}\bar{f} + \partial_j T_{g^{jk}}^w \partial_k \mathcal{R}\bar{w} \\ &\quad + \mathcal{R}\partial_j T_{g^{jk}}^w \partial_k \bar{w} - T_{\tilde{b}}^w \cdot \nabla \bar{w} + e[w]. \end{aligned} \quad (6.19)$$

Here the lower order terms $e[w]$ are given by

$$e[w] = T_b^w \cdot \nabla \mathcal{R}\bar{w} + \mathcal{R}T_b^w \cdot \nabla \bar{w} + \mathcal{R}T_{\tilde{b}}^w \cdot \nabla w + i\mathcal{R}_t \bar{w}. \quad (6.20)$$

To cancel the $\bar{w}$ terms on the right in (6.19), we use the ellipticity of g . Namely, our assumptions on g ensure that we have

$$g^{jk}\xi_j\xi_k \geq c_0|\xi|^2, \quad c_0 > 0.$$

We then select the symbol $r(x, \xi)$ of $\mathcal{R}$ to be

$$r(x, \xi; t) = (1 - \chi(|\xi|)) \frac{i\tilde{b}^j \xi_j}{2g^{jk}\xi_j\xi_k}, \quad (6.21)$$

where the time dependence is implicit in the b and g terms. Here χ is a smooth compactly supported bump function so that $\chi(\xi) = 1$ for ξ in a large neighborhood of 0, depending on M .

In view of the regularity properties (6.3) for $\tilde{b}$ and g and using the multiplicative bounds and Moser estimates in Proposition 4.2, it follows that the symbol r has regularity

$$r \in l^1 X^{s_0-1} S^{-1}, \quad r_t \in l^1 X^{s_0-3} S^{-1}. \quad (6.22)$$

Here, for symbol classes we follow the notations in [42]. For instance, by $r \in l^1 X^{s_0-1} S^{-1}$, we mean that for each $\xi \in \mathbb{R}^n$, we have $\|r(x, \xi; t)\|_{l^1 X^{s_0-1}} \lesssim (1 + |\xi|)^{-1}$, and that each ξ derivative gains one order of decay in ξ with x regularity that remains in $l^1 X^{s_0-1}$.

The paradifferential implementation of the symbol

$$r^P(x, \xi; t) = \sum_{\ell} S_{<\ell-4}(D_x) r(x, \xi; t) S_{\ell}(\xi)$$

is chosen such that the action of the operator only involves low-high interactions of the symbol with the function on which it is acting. This will allow us to prove mapping properties of $\mathcal{R} = r^P(t, x, D)$ at any regularity. Since without the χ cutoff r is a homogeneous symbol of order -1 , restricting to large frequencies helps to insure invertibility of the map $Sw = w + \mathcal{R}\bar{w}$. With this choice for r , we have

Lemma 6.6. *Suppose that w solves (6.2) with $g \in l^1 X^{s_0}$ that is uniformly elliptic, $b, \tilde{b} \in l^1 X^{s_0-1}$ as in (6.3) and $s_0 > d/2 + 2$. Let r be given by (6.21). Then the transformation S in (6.18) is invertible² in $L^2, l^1 H^s$ and $l^1 X^s$ for $0 \leq s \leq \infty$, and $\tilde{w}$ defined by (6.18) solves an equation of the form*

$$(i\partial_t + \partial_k T_{g^{kl}}^w \partial_l + T_b^w \cdot \nabla) \tilde{w} = \tilde{f}, \quad \tilde{w}(0, x) = \tilde{w}_0, \quad (6.23)$$

where

$$\|\tilde{f}\|_{Y^0} \lesssim_M \|f\|_{Y^0} + \|w\|_{L^2 L^2}, \quad \|\tilde{f}\|_{l^1 Y^s} \lesssim_M \|f\|_{l^1 Y^s} + \|w\|_{l^1 L^2 H^s}. \quad (6.24)$$

This lemma allows us to reduce the proof of Proposition 6.4 to the case when $\tilde{b} = 0$.

² With implicit bounds depending on M .

Proof. We recall the regularity of the symbol r in (6.22).

We begin with the mapping properties of $\mathcal{R}$ and the invertibility of the renormalization operator S in (6.18). It suffices to show that $\mathcal{R}$ is bounded with small norm in L^2 , $l^1 H^s$ and $l^1 X^s$. The proof is the same for all these spaces, so to fix the notations we consider $l^1 X^s$.

Since by construction we have $r \in l^1 X^{s_0-1} S^{-1}$, by separating variables it follows that we can represent the operator $\mathcal{R}$ as a rapidly convergent series of operators of the form

$$\mathcal{R} = \sum_{l>l_0} 2^{-l} \sum_{m=1}^{\infty} r_{<l-4}^m(t, x) S_l^m(D_x).$$

Here the subscript indicates the frequency localizations and m is the summation index for separation of variables. Choosing S_l^m to be uniformly bounded, we include the rapid decay in $r^m \in l^1 X^{s_0-1}$, which can then be assumed to satisfy bounds of the form

$$\|r^m\|_{l^1 X^{s_0-1}} \lesssim_M m^{-N}.$$

The dyadic multipliers S_l^k are bounded in all of our function spaces, so in order to obtain bounds for $\mathcal{R}$ it suffices to consider bilinear multiplicative bounds, as given in Proposition 4.2. More precisely, the bound (4.5) shows that

$$\|r_{<l-4}^m(t, x) S_l^m(D_x) u\|_{l^1 X^s} \lesssim \|r^m\|_{l^1 X^{s_0-1}} \|S_l^m u\|_{l^1 X^s},$$

which after summation in m and also in the dyadic index l yields

$$\|\mathcal{R}u\|_{l^1 X^s} \lesssim_M 2^{-l_0} \|u\|_{l^1 X^s}. \quad (6.25)$$

Here smallness is gained by making l_0 large enough, which in turn is accomplished through the choice of the cutoff χ in the definition of the symbol r . If we instead use the 2^{-l} factor to gain Sobolev regularity rather than smallness, then we obtain

$$\|\mathcal{R}w\|_{l^1 X^s} \lesssim_M \|w\|_{l^1 X^{s-1}}, \quad (6.26)$$

as well as the related fixed time L^2 type bounds

$$\|\mathcal{R}w\|_{H^\sigma} \lesssim_M \|w\|_{H^{\sigma-1}}, \quad \sigma \in \mathbb{R}. \quad (6.27)$$

We similarly have the closely related fixed time bound for $\mathcal{R}_t$,

$$\|\mathcal{R}_t w\|_{L^2} \lesssim_M \|w\|_{L^2} \quad (6.28)$$

as $r_t \in l^1 X^{s_0-3} S^{-1} \subset C^{-1+} S^{-1}$ (provided that $s_0 > d/2 + 2$).

In view of the bound (6.25), the operator S is invertible on $l^1 X^s$ for $0 \leq s \leq \infty$. A similar analysis holds for L^2 and $l^1 H^s$.

Now we consider the source term bound (6.24). From (6.19), (6.20) we have (6.23) with

$$\tilde{f} = f - \mathcal{R}\tilde{f} + E\tilde{w} + e[w]$$

where

$$\begin{aligned} E &= \partial_j T_{g^{jk}}^w \partial_k \mathcal{R} + \mathcal{R} \partial_j T_{g^{jk}}^w \partial_k - T_{\tilde{b}}^w \cdot \nabla, \\ e[w] &= T_{\tilde{b}}^w \cdot \nabla \mathcal{R} \bar{w} + \mathcal{R} T_{\tilde{b}}^w \cdot \nabla \bar{w} + \mathcal{R} T_{\tilde{b}}^w \cdot \nabla w + i \mathcal{R}_t \bar{w}. \end{aligned}$$

We need to establish the appropriate bounds for each of the terms on the right.

For $\mathcal{R}\tilde{f}$ it suffices to see that $\mathcal{R}$ is bounded on Y^0 , and respectively on $l^1 Y^s$,

$$\|\mathcal{R}\tilde{f}\|_{Y^0} \lesssim_M \|f\|_{Y^0}, \quad \|\mathcal{R}\tilde{f}\|_{l^1 Y^s} \lesssim_M \|f\|_{l^1 Y^s}.$$

This follows in the same way as (6.25), but using the high-low frequency interactions we have the straightforward estimate

$$\|S_{<\ell-4} u S_{\ell} v\|_{Y^0} \lesssim \|u\|_{L^\infty} \|S_{\ell} v\|_{Y^0} \lesssim \|u\|_{l^1 X^{s_0}} \|S_{\ell} v\|_{Y^0} \quad (6.29)$$

instead of Proposition 4.2. A similar computation applies with Y^0 replaced by $l^1 Y^s$.

For the remaining terms in $\tilde{f}$, it suffices to prove L^2 type bounds, namely,

$$\|E\bar{w}\|_{L^2 L^2} + \|e[w]\|_{L^2 L^2} \lesssim_M \|w\|_{L^2 L^2}, \quad (6.30)$$

and

$$\|E\bar{w}\|_{l^1 L^2 H^s} + \|e[w]\|_{l^1 L^2 H^s} \lesssim_M \|w\|_{l^1 L^2 H^s}. \quad (6.31)$$

Since all operators involved are paradifferential, the analysis happens at fixed frequency so the two bounds are virtually identical. Thus we will focus on (6.30).

After a dyadic Littlewood-Paley localization, the bound for E in (6.30) reduces to a paraproduct type product formula for scalar multiplications

$$\begin{aligned} \|[r_{<l-4} g_{<l-4} - (rg)_{<l-4}] w_l\|_{L^2} &\lesssim 2^{-l} \|r\|_{C^1} \|g\|_{C^1} \|w_l\|_{L^2} \\ &\lesssim \|r\|_{l^1 X^{s_0-1}} \|g\|_{l^1 X^{s_0}} \|w_l\|_{L^2}. \end{aligned}$$

Here, the two terms $(rg)_{<l-4}$ cancel the high frequency contributions of $T_{\tilde{b}}^w \cdot \nabla w$ and the remaining low frequency bound on $\chi T_{\tilde{b}}^w \cdot \nabla w$ is easily bounded.

Finally, we consider the bound for $e[w]$ in (6.30). For the $\mathcal{R}$ terms in $e[w]$ it suffices to use the bound in (6.27), while for the $\mathcal{R}_t$ term in $e[w]$ we need the bound in (6.28). $\square$

6.3. The Incoming Estimate

To motivate the definitions that follow we briefly consider the constant coefficient case, where $g = I_d$. Then the Hamilton flow associated to the linear constant coefficient Schrödinger equation has the form

$$\dot{x} = 2\xi.$$

Hence rays are straight lines, which approach the origin when $x \cdot \xi < 0$, or, in other words, as long as the angle $\theta = \angle(x, \xi)$ satisfies $\cos \theta < 0$.

In the problem we are considering, the coefficients are not constant but are small in the exterior region, and, as seen in Section 5, the direction for the bicharacteristics does not deviate much from being constant. Thus we can define the incoming part of a solution w using a pseudodifferential truncation,

$$w_{in} = P_{in}(x, D)w,$$

where the symbol of P_{in} is chosen as

$$p_{in}(x, \xi) = \chi_{in}(\cos \theta) \chi_{>5R}(|x|).$$

Here χ_{in} is a nonincreasing cutoff which selects the interval $[-\infty, -\frac{3}{8})$, and $\chi_{>5R}$ is nondecreasing and selects the exterior of $5B_R$. Then the aim of this section is to establish the bound

$$\|\chi_{<100R} w_{in}\|_{X^0} \lesssim \|w_0\|_{L^2} + \|f\|_{Y^0} + C(M)\|w\|_{L^2 L^2}. \quad (6.32)$$

We remark here that one could remove the $\chi_{<100R}$ truncation, as well as the M dependence in the constant in the last term. But this would require extra work, and the result is not needed in our sequence of steps.

As a first minor simplification, we can use the small data results of [26] (see Theorem 6.5 above) to reduce the problem to the case when the portion of f outside of B_{3R} need only be measured in $L^2 L^2$:

$$\chi_{>3R} f \in L^2 L^2. \quad (6.33)$$

By truncating the coefficients $g - I_d$ and b outside B_R ,

$$g_{ext} - I_d = \chi_{>R}(g - I_d), \quad b_{ext} = \chi_{>R}b$$

we obtain coefficients $g_{ext} - I_d, b_{ext}$ which are small overall in the norms of (6.3) and coincide with $g - I_d, b$ in the exterior region. Then we solve the auxiliary problem

$$(i\partial_t + \partial_k T_{g_{ext}}^{w} \partial_l + T_{b_{ext}}^w \cdot \nabla) \tilde{w} = f, \quad \tilde{w}(0) = 0. \quad (6.34)$$

By Theorem 6.5, $\tilde{w}$ satisfies a global favorable estimate,

$$\|\tilde{w}\|_{X^0} \lesssim \|f\|_{Y^0}. \quad (6.35)$$

On the other hand, for the difference

$$w_1 = w - \chi_{>2R} \tilde{w},$$

we have the equation

$$(i\partial_t + \partial_k T_{g_{ext}}^w \partial_l + T_b^w \cdot \nabla) w_1 = (1 - \chi_{>2R})f + f_1 + f_2 + f_3$$

where f_1 arises from the change of metric,

$$f_1 = -(\partial_k T_{g_{ext}}^{kl} - g_{ext}^{kl} \partial_l + T_{b-b_{ext}} \cdot \nabla) \chi_{>2R} \tilde{w}$$

and f_2, f_3 from the localization of $\tilde{w}$ to the exterior region,

$$\begin{aligned} f_2 &= (1 - \chi_{>3R})[\chi_{>2R}, (\partial_k T_{g_{ext}^{kl}} \partial_l + T_{b_{ext}} \cdot \nabla)](1 - \chi_{>3R})\tilde{w}. \\ f_3 &= \chi_{>3R}[\chi_{>2R}, (\partial_k T_{g_{ext}^{kl}} \partial_l + T_{b_{ext}} \cdot \nabla)](1 - \chi_{>3R})\tilde{w} \\ &\quad + [\chi_{>2R}, (\partial_k T_{g_{ext}^{kl}} \partial_l + T_{b_{ext}} \cdot \nabla)]\chi_{>3R}\tilde{w}. \end{aligned}$$

It is easily seen that f_1 is a Schwartz function in x since the functions $\chi_{>2R}$ and $g^{kl} - g_{ext}^{kl}, b - b_{ext}$ have separated localizations. In particular we have

$$\|f_1\|_{L^2} \lesssim \|\tilde{w}\|_{L^2 L^2}. \quad (6.36)$$

A similar bound holds for f_3 due to the separated localizations of $\chi_{>3R}$ and $\nabla \chi_{>2R}$. In particular we have

$$\|f_3\|_{L^2} \lesssim \|\tilde{w}\|_{L^2 L^2}. \quad (6.37)$$

The function f_2 , on the other hand, is localized and satisfies

$$\|f_2\|_{Y^0} \lesssim \|\chi_{<4R}\tilde{w}\|_{L^2 H^{1/2}} \quad (6.38)$$

to which we may subsequently apply (6.35). This completes our reduction to the case when (6.33) holds.

Now we return to the estimate for w_{in} , under the additional assumption (6.33) on f . The main step in the proof of this bound is based on the positive commutator method, using a well chosen order zero formally self-adjoint pseudodifferential multiplier $Q_{in} \in OPS^0$ with symbol supported in the incoming region.

Writing the second order part of the Schrödinger operator in divergence form, we will use the notations

$$\partial_k T_{g^{kl}}^w \partial_l + T_{b^j}^w \partial_j = P + B$$

where the principal part P is self-adjoint. To understand the choice of Q_{in} , which we take to be a self-adjoint operator of order zero that is independent of t , we first compute formally using the L^2 inner product in $\mathbb{R}^d$:

$$\begin{aligned} &\frac{d}{dt} \operatorname{Re} \langle Q_{in} w, w \rangle \\ &= i \langle Q_{in} (P + B) w, w \rangle - i \langle Q_{in} w, (P + B) w \rangle + 2 \operatorname{Re} \langle Q_{in} w, -if \rangle \\ &= \langle i([Q_{in}, P] + Q_{in} B - B^* Q_{in}) w, w \rangle + 2 \operatorname{Re} \langle Q_{in} w, -if \rangle \\ &:= -\langle C w, w \rangle + 2 \operatorname{Re} \langle Q_{in} w, -if \rangle. \end{aligned}$$

Here the operator C is an order one self-adjoint pseudodifferential operator. Integrating between 0 and T with $T \leq 1$ we obtain

$$\begin{aligned} \operatorname{Re} \langle Q_{in} w, w \rangle(T) + \int_0^T \langle C w, w \rangle dt &= \operatorname{Re} \langle Q_{in} w, w \rangle(0) \\ &\quad + 2 \int_0^T \operatorname{Re} \langle Q_{in} w, -if \rangle dt. \end{aligned} \quad (6.39)$$

The first term on the right is easily estimated in terms of the L^2 norm of the data,

$$\operatorname{Re}\langle Q_{in}w, w\rangle(0) \leq \|w(0)\|_{L^2}^2.$$

For the second term on the right we can take advantage of (6.33) and use the fact that we shall take Q_{in} to be supported³ where $|x| > 4R$ to obtain

$$\operatorname{Re}\langle Q_{in}w, if\rangle \lesssim \|w\|_{L^2}(\|\chi_{>3R}f\|_{L^2} + \|\chi_{<4R}f\|_{Y^0}),$$

where the interior Y^0 bound arises to account for smoothing tails.

Our goal therefore is to choose the operator Q_{in} favorably so that we can prove a good bound from below for the left hand side of (6.39). This is easily done for the first term, where it suffices to impose the conditions

$$q(x, \xi) \geq 0, \quad q(x, \xi) \gtrsim 1 \quad \text{in } \{\cos \theta < -\frac{1}{4}, |x| > 4R\}. \quad (6.40)$$

Then Gårding's inequality shows that

$$\operatorname{Re}\langle Q_{in}w, w\rangle(T) \gtrsim \|w_{in}(T)\|_{L^2}^2 - C\|w(T)\|_{H^{-\frac{1}{2}}}^2, \quad (6.41)$$

where the last term can be further estimated in a naive fashion by the energy type relation

$$\|w(T)\|_{H^{-\frac{1}{2}}}^2 \lesssim \|w(0)\|_{H^{-\frac{1}{2}}}^2 + \|f\|_{L^2H^{-1}}^2 + C(M)\|w\|_{L^2L^2}^2. \quad (6.42)$$

Indeed, (6.42) results from considering $2\operatorname{Im}\langle (i\partial_t + \partial_k T_{gkl}^w \partial_l + T_b^w \cdot \nabla)w, \langle D \rangle^{-1}w \rangle$. Note, in this estimate there will be implicit dependence upon $\|b\|_{L^\infty}$ and hence on our large constant M .

We next consider the time integral involving the operator C . For this we will seek to carefully apply Gårding's inequality in order to prove the bound

$$\int_0^T \langle Cw, w \rangle dt \gtrsim \|\chi_{<100R}w_{in}\|_{X^0}^2 - \|w\|_{L^2L^2}^2. \quad (6.43)$$

Together with (6.41), this would complete the proof of the desired estimate (6.32). The rest of this section is concerned with the choice of Q_{in} so that the above estimate holds.

Within the ball B_{100R} , the X norm is equivalent to the $L^2H^{\frac{1}{2}}$ norm, therefore at least heuristically the principal symbol $c_0(x, \xi)$ of C should satisfy

$$c_0(x, \xi) \geq 0, \quad c_0(x, \xi) \gtrsim |\xi| \quad \text{in } \{\cos \theta < -\frac{1}{4}, 100R > |x| > 4R\}. \quad (6.44)$$

However, the matters are a bit more delicate because of the coefficients g and b which have limited regularity.

³ Here we harmlessly gloss over the difference between the support of the symbol and that of the kernel; this can be readily rectified with an $OPS^{-\infty}$ adjustment to Q_{in} .

We write C in the form

$$C = i[\Delta, Q_{in}] + (i[(P - \Delta), Q_{in}] + iQ_{in}B - iB^*Q_{in}) := C^{main} + C^{err}$$

and begin with computing the principal symbol of C , which is

$$\begin{aligned} c_0(x, \xi; t) = & -2\xi \cdot \partial_x q_{in}(x, \xi) - \{(g^{ij} - \delta^{ij})\xi_i \xi_j, q_{in}(x, \xi)\} \\ & - 2 \operatorname{Im} b(t, x, \xi) q_{in}(x, \xi) := c_0^{main} + c_0^{err}. \end{aligned}$$

Note that, as g, b depend upon both x and t , c_0 also has some time dependence. However, as we will only rely upon the spatial regularity of those terms for the analysis of this operator, we will drop the explicit reference to the t dependence below for convenience. We will choose q_{in} so that the first term yields a bounded nonnegative contribution that also controls the remaining terms.

Precisely, our assumption (6.4) for the coefficients $g - I_d$ and b of our Schrödinger operator guarantees that we have a uniform pointwise bound

$$|g - I_d| + |\nabla g| + |b| \lesssim \varepsilon \mu_k^2, \quad R \leq k < |x| < k + 1,$$

where the sequence $\{\mu_k\}$ is square summable,

$$\sum \mu_k^2 \lesssim 1, \quad \mu_R \approx 1.$$

Without any restriction in generality we also assume that the sequence μ_k is slowly varying. Then, we can find a increasing function

$$\rho_R: [R, \infty) \rightarrow [1, 2]$$

so that

$$\rho'_R(r) \gtrsim \mu_k^2, \quad \rho_R^{(j)}(r) \lesssim_j \mu_k^2, \quad j \geq 2, \quad r \in [k, k + 1].$$

Using this weight ρ_R , we define the symbol q_{in} as

$$q_{in}(x, \xi) = \rho_R(r) \chi_{>5R}(r) \chi_{in}(\cos \theta - c \rho_R(r)),$$

where $\chi_{in}(\rho)$ is nonincreasing, supported in $\{\rho < -1/4\}$, and identically 1 on $\{\rho < -1/2\}$. The constant c is small and satisfies $\varepsilon \ll c \ll 1$. With this choice, it suffices to prove the following:

Lemma 6.7. *With q_{in} chosen as above, the estimate (6.43) holds.*

Proof. The leading part c_0^{main} of c_0 is given by

$$\begin{aligned} \frac{1}{2} c_0^{main}(x, \xi) = & -|\xi| \cos \theta \chi_{in}(\cos \theta - c \rho_R(r)) (\rho_R \chi_{>5R})'(r) \\ & - \chi'_{in}(\cos \theta - c \rho_R(r)) \frac{|\xi|}{r} \sin^2 \theta (\rho_R \chi_{>5R})(r) \\ & + c \chi'_{in}(\cos \theta - c \rho_R(r)) |\xi| \cos \theta (\rho_R \chi_{>5R})(r) \rho'_R(r). \end{aligned}$$

Provided c is sufficiently small, all three terms are nonnegative, and we get the size

$$\begin{aligned} c_0^{main}(x, \xi) &\approx |\xi| \chi_{in}(\cos \theta - c\rho_R(r))(\rho_R \chi_{>5R})'(r) \\ &\quad - \chi'_{in}(\cos \theta - c\rho_R(r)) \frac{|\xi|}{r} (\rho_R \chi_{>5R})(r) \\ &\quad - c \chi'_{in}(\cos \theta - c\rho_R(r)) |\xi| (\rho_R \chi_{>5R})(r) \rho'_R(r). \end{aligned}$$

On the other hand, for the remaining terms in c_0 we have a favorable bound

$$|c_0^{err}| \lesssim \varepsilon (|\xi| \chi_{in}(\cos \theta - c\rho_R(r))(\rho_R \chi_{>5R})'(r) d - \rho'_R(r) \chi'_{in}(\cos \theta - c\rho_R(r)) |\xi| (\rho_R \chi_{>5R})(r)). \quad (6.45)$$

It is easier to argue in the case of C^{main} , which belongs to OPS^1 . Since its principal symbol is nonnegative and of size ξ within the region

$$\left\{ 4R < |x| < 100R, \cos \theta < -\frac{1}{2} \right\},$$

by the classical Gårding inequality for S^1 symbols we have the fixed time bound

$$\begin{aligned} \langle C^{main} w, w \rangle &= \langle (c_0^{main})^w(x, D)w, w \rangle + O(\|w\|_{L^2}^2) \gtrsim \|\chi_{<100R} w_{in}\|_{H^{\frac{1}{2}}}^2 - C\|w\|_{L^2}^2. \end{aligned} \quad (6.46)$$

The similar bound for C^{err} is slightly more delicate. We will show that it satisfies the fixed time bound

$$\langle C^{err} w, w \rangle \lesssim \varepsilon (\langle (c_0^{main})^w(x, D)w, w \rangle + \|w\|_{L^2}^2). \quad (6.47)$$

If we have this, then combining the bounds (6.46) and (6.47) we obtain

$$\langle Cw, w \rangle \gtrsim \langle (c_0^{main})^w(x, D)w, w \rangle - C\|w\|_{L^2}^2 \gtrsim \|\chi_{<100R} w_{in}\|_{H^{\frac{1}{2}}}^2 - C\|w\|_{L^2}^2,$$

which after time integration yields (6.43) and in turn gives (6.32). It remains to prove (6.47).

The difficulty here is that the symbol $c^{err} \in C^1 S^1$ since it involves spatial coefficients depending upon b and the Poisson bracket with coefficients in g (and hence involves derivatives of g), which means we only have bounds in X^{s_0-1} and cannot guarantee enough regularity to allow us to directly use Gårding's inequality, see [38]. Instead we will make a more careful symbol analysis. We first reconsider c_0^{main} , for which we write a sum of squares decomposition,

$$c_0^{main} = \phi_1^2 + \phi_2^2 + \phi_3^2,$$

where $\phi_1, \phi_2, \phi_3 \in S^{\frac{1}{2}}$ are smooth nonnegative symbols given by

$$\begin{aligned} \phi_1^2 &= -|\xi| \cos \theta \chi_{in}(\cos \theta - c\rho_R(r))(\rho_R \chi_{>5R})'(r), \\ \phi_2^2 &= -\chi'_{in}(\cos \theta - c\rho_R(r)) \frac{|\xi|}{r} \sin^2 \theta (\rho_R \chi_{>5R})(r), \\ \phi_3^2 &= c|\xi| \cos \theta \chi'_{in}(\cos \theta - c\rho_R(r))(\rho_R \chi_{>5R})(r) \rho'_R(r). \end{aligned}$$

We also consider a fourth nonnegative symbol $\phi_4 \in S^{\frac{1}{2}}$ given by

$$\phi_4^2 = |\xi| \rho'_R(r) q_{in}(x, \xi).$$

By pseudodifferential calculus we have the fixed time bound

$$\begin{aligned} \langle C^{main} w, w \rangle \\ = \|\Phi_1(x, D)w\|_{L^2}^2 + \|\Phi_2(x, D)w\|_{L^2}^2 + \|\Phi_3(x, D)w\|_{L^2}^2 + O(\|w\|_{L^2}^2). \end{aligned}$$

On the other hand,

$$\phi_4^2 \lesssim \phi_1^2 + \phi_2^2 + \phi_3^2.$$

Therefore, by Gårding's inequality, we have

$$\begin{aligned} \|\Phi_4(x, D)w\|_{L^2}^2 \\ \lesssim \|\Phi_1(x, D)w\|_{L^2}^2 + \|\Phi_2(x, D)w\|_{L^2}^2 + \|\Phi_3(x, D)w\|_{L^2}^2 + O(\|w\|_{L^2}^2). \end{aligned}$$

Now we consider the symbol $c^{err} \in C^1 S^1$. Modulo an L^2 bounded $C^0 S^0$ component we can replace it with its principal part c_0^{err} . Given its expression, it is easily seen that we can use the above ϕ_j 's to represent the principal part c_0^{err} in the form

$$c_0^{err} = \sum_{j=1}^4 d_j(x, \xi) \phi_j^2(x, \xi), \quad d_j \in \varepsilon C^1 S^0,$$

which is a more careful substitute for (6.45). Then at the operator level we can write

$$c_0^{err}(x, D) = \sum_{j=1}^4 \Phi_j(x, D)^* D_j(x, D) \Phi_j(x, D) + \varepsilon OPC^0 S^0,$$

which yields the bound

$$\langle C^{err} w, w \rangle \lesssim \sum_{j=1}^4 \|\Phi_j(x, D)w\|_{L^2}^2 + \|w\|_{L^2}^2.$$

Thus (6.47) follows, and the proof is complete. $\square$

6.4. Nontrapping Estimates on $B(0, R)$

Here we use the nontrapping condition to produce a high frequency bound for w within the compact set $B(0, R)$, in terms of the incoming part w_{in} estimated in the previous subsection. Precisely, we will show that

$$\|\chi_{<R} w\|_{X^0} \lesssim e^{C(M)L} (\|w_0\|_{L^2} + \|\chi_{<100R} f\|_{Y^0} + \|\chi_{<100R} w_{in}\|_{X^0} + \|w\|_{L^2 L^2}). \quad (6.48)$$

To clarify the meaning of the norms in (6.48), we recall that within a compact set (e.g. $100B_R$) the X^0 norm is equivalent to the $L_t^2 H^{\frac{1}{2}}$ norm, while the Y^0 norm is equivalent to the $L^2 H^{-\frac{1}{2}}$ norm.

This proof also uses a positive commutator argument, based on propagation of singularities in $B(0, 100R)$. The key idea is that any bicharacteristic ray which enters $2B_R$ is coming from the phase space support of $\chi_{<100R} w_{in}$. From a purely qualitative perspective, the nontrapping condition implies that such an estimate must hold, with a suitable implicit constant. The challenge is to carefully track the constant.

For the positive commutator argument we use again a nonnegative pseudo-differential multiplier $Q_{comp} \in OPS^0$, whose symbol is this time supported in $100B(0, R)$, and repeat the computation leading to (6.39). We split the analysis into an ode part, where we construct the symbol for Q_{comp} , and a microlocal part, where we use the properties of symbol in order to prove the desired estimate (6.48). We begin with some heuristic considerations.

- The size of the frequency ξ may vary considerably along the Hamilton flow of the operator $g^{ij} \xi_i \xi_j$. To avoid difficulties arising from this, we will take the symbol q_{comp} to be homogeneous in ξ for $|\xi| \gg 1$.
- The time t also varies along the Hamilton flow, which would seem to require a time dependent construction of q_{comp} . However, the propagation speed is proportional to the frequency size $|\xi|$, therefore in the high frequency limit the time is constant along the flow. Because of this, we will only use the fixed time flow in the q construction.

To construct q_{comp} we will only use the principal symbol for the Schrödinger operator,

$$a(x, \xi) = g^{ij} \xi_i \xi_j$$

and

$$H_a = \sum_{j=1}^d \partial_{\xi_j} a \partial_{x_j} - \partial_{x_j} a \partial_{\xi_j}$$

its Hamiltonian vector field. We summarize our result as follows:

Proposition 6.8. *Assume that the coefficients g^{ij} satisfy the conditions in (6.3), (6.4). Moreover, we assume that the flow (5.10) permits no trapped geodesics, with the longest within B_{2R} with $|\xi| = 1$ being of length L . Let $C \gg 1$ be a large universal constant.*

Then there exists a smooth, homogeneous, real-valued, nonnegative symbol $q \in S^0$ in the region $|x| \leq 100R$ with the following properties:

(1) *Support property:*

$$\text{supp } q \cap \{|x| > 50R\} \subset \{\cos \theta < -\frac{1}{2}\}.$$

(2) *Size:*

$$|\partial_t^j \partial_x^\alpha \partial_\xi^\beta q(x, \xi)| \lesssim e^{CML} |\xi|^{-|\beta|}. \quad (6.49)$$

(3) *Positive commutator:*

$$-H_a q \geq CM|\xi|q \quad (6.50)$$

and

$$-H_a q \geq |\xi| \quad \text{inside } \{|x| < 2R\} \quad (6.51)$$

(4) *Bounded gradient,*

$$|q_x| + |\xi||q_\xi| \lesssim e^{CML} H_a q. \quad (6.52)$$

Proof. From the conditions (6.3) we have the following regularity properties for g in a compact set:

$$\|g\|_{C^{2,\delta}} + \|g_t\|_{C^\delta} \lesssim M \quad \text{in } B(0, R)$$

respectively

$$\|g\|_{C^{2,\delta}} + \|g_t\|_{C^\delta} \lesssim \varepsilon \quad \text{outside } B(0, R).$$

We will only use these properties in the proof of the proposition.

This would be a standard construction for smooth g ; similar constructions have already been done in [35] as well as in Proposition 3.6 in [25]. The difficulty we encounter in the nonsmooth case is that a direct construction based on the Hamilton flow of a will yield a nonsmooth q . There are two possible strategies here, to regularize g and then construct q or vice versa; both work, but we choose the former.

Since we seek q homogeneous, we will work on the cosphere bundle $S^*\mathbb{R}^d = \{(x, \xi); |\xi| = 1\}$ using the notion of flow in (5.10).

We begin by regularizing g . Given a frequency scale λ , to be chosen later, we regularize g to g_λ at scale λ in x and at scale λ^2 in t . Then our uniform bounds above imply that

$$|g - g_\lambda| \lesssim \lambda^{-2}, \quad |\nabla g - \nabla g_\lambda| \lesssim \lambda^{-1}.$$

Next we compare their Hamilton flows $\Phi(s; x_0, \xi_0)$, respectively $\Phi_\lambda(s; x_0, \xi_0)$ starting at any point (x_0, ξ_0) with $x_0 \in 100B_R$.

Lemma 6.9. *Assume that*

$$\lambda \geq e^{2CML}, \quad C \gg 1.$$

Then the flows of $\tilde{H}_a$ and $\tilde{H}_{a_\lambda}$, defined as in (5.10), stay close

$$|\Phi(s; x_0, \xi_0) - \Phi_\lambda(s; x_0, \xi_0)| \lesssim e^{-CML}$$

until exiting $100B_R$.

Proof. Recall from (5.10), that the flow $\tilde{H}_a$ can be written

$$\tilde{H}_a = \sum_{j=1}^d \partial_{\xi_j} a \partial_{x_j} - (\partial_{x_j} a - (\nabla a \cdot \xi) \xi_j) \partial_{\xi_j}. \quad (6.53)$$

This then follows directly from a small modification of the arguments in Section 5, Proposition 5.2. $\square$

In particular this shows that for such λ the $\tilde{H}_{a_\lambda}$ flow is also uniformly nontrapping, with comparable parameters and hence we may choose L such that it is a bound on the longest geodesic for both flows within B_R .

The first element of our construction is a smooth 0 homogeneous nonnegative symbol χ in $\{|x| < 100R\}$ that is incoming relative to the flat metric, i.e. $a_0 = |\xi|^2$. This is chosen akin to the incoming localization in the previous subsection, namely q_{in} , but shifted so that its support includes $B(0, 2R)$, i.e. for instance

$$\chi(x, \xi) = \chi_{>2R}(|x - 8R\xi|) \chi_{<-1/2}(\cos \angle(x - 8R\xi, \xi)), \quad |\xi| = 1.$$

The three important properties of this symbol are as follows:

- Incoming relative to a_0 ,

$$-H_{a_0} \chi \gtrsim |\nabla \chi| \text{ in } 100B_R. \quad (6.54)$$

- Covers $2B_R$,

$$\chi \gtrsim 1 \quad \text{in } 2B_R.$$

- Narrower than the previous incoming multiplier q_{in} ,

$$\text{supp } \chi \cap \{|x| > 50R\} \subset \left\{ \cos \theta < -\frac{1}{2} \right\}.$$

We use χ to construct our smooth nonnegative symbol q , by solving the ODE

$$-\tilde{H}_{a_\lambda} q = CMq + \chi$$

with initial data set by the condition

$$\text{supp } q \subset \text{supp } \chi.$$

In other words, we solve the ode backwards along the $\tilde{H}_{a_\lambda}$ flow, beginning when χ is first encountered. Note that here q will depend upon t since a does. Since a is of class C^2 , the Hamilton flow of $\tilde{H}_{a_\lambda}$ is well-posed. Since the C^2 norm of a_λ is bounded by M and the longest trajectory has length L (within $100B_R$), by Grönwall it follows that

$$|q| + |\nabla q| \lesssim e^{CML}$$

for $x \in 100B_R$. We can also estimate higher regularity for q , losing only powers of λ when differentiating in x, ξ , respectively λ^2 for time derivatives. Thus, if λ is chosen so that $\lambda \approx e^{CML}$, then q already has all the properties in the Proposition 6.8, but relative to a_λ . In order to switch from a_λ to a we need one more piece of information, namely that

$$-\tilde{H}_{a_\lambda} q \gtrsim e^{-CML} |\nabla q|. \quad (6.55)$$

To prove (6.55), we compute the ODEs for the two quantities,

$$-\tilde{H}_{a_\lambda}(\nabla q) = CM \nabla q + O(|\nabla^2 a|) \nabla q + \nabla \chi,$$

whereas

$$-\tilde{H}_{a_\lambda} \tilde{H}_{a_\lambda} q = CM \tilde{H}_{a_\lambda} q - \tilde{H}_{a_\lambda} \chi,$$

and (6.55) is easily seen to follow by the comparison principle for ODE's in view of (6.54).

Now we proceed to the final step of our construction, which is to replace a_λ by a . We have

$$|\tilde{H}_{a_\lambda} q - \tilde{H}_a q| \lesssim |\nabla a - \nabla a_\lambda| |\nabla q| \lesssim M \lambda^{-1} e^{CML} |\nabla q| \lesssim -e^{-CML} \tilde{H}_{a_\lambda} q$$

which suffices. $\square$

We define the symbol q_{comp} by fine tuning the symbol q constructed above in the exterior region. First we consider a symbol $\tilde{\chi}$ which is akin to χ but with slightly larger support, so that

$$\tilde{\chi} \gtrsim 1 \quad \text{in supp } q.$$

As in the previous subsection, we have $-H_{a_0} \tilde{\chi} \gtrsim \chi$, and also the sums of squares representation

$$-H_{a_0} \tilde{\chi} = \phi_1^2 + \phi_2^2, \quad \chi = \phi_3^2$$

for smooth nonnegative symbols $\phi_j \in S^{\frac{1}{2}}$.

Now we define

$$\tilde{q} = q + \tilde{\chi}, \quad q_{comp} = \chi_{<75R}(|x|) \tilde{q}.$$

Here the symbol $\tilde{q}$ inherits from q all the properties in Proposition 6.8, and we note that $\tilde{q}$ and q_{comp} depend upon t through q and hence through a as above. This is because outside $2B_R$ the symbol $\tilde{\chi}$ has similar properties, while inside $2B_R$

the contribution of the symbol $\tilde{\chi}$ is small compared to that of q . The reason we introduce $\tilde{q}$ is related to the low regularity of the coefficients, which causes us to once again replace Gårding's inequality with more robust sums of squares methods. Precisely, we will have a representation of $\tilde{q}$ and $\nabla \tilde{q}$ in terms of the above squares,

$$\tilde{q} = d_3 \phi_3^2, \quad (q_x, |\xi| q_\xi) = d_1 \phi_1^2 + d_2 \phi_2^2$$

with $d_j \in e^{CML} S^0$. The bound on d_j 's is exponentially large, but that will suffice later on. One consequence of it is that

$$-H_{a_\mu} q - CMq \gtrsim e^{-CML} \sum \phi_j^2, \quad \mu \geq \lambda. \quad (6.56)$$

On the other hand the cutoff $\chi_{<75R}(|x|)$ achieves the goal of having q_{comp} compactly supported, at the price of violating the positive commutator condition (3) in Proposition 6.8 in the region $\{50R < |x| < 100R\}$.

We now show that the above choice for the symbol q_{comp} yields the bound (6.48). We start with the counterpart of (6.39), namely

$$\begin{aligned} \operatorname{Re} \langle Q_{comp} w, w \rangle(T) + \int_0^T \langle Cw, w \rangle dt &= \operatorname{Re} \langle Q_{comp} w, w \rangle(0) \\ &+ 2 \operatorname{Re} \int_0^T \langle Q_{comp} w, -if \rangle + \langle Q_{comp,t} w, w \rangle dt \end{aligned} \quad (6.57)$$

where C is now given by

$$C = i[P, Q_{comp}] + iQ_{comp}B - iB^*Q_{comp}.$$

The terms on the right are easily bounded as

$$\operatorname{Re} \langle Q_{comp} w, w \rangle(0) \lesssim e^{CML} \|w(0)\|_{L^2}^2,$$

and

$$\operatorname{Re} \int_0^T \langle Q_{comp,t} w, w \rangle dt \lesssim e^{CML} \|w\|_{L^2 L^2}^2,$$

and

$$\int_0^T \langle Q_{comp} w, -if \rangle dt \lesssim \|Q_{comp} w\|_{L^2 H^{\frac{1}{2}}} \|\chi_{<100R} f\|_{L^2 H^{-\frac{1}{2}}} + T^{\frac{1}{2}} \|w\|_{L^\infty L^2} \|f\|_{Y^0},$$

where the second term on the right accounts for the smoothing, rapidly decreasing tails arising from the contribution of $\chi_{>100R} f$. The first term on the left is estimated by Gårding's inequality,

$$\langle Q_{comp} w, w \rangle(T) \gtrsim \|\chi_{<2R} w(T)\|_{L^2}^2 - e^{CML} \|w(T)\|_{H^{-1/2}}^2,$$

where the right hand side is further estimated by (6.42).

It remains to consider the contribution of C , where we again have the difficulty of having to deal with low regularity coefficients. In order to deal with this, we first truncate the coefficients at the scale $\mu = 2^{k_0} > \lambda$, where $\lambda = e^{CML}$ is the scale

used earlier in the proof of the Proposition. Then we consider the contribution of P_μ and B_μ . This is given by

$$C_\mu = i[P_\mu, Q_{comp}] + iQ_{comp}B_\mu - iB_\mu^*Q_{comp}.$$

Then $C_\mu \in OPC^1S^1 \cap \mu OPC^2S^1$, with principal symbol

$$c_{\mu,0} = -\{a_\mu, q_{comp}\} + 2q_{comp} \operatorname{Im} b_\mu.$$

By the above proposition, this satisfies the bound from below

$$|\xi|^{-1}c_{\mu,0}(x, \xi) \geq c\chi_{<100R}(|x|)\chi(x, \xi) - e^{CML}\chi_{<100R}p_{in}(x, \xi).$$

Then Gårding's inequality yields the fixed time bound (see [38])

$$\langle C_\mu w, w \rangle \gtrsim c\|\chi_{<2R}w\|_{H^{\frac{1}{2}}}^2 - e^{CML}\|\chi_{<100R}w_{in}\|_{H^{\frac{1}{2}}}^2 - \mu e^{CML}\|w\|_{L^2L^2}^2.$$

Here the μ factor in the last term arises due to the fact that the C^2S^1 symbol regularity is needed for Gårding's inequality. For later use, we record another consequence of Gårding's inequality. Precisely, by (6.56) we get

$$c_{\mu,0}q \gtrsim e^{-CML} \sum \phi_j^2$$

which gives

$$\langle C_\mu w, w \rangle \gtrsim e^{-CML} \sum \|\Phi_j w\|_{L^2}^2 - \mu e^{CML}\|w\|_{L^2}^2. \quad (6.58)$$

It remains to consider the contribution of $C - C_\mu$, for which it suffices to prove the bound

$$\langle (C - C_\mu)w, w \rangle \lesssim \mu^{-\delta}e^{CML}(\langle C_\mu w, w \rangle + \|\chi_{<100R}w_{in}\|_{H^{\frac{1}{2}}}^2) + \mu e^{CML}\|w\|_{L^2}^2. \quad (6.59)$$

This suffices provided that μ is large enough, $\mu = e^{C_1ML}$ with $C_1 \gg C$.

We first directly compute the regularity

$$C - C_\mu \in \mu^{-\delta}OPC^1S^1,$$

which shows that only the principal symbol of $C - C_\mu$ matters.

We then use the squares representation for q and ∇q to write

$$(c - c_\mu)_0 = \sum e_j \phi_j^2 + e_0 |\xi| (\chi_{<100R}p_{in})^2, \quad e_j \in \mu^{-\delta}e^{CML}C^1S^0,$$

which implies the bound

$$\langle (C - C_\mu)w, w \rangle \lesssim \mu^{-\delta}e^{CML} \left(\sum \|\Phi_j w\|_{L^2}^2 + \|\chi_{<100R}w_{in}\|_{H^{\frac{1}{2}}}^2 + \|w\|_{L^2}^2 \right).$$

Combining this with (6.58) yields (6.59) and completes the proof of (6.48).

6.5. The High Frequency Local Energy Decay Bound

Our objective here is to use the results from Sections 6.1, 6.3 and 6.4 in order to complete the proof of the high frequency local energy decay bound for the paradifferential equation in Proposition 6.4.

Combining the bounds in (6.32) and (6.48) we obtain a local energy decay bound

$$\|\chi_{<4R} w\|_{L^2 H^{\frac{1}{2}}} \lesssim e^{C(M)L} (\|w_0\|_{L^2} + \|f\|_{Y^0} + \|w\|_{L^2 L^2}). \quad (6.60)$$

It remains to estimate the exterior part of w . For that we simply truncate w , setting

$$w_{ext} = \chi_{>2R} w.$$

Then we write the paradifferential equation for w_{ext} , but using the truncated coefficients g_{ext} and b_{ext} . This takes the form

$$\begin{cases} (i\partial_t + \partial_k T_{g_{ext}}^w \partial_l + T_{b_{ext}}^w \cdot \nabla) w_{ext} = f_{ext} \\ w_{ext}(0) = \chi_{>2R} w_0, \end{cases} \quad (6.61)$$

where

$$\begin{aligned} f_{ext} = & \chi_{>2R} f + [(\partial_k T_{g_{ext}}^w \partial_l + T_{b_{ext}}^w \cdot \nabla), \chi_{>2R}] w + (\partial_k T_{g_{ext}}^w \partial_l \\ & - \partial_k T_{g_{ext}}^w \partial_l) w_{ext} + (T_{b_{ext}}^w \cdot \nabla - T_b^w \cdot \nabla) w_{ext}. \end{aligned}$$

Using (2.4) and the disjointness of the supports of w_{ext} and $g_{ext} - g$, $b_{ext} - b$, it follows that

$$\|(\partial_k T_{g_{ext}}^w \partial_l - \partial_k T_{g_{ext}}^w \partial_l) w_{ext} + (T_{b_{ext}}^w \cdot \nabla - T_b^w \cdot \nabla) w_{ext}\|_{Y^0} \lesssim \|w\|_{L^2 L^2}.$$

Moreover,

$$\|[(\partial_k T_{g_{ext}}^w \partial_l + T_{b_{ext}}^w \cdot \nabla), \chi_{>2R}] w\|_{Y^0} \lesssim \|\chi_{<4R} w\|_{L^2 H^{1/2}} + \|w\|_{L^2 L^2}.$$

Thus we can apply the small data result (6.16) to bound w_{ext} ,

$$\|w_{ext}\|_{X^0} \lesssim_M \|w_0\|_{L^2} + \|f\|_{Y^0} + \|w\|_{L^2 L^2} + \|\chi_{<4R} w\|_{L^2 H^{1/2}}. \quad (6.62)$$

Combined with (6.60), this yields the conclusion of Proposition 6.4 for the case of the paradifferential equation.

6.6. Higher Regularity Bounds for the Paradifferential Equation

Here we extend the L^2 high frequency bounds (6.12) for the paradifferential equation (6.2) to H^s and $l^1 H^s$. We begin by proving the $l^1 L^2$ bound,

$$\|w\|_{l^1 X^0} \lesssim e^{C(M)L} (\|w_0\|_{l^1 L^2} + \|f\|_{l^1 Y^0} + \|w\|_{l^1 L^2 L^2}). \quad (6.63)$$

Here the l^1 norms control the ℓ^2 norms, so by (6.12) we can bound $\|w\|_{X^0}$ by the right side of (6.63). This suffices to establish (6.63) in B_{4R} , and it remains to consider the exterior part of w . Here we apply (6.17) to w_{ext} , which solves (6.61). Arguing in a fashion analogous to that in the preceding subsection yields (6.63).

Next we consider (6.13). By interpolation, it suffices to consider the case when σ is a positive integer. But this case can be obtained simply by differentiating the paradifferential equation and applying (6.63) to $\partial^\sigma w$. Here we note that $\partial^\sigma w$ can be viewed as a vector valued function, which solves a system which is diagonal at leading order but coupled through the first order terms. This makes no difference, as the proof of (6.63) equally applies to this case without any changes.

Finally we prove (6.14). Here we apply the same reasoning as in the previous paragraph but starting with (6.12) instead of (6.63).

6.7. Bounds for the Original Equation

Here we transfer the high frequency bounds from the paradifferential equation (6.2) to the original equation (6.1), and prove Proposition 6.4 (which immediately implies Theorem 6.1), as well as the bounds (6.13) and (6.14) (which immediately imply Corollary 6.2).

We begin with (6.13) and (6.14). This requires bounds for the operator

$$E = (g^{kl} - T_{g^{kl}}^w) \partial_k \partial_l + (b^j - T_{b^j}^w) \partial_j$$

namely

$$\|Ew\|_{l^2 Y^\sigma} \lesssim M e^{-CML} \|w\|_{l^2 X^\sigma} + e^{CML} \|w\|_{l^2 L^2 H^\sigma}, \quad (6.64)$$

$$\|Ew\|_{l^1 Y^\sigma} \lesssim M e^{-CML} \|w\|_{l^1 X^\sigma} + e^{CML} \|w\|_{l^1 L^2 H^\sigma}. \quad (6.65)$$

Using a Littlewood–Paley decomposition we identify two distinct cases:

- (i) High $\times$ high $\rightarrow$ low interactions. If the high frequency is $k \gg ML$ then we use (4.15), gaining a factor of $2^{-\tilde{\delta}k}$. Else we bound the $l^1 X^\sigma$ norm by the $l^1 L^2 H^\sigma$ norm, losing an e^{CML} factor. Here $\tilde{\delta} = s - s_0$.
- (ii) High $\times$ low $\rightarrow$ high interactions, where we can still apply the same strategy as above. If the low frequency is $k \gg ML$ then we use (4.15), gaining a factor of $2^{-\delta k}$. Else we bound the $l^1 X^\sigma$ norm by the $l^1 H^\sigma$ norm, losing an e^{CML} factor.

7. Proof of Theorem 1.2

We recall that the Eq. (1.1) turns into an equation of the form (1.4) by differentiation. Hence it suffices to prove part (b) of the theorem. Thus, we are working with the equation

$$\begin{cases} iu_t + \partial_j g^{jk}(u, \bar{u}) \partial_k u = F(u, \bar{u}, \nabla u, \nabla \bar{u}), & u : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{C}^m \\ u(0, x) = u_0(x) \end{cases} \quad (7.1)$$

which we rewrite in the paradifferential form

$$\begin{cases} i \partial_t u + \partial_j T_{g^{jk}}^w \partial_k u + T_{bj}^w \partial_j u + T_{\bar{b}j}^w \partial_j \bar{u} = G \\ u(0, x) = u_0(x). \end{cases} \quad (7.2)$$

Here the nonlinearity $G = G(u, \bar{u}, \nabla u, \nabla \bar{u})$ is no longer purely algebraic, as it involves frequency localizations. However, G plays a perturbative role, due to the estimates in Section 4.2.

7.1. The Iteration Scheme

To construct a local solution u for our nonlinear equation (7.1) we introduce an iterative scheme as follows:

- Our starting point is the function $u^{(0)} = 0$.
- The iteration step is as follows. Given $u^{(n)}$, we construct $u^{(n+1)}$ as the solution to the linear paradifferential equation

$$\begin{cases} (i \partial_t u^{(n+1)} + \partial_j T_{g^{jk}(u^{(n)})}^w \partial_k u^{(n+1)} + T_{bj(u^{(n)})}^w \partial_j u^{(n+1)} + T_{\bar{b}j(u^{(n)})}^w \partial_j \bar{u}^{(n+1)} = G(u^{(n)})) \\ u^{(n+1)}(0) = u_0. \end{cases} \quad (7.3)$$

A priori we do not even know whether the sequence $u^{(n)}$ is well defined for all n . Even if it is defined locally in time, we do not know whether their lifespans are uniformly bounded from below away from zero. Thus, our objectives will be, in order, as follows:

- Establish uniform bounds for the sequence $u^{(n)}$ on a fixed time interval $[0, T]$ *not depending on n* .
- Prove convergence for the sequence $u^{(n)}$.

To achieve this we first need to establish the main parameters which will be used to control the sequence $u^{(n)}$, which should depend only on the initial data u_0 . A priori we know that $u_0 \in l^1 H^s$. We introduce a second Sobolev index s_0 so that

$$\frac{d}{2} + 2 < s_0 < s.$$

Then we consider the coefficients $g(u_0)$, $b(u_0)$ and $\tilde{b}(u_0)$ in the linearized equation, and measure them at both regularities, denoting

$$M_s = \|u_0\|_{l^1 H^s}, \quad M = \|u_0\|_{l^1 H^{s_0}}. \quad (7.4)$$

The next step is to choose a large enough ball B_R of radius R so that u_0 is small outside B_R ,

$$\|u_0\|_{l^1 H^{s_0}(B_R^c)} \leq \varepsilon \ll 1. \quad (7.5)$$

Here ε is a small universal constant.

Finally, the metric $g(u_0)$ is nontrapping. Then we denote by L the length of the longest bicharacteristic for $g(u_0)$, measured on the cosphere bundle, from the entry to the exit from $2B_R$.

Given the initial data parameters, M , R , L and M_s we seek to use them in order to uniformly describe the sequence $u^{(n)}$:

Proposition 7.1. *Assume that the time T is small enough,*

$$T \ll_{M_s} e^{-C(M)L}. \quad (7.6)$$

Then the sequence $u^{(n)}$ is well defined in $[0, T]$ for all n and satisfies the following uniform properties:

(1) *Uniform H^s bounds:*

$$\|u^{(n)}\|_{l^1 X^s[0, T]} \leq e^{C(M)L} M_s. \quad (7.7)$$

(2) *Uniform H^{s_0} bounds:*

$$\|u^{(n)}\|_{l^1 X^{s_0}[0, T]} \leq 2M. \quad (7.8)$$

(3) *Uniform exterior size:*

$$\|u^{(n)}\|_{l^1 X^{s_0}([0, T] \times B_R^c)} \leq 2\varepsilon. \quad (7.9)$$

(4) *Uniform nontrapping:*

$$L(u^{(n)}) \leq 2L. \quad (7.10)$$

Once we have the uniform bounds on the iterations, the next goal is to prove convergence in a weaker topology:

Proposition 7.2. *Assume that the time T is small enough,*

$$T \ll_{M_s} e^{-C(M)L}. \quad (7.11)$$

Then the sequence $u^{(n)}$ converges in $l^1 X^\sigma$ for $0 \leq \sigma < s_0 - 1$.

After these two propositions are proved, it follows that the sequence $u^{(n)}$ is uniformly bounded in $l^1 X^s$ and convergent in $l^1 X^{s-1}$. Then it is convergent in all intermediate topologies. This suffices in order to pass to the limit in the equation and conclude that in the limit we obtain a solution $u \in l^1 X^s$ for the original equation:

Proposition 7.3. *Let $u_0 \in l^1 H^s$ be a nontrapping initial datum with parameters R, M, L, M_s . Assume that the time T is small enough,*

$$T \ll_{M_s} e^{-C(M)L}. \quad (7.12)$$

Then there exists a solution $u \in l^1 X^s$ with the following properties:

(1) *H^s bound:*

$$\|u\|_{l^1 X^s[0,T]} \leq e^{C(M)L} M_s. \quad (7.13)$$

(2) *H^{s_0} bounds:*

$$\|u\|_{l^1 X^{s_0}[0,T]} \leq 2M. \quad (7.14)$$

(3) *Small exterior size:*

$$\|u\|_{l^1 X^s([0,T] \times B_R^c)} \leq 2\varepsilon. \quad (7.15)$$

(4) *Nontrapping:*

$$L(u) \leq 2L. \quad (7.16)$$

The aim of the next two subsections is to prove the first two propositions.

7.2. The Iteration Scheme: Uniform Bounds

Our aim here is to prove Proposition 7.1. We use induction on n .

Proof of (7.7). Here we use the bounds for the paradifferential equation in Corollary 6.2 together with the bounds for G in Proposition 4.5 to get

$$\begin{aligned} \|u^{(n+1)}\|_{l^1 X^s[0,T]} &\lesssim e^{C(M)L} (\|u_0\|_{l^1 H^s} + \|G(u^{(n)})\|_{l^1 Y^s}) \\ &\lesssim e^{C(M)L} (M_s + T^\delta C(M) \|u^{(n)}\|_{l^1 X^s[0,T]}). \end{aligned}$$

Hence for small enough T as in (7.12) the bound (7.7) follows.

Proof of (7.8). A direct time integration in our iteration yields

$$\|u^{(n+1)}\|_{l^1 X^0} \lesssim \sqrt{T} e^{C(M)L} M_s.$$

Interpolating with the $l^1 X^s$ bound above yields

$$\|u^{(n+1)}\|_{l^1 X^{s_0}} \lesssim T^{\delta/2} e^{C(M)L} M_s, \quad (7.17)$$

which suffices for T as in (7.12).

Proof of (7.9). This follows directly from the bound (7.17) in view of the choice of R in (7.5).

Proof of (7.10). This is a consequence of Proposition 5.2 due to (7.17).

7.3. The Iteration Scheme: Weak Convergence

Here we prove Proposition 7.2, which asserts that our iteration scheme converges in the weaker $l^1 X^\sigma$ topology. We recall the range for σ , namely,

$$0 \leq \sigma < s_0 - 1. \quad (7.18)$$

For this we write an equation for the difference $v^{(n)} = u^{(n+1)} - u^{(n)}$:

$$\begin{cases} (i\partial_t + \partial_j T_{g^{jk,(n)}}^w \partial_k + T_{b^{(n)}}^w \nabla) v^{(n)} + T_{\tilde{b}^{(n)}}^w \nabla \bar{v}^{(n)} = G(u^{(n)}) - G(u^{(n-1)}) + H^{(n)} \\ v^{(n)}(0, x) = 0, \end{cases} \quad (7.19)$$

where

$$H^{(n)} = \partial_j T_{g^{jk,(n)} - g^{jk,(n-1)}}^w \partial_k u^{(n)} + T_{b^{(n)} - b^{(n-1)}}^w \nabla u^{(n)} + T_{\tilde{b}^{(n)} - \tilde{b}^{(n-1)}}^w \nabla \bar{u}^{(n)}.$$

For the G difference we apply Proposition 4.5, which yields

$$\|G(u^{(n)}) - G(u^{(n-1)})\|_{l^1 Y^\sigma} \lesssim_M T^\delta \|v^{(n-1)}\|_{l^1 X^\sigma}. \quad (7.20)$$

Similarly, for $H^{(n)}$, we claim the bound

$$\|H^{(n)}\|_{l^1 Y^\sigma} \lesssim_M T^\delta \|v^{(n-1)}\|_{l^1 X^\sigma}. \quad (7.21)$$

Assume this holds, then for the paradifferential equation we use Corollary 6.2. We obtain

$$\|v^{(n+1)}\|_{l^1 X^\sigma} \lesssim T^\delta e^{C(M)L} \|v^{(n)}\|_{l^1 X^\sigma},$$

which for T as in (7.12) yields

$$\|v^{(n+1)}\|_{l^1 X^\sigma} \lesssim \frac{1}{2} \|v^{(n)}\|_{l^1 X^\sigma}.$$

The desired convergence follows. Thus we have established the existence part of our main theorem.

It remains to prove the bound (7.21). We write the coefficients above in the form

$$g^{jk,(n)} - g^{jk,(n-1)} = v^{(n-1)} h_0(u)$$

and

$$b^{(n)} - b^{(n-1)} = h_1(u, \nabla u) \nabla v^{(n-1)} + h_2(u, \nabla u) v^{(n-1)},$$

where u stands for $(u^{(n)}, u^{(n-1)})$. For the functions h_j we can apply the Moser estimates in (4.6). Then we are left with proving the following two trilinear bounds,

$$\|S_{<k-4}(wv)u_k\|_{l^1 Y^\sigma} \lesssim T^\delta \|w\|_{l^1 X^{s_0-1}} \|v\|_{l^1 X^\sigma} \|u\|_{l^1 X^{s_0-2}}, \quad (7.22)$$

respectively

$$\|S_{<k-4}(wv)u_k\|_{l^1 Y^\sigma} \lesssim T^\delta \|w\|_{l^1 X^{s_0-1}} \|v\|_{l^1 X^{\sigma-1}} \|u\|_{l^1 X^{s_0-1}}. \quad (7.23)$$

We consider two cases:

- (a) The v frequency is $\leq 2^{k-2}$. Then the frequency of w is similar, and the first bound (7.22) is worse. We harmlessly drop the multiplier, use (4.16) for the vu product and bound w in L^∞ .
- (b) The v frequency is $\geq 2^{k-2}$. Then the frequency of w is also similar, and the worst case is (7.23) with $\sigma = 0$. But then this is exactly the bound (4.25), proved earlier.

7.4. Uniqueness Via Weak Lipschitz Dependence

Our aim here is to prove the following estimate for the difference of two solutions:

Proposition 7.4. *Let $u_0^{(1)} \in l^1 H^s$ be a nontrapping initial datum with parameters M, R, L and $M(s)$. Let $u_0^{(2)} \in l^1 H^s$ be another initial datum with comparable size,*

$$\|u_0^{(2)}\|_{l^1 H^s} \lesssim M_s$$

and close to $u_0^{(1)}$ in a weak topology,

$$\|u_0^{(1)} - u_0^{(2)}\|_{l^1 L^2} \ll_{M_s} e^{-C(M)L}.$$

Then the following hold:

- (a) $u_0^{(2)}$ has comparable parameters, and the same associated ball B .
- (b) The associated solutions exist on the time interval $[0, T]$ with T as in (7.12).
- (c) The following difference estimate holds in $[0, T]$ for σ as in (7.18):

$$\|u^{(1)} - u^{(2)}\|_{l^1 X^\sigma} \lesssim e^{C(M)L} \|u^{(1)}(0) - u^{(2)}(0)\|_{l^1 H^\sigma}. \quad (7.24)$$

Uniqueness follows as a corollary of this result.

Proof. (a) Interpolating between the $l^1 L^2$ and $l^1 H^s$ bounds we get

$$\|u_0^{(1)} - u_0^{(2)}\|_{l^1 H^{s_0}} \ll e^{-C(M)L}.$$

Thus M and B_R remain the same. Finally, L remains the same by Proposition 5.2.

- (b) This is a consequence of Proposition 7.3.
- (c) This repeats the arguments in the proof of Proposition 7.2. $\square$

7.5. Frequency Envelopes and Higher Regularity

Here we consider solutions with initial data u_0 as in Proposition 7.3 and provide bounds for the frequency envelope of the solutions in terms of the frequency envelope of the initial data. As a corollary, we show that higher regularity for the data implies higher regularity for the solution.

Our starting point is the initial data $u_0 \in l^1 H^s$, for which we consider an admissible frequency envelope c_k . Then our frequency envelope bound for the solutions is

Proposition 7.5. *Let u be a solution to (3.1) as in Proposition 7.3, with initial data $u_0 \in l^1 H^s$. Let c_k be an admissible frequency envelope for u_0 in $l^1 H^s$. Then the solution u satisfies the dyadic bounds*

$$\|S_k u\|_{l^1 X^s} \lesssim e^{C(M)L} c_k. \quad (7.25)$$

As a consequence, we obtain the following higher regularity statement:

Corollary 7.6. *Let u be a solution to (3.1) as in Proposition 7.3, with initial data u_0 . Assume in addition that $u_0 \in l^1 H^\sigma$ for some $\sigma \geq s$. Then the solution u satisfies the bound*

$$\|u\|_{l^1 X^\sigma} \lesssim e^{C(M)L} \|u_0\|_{l^1 H^\sigma}, \quad \sigma \geq s. \quad (7.26)$$

Proof of Proposition 7.5. We denote by d_k a *minimal* admissible frequency envelope for u in $l^1 X^s$. By Corollary 4.6 we have a corresponding bound for $G(u)$, namely,

$$\|S_k G(u)\|_{l^1 X^s} \lesssim_M T^\delta d_k.$$

Now we apply to u the bound for the linear paradifferential equation in Corollary 6.3. This gives

$$\|S_k u\|_{l^1 X^s} \lesssim e^{C(M)L} (c_k + T^\delta d_k).$$

The envelope on the right is admissible, so by the minimality of d_k we obtain

$$d_k \lesssim e^{C(M)L} (c_k + T^\delta d_k).$$

Now the choice of T as in (7.12) guarantees that

$$d_k \lesssim e^{C(M)L} c_k,$$

as needed. $\square$

Proof of Corollary 7.6. Let c_k be a minimal admissible frequency envelope for the initial data u_0 in $l^1 H^s$. Then

$$\|u_0\|_{l^1 H^\sigma}^2 \approx \sum 2^{2(\sigma-s)k} c_k^2.$$

On the other hand, the previous proposition implies that

$$\|u\|_{l^1 X^\sigma}^2 \lesssim e^{C(M)L} \sum 2^{2(\sigma-s)k} c_k^2 \lesssim e^{C(M)L} \|u_0\|_{l^1 H^\sigma}^2.$$

$\square$

7.6. Continuous Dependence on the Initial Data

Here we want to show that $u_0 \rightarrow u$ is continuous from $l^1 H^s$ into $l^1 X^s$. The argument is exactly as in [26].

Let $a_j^{(n)}$ and a_j be minimal frequency envelopes given by (2.5) for $u_0^{(n)}$ and u_0 in $l^1 H^s$. If $u_0^{(n)} \rightarrow u_0$ in $l^1 H^s$, then $(a_j^{(n)}) \rightarrow (a_j)$ in l^2 . So for any $\varepsilon > 0$, there is N_ε so

$$\|a_{>N_\varepsilon}^{(n)}\|_{l^2} \leq \varepsilon \quad \text{for all } n.$$

We remark that our initial data convergence in $l^1 H^s$ guarantees that, for large enough n , the control parameters R, L can be uniformly chosen independently of n . Then the associated solutions $u^{(n)}$ exist on a uniform time interval $[0, T]$. By Proposition 7.5, the envelopes $a_j^{(n)}$ carry over to the solutions $u^{(n)}$ measured in $l^1 X^s[0, T]$. In particular, we conclude that

$$\|u_{>N_\varepsilon}^{(n)}\|_{l^1 X^s} \leq \varepsilon \quad \text{for all } n. \quad (7.27)$$

Using (7.24) for low frequencies and (7.27) for the high frequencies, we obtain

$$\begin{aligned} \|u^{(n)} - u\|_{l^1 X^s} &\lesssim \|S_{<N_\varepsilon}(u^{(n)} - u)\|_{l^1 X^s} + \|S_{>N_\varepsilon}u^{(n)}\|_{l^1 X^s} + \|S_{>N_\varepsilon}u\|_{l^1 X^s} \\ &\lesssim 2^{N_\varepsilon} \|S_{<N_\varepsilon}(u^{(n)} - u)\|_{l^1 X^{s-1}} + 2\varepsilon \\ &\lesssim 2^{N_\varepsilon} \|S_{<N_\varepsilon}(u_0^{(n)} - u_0)\|_{l^1 H^{s-1}} + 2\varepsilon. \end{aligned}$$

As $n \rightarrow \infty$ we see that

$$\limsup_{n \rightarrow \infty} \|u^{(n)} - u\|_{l^1 X^s} \lesssim \varepsilon.$$

There, upon letting $\varepsilon \rightarrow 0$,

$$\lim_{n \rightarrow \infty} \|u^{(n)} - u\|_{l^1 X^s} = 0,$$

which gives the desired result.

Acknowledgements. JEREMY L. MARZUOLA was supported in part by U.S. NSF Grant DMS-1312874 and NSF CAREER Grant DMS-1352353. The third author is supported in part by the NSF Grant DMS-1800294 as well as by a Simons Investigator grant from the Simons Foundation. The authors also wish to thank the Mathematical Sciences Research Institute for hosting two of them during part of this work. We thank the anonymous referees for a careful reading of the result and for making helpful suggestions to improve the exposition.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. BEJENARU, I.: Quadratic nonlinear derivative Schrödinger equations. I. *IMRP Int. Math. Res. Pap.*, 2006
2. BEJENARU, I.: Quadratic nonlinear derivative Schrödinger equations. II. *Trans. Am. Math. Soc.* **360**(11), 5925–5957, 2008
3. BEJENARU, I., TATARU, D.: Large data local solutions for the derivative NLS equation. *J. Eur. Math. Soc.* **10**, 957–985, 2008
4. BLAIR, M.: Strichartz estimates for wave equations with coefficients of Sobolev regularity. *Commun. Partial Differ. Equ.* **31**(5), 649–688, 2006
5. BLASCO, O.; RUIZ, A.; VEGA, L.: Non interpolation in Morrey–Campanato and block spaces. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **28**, 31–40, 1999
6. CHIHARA, H.: Local existence for semilinear Schrödinger equations. *Math. Jpn.* **42**, 35–52, 1995
7. CONSTANTIN, P., SAUT, J.-C.: Local smoothing properties of dispersive equations. *J. Am. Math. Soc.* **1**, 413–446, 1989
8. CRAIG, W., KAPPELER, T., STRAUSS, W.: Microlocal dispersive smoothing for the Schrödinger equation. *Commun. Pure Appl. Math.* **48**(8), 769–860, 1995
9. DOI, S.: Remarks on the Cauchy problem for Schrödinger-type equations. *Commun. Partial Differ. Equ.* **21**, 163–178, 1996
10. DOI, S.: Smoothing effects for Schrödinger evolution equation and global behavior of geodesic flow. *Math. Ann.* **318**, 355–389, 2000
11. HAYASHI, N., OZAWA, T.: Remarks on nonlinear Schrödinger equations in one space dimension. *Differ. Integral Equ.* **7**, 453–461, 1994
12. HÖRMANDER, L.: *Lectures on Nonlinear Hyperbolic Differential Equations*. Mathematics & Applications, Berlin 1997
13. ICHINOSE, W.: On L^2 well-posedness of the Cauchy problem for Schrödinger type equations on a Riemannian manifold and Maslov theory. *Duke Math. J.* **56**, 549–588, 1988
14. KATO, T.: On the Cauchy problem for the (generalized) Korteweg–de Vries equation. *Adv. Math. Suppl. Stud. Appl. Math.* **8**, 93–128, 1983
15. KENIG, C.E.; PONCE, G.; VEGA, L.: Small solutions to nonlinear Schrödinger equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **10**, 255–288, 1993
16. KENIG, C.E., PONCE, G., VEGA, L.: Smoothing effects and local existence theory for the generalized nonlinear Schrödinger equations. *Invent. Math.* **134**, 489–545, 1998
17. KENIG, C.E., PONCE, G., VEGA, L.: The Cauchy problem for quasi-linear Schrödinger equations. *Invent. Math.* **158**, 343–388, 2004
18. KENIG, C.E., PONCE, G., ROLVUNG, C., VEGA, L.: The general quasilinear ultrahyperbolic Schrödinger equation. *Adv. Math.* **196**(2), 402–433, 2005
19. KENIG, C.E., PONCE, G., ROLVUNG, C., VEGA, L.: Variable coefficient Schrödinger flows for ultrahyperbolic operators. *Adv. Math.* **206**(2), 373–486, 2006
20. KOCH, H., TATARU, D.: Dispersive estimates for principally normal pseudodifferential operators. *Commun. Pure Appl. Math.* **58**, 217–284, 2005
21. KRUSHKOV, S.N., FAMINSKII, A.V.: Generalized solutions of the Cauchy problem for the Korteweg–de Vries equation. *Mat. Sh. USSR* **48**, 93–138, 1984
22. LIM, W.-K., PONCE, G.: On the initial value problem for the one dimensional quasilinear Schrödinger equation. *SIAM J. Math. Anal.* **34**, 435–459, 2003
23. LINAES, F., PONCE, G.: *Introduction to Nonlinear Dispersive Equations*. Universitext, Springer, New York 2009
24. MARZUOLA, J.; METCALFE, J.; TATARU, D.: Wave packet parametrices for evolutions governed by PDO's with rough symbols. *Proc. Am. Math. Soc.* **136**(2), 597–604, 2007
25. MARZUOLA, J., METCALFE, J., TATARU, D.: Strichartz estimates and local smoothing estimates for asymptotically flat Schrödinger equations. *J. Funct. Anal.* **255**(6), 1497–1553, 2008

26. MARZUOLA, J., METCALFE, J., TATARU, D.: Quasilinear Schrödinger equations I: small data and quadratic interactions. *Adv. Math.* **231**(2), 1151–1172, 2012
27. MARZUOLA, J., METCALFE, J., TATARU, D.: Quasilinear Schrödinger equations, II: small data and cubic nonlinearities. *Kyoto J. Math.* **54**(3), 529–546, 2014
28. MICHALOWSKI, N.: Quasilinear Schrödinger equations. *Preprint* [arXiv:1410.0057](https://arxiv.org/abs/1410.0057), 2014
29. MIZOHATA, S.: Some remarks on the Cauchy problem. *J. Math. Kyoto Univ.* **1**, 109–127, 1961
30. MIZOHATA, S.: *Sur quelques equations du type Schrödinger*. Journées Equations aux dérivées partielles, Saint-Jean de Monts 1981
31. MIZOHATA, S.: On the Cauchy Problem. Notes and Reports in Mathematics in Science and Engineering, 3. Science Press and Academic Press, 1985
32. SCHOTTDORF, T.: Ill-Posedness for the quadratic D-NLS Equation. Master's Thesis. University of California at Berkeley, 2010
33. SJÖLIN, P.: Regularity of solutions to the Schrödinger equations. *Duke Math. J.* **55**, 699–715, 1987
34. SOGGE, C.D.: *Lectures on Nonlinear Wave Equations, Monographs in Analysis. II*. International Press, Boston 1995
35. STAFFILANI, G., TATARU, D.: Strichartz estimates for a Schrödinger operator with non-smooth coefficients. *Commun. PDE* **27**(7–8), 1337–1372, 2002
36. TAKEUCHI, J.: On the Cauchy problem for some non-Kowalewskian equations with distinct characteristic roots. *J. Math. Kyoto Univ.* **20**(1), 105–124, 1980
37. TATARU, D.: Strichartz estimates for second order hyperbolic operators with nonsmooth coefficients III. *J. Am. Math. Soc.* **15**, 419–442, 2002
38. TATARU, D.: On the Fefferman–Phong inequality and related problems. *Commun. PDE* **27**(11–12), 2101–2138, 2002
39. TATARU, D.: Phase space transforms and microlocal analysis. Phase space analysis of partial differential equations. Vol. II. Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup. Pisa, pp. 505–524, 2004
40. TATARU, D.: Rough solutions for the wave maps equation. *Am. J. Math.* **127**, 293–377, 2005
41. TATARU, D.: Parametrics and dispersive estimates for Schrödinger operators with variable coefficients. *Am. J. Math.* **130**, 571–634, 2008
42. TAYLOR, M.E.: Tools for PDE: Pseudodifferential Operators, Paradifferential Operators, and Layer potentials, vol. 81. American Mathematical Soc, 2007
43. VARGAS, A., VEGA, L.: Global wellposedness for 1D non-linear Schrödinger equation for data with an infinite L^2 norm. *J. Math. Pures Appl.* **80**, 1029–1044, 2001
44. VEGA, L.: The Schrödinger equation: pointwise convergence to the initial data. *Proc. Am. Math. Soc.* **102**, 874–878, 1988

JEREMY L. MARZUOL & JASON METCALFE
Department of Mathematics,
University of North Carolina-Chapel Hill,
Phillips Hall,
Chapel Hill
NC
27599 USA.
e-mail: marzuola@email.unc.edu

JASON METCALFE
e-mail: metcalfe@email.unc.edu

and

DANIEL TATARU
Mathematics Department,
University of California,
Evans Hall,
Berkeley
CA
94720 USA.
e-mail: tataru@math.berkeley.edu

(Received March 4, 2020 / Accepted August 2, 2021)

Published online August 28, 2021

© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE, part of Springer
Nature (2021)