ANALYSIS & PDE

Volume 6 No. 3 2013

PAUL SMITH

CONDITIONAL GLOBAL REGULARITY OF SCHRÖDINGER MAPS: SUBTHRESHOLD DISPERSED ENERGY

msp

ANALYSIS AND PDE
Vol. 6, No. 3, 2013
dx.doi.org/10.2140/apde.2013.6.601

CONDITIONAL GLOBAL REGULARITY OF SCHRÖDINGER MAPS: SUBTHRESHOLD DISPERSED ENERGY

PAUL SMITH

We consider the Schrödinger map initial value problem

$$\begin{cases} \partial_t \varphi = \varphi \times \Delta \varphi, \\ \varphi(x, 0) = \varphi_0(x), \end{cases}$$

with $\varphi_0 : \mathbb{R}^2 \to \mathbb{S}^2 \hookrightarrow \mathbb{R}^3$ a smooth H_Q^∞ map from the Euclidean space $\mathbb{R}^2$ to the sphere $\mathbb{S}^2$ with subthreshold ($< 4\pi$) energy. Assuming an a priori L^4 boundedness condition on the solution φ, we prove that the Schrödinger map system admits a unique global smooth solution $\varphi \in C(\mathbb{R} \to H_Q^\infty)$ provided that the initial data φ_0 is sufficiently energy-dispersed, i.e., sufficiently small in the critical Besov space $\dot{B}^1_{2,\infty}$. Also shown are global-in-time bounds on certain Sobolev norms of φ. Toward these ends we establish improved local smoothing and bilinear Strichartz estimates, adapting the Planchon–Vega approach to such estimates to the nonlinear setting of Schrödinger maps.

1. Introduction

We consider the Schrödinger map initial value problem

$$\begin{cases} \partial_t \varphi = \varphi \times \Delta \varphi, \\ \varphi(x, 0) = \varphi_0(x), \end{cases} \tag{1-1}$$

with $\varphi_0 : \mathbb{R}^d \to \mathbb{S}^2 \hookrightarrow \mathbb{R}^3$.

The author is supported in part by NSF grant number DMS-1103877. This research was carried out at UCLA with their support and at UC Berkeley with NSF support.
MSC2010: primary 35Q55; secondary 35B33.
Keywords: Schrödinger maps, global regularity, energy-critical, critical Besov spaces, subthreshold.

The system (1-1) enjoys conservation of energy,

$$E(\varphi(t)) := \frac{1}{2}\int_{\mathbb{R}^d} |\partial_x \varphi(t)|^2 \, dx, \tag{1-2}$$

and mass,

$$M(\varphi(t)) := \int_{\mathbb{R}^d} |\varphi(t) - Q|^2 \, dx,$$

where $Q \in \mathbb{S}^2$ is some fixed base point. When $d = 2$, both (1-1) and (1-2) are invariant with respect to the scaling

$$\varphi(x, t) \to \varphi(\lambda x, \lambda^2 t), \qquad \lambda > 0, \tag{1-3}$$

and in this case we call the equation (1-1) *energy-critical*. In this article we restrict ourselves to the energy-critical setting.

For the physical significance of (1-1), see [Chang et al. 2000; Nahmod et al. 2003; Papanicolaou and Tomaras 1991; Landau 1967]. The system also arises naturally from the (scalar-valued) free linear Schrödinger equation

$$(\partial_t + i\Delta)u = 0$$

by replacing the target manifold $\mathbb{C}$ with the sphere $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$, which then requires replacing Δu with $(u^*\nabla)_j \partial_j u = \Delta u - \perp(\Delta u)$ and i with the complex structure $u \times \cdot$. Here $\perp$ denotes orthogonal projection onto the normal bundle, which, for a given point (x, t), is spanned by $u(x, t)$. For more general analogues of (1-1), e.g., for Kähler targets other than $\mathbb{S}^2$, see [Ding and Wang 2001; McGahagan 2007; Nahmod et al. 2007]. See also [Kenig et al. 2000; Kenig and Nahmod 2005; Bejenaru et al. 2011b] for connections with other spin systems. The local theory for Schrödinger maps is developed in [Sulem et al. 1986; Chang et al. 2000; Ding and Wang 2001; McGahagan 2007]. For global results in the $d = 1$ setting, see [Chang et al. 2000; Rodnianski et al. 2009]. For $d \geq 3$, see [Bejenaru 2008a; 2008b; Bejenaru et al. 2007; 2011c; Ionescu and Kenig 2006; 2007b]. Concerning the related modified Schrödinger map system, see [Kato 2005; Kato and Koch 2007; Nahmod et al. 2007].

The small-energy (take $d = 2$) theory for (1-1) is now well-understood: building upon previous work (see below or [Bejenaru et al. 2011c, §1] for a brief history), global well-posedness and global-in-time bounds on certain Sobolev norms are shown in [Bejenaru et al. 2011c] given initial data with sufficiently small energy. The high-energy theory, however, is still very much in development. One of the main goals is to establish what is known as the *threshold conjecture*, which asserts that global well-posedness holds for (1-1) given initial data with energy below a certain energy threshold, and that finite-time blowup is possible for certain initial data with energy above this threshold. The threshold is directly tied to the nontrivial stationary solutions of (1-1), i.e., maps ϕ into $\mathbb{S}^2$ that satisfy

$$\phi \times \Delta\phi \equiv 0$$

and that do not send all of $\mathbb{R}^2$ to a single point of $\mathbb{S}^2$. Therefore we identify such stationary solutions with nontrivial harmonic maps $\mathbb{R}^2 \to \mathbb{S}^2$, which we refer to as *solitons* for (1-1). It turns out that there exist no nontrivial harmonic maps into the sphere $\mathbb{S}^2$ with energy less than 4π, and that the harmonic map given by the inverse of stereographic projection has energy precisely equal to $4\pi =: E_{\mathrm{crit}}$. We therefore refer to the range of energies $[0, E_{\mathrm{crit}})$ as *subthreshold*, and call E_{crit} the *critical* or *threshold energy*.

Recently, an analogous threshold conjecture was established for wave maps (see [Krieger et al. 2008; Rodnianski and Sterbenz 2010; Sterbenz and Tataru 2010a; 2010b] and, for hyperbolic space, [Krieger and Schlag 2012; Tao 2008a; 2008b; 2008c; 2009a; 2009b]). When $\mathcal{M}$ is a hyperbolic space, or, as in [Sterbenz and Tataru 2010a; 2010b], a generic compact manifold, we may define the associated energy threshold $E_{\mathrm{crit}} = E_{\mathrm{crit}}(\mathcal{M})$ as follows. Given a target manifold $\mathcal{M}$, consider the collection $\mathcal{S}$ of all nonconstant finite-energy harmonic maps $\phi : \mathbb{R}^2 \to \mathcal{M}$. If this set is empty, as is, for instance, the case when $\mathcal{M}$ is equal to a hyperbolic space $\mathbb{H}^m$, then we formally set $E_{\mathrm{crit}} = +\infty$. If $\mathcal{S}$ is nonempty, then it turns out that the set $\{E(\phi) : \phi \in \mathcal{S}\}$ has a least element and that, moreover, this energy value is positive. In such case we call this least energy E_{crit}. The threshold E_{crit} depends upon geometric and topological properties of the target manifold $\mathcal{M}$; see [Lin and Wang 2008, Chapter 6] for further discussion. This definition yields $E_{\mathrm{crit}} = 4\pi$ in the case of the sphere $\mathbb{S}^2$. For further discussion of the critical energy level in the wave maps setting, see [Sterbenz and Tataru 2010b; Tao 2008a].

We now summarize what is known for Schrödinger maps in $d = 2$. Asymptotic stability of harmonic maps of topological degree $|m| \geq 4$ under the Schrödinger flow is established in [Gustafson et al. 2008]. The result is extended to maps of degree $|m| \geq 3$ in [Gustafson et al. 2010]. A certain energy-class instability for degree-1 solitons of (1-1) is shown in [Bejenaru and Tataru 2010], where it is also shown that global solutions always exist for small localized equivariant perturbations of degree-1 solitons. Finite-time blowup for (1-1) is demonstrated in [Merle et al. 2011a; 2011b], using less-localized equivariant perturbations of degree-1 solitons, thus resolving the blowup assertion of the threshold conjecture. Blow-up dynamics for equivariant critical Schrödinger maps are studied in [Perelman 2012]. Global well-posedness given data with small critical Sobolev norm (in all dimensions $d \geq 2$) is shown in [Bejenaru et al. 2011c]. Recent work of the author [Smith 2012b] extends the result of Bejenaru et al. and the present conditional result to global regularity (in $d = 2$) assuming small critical Besov norm $\dot{B}^1_{2,\infty}$. In a different direction, [Dodson and Smith 2013] shows that the L^4 norm considered in this paper is in fact a controlling norm for critical Schrödinger maps. In the radial setting (which excludes harmonic maps), Gustafson and Koo [2011] established global well-posedness at any energy level. In the equivariant setting, Bejenaru et al. [2011a] established global existence and uniqueness as well as scattering given 1-equivariant data with energy less than 4π. They note that, although these results are stated only for data with energy less than 4π, their proofs remain valid for maps with energy slightly larger than 4π, suggesting that the "right" threshold conjecture for equivariant Schrödinger maps should be stated also in terms of homotopy class, leading to a threshold of 8π rather than 4π in the case where the target is $\mathbb{S}^2$. See the introduction of [Bejenaru et al. 2011a] for further discussion of this point. This global result has been extended to the $\mathbb{H}^2$ target in [Bejenaru et al. 2012], under the assumption that the initial data has finite energy.

The main purpose of this paper is to show that (1-1) admits a unique smooth global solution φ given smooth initial data φ_0 satisfying appropriate energy conditions and assuming a priori boundedness of a certain L^4 spacetime norm of the spatial gradient of the solution φ. In particular, we admit a restricted class of initial data with energy ranging over the entire subthreshold range.

In order to go beyond the small-energy results of [Bejenaru et al. 2011c], we introduce physical-space proofs of local smoothing and bilinear Strichartz estimates, in the spirit of [Planchon and Vega 2009; Planchon 2012, p. 1042-08; Tao 2010], that do not heavily depend upon perturbative methods. The local smoothing estimate that we establish is a nonlinear analogue of that shown in [Ionescu and Kenig 2006]. The bilinear Strichartz estimate is a nonlinear analogue of the improved bilinear Strichartz estimate of [Bourgain 1998]. These proofs more naturally account for magnetic nonlinearities, and we believe the technique developed here to be of independent interest and applicable to other settings. For local smoothing in the context of Schrödinger equations, see [Kenig et al. 1993; 1998; 2004; Ionescu and Kenig 2005; 2006; 2007b]. For other Strichartz and smoothing results for magnetic Schrödinger equations, see [Stefanov 2007; D'Ancona and Fanelli 2008; D'Ancona et al. 2010; Erdoğan et al. 2008; 2009; Fanelli and Vega 2009] and the references therein. We also use in a fundamental way the subthreshold *caloric gauge* of [Smith 2012a], which is an extension of a construction introduced in [Tao 2004].

To make these statements more precise, we now turn to some basic definitions and observations.

1A. ***Preliminaries.*** First we establish some basic notation. The boldfaced letters $\mathbb{Z}$ and $\mathbb{R}$ respectively denote the integers and real numbers. We use $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ to denote the nonnegative integers. Usual Lebesgue function spaces are denoted by L^p, and these sometimes include a subscript to indicate the variable or variables of integration. When function spaces are iterated, e.g., $L_t^\infty L_x^2$, the norms are applied starting with the rightmost one. When we use L^4 without subscripts, we mean $L^4_{t,x}$.

We use $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$ to denote the standard 2-sphere embedded in 3-dimensional Euclidean space. The ambient space $\mathbb{R}^3$ carries the usual metric and $\mathbb{S}^2$ the inherited one. Throughout, $\mathbb{S}^1$ denotes the unit circle.

We use $\partial_x = (\partial_{x_1}, \partial_{x_2}) = (\partial_1, \partial_2)$ to denote the gradient operator, as throughout "∇" will stand for the Riemannian connection on $\mathbb{S}^2$. As usual, "Δ" denotes the (flat) spatial Laplacian.

The symbol $|\partial_x|^\sigma$ denotes the Fourier multiplier with symbol $|\xi|^\sigma$. We also use standard Littlewood–Paley Fourier multipliers P_k and $P_{\le k}$, respectively denoting restrictions to frequencies $\sim 2^k$ and $\lesssim 2^k$; see Section 3 for details. We use $\hat{f}$ to denote the Fourier transform of a function f in the spatial variables.

We also employ without further comment (finite-dimensional) vector-valued analogues of the above.

We use $f \lesssim g$ to denote the estimate $|f| \le C|g|$ for an absolute constant $C > 0$. As usual, the constant is allowed to change from line to line. To indicate dependence of the implicit constant upon parameters (which, for instance, can include functions), we use subscripts, e.g., $f \lesssim_k g$. As an equivalent alternative we write $f = O(g)$ (or, with subscripts, $f = O_k(g)$, for instance) to denote $|f| \le C|g|$. If both $f \lesssim g$ and $g \lesssim f$, then we indicate this by writing $f \sim g$.

Now we introduce the notion of Sobolev spaces of functions mapping from Euclidean space into $\mathbb{S}^2$. The spaces are constructed with respect to a choice of base point $Q \in \mathbb{S}^2$, the purpose of which is to define a notion of decay: instead of decaying to zero at infinity, our Sobolev class functions decay to Q.

For $\sigma \in [0, \infty)$, let $H^\sigma = H^\sigma(\mathbb{R}^2)$ denote the usual Sobolev space of complex-valued functions on $\mathbb{R}^2$. For any $Q \in \mathbb{S}^2$, set

$$H_Q^\sigma := \{f : \mathbb{R}^2 \to \mathbb{R}^3 \text{ such that } |f(x)| \equiv 1 \text{ a.e. and } f - Q \in H^\sigma\}.$$

This is a metric space with induced distance $d_Q^\sigma(f,g)=\|f-g\|_{H^\sigma}$. For $f\in H_Q^\sigma$ we set $\|f\|_{H_Q^\sigma}=d_Q^\sigma(f,Q)$ for short. We also define the spaces

$$H^\infty := \bigcap_{\sigma\in\mathbb{Z}_+} H^\sigma \quad \text{and} \quad H_Q^\infty := \bigcap_{\sigma\in\mathbb{Z}_+} H_Q^\sigma.$$

For any time $T\in(0,\infty)$, these definitions may be extended to the spacetime slab $\mathbb{R}^2\times(-T,T)$ (or $\mathbb{R}^2\times[-T,T]$). For any $\sigma,\rho\in\mathbb{Z}_+$, let $H^{\sigma,\rho}(T)$ denote the Sobolev space of complex-valued functions on $\mathbb{R}^2\times(-T,T)$ with the norm

$$\|f\|_{H^{\sigma,\rho}(T)} := \sup_{t\in(-T,T)} \sum_{\rho'=0}^{\rho} \|\partial_t^{\rho'} f(\cdot,t)\|_{H^\sigma},$$

and for $Q\in\mathbb{S}^2$ endow

$$H_Q^{\sigma,\rho} := \{f:\mathbb{R}^2\times(-T,T)\to\mathbb{R}^3 \text{ such that } |f(x,t)|\equiv 1 \text{ a.e. and } f-Q\in H^{\sigma,\rho}(T)\}$$

with the metric induced by the $H^{\sigma,\rho}(T)$ norm. Also, define the spaces

$$H^{\infty,\infty}(T)=\bigcap_{\sigma,\rho\in\mathbb{Z}_+} H^{\sigma,\rho}(T) \quad \text{and} \quad H_Q^{\infty,\infty}(T)=\bigcap_{\sigma,\rho\in\mathbb{Z}_+} H_Q^{\sigma,\rho}(T).$$

For $f\in H^\infty$ and $\sigma\geq 0$ we define the homogeneous Sobolev norms as

$$\|f\|_{\dot H^\sigma}=\|\hat f(\xi)\cdot|\xi|^\sigma\|_{L^2}.$$

We mention two important conservation laws obeyed by solutions of the Schrödinger map system (1-1). In particular, if $\varphi\in C((T_1,T_2)\to H_Q^\infty)$ solves (1-1) on a time interval (T_1,T_2), then both

$$\int_{\mathbb{R}^2}|\varphi(t)-Q|^2\,dx \quad \text{and} \quad \int_{\mathbb{R}^2}|\partial_x\varphi(t)|^2\,dx$$

are conserved. Hence the Sobolev norms H_Q^0 and H_Q^1 are conserved, as well as the energy (1-2). Note also the time-reversibility obeyed by (1-1), which in particular permits the smooth extension to $(-T,T)$ of a smooth solution on $[0,T)$.

According to our conventions,

$$|\partial_x\varphi(t)|^2 := \sum_{m=1,2}|\partial_m\varphi(t)|^2.$$

We can now give a precise statement of a key known local result.

Theorem 1.1 (local existence and uniqueness). *If the initial data φ_0 is such that $\varphi_0\in H_Q^\infty$ for some $Q\in\mathbb{S}^2$, then there exists a time $T=T(\|\varphi_0\|_{H_Q^{25}})>0$ for which there exists a unique solution φ in $C([-T,T]\to H_Q^\infty)$ of the initial value problem (1-1).*

Proof. See [Sulem et al. 1986; Chang et al. 2000; Ding and Wang 2001; McGahagan 2007] and the references therein. □

1B. ***Global theory.*** Theorem 1.1 yields short-time existence and uniqueness as well as a blowup criterion; as such it is central to the continuity arguments used for global results. In the small-energy setting, global regularity (and more) was proved for (1-1) by Bejenaru, Ionescu, Kenig, and Tataru [Bejenaru et al. 2011c]. We now state a special case of their main result, omitting for the sake of brevity the consideration of higher spatial dimensions and continuity of the solution map.

Theorem 1.2 (global regularity). *Let $Q \in \mathbb{S}^2$. Then there exists an $\varepsilon_0 > 0$ such that, for any $\varphi_0 \in H_Q^\infty$ with $\|\partial_x \varphi_0\|_{L_x^2} \leq \varepsilon_0$, there is a unique solution $\varphi \in C(\mathbb{R} \to H_Q^\infty)$ of the initial value problem (1-1). Moreover, for any $T \in [0, \infty)$ and $\sigma \in \mathbb{Z}_+$,*

$$\sup_{t \in (-T,T)} \|\varphi(t)\|_{H_Q^\sigma} \lesssim_{\sigma, T, \|\varphi_0\|_{H_Q^\sigma}} 1.$$

Also, given any $\sigma_1 \in \mathbb{Z}_+$, there exists a positive $\varepsilon_1 = \varepsilon_1(\sigma_1) \leq \varepsilon_0$ such that the uniform bounds

$$\sup_{t \in \mathbb{R}} \|\varphi(t)\|_{H_Q^\sigma} \lesssim_\sigma \|\varphi_0\|_{H_Q^\sigma}$$

hold for all $1 \leq \sigma \leq \sigma_1$, provided $\|\partial_x \varphi_0\|_{L_x^2} \leq \varepsilon_1$.

A complete proof may be found in [Bejenaru et al. 2011c]. Among the key contributions of that work are the construction of the main function spaces and the completion of the linear estimate relating them, which includes an important maximal function estimate. A significant observation made in the same paper is that it is important that these spaces take into account a local smoothing effect; the authors crucially use this effect to help bring under control the worst term of the nonlinearity. Another novelty of [Bejenaru et al. 2011c] is its implementation of the caloric gauge, which was first introduced by Tao [2004] and subsequently recommended by him for use in studying Schrödinger maps [Tao 2006a]. As the caloric gauge is defined using harmonic map heat flow, it can be thought of as an intrinsic and nonlinear analogue of classical Littlewood–Paley theory. In [Bejenaru et al. 2011c], both the intrinsic caloric gauge and the extrinsic (and modern) Littlewood–Paley theory are used simultaneously.

Our main result extends Theorem 1.2.

Theorem 1.3. *Let $T > 0$ and $Q \in S^2$. Let $\varepsilon_0 > 0$ and let $\varphi \in H_Q^{\infty,\infty}(T)$ be a solution of the Schrödinger map system (1-1) whose initial data φ_0 has energy $E_0 := E(\varphi_0) < E_{\text{crit}}$ and satisfies the energy dispersion condition*

$$\sup_{k \in \mathbb{Z}} \|P_k \partial_x \varphi_0\|_{L_x^2} \leq \varepsilon_0. \tag{1-4}$$

Let $I \supset (-T, T)$ denote the maximal time interval for which there exists a smooth (necessarily unique) extension of φ satisfying (1-1). Suppose a priori that

$$\sum_{k \in \mathbb{Z}} \|P_k \partial_x \varphi\|^2_{L^4_{t,x}(I \times \mathbb{R}^2)} \leq \varepsilon_0^2. \tag{1-5}$$

Then, for ε_0 sufficiently small,

$$\sup_{t \in (-T,T)} \|\varphi(t)\|_{H_Q^\sigma} \lesssim_{\sigma, T, \|\varphi_0\|_{H_Q^\sigma}} 1, \tag{1-6}$$

for all $\sigma \in \mathbb{Z}_+$. Additionally, $I = \mathbb{R}$, so that, in particular, φ admits a unique smooth global extension $\varphi \in C(\mathbb{R} \to H_Q^\infty)$. Moreover, for any $\sigma_1 \in \mathbb{Z}_+$, there exists a positive $\varepsilon_1 = \varepsilon_1(\sigma_1) \leq \varepsilon_0$ such that

$$\|\varphi\|_{L_t^\infty H_Q^\sigma(\mathbb{R}\times\mathbb{R}^2)} \lesssim_\sigma \|\varphi\|_{H_Q^\sigma(\mathbb{R}^2)} \tag{1-7}$$

holds for all $0 \leq \sigma \leq \sigma_1$ provided (1-4) and (1-5) hold with ε_1 in place of ε_0.

Note that the energy dispersion condition (1-4) holds automatically in the case of small energy. In such case, our proofs may be modified (essentially by collapsing to or reverting to the arguments of [Bejenaru et al. 2011c]) so that the a priori L^4 bound is not required. Such an L^4 bound, however, can then be seen to hold a posteriori.

Using time divisibility of the L^4 norm, we can replace (1-5) with

$$\sum_{k\in\mathbb{Z}} \|P_k \partial_x \varphi\|^2_{L^4_{t,x}(I\times\mathbb{R}^2)} \leq K$$

for any $K > 0$ provided we allow the threshold for ε_0 and the implicit constant in (1-7) to depend upon $K > 0$. We work with (1-5) as stated so as to avoid the additional technicalities that would arise otherwise.

We now turn to a very rough sketch of the proof of Theorem 1.3; for a detailed outline, see Section 4.

Basic setup and gauge selection. It suffices to prove homogeneous Sobolev variants of (1-6) and (1-7) over a suitable range. Thanks to mass and energy conservation, we need only consider $\sigma > 1$. For $\sigma \geq 1$, controlling $\|\varphi(t)\|_{\dot{H}^\sigma}$ is equivalent to controlling $\|\partial_x \varphi(t)\|_{\dot{H}^{\sigma-1}}$. We therefore consider the time evolution of $\partial_x \varphi$, which may be written entirely in terms of derivatives of the map φ. A more intrinsic way of expressing these equations is to select a *gauge* rather than an extrinsic embedding and coordinate system. We employ the caloric gauge, which is geometrically natural and is analytically well-suited for studying Schrödinger maps. See [Smith 2012a] for the complete details of the construction. It turns out that Sobolev bounds for the gauged derivative map imply corresponding Sobolev bounds for the ungauged derivative map. We schematically write the gauged equation as

$$(\partial_t - \Delta)\psi = \mathcal{N},$$

where ψ is $\partial_x \varphi$ placed in the caloric gauge and $\mathcal{N}$ is a nonlinearity constructed in part from ψ and $\partial_x \psi$.

Function spaces and their interrelation. To prove global results in the energy-critical setting, we of course must look for bounds other than energy estimates to control the solution. Local smoothing estimates and Strichartz estimates will be among the most important required. Our goal is to prove control over ψ within a suitable space through the use of a bootstrap argument. A standard setup requires a space, say G, for the functions ψ and a space, say N, for the nonlinearity $\mathcal{N}$. In fact, we work with stronger, frequency-localized spaces, G_k and N_k, to respectively hold $P_k \psi$ and $P_k \mathcal{N}$. We want them to be related at least by the linear estimate

$$\|P_k \psi\|_{G_k} \lesssim \|P_k \psi(t = 0)\|_{L_x^2} + \|P_k \mathcal{N}\|_{N_k}.$$

The hope, then, is to control $\|P_k\mathcal{N}\|_{N_k}$ in terms of $\|P_k\psi(t=0)\|_{L^2_x}$ and $\varepsilon\|P_k\psi\|_{G_k}$ (with ε small), so that, by proving (under a bootstrap hypothesis) a statement such as

$$\|P_k\psi\|_{G_k} \lesssim \|P_k\psi(t=0)\|_{L^2_x} + \varepsilon\|P_k\psi\|_{G_k},$$

we may conclude

$$\|P_k\psi\|_{G_k} \lesssim \|P_k\psi(t=0)\|_{L^2_x}. \tag{1-8}$$

Once (1-8) is proved, showing (1-6) and (1-7) is reduced to the comparatively easy tasks of unwinding the gauging and frequency localization steps so as to conclude with a standard continuity argument.

Controlling the nonlinearity. In this context, the main contribution of this paper lies in showing that we may conclude (1-8) without assuming small energy. The most difficult-to-control terms in the nonlinearity $P_k\mathcal{N}$ are those involving a derivative landing on high-frequency pieces of the derivative fields; we represent them schematically as $A_{\mathrm{lo}}\partial\psi_{\mathrm{hi}}$. Local smoothing estimates controlling the linear evolution (introduced in [Ionescu and Kenig 2006; 2007b]) were successfully used in [Bejenaru et al. 2011c] to handle $A_{\mathrm{lo}}\partial_x\psi_{\mathrm{hi}}$. These are not strong enough to control $A_{\mathrm{lo}}\partial_x\psi_{\mathrm{hi}}$ in the subthreshold energy setting. We instead pursue a more covariant approach, working directly with a certain covariant frequency-localized Schrödinger equation (see Section 5). Our approach is also physical-space based, in the vein of [Planchon and Vega 2009; 2012; Tao 2010], and modular.

2. Gauge field equations

In Section 2A we pass to the derivative formulation of the Schrödinger map system (1-1). All of the main arguments of our subsequent analysis take place at this level. The derivative formulation is at once both overdetermined, reflecting geometric constraints, and underdetermined, exhibiting *gauge invariance*. Section 2B introduces the caloric gauge, which is the gauge we select and work with throughout. Both Tao [2006a] and Bejenaru et al. [2011c] give good explanations justifying the use of the caloric gauge in our setting as opposed to alternative gauges. The reader is referred to [Smith 2012a] for the requisite construction of the caloric gauge for maps with energy up to E_{crit}. Section 2C deals with frequency localizing components of the caloric gauge. Proofs are postponed to Section 6 so that we can more quickly turn our attention to the gauged Schrödinger map system.

2A. ***Derivative equations.*** We begin with some constructions that are valid for any smooth function $\phi : \mathbb{R}^2 \times (-T, T) \to \mathbb{S}^2$. For a more general and extensive introduction to the gauge formalism we now introduce, see [Tao 2004]. Space and time derivatives of ϕ are denoted by $\partial_\alpha\phi(x,t)$, where $\alpha = 1, 2, 3$ ranges over the spatial variables x_1, x_2 and time t with $\partial_3 = \partial_t$.

Select a (smooth) orthonormal frame $(v(x,t), w(x,t))$ for the bundle $T_{\phi(x,t)}\mathbb{S}^2$, that is, smooth functions $v, w : \mathbb{R}^2 \times (-T, T) \to T_{\phi(x,t)}\mathbb{S}^2$ such that at each point (x,t) in the domain the vectors $v(x,t), w(x,t)$ form an orthonormal basis for $T_{\phi(x,t)}\mathbb{S}^2$. As a matter of convention we assume that v and w are chosen so that $v \times w = \phi$.

With respect to this chosen frame we then introduce the derivative fields ψ_α, setting

$$\psi_\alpha := v \cdot \partial_\alpha \phi + i w \cdot \partial_\alpha \phi. \tag{2-1}$$

Then $\partial_\alpha \phi$ admits the representation

$$\partial_\alpha \phi = v \operatorname{Re} \psi_\alpha + w \operatorname{Im} \psi_\alpha \tag{2-2}$$

with respect to the frame (v, w). The derivative fields can be thought of as arising from the following process: First, rewrite the vector $\partial_\alpha \phi$ with respect to the orthonormal basis (v, w); then, identify $\mathbb{R}^2$ with the complex numbers $\mathbb{C}$ according to $v \leftrightarrow 1$, $w \leftrightarrow i$. Note that this identification respects the complex structure of the target manifold.

Through this identification the Riemannian connection on $\mathbb{S}^2$ pulls back to a covariant derivative on $\mathbb{C}$, which we denote by

$$D_\alpha := \partial_\alpha + i A_\alpha.$$

The real-valued connection coefficients A_α are defined via

$$A_\alpha := w \cdot \partial_\alpha v, \tag{2-3}$$

so that in particular

$$\partial_\alpha v = -\phi \operatorname{Re} \psi_\alpha + w A_\alpha \quad \text{and} \quad \partial_\alpha w = -\phi \operatorname{Im} \psi_\alpha - v A_\alpha.$$

Due to the fact that the Riemannian connection on $\mathbb{S}^2$ is torsion-free, the derivative fields satisfy the relations

$$D_\beta \psi_\alpha = D_\alpha \psi_\beta. \tag{2-4}$$

or equivalently,

$$\partial_\beta A_\alpha - \partial_\alpha A_\beta = \operatorname{Im}(\psi_\beta \overline{\psi}_\alpha) =: q_{\beta\alpha}.$$

The curvature of the connection is therefore given by

$$[D_\beta, D_\alpha] := D_\beta D_\alpha - D_\alpha D_\beta = i q_{\beta\alpha}. \tag{2-5}$$

Assuming now that we are given a smooth solution φ of the Schrödinger map system (1-1), we derive the equations satisfied by the derivative fields ψ_α. The system (1-1) directly translates to

$$\psi_t = i D_l \psi_l \tag{2-6}$$

because

$$\varphi \times \Delta \varphi = J(\varphi)(\varphi^* \nabla)_j \partial_j \varphi,$$

where $J(\varphi)$ denotes the complex structure $\varphi \times$ and $(\varphi^* \nabla)_j$ the pullback of the Levi-Civita connection ∇ on the sphere.

Let us pause to note the following conventions regarding indices. Roman typeface letters are used to index spatial variables. Greek typeface letters are used to index the spatial variables along with time.

Repeated lettered indices within the same subscript or occurring in juxtaposed terms indicate an implicit summation over the appropriate set of indices.

Using (2-4) and (2-5) in (2-6) yields

$$D_t \psi_m = i D_l D_l \psi_m + q_{lm} \psi_l,$$

which is equivalent to the nonlinear Schrödinger equation

$$(i\partial_t + \Delta)\psi_m = \mathcal{N}_m, \tag{2-7}$$

where the nonlinearity $\mathcal{N}_m$ is defined by the formula

$$\mathcal{N}_m := -i A_l \partial_l \psi_m - i\partial_l (A_l \psi_m) + (A_t + A_x^2)\psi_m - i\psi_l \operatorname{Im}(\overline{\psi_l}\psi_m).$$

We split this nonlinearity as a sum $\mathcal{N}_m = B_m + V_m$, with B_m and V_m defined by

$$B_m := -i\partial_l (A_l \psi_m) - i A_l \partial_l \psi_m \tag{2-8}$$

and

$$V_m := (A_t + A_x^2)\psi_m - i\psi_l \operatorname{Im}(\overline{\psi_l}\psi_m), \tag{2-9}$$

thus separating the essentially semilinear magnetic potential terms and the essentially semilinear electric potential terms from each other.

We now state the gauge formulation of the differentiated Schrödinger map system:

$$\begin{cases} D_t \psi_m = i D_l D_l \psi_m + \operatorname{Im}(\psi_l \overline{\psi_m})\psi_l, \\ D_\alpha \psi_\beta = D_\beta \psi_\alpha, \\ \operatorname{Im}(\psi_\alpha \overline{\psi_\beta}) = \partial_\alpha A_\beta - \partial_\beta A_\alpha. \end{cases} \tag{2-10}$$

A solution ψ_m to (2-10) cannot be determined uniquely without first choosing an orthonormal frame (v, w). Changing a given choice of orthonormal frame induces a gauge transformation and may be represented as

$$\psi_m \to e^{-i\theta}\psi_m \quad \text{and} \quad A_m \to A_m + \partial_m \theta$$

in terms of the gauge components. The system (2-10) is invariant with respect to such gauge transformations.

The advantage of working with this gauge formalism rather than the Schrödinger map system or the derivative equations directly is that a carefully selected choice of gauge tames the nonlinearity. In particular, when the caloric gauge is employed, the nonlinearity in (2-7) is nearly perturbative.

2B. ***Introduction to the caloric gauge.*** In this section we introduce the caloric gauge, which is the gauge we shall employ throughout the remainder of the paper. Gauges were first used to study (1-1) in the context of proving local wellposedness in [Chang et al. 2000]. We note here that the while the Coulomb gauge would seem an attractive choice, it turns out that this gauge is not well-suited to the study of Schrödinger maps in low dimension, as in low dimension parallel interactions of waves are more probable than in high dimension, resulting in unfavorable high $\times$ high $\to$ low cascades. See [Tao 2006a] and

[Bejenaru et al. 2011c] for further discussion and a comparison of the Coulomb and caloric gauges. Also see [Tao 2006b, Chapter 6] for a discussion of various gauges that have been used in the study of wave maps.

The caloric gauge was introduced by Tao [2004] in the setting of wave maps into hyperbolic space. In a series of unpublished papers [2008a; 2008b; 2008c; 2009a; 2009b], Tao used this gauge in establishing global regularity of wave maps into hyperbolic space. In his unpublished note [Tao 2006a], Tao also suggested the caloric gauge as a suitable gauge for the study of Schrödinger maps. The caloric gauge was first used in the Schrödinger maps problem by Bejenaru, Ionescu, Kenig, and Tataru [2011c] to establish global well-posedness in the setting of initial data with sufficiently small critical norm. We recommend [Tao 2004; 2006a; 2008b; Bejenaru et al. 2011c] for background on the caloric gauge and for helpful heuristics.

Theorem 2.1 (the caloric gauge). *Let $T \in (0, \infty)$, $Q \in S^2$, and let $\phi(x,t) \in H_Q^{\infty,\infty}(T)$ be such that $\sup_{t \in (-T,T)} E(\phi(t)) < E_{\text{crit}}$. There exists a unique smooth extension $\phi(s,x,t) \in C([0,\infty) \to H_Q^{\infty,\infty}(T))$ solving the covariant heat equation*

$$\partial_s \phi = \Delta\phi + \phi \cdot |\partial_x \phi|^2 \tag{2-11}$$

and with $\phi(0,x,t) = \phi(x,t)$. Moreover, for any given choice of a (constant) orthonormal basis (v_∞, w_∞) of $T_Q\mathbb{S}^2$, there exist smooth functions $v, w : [0,\infty) \times \mathbb{R}^2 \times (-T,T) \to S^2$ such that at each point (s,x,t), the set $\{v, w, \phi\}$ naturally forms an orthonormal basis for $\mathbb{R}^3$, the gauge condition

$$w \cdot \partial_s v \equiv 0, \tag{2-12}$$

is satisfied, and

$$|\partial_x^\rho f(s)| \lesssim_\rho \langle s \rangle^{-(|\rho|+1)/2} \tag{2-13}$$

for each $f \in \{\phi - Q, v - v_\infty, w - w_\infty\}$, multiindex ρ, and $s \geq 0$.

Proof. This is a special case of the more general result [Smith 2012a, Theorem 7.6]. Whereas in [Smith 2012a] everything is stated in terms of the category of Schwartz functions, in fact this requirement may be relaxed to $H_Q^{\infty,\infty}(T)$ without difficulty (at least in the case of compact target manifolds) since weighted decay in L^2-based Sobolev spaces is not used in any proofs. □

In our application in this paper, $E(\varphi(t))$ is conserved. Therefore, we set $E_0 := E(\varphi_0)$.

Having extended v, w along the heat flow, we may likewise extend A_x along the flow. We record here for reference a technical bound that proves useful; for the proof, see [Smith 2012a, §7.1].

Theorem 2.2. *Assume the conditions of Theorem 2.1 are in force. Then we have the bound*

$$\|A_x(s)\|_{L_x^2(\mathbb{R}^2)} \lesssim_{E_0} 1. \tag{2-14}$$

Corollary 2.3 (energy bounds for the frame). *Let φ be a Schrödinger map with energy $E_0 < E_{\text{crit}}$. Then*

$$\|\partial_x v\|_{L_t^\infty L_x^2} \lesssim_{E_0} 1. \tag{2-15}$$

Proof. Because $|v| \equiv 1$, we have $v \cdot \partial_m v \equiv 0$. Therefore, with respect to the orthonormal frame (v, w, φ), the vector $\partial_m v$ admits the representation

$$\partial_m v = A_m \cdot w - \mathrm{Re}\, \psi_m \cdot \varphi. \tag{2-16}$$

The bound (2-15) then follows from using $|w| \equiv 1 \equiv |\varphi|$, $\|\psi_m\|_{L_x^2} \equiv \|\partial_m \varphi\|_{L_x^2}$, energy conservation, and (2-14) all in (2-16). □

Adopting the convention $\partial_0 = \partial_s$, and now and hereafter allowing all Greek indices to range over heat time, spatial variables, and time, we define for all $(s, x, t) \in [0, \infty) \times \mathbb{R}^2 \times (-T, T)$ the various gauge components

$$\begin{aligned} \psi_\alpha &:= v \cdot \partial_\alpha \varphi + i w \cdot \partial_\alpha \varphi, \\ A_\alpha &:= w \cdot \partial_\alpha v, \\ D_\alpha &:= \partial_\alpha + A_\alpha, \\ q_{\alpha\beta} &:= \partial_\alpha A_\beta - \partial_\beta A_\alpha. \end{aligned}$$

For $\alpha = 0, 1, 2, 3$ we have

$$\partial_\alpha \varphi = v\, \mathrm{Re}\, \psi_\alpha + w\, \mathrm{Im}\, \psi_\alpha.$$

The parallel transport condition $w \cdot \partial_s v \equiv 0$ is equivalently expressed in terms of the connection coefficients as

$$A_s \equiv 0. \tag{2-17}$$

Expressed in terms of the gauge, the heat flow (2-11) lifts to

$$\psi_s = D_l \psi_l. \tag{2-18}$$

Using (2-4) and (2-5), we may rewrite the D_m covariant derivative of (2-18) as

$$\partial_s \psi_m = D_l D_l \psi_m + i\, \mathrm{Im}(\psi_m \overline{\psi_l}) \psi_l,$$

or equivalently

$$(\partial_s - \Delta)\psi_m = i A_l \partial_l \psi_m + i \partial_l (A_l \psi_m) - A_x^2 \psi_m + i \psi_l\, \mathrm{Im}(\overline{\psi_l} \psi_m). \tag{2-19}$$

More generally, taking the D_α covariant derivative, we obtain

$$(\partial_s - \Delta)\psi_\alpha = U_\alpha, \tag{2-20}$$

where we set

$$U_\alpha := i A_l \partial_l \psi_\alpha + i \partial_l (A_l \psi_\alpha) - A_x^2 \psi_\alpha + i \psi_l\, \mathrm{Im}(\overline{\psi_l} \psi_\alpha), \tag{2-21}$$

which admits the alternative representation

$$U_\alpha = 2i A_l \partial_l \psi_\alpha + i (\partial_l A_l) \psi_\alpha - A_x^2 \psi_\alpha + i \psi_l\, \mathrm{Im}(\overline{\psi_l} \psi_\alpha). \tag{2-22}$$

From (2-5) and (2-17) it follows that

$$\partial_s A_\alpha = \mathrm{Im}(\psi_s \overline{\psi_\alpha}).$$

Integrating back from $s = \infty$ (justified using (2-13)) yields

$$A_\alpha(s) = -\int_s^\infty \operatorname{Im}(\overline{\psi_\alpha}\psi_s)(s')\,ds'. \tag{2-23}$$

At $s = 0$, φ satisfies both (1-1) and (2-11), or equivalently, $\psi_t(s = 0) = i\psi_s(s = 0)$. While for $s > 0$ it continues to be the case that $\psi_s = D_l\psi_l$ by construction, we no longer necessarily have $\psi_t(s) = iD_l(s)\psi_l(s)$, i.e., $\varphi(s, x, t)$ is not necessarily a Schrödinger map at fixed $s > 0$. In the following lemma we derive an evolution equation for the commutator $\Psi = \psi_t - i\psi_s$.

Lemma 2.4 (flows do not commute). *Set $\Psi := \psi_t - i\psi_s$. Then*

$$\partial_s\Psi = D_lD_l\Psi + i\operatorname{Im}(\psi_t\overline{\psi_l})\psi_l - \operatorname{Im}(\psi_s\overline{\psi_l})\psi_l \tag{2-24}$$

$$= D_lD_l\Psi + i\operatorname{Im}(\Psi\overline{\psi_l})\psi_l + i\operatorname{Im}(i\psi_s\overline{\psi_l})\psi_l - \operatorname{Im}(\psi_s\overline{\psi_l})\psi_l. \tag{2-25}$$

Proof. We prove (2-24), since (2-25) is a trivial consequence of it.

Applying (2-19) and (2-20) to ψ_s and ψ_t and collapsing the covariant derivative terms yields

$$\partial_s\psi_t = D_lD_l\psi_t + i\operatorname{Im}(\psi_t\overline{\psi_l})\psi_l, \tag{2-26}$$

$$\partial_s\psi_s = D_lD_l\psi_s + i\operatorname{Im}(\psi_s\overline{\psi_l})\psi_l. \tag{2-27}$$

Multiply (2-27) by i to obtain the s-evolution of $i\psi_s$. Multiplication by i commutes with D_l, but fails to do so with $\operatorname{Im}(\,\cdot\,)$, and thus we obtain

$$\partial_s i\psi_m = D_lD_l i\psi_s - \operatorname{Im}(\psi_s\overline{\psi_l})\psi_l. \tag{2-28}$$

Together (2-26) and (2-28) imply (2-24). □

2C. *Frequency localization.* Frequency localization plays an indispensable role in our analysis. In this subsection we establish some basic concepts and then state some basic results for the caloric gauge.

Our notation for a standard Littlewood–Paley frequency localization of a function f to frequencies $\sim 2^k$ is P_kf and to frequencies $\lesssim 2^k$ is $P_{\le k}f$. The particular localization chosen is of course immaterial to our analysis, but for definiteness is specified in the next section and chosen for convenience to coincide with that in [Bejenaru et al. 2011c].

We shall frequently make use of the following standard *Bernstein inequalities* for $\mathbb{R}^2$ with $\sigma \ge 0$ and $1 \le p \le q \le \infty$:

$$\|P_{\le k}|\partial_x|^\sigma f\|_{L^p_x(\mathbb{R}^2)} \lesssim_{p,\sigma} 2^{\sigma k}\|P_{\le k}f\|_{L^p_x(\mathbb{R}^2)}.$$

$$\|P_k|\partial_x|^{\pm\sigma} f\|_{L^p_x(\mathbb{R}^2)} \lesssim_{p,\sigma} 2^{\pm\sigma k}\|P_kf\|_{L^p_x(\mathbb{R}^2)}.$$

$$\|P_{\le k}f\|_{L^q_x(\mathbb{R}^2)} \lesssim_{p,q} 2^{2k(1/p-1/q)}\|P_{\le k}f\|_{L^p_x(\mathbb{R}^2)}.$$

$$\|P_kf\|_{L^q_x(\mathbb{R}^2)} \lesssim_{p,q} 2^{2k(1/p-1/q)}\|P_kf\|_{L^p_x(\mathbb{R}^2)}.$$

A particularly important notion for us is that of a frequency envelope, as it provides a way to rigorously manage the “frequency leakage” phenomenon and the frequency cascades produced by nonlinear interactions. We introduce a parameter δ in the definition; for the purposes of this paper $\delta = \frac{1}{40}$ suffices.

Definition 2.5 (frequency envelopes). A positive sequence $\{a_k\}_{k\in\mathbb{Z}}$ is a frequency envelope if it belongs to l^2 and is slowly varying:

$$a_k \le a_j 2^{\delta|k-j|}, \quad j,k\in\mathbb{Z}. \tag{2-29}$$

A frequency envelope $\{a_k\}_{k\in\mathbb{Z}}$ is ε-energy dispersed if it satisfies the additional condition

$$\sup_{k\in\mathbb{Z}} a_k \le \varepsilon.$$

Note in particular that frequency envelopes satisfy the summation rules

$$\sum_{k'\le k} 2^{pk'} a_{k'} \lesssim (p-\delta)^{-1} 2^{pk} a_k, \qquad p>\delta, \tag{2-30}$$

$$\sum_{k'\ge k} 2^{-pk'} a_{k'} \lesssim (p-\delta)^{-1} 2^{-pk} a_k, \quad p>\delta. \tag{2-31}$$

In practice we work with p bounded away from δ — for instance, $p>2\delta$ suffices — and iterate these inequalities only $O(1)$ times. Therefore, in applications we drop the factors $(p-\delta)^{-1}$ appearing in (2-30) and (2-31).

Finally, pick a positive integer σ_1 and hold it fixed throughout the remainder of this section. Results in this section hold for any such σ_1, though implicit constants are allowed to depend upon this choice.

Given initial data $\varphi_0 \in H_Q^\infty$, define for all $\sigma\ge 0$ and $k\in\mathbb{Z}$

$$c_k(\sigma) := \sup_{k'\in\mathbb{Z}} 2^{-\delta|k-k'|} 2^{\sigma k'} \|P_{k'}\partial_x\varphi_0\|_{L_x^2}. \tag{2-32}$$

Set $c_k := c_k(0)$ for short. For $\sigma\in[0,\sigma_1]$ we then have that

$$\|\partial_x\varphi_0\|_{\dot H_x^\sigma}^2 \sim \sum_{k\in\mathbb{Z}} c_k^2(\sigma) \quad \text{and} \quad \|P_k\partial_x\varphi_0\|_{L_x^2} \le c_k(\sigma)2^{-\sigma k}. \tag{2-33}$$

Similarly, for $\varphi\in H_Q^{\infty,\infty}(T)$, define for all $\sigma\ge 0$ and $k\in\mathbb{Z}$

$$\gamma_k(\sigma) := \sup_{k'\in\mathbb{Z}} 2^{-\delta|k-k'|} 2^{\sigma k'} \|P_{k'}\varphi\|_{L_t^\infty L_x^2}. \tag{2-34}$$

Set $\gamma_k := \gamma_k(1)$.

Theorem 2.6 (frequency-localized energy bounds for heat flow). *Let $f\in\{\varphi, v, w\}$. Then for $\sigma\in[1,\sigma_1]$ the bound*

$$\|P_k f(s)\|_{L_t^\infty L_x^2} \lesssim 2^{-\sigma k}\gamma_k(\sigma)(1+s2^{2k})^{-20} \tag{2-35}$$

holds and for any $\sigma,\rho\in\mathbb{Z}_+$ we have that

$$\sup_{k\in\mathbb{Z}} \sup_{s\in[0,\infty)} (1+s)^{\sigma/2} 2^{\sigma k} \|P_k\partial_t^\rho f(s)\|_{L_t^\infty L_x^2} < \infty. \tag{2-36}$$

Corollary 2.7 (frequency-localized energy bounds for the caloric gauge). *For $\sigma\in[0,\sigma_1-1]$, we have*

$$\|P_k\psi_x(s)\|_{L_t^\infty L_x^2} + \|P_k A_m(s)\|_{L_t^\infty L_x^2} \lesssim 2^k 2^{-\sigma k}\gamma_k(\sigma)(1+s2^{2k})^{-20}. \tag{2-37}$$

Moreover, for any $\sigma \in \mathbb{Z}_+$,

$$\sup_{k\in\mathbb{Z}} \sup_{s\in[0,\infty)} (1+s)^{\sigma/2} 2^{\sigma k} 2^{-k} \big(\| P_k(\partial_t^\rho \psi_x(s)) \|_{L_t^\infty L_x^2} + \| P_k(\partial_t^\rho A_x(s)) \|_{L_t^\infty L_x^2} \big) < \infty \tag{2-38}$$

and

$$\sup_{k\in\mathbb{Z}} \sup_{s\in[0,\infty)} (1+s)^{\sigma/2} 2^{\sigma k} \big(\| P_k(\partial_t^\rho \psi_t(s)) \|_{L_t^\infty L_x^2} + \| P_k(\partial_t^\rho A_t(s)) \|_{L_t^\infty L_x^2} \big) < \infty. \tag{2-39}$$

We prove Theorem 2.6 and its corollary in Section 6. Corollary 2.7 has an elementary consequence:

Corollary 2.8. *For* $\sigma \in [0, \sigma_1 - 1]$ *we have*

$$\| P_k \psi_x(0, \cdot, 0) \|_{L_x^2} \lesssim 2^{-\sigma k} c_k(\sigma). \tag{2-40}$$

3. Function spaces and basic estimates

3A. *Definitions.*

Definition 3.1 (Littlewood–Paley multipliers). Let $\eta_0 : \mathbb{R} \to [0, 1]$ be a smooth even function vanishing outside the interval $[-8/5, 8/5]$ and equal to 1 on $[-5/4, 5/4]$. For $j \in \mathbb{Z}$, set

$$\chi_j(\cdot) = \eta_0(\cdot/2^j) - \eta_0(\cdot/2^{j-1}), \qquad \chi_{\le j}(\cdot) = \eta_0(\cdot/2^j).$$

Let P_k denote the operator on $L^\infty(\mathbb{R}^2)$ defined by the Fourier multiplier $\xi \to \chi_k(|\xi|)$. For any interval $I \subset \mathbb{R}$, define the Fourier multiplier

$$\chi_I = \sum_{j \in I \cap \mathbb{Z}} \chi_j$$

and let P_I denote its corresponding operator on $L^\infty(\mathbb{R}^2)$. We shall denote $P_{(-\infty,k]}$ by $P_{\le k}$ for short. For $\theta \in \mathbb{S}^1$ and $k \in \mathbb{Z}$, we define the operators $P_{k,\theta}$ by the Fourier multipliers $\xi \to \chi_k(\xi \cdot \theta)$.

Some frequency interactions in the nonlinearity of (2-7) can be controlled using the following lemma.

Lemma 3.2 (Strichartz estimate). *Let* $f \in L_x^2(\mathbb{R}^2)$ *and* $k \in \mathbb{Z}$. *Then the Strichartz estimate*

$$\| e^{it\Delta} f \|_{L_{t,x}^4} \lesssim \| f \|_{L_x^2}$$

holds, as does the maximal function bound

$$\| e^{it\Delta} P_k f \|_{L_x^4 L_t^\infty} \lesssim 2^{k/2} \| f \|_{L_x^2}.$$

The first bound is the original Strichartz estimate [1977] and the second follows from scaling. These will be augmented with certain lateral Strichartz estimates to be introduced shortly. Strichartz estimates alone are not sufficient for controlling the nonlinearity in (2-7). The additional control required comes from local smoothing and maximal function estimates. Certain local smoothing spaces localized to cubes were introduced in [Kenig et al. 1993] to study the local well-posedness of Schrödinger equations with general derivative nonlinearities. Stronger spaces were introduced in [Ionescu and Kenig 2007a] to prove a low-regularity global result. In the Schrödinger map setting, local smoothing spaces were first used in [Ionescu and Kenig 2006] and subsequently in [Ionescu and Kenig 2007b; Bejenaru et al. 2007;

Bejenaru 2008a]. The particular local smoothing/maximal function spaces we shall use were introduced in [Bejenaru et al. 2011c].

For a unit length $\theta \in \mathbb{S}^1$, we denote by H_θ its orthogonal complement in $\mathbb{R}^2$ with the induced measure. Define the lateral spaces $L_\theta^{p,q}$ as those consisting of all measurable f for which the norm

$$\|h\|_{L_\theta^{p,q}} = \left(\int_{\mathbb{R}} \left(\int_{H_\theta \times \mathbb{R}} |h(x_1\theta + x_2, t)|^q \, dx_2 \, dt \right)^{p/q} dx_1 \right)^{1/p},$$

is finite. We make the usual modifications when $p = \infty$ or $q = \infty$. The most important spaces for our analysis are the local smoothing space $L_\theta^{\infty,2}$ and the inhomogeneous local smoothing space $L_\theta^{1,2}$. To move between these spaces we use the maximal function space $L_\theta^{2,\infty}$.

Lemma 3.3 (local smoothing [Ionescu and Kenig 2006; 2007b]). *Let $f \in L_x^2(\mathbb{R}^2)$, $k \in \mathbb{Z}$, and $\theta \in \mathbb{S}^1$. Then*

$$\|e^{it\Delta} P_{k,\theta} f\|_{L_\theta^{\infty,2}} \lesssim 2^{-k/2} \|f\|_{L_x^2}.$$

For $f \in L_x^2(\mathbb{R}^d)$, the maximal function space bound

$$\|e^{it\Delta} P_k f\|_{L_\theta^{2,\infty}} \lesssim 2^{k(d-1)/2} \|f\|_{L_x^2}$$

holds for dimension $d \geq 3$.

In $d = 2$, the maximal function bound fails due to a logarithmic divergence. In order to overcome this, we exploit Galilean invariance as in [Bejenaru et al. 2011c] (the idea goes back to [Tataru 2001] in the setting of wave maps).

For $p, q \in [1, \infty]$, $\theta \in \mathbb{S}^1$, $\lambda \in \mathbb{R}$, define $L_{\theta,\lambda}^{p,q}$ using the norm

$$\|h\|_{L_{\theta,\lambda}^{p,q}} = \|T_{\lambda\theta}(h)\|_{L_\theta^{p,q}} = \left(\int_{\mathbb{R}} \left(\int_{H_\theta \times \mathbb{R}} |h((x_1 + t\lambda)\theta + x_2, t)|^q \, dx_2 \, dt \right)^{p/q} dx_1 \right)^{1/p},$$

where T_w denotes the Galilean transformation

$$T_w(f)(x,t) = e^{-ix\cdot w/2} e^{-it|w|^2/4} f(x + tw, t).$$

With $W \subset \mathbb{R}$ finite we define the spaces $L_{\theta,W}^{p,q}$ by

$$L_{\theta,W}^{p,q} = \sum_{\lambda \in W} L_{\theta,\lambda}^{p,q}, \qquad \|f\|_{L_{\theta,W}^{p,q}} = \inf_{f = \sum_{\lambda \in W} f_\lambda} \sum_{\lambda \in W} \|f_\lambda\|_{L_{\theta,\lambda}^{p,q}}.$$

For $k \in \mathbb{Z}$, $\mathcal{K} \in \mathbb{Z}_+$, set

$$W_k := \{\lambda \in [-2^k, 2^k] : 2^{k+2\mathcal{K}} \lambda \in \mathbb{Z}\}.$$

In our application we shall work on a finite time interval $[-2^{2\mathcal{K}}, 2^{2\mathcal{K}}]$ in order to ensure that the W_k are finite. This still suffices for proving global results so long as our effective bounds are proved with constants independent of $T, \mathcal{K}$. As discussed in [Bejenaru et al. 2011c, §3], restricting T to a finite time interval avoids introducing additional technicalities.

Lemma 3.4 (local smoothing/maximal function estimates). *Let $f \in L^2_x(\mathbb{R}^2)$, $k \in \mathbb{Z}$, and $\theta \in \mathbb{S}^1$. Then*

$$\|e^{it\Delta} P_{k,\theta} f\|_{L^{\infty,2}_{\theta,\lambda}} \lesssim 2^{-k/2} \|f\|_{L^2_x}, \quad |\lambda| \le 2^{k-40},$$

and if $T \in (0, 2^{2\mathcal{K}}]$, then

$$\|1_{[-T,T]}(t) e^{it\Delta} P_k f\|_{L^{2,\infty}_{\theta,W_{k+40}}} \lesssim 2^{k/2} \|f\|_{L^2_x}.$$

Proof. The first bound follows from Lemma 3.3 via a Galilean boost. The second is more involved and is proven in [Bejenaru et al. 2011c, §7]. □

Lemma 3.5 (lateral Strichartz estimates). *Let $f \in L^2_x(\mathbb{R}^2)$, $k \in \mathbb{Z}$, and $\theta \in \mathbb{S}^1$. Let $2 < p \le \infty$, $2 \le q \le \infty$ and $1/p + 1/q = 1/2$. Then*

$$\|e^{it\Delta} P_{k,\theta} f\|_{L^{p,q}_\theta} \lesssim 2^{k(2/p-1/2)} \|f\|_{L^2_x}, \qquad p \ge q,$$
$$\|e^{it\Delta} P_k f\|_{L^{p,q}_\theta} \lesssim_p 2^{k(2/p-1/2)} \|f\|_{L^2_x}, \qquad p \le q.$$

Proof. Informally speaking, these bounds follow from interpolating between the L^4 Strichartz estimate and the local smoothing/maximal function estimates of Lemma 3.4. See [Bejenaru et al. 2011c, Lemma 7.1] for the rigorous argument. □

We now introduce the main function spaces. Let $T > 0$. For $k \in \mathbb{Z}$, let $I_k = \{\xi \in \mathbb{R}^2 : |\xi| \in [2^{k-1}, 2^{k+1}]\}$. Let

$$L^2_k(T) := \{f \in L^2(\mathbb{R}^2 \times [-T, T]) : \operatorname{supp} \hat{f}(\xi, t) \subset I_k \times [-T, T]\}.$$

For $f \in L^2(\mathbb{R}^2 \times [-T, T])$, let

$$\|f\|_{F^0_k(T)} := \|f\|_{L^\infty_t L^2_x} + \|f\|_{L^4_{t,x}} + 2^{-k/2} \|f\|_{L^4_x L^\infty_t} + 2^{-k/6} \sup_{\theta \in \mathbb{S}^1} \|f\|_{L^{3,6}_\theta} + +2^{-k/2} \sup_{\theta \in \mathbb{S}^1} \|f\|_{L^{2,\infty}_{\theta,W_{k+40}}}.$$

We then define, similarly to what is done in [Bejenaru et al. 2011c], $F_k(T)$, $G_k(T)$, $N_k(T)$ as the normed spaces of functions in $L^2_k(T)$ for which the corresponding norms

$$\|f\|_{F_k(T)} := \inf_{J, m_1, \dots, m_J \in \mathbb{Z}_+} \inf_{f = f_{m_1} + \dots + f_{m_J}} \sum_{j=1}^{J} 2^{m_j} \|f_{m_j}\|_{F^0_{k+m_j}},$$

$$\|f\|_{G_k(T)} := \|f\|_{F^0_k(T)} + 2^{k/6} \sup_{|j-k| \le 20} \sup_{\theta \in \mathbb{S}^1} \|P_{j,\theta} f\|_{L^{6,3}_\theta} + 2^{k/2} \sup_{|j-k| \le 20} \sup_{\theta \in \mathbb{S}^1} \sup_{|\lambda| < 2^{k-40}} \|P_{j,\theta} f\|_{L^{\infty,2}_{\theta,\lambda}},$$

$$\begin{aligned}\|f\|_{N_k(T)} := \inf_{f = f_1 + f_2 + f_3 + f_4 + f_5 + f_6} &\|f_1\|_{L^{4/3}_{t,x}} + 2^{k/6} \|f_2\|_{L^{3/2,6/5}_{\hat{\theta}_1}} + 2^{k/6} \|f_3\|_{L^{3/2,6/5}_{\hat{\theta}_2}} \\ &+ 2^{-k/6} \|f_4\|_{L^{6/5,3/2}_{\hat{\theta}_1}} + 2^{-k/6} \|f_5\|_{L^{6/5,3/2}_{\hat{\theta}_2}} + 2^{-k/2} \sup_{\theta \in \mathbb{S}^1} \|f_6\|_{L^{1,2}_{\theta,W_{k-40}}},\end{aligned}$$

are finite, where $(\hat{\theta}_1, \hat{\theta}_2)$ denotes the canonical basis in $\mathbb{R}^2$.

There are a few minor differences between these spaces and those appearing in [Bejenaru et al. 2011c]. The space F^0_k now includes the lateral Strichartz space $L^{3,6}_\theta$, whereas in that reference, only G_k was endowed with this norm. The net effect on the space G_k is that it is left unchanged. The space F_k, however, now explicitly incorporates this particular lateral Strichartz structure. Note though, that for fixed $\theta \in \mathbb{S}^1$, we have by enough applications of Young's and Hölder's inequalities that

$$\begin{aligned}2^{-k/6}\|f\|_{L^{3,6}_\theta} &= 2^{-k/6}\bigg(\int_{\mathbb{R}}\bigg(\int_{H_\theta\times\mathbb{R}}|f(x_1\theta+x_2,t)|^6\,dx_2\,dt\bigg)^{1/2}dx_1\bigg)^{1/3}\\ &\lesssim 2^{-k/6}\bigg(\int_{\mathbb{R}}\|f\|^2_{L^4_{\theta,t}}\|f\|_{L^\infty_{\theta,t}}\,dx_1\bigg)^{1/3}\\ &\lesssim 2^{-k/6}\bigg(\int_{\mathbb{R}}\|f\|^4_{L^4_{\theta,t}}\,dx_1\bigg)^{1/6}\bigg(\int_{\mathbb{R}}\|f\|^2_{L^\infty_{\theta,t}}dx_1\bigg)^{1/6}\\ &\lesssim \|f\|^{2/3}_{L^4}\cdot 2^{-k/6}\|f\|^{1/3}_{L^{2,\infty}_\theta} \lesssim \|f\|_{L^4}+2^{-k/2}\|f\|_{L^{2,\infty}_\theta}.\end{aligned}$$

We also make one change to the N_k space: We explicitly incorporate $L^{6/5,3/2}_\theta$.

Incorporating these extra lateral Strichartz spaces affords us greater flexibility in certain estimates: We can avoid having to use local smoothing/maximal function spaces if we are willing to give up some decay. This tradeoff pays off in Section 5, where as a consequence we can prove a stronger local smoothing estimate for a certain magnetic nonlinear Schrödinger equation in the one regime where this improvement is absolutely essential.

Proposition 3.6 (main linear estimate). *Assume $\mathcal{K} \in \mathbb{Z}_+$, $T \in (0, 2^{2\mathcal{K}}]$ and $k \in \mathbb{Z}$. Then for each $u_0 \in L^2$ that is frequency-localized to I_k and for any $h \in N_k(T)$, the solution u of*

$$(i\partial_t + \Delta_x)u = h, \quad u(0) = u_0,$$

satisfies

$$\|u\|_{G_k(T)} \lesssim \|u(0)\|_{L^2_x} + \|h\|_{N_k(T)}.$$

Proof. See [Bejenaru et al. 2011c, Proposition 7.2] for details. Our changes to the spaces necessitate only minor changes in their proof, as we must incorporate $L^{6/5,3/2}_{\hat{\theta}_1}$ and $L^{6/5,3/2}_{\hat{\theta}_2}$ into the space $N^0_k(T)$. □

The spaces $G_k(T)$ are used to hold projections $P_k\psi_m$ of the derivative fields ψ_m satisfying (2-7). The main components of $G_k(T)$ are the local smoothing/maximal function spaces $L^{\infty,2}_{\theta,\lambda}$, $L^{2,\infty}_{\theta,W_{k+40}}$, and the lateral Strichartz spaces. The local smoothing and maximal function space components play an essential role in recovering the derivative loss that is due to the magnetic nonlinearity.

The spaces $N_k(T)$ hold frequency projections of the nonlinearities in (2-7). Here the main spaces are the inhomogeneous local smoothing spaces $L^{1,2}_{\theta,W_{k-40}}$ and the Strichartz spaces, both chosen to match those of $G_k(T)$.

The spaces $G_k(T)$ clearly embed in $F_k(T)$. Two key properties enjoyed only by the larger spaces $F_k(T)$ are

$$\|f\|_{F_k(T)} \approx \|f\|_{F_{k+1}(T)},$$

for $k \in \mathbb{Z}$ and $f \in F_k(T) \cap F_{k+1}(T)$, and

$$\|P_k(uv)\|_{F_k(T)} \lesssim \|u\|_{F_{k'}(T)}\|v\|_{L^\infty_{t,x}}$$

for $k, k' \in \mathbb{Z}$, $|k-k'| \le 20$, $u \in F_{k'}(T)$, $v \in L^\infty(\mathbb{R}^2 \times [-T, T])$. Both of these properties follow readily from the definitions.

In order to bound the nonlinearity of (2-7) in $N_k(T)$, it is important to gain regularity from the parabolic heat-time smoothing effect. The desired frequency-localized bounds do not (or at least not so readily) propagate in heat-time in the spaces $G_k(T)$, whereas these bounds do propagate with decay in the larger spaces $F_k(T)$. Note that since the $F_k(T)$ norm is translation invariant, we have

$$\|e^{s\Delta}h\|_{F_k(T)} \lesssim (1+s2^{2k})^{-20}\|h\|_{F_k(T)}, \quad s \ge 0,$$

for $h \in F_k(T)$. In certain bilinear estimates we do not need the full strength of the spaces $F_k(T)$ and instead can use the bound

$$\|f\|_{F_k(T)} \lesssim \|f\|_{L^2_x L^\infty_t} + \|f\|_{L^4_{t,x}}, \tag{3-1}$$

which follows from

$$\|f\|_{L^{2,\infty}_{\theta, W_{k+m_j}}} \le \|f\|_{L^{2,\infty}_\theta} \lesssim 2^{k/2}\|f\|_{L^2_x L^\infty_t}.$$

We introduce one more class of function spaces. These can be viewed as a refinement of the Strichartz part of $F_k(T)$. For $k \in \mathbb{Z}$ and $\omega \in [0, 1/2]$ we define $S_k^\omega(T)$ to be the normed space of functions belonging to $L^2_k(T)$ whose norm

$$\|f\|_{S_k^\omega(T)} = 2^{\omega k}\big(\|f\|_{L^\infty_t L^{2_\omega}_x} + \|f\|_{L^4_t L^{p_\omega}_x} + 2^{-k/2}\|f\|_{L^{p_\omega}_x L^\infty_t}\big) \tag{3-2}$$

is finite, where the exponents 2_ω and p_ω are determined by

$$\frac{1}{2_\omega} - \frac{1}{2} = \frac{1}{p_\omega} - \frac{1}{4} = \frac{\omega}{2}.$$

Note that $F_k(T) \hookrightarrow S_k^0(T)$ and that by Bernstein we have

$$\|f\|_{S_k^{\omega'}(T)} \lesssim \|f\|_{S_k^\omega(T)}, \quad \omega' \le \omega.$$

3B. *Bilinear estimates.*

Lemma 3.7 (bilinear estimates on $N_k(T)$). *For $k, k_1, k_3 \in \mathbb{Z}$, $h \in L^2_{t,x}$, $f \in F_{k_1}(T)$, and $g \in G_{k_3}(T)$, we have the following inequalities under the given restrictions on k_1, k_3:*

$$\|P_k(hf)\|_{N_k(T)} \lesssim \|h\|_{L^2_{t,x}}\|f\|_{F_{k_1}(T)} \qquad \text{if } |k_1-k| \le 80. \tag{3-3}$$

$$\|P_k(hf)\|_{N_k(T)} \lesssim 2^{-|k-k_1|/6}\|h\|_{L^2_{t,x}}\|f\|_{F_{k_1}(T)} \qquad \text{if } k_1 \le k-80. \tag{3-4}$$

$$\|P_k(hg)\|_{N_k(T)} \lesssim 2^{-|k-k_3|/6}\|h\|_{L^2_{t,x}}\|g\|_{G_{k_3}(T)} \qquad \text{if } k \le k_3-80. \tag{3-5}$$

Proof. Estimate (3-3) follows from Hölder's inequality and the definition of $F_k(T)$, $N_k(T)$:

$$\|Ff\|_{L^{4/3}} \le \|F\|_{L^2}\|f\|_{L^4}.$$

For (3-4) and (3-5), we use an angular partition of unity in frequency to write

$$f = f_1 + f_2, \quad \|f_1\|_{L^{3,6}_{\hat\theta_1}} + \|g_1\|_{L^{3,6}_{\hat\theta_2}} \lesssim 2^{k_1/6}\|f\|_{F_k(T)},$$
$$g = g_1 + g_2, \quad \|g_1\|_{L^{6,3}_{\hat\theta_1}} + \|g_1\|_{L^{6,3}_{\hat\theta_2}} \lesssim 2^{-k_1/6}\|g\|_{G_k(T)}.$$

Then

$$\begin{aligned}\|P_k(Ff)\|_{N_k(T)} &\lesssim 2^{-k/6}\big(\|Ff_1\|_{L^{6/5,3/2}_{\hat\theta_1}} + \|Ff_2\|_{L^{6/5,3/2}_{\hat\theta_2}}\big) \lesssim 2^{-k/6}\|F\|_{L^2}\big(\|f_1\|_{L^{3,6}_{\hat\theta_1}} + \|f_1\|_{L^{3,6}_{\hat\theta_2}}\big)\\ &\lesssim 2^{(k_1-k)/6}\|F\|_{L^2}\|f\|_{F_{k_1}(T)},\\ \|P_k(Fg)\|_{N_k(T)} &\lesssim 2^{k/6}\big(\|Fg_1\|_{L^{3/2,6/5}_{\hat\theta_1}} + \|Fg_2\|_{L^{3/2,6/5}_{\hat\theta_2}}\big) \lesssim 2^{k/6}\|F\|_{L^2}\big(\|g_1\|_{L^{6,3}_{\hat\theta_1}} + \|g_1\|_{L^{6,3}_{\hat\theta_2}}\big)\\ &\lesssim 2^{(k-k_1)/6}\|F\|_{L^2}\|g\|_{G_{k_3}(T)}.\end{aligned}$$

□

Lemma 3.8 (bilinear estimates on $L^2_{t,x}$). *For $k_1, k_2, k_3 \in \mathbb{Z}$, $f_1 \in F_{k_1}(T)$, $f_2 \in F_{k_2}(T)$, and $g \in G_{k_3}(T)$, we have*

$$\|f_1 \cdot f_2\|_{L^2_{t,x}} \lesssim \|f_1\|_{F_{k_1}(T)}\|f_2\|_{F_{k_2}(T)}, \tag{3-6}$$

$$\|f \cdot g\|_{L^2_{t,x}} \lesssim 2^{-|k_1-k_3|/6}\|f\|_{F_{k_1}(T)}\|g\|_{G_{k_3}(T)} \quad \textit{for } k_1 \le k_3. \tag{3-7}$$

Proof. It suffices to show that

$$\|fg\|_{L^2} \lesssim \|f\|_{F^0_{k_1}(T)}\|g\|_{G_{k_2}(T)} \qquad \text{for } k_1 \ge k_2 - 100, \tag{3-8}$$

$$\|fg\|_{L^2} \lesssim 2^{(k_1-k_2)/6}\|f\|_{F^0_{k_1}(T)}\|g\|_{G_{k_2}(T)} \qquad \text{for } k_1 < k_2 - 100. \tag{3-9}$$

Estimate (3-8) follows from estimating each factor in L^4. For (3-9), we first observe that, using a smooth partition of unity in frequency space, we may assume that $\hat g$ is supported in the set

$$\big\{\xi : |\xi| \in [2^{k_2-1}, 2^{k_2+1}] \text{ and } \xi \cdot \theta_0 \ge 2^{k_2-5}\big\}$$

for some direction $\theta_0 \in \mathbb{S}^1$. Then $\|fg\|_{L^2} \lesssim \|f\|_{L^{3,6}_{\theta_0}}\|g\|_{L^{6,3}_{\theta_0}} \lesssim 2^{(k_1-k_2)/6}\|f\|_{F^0_{k_1}(T)}\|g\|_{G_{k_2}(T)}$. □

We also have the following stronger estimates, which rely upon the local smoothing and maximal function spaces.

Lemma 3.9 (bilinear estimates using local smoothing/maximal function bounds). *For $k, k_1, k_2 \in \mathbb{Z}$, $h \in L^2_{t,x}$, $f \in F_{k_1}(T)$, $g \in G_{k_2}(T)$, we have, under the given restrictions on k_1, k_2:*

$$\|P_k(hf)\|_{N_k(T)} \lesssim 2^{-|k-k_1|/2}\|h\|_{L^2_{t,x}}\|f\|_{F_{k_1}(T)} \qquad \textit{if } k_1 \le k - 80. \tag{3-10}$$

$$\|f \cdot g\|_{L^2_{t,x}} \lesssim 2^{-|k_1-k_2|/2}\|f\|_{F_{k_1}(T)}\|g\|_{G_{k_2}(T)} \qquad \textit{if } k_1 \le k_2. \tag{3-11}$$

Proof. Estimate (3-10) follows from the definitions since

$$\|P_k(hf)\|_{N_k(T)} \lesssim 2^{-k/2}\sup_{\theta\in\mathbb{S}^1}\|hf\|_{L^{1,2}_{\theta,W_{k-40}}} \lesssim 2^{-k/2}\sup_{\theta\in\mathbb{S}^1}\|f\|_{L^{2,\infty}_{\theta,W_{k_1+40}}}\|h\|_{L^2_{t,x}}.$$

The proof of (3-11) parallels that of (3-7) and is omitted (see [Bejenaru et al. 2011c, Lemma 6.5] for details). □

3C. ***Trilinear estimates and summation.*** We combine the bilinear estimates to establish some trilinear estimates. As we do not control local smoothing norms along the heat flow, we will oftentimes be able to put only one term in a G_k space. Nonetheless, such estimates still exhibit good off-diagonal decay.

Define the sets $Z_1(k), Z_2(k), Z_3(k) \subset \mathbb{Z}^3$ as follows:

$$\begin{aligned} Z_1(k) &:= \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k_1, k_2 \le k - 40 \text{ and } |k_3 - k| \le 4\}. \\ Z_2(k) &:= \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k, k_3 \le k_1 - 40 \text{ and } |k_2 - k_1| \le 45\}. \\ Z_3(k) &:= \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : |\max\{k, k_3\} - \max\{k_1, k_2\}| \le 40\}. \end{aligned} \tag{3-12}$$

In our main trilinear estimate, we avoid using local smoothing/maximal function spaces.

Lemma 3.10 (main trilinear estimate). *Let C_{k,k_1,k_2,k_3} denote the best constant C in the estimate*

$$\|P_k\left(P_{k_1} f_1 P_{k_2} f_2 P_{k_3} g\right)\|_{N_k(T)} \lesssim C \|P_{k_1} f_1\|_{F_{k_1}(T)} \|P_{k_2} f_2\|_{F_{k_2}(T)} \|P_{k_3} g\|_{G_{k_3}(T)}. \tag{3-13}$$

The best constant C_{k,k_1,k_2,k_3} satisfies the bounds

$$C_{k,k_1,k_2,k_3} \lesssim \begin{cases} 2^{-|(k_1+k_2)/6 - k/3|} & \textit{if } (k_1, k_2, k_3) \in Z_1(k), \\ 2^{-|k-k_3|/6} & \textit{if } (k_1, k_2, k_3) \in Z_2(k), \\ 2^{-|\Delta k|/6} & \textit{if } (k_1, k_2, k_3) \in Z_3(k), \\ 0 & \textit{if } (k_1, k_2, k_3) \in \mathbb{Z}^3 \setminus \{Z_1(k) \cup Z_2(k) \cup Z_3(k)\}, \end{cases}$$

where $\Delta k = \max\{k, k_1, k_2, k_3\} - \min\{k, k_1, k_2, k_3\} \ge 0$.

Proof. After placing the term $P_k(P_{k_1} f_1 P_{k_2} f_2 P_{k_3} g)$ in $L^{4/3}_{t,x}$ and then using Hölder's inequality to bound each factor in $L^4_{t,x}$, it follows from Bernstein that

$$C_{k,k_1,k_2,k_3} \lesssim 1, \tag{3-14}$$

and so, in particular, for any choice of integers k, k_1, k_2, k_3, such a constant C_{k,k_1,k_2,k_3} exists.

Frequencies not represented in one of $Z_1(k)$, $Z_2(k)$, $Z_3(k)$ cannot interact so as to yield a frequency in I_k. Over $Z_1(k)$, we apply (3-4) and (3-7).

On $Z_2(k)$ we apply (3-4) if $k > k_3$ and (3-5) if $k \le k_3$. We conclude with (3-6).

On $Z_3(k)$ we may assume without loss of generality that $k_1 \le k_2$. First suppose that $k_3 \le k$ and $|k - k_2| \le 40$. If $k_1 \le k_3$, then use (3-4), applying (3-6) to $P_{k_2} f_2 P_{k_3} g$. If $k_3 < k_1$, then use (3-6) on $P_{k_1} f_1 P_{k_2} f_2$ instead.

Now suppose that $k_3 > k$ and $|k_3 - k_2| \le 40$. If $k_1 \le k$, then use (3-3), applying (3-7) to $P_{k_1} f_1 P_{k_3} g$. If $k_{\min} = k$, then use (3-5) and (3-6). □

Corollary 3.11. *Let $\{a_k\}$, $\{b_k\}$, $\{c_k\}$ be δ-frequency envelopes. Let C_{k,k_1,k_2,k_3} be as in Lemma 3.10. Then*

$$\sum_{(k_1,k_2,k_3) \in \mathbb{Z}^3 \setminus Z_2(k)} C_{k,k_1,k_2,k_3} a_{k_1} b_{k_2} c_{k_3} \lesssim a_k b_k c_k.$$

Proof. By Lemma 3.10, it suffices to restrict the sum to (k_1,k_2,k_3) lying in $Z_1(k)\cup Z_3(k)$. On $Z_1(k)$, the sum is bounded by

$$\begin{aligned}\sum_{(k_1,k_2,k_3)\in Z_1(k)} 2^{-|(k_1+k_2)/6-k/3|}a_{k_1}b_{k_2}c_{k_3} &\lesssim \sum_{k_1,k_2\le k-40} 2^{-|(k_1+k_2)/6-k/3|}2^{\delta|2k-k_1-k_2|}a_kb_kc_k\\ &\lesssim a_kb_kc_k.\end{aligned}$$

On Z_3, we may assume without loss of generality that $k_2\le k_1$. The sum is then controlled by

$$\begin{aligned}\sum_{(k_1,k_2,k_3)\in Z_3(k)} 2^{-|\Delta k|/6}a_{k_1}b_{k_2}c_{k_3} &\lesssim \sum_{\substack{k_2\le k\\ k_3\le k\\ |k_1-k|\le 40}} 2^{-|k-\min\{k_2,k_3\}|/6}a_{k_1}b_{k_2}c_{k_3} + \sum_{\substack{k_2\le k_1\\ k_1>k\\ |k_3-k_1|\le 40}} 2^{-|k_1-\min\{k_2,k\}|/6}a_{k_1}b_{k_2}c_{k_3}\\ &\lesssim \sum_{\substack{k_2\le k\\ k_3\le k}} 2^{-|k-\min\{k_2,k_3\}|/6}a_kb_{k_2}c_{k_3} + \sum_{\substack{k_2\le k_1\\ k_1>k}} 2^{-|k_1-\min\{k_2,k\}|/6}a_{k_1}b_{k_2}c_{k_1}.\end{aligned}$$

The first of these summands is controlled by

$$\begin{aligned}&\sum_{k_3\le k_2\le k} 2^{-|k-k_3|/6}a_kb_{k_2}c_{k_3} + \sum_{k_2<k_3\le k} 2^{-|k-k_2|/6}a_kb_{k_2}c_{k_3}\\ &\qquad\lesssim \sum_{k_3\le k_2\le k} 2^{-|k-k_3|/6}2^{\delta|k-k_2|}a_kb_kc_{k_3} + \sum_{k_2<k_3\le k} 2^{-|k-k_2|/6}2^{\delta|k-k_3|}a_kb_{k_2}c_k\\ &\qquad\lesssim \sum_{k_3\le k} 2^{(\delta-1/6)|k-k_3|}a_kb_kc_{k_3} + \sum_{k_2<k} 2^{(\delta-1/6)|k-k_2|}a_kb_{k_2}c_k\\ &\qquad\lesssim \sum_{k_3\le k} 2^{(2\delta-1/6)|k-k_3|}a_kb_kc_k + \sum_{k_2<k} 2^{(2\delta-1/6)|k-k_2|}a_kb_kc_k\\ &\qquad\lesssim a_kb_kc_k.\end{aligned}$$

The second is controlled by

$$\begin{aligned}&\sum_{k\le k_2\le k_1} 2^{-|k_1-k|/6}a_{k_1}b_{k_2}c_{k_1} + \sum_{k_2<k\le k_1} 2^{-|k_1-k_2|/6}a_{k_1}b_{k_2}c_{k_1}\\ &\qquad\lesssim \sum_{k\le k_2\le k_1} 2^{-|k_1-k|/6}2^{\delta|k_2-k|}a_{k_1}b_kc_{k_1} + \sum_{k_2<k\le k_1} 2^{|k_1-k_2|/6}2^{\delta|k_2-k|}a_{k_1}b_kc_{k_1}\\ &\qquad\lesssim \sum_{k\le k_1} 2^{(\delta-1/6)|k_1-k|}a_{k_1}b_kc_{k_1} + \sum_{k_2<k\le k_1} 2^{(\delta-1/6)|k_1-k_2|}a_{k_1}b_kc_{k_1}\\ &\qquad\lesssim \sum_{k\le k_1} 2^{(3\delta-1/6)|k_1-k|}a_kb_kc_k + \sum_{k_2<k\le k_1} 2^{(3\delta-1/6)|k_1-k_2|}a_kb_kc_k\\ &\qquad\lesssim a_kb_kc_k.\end{aligned}$$

□

Corollary 3.12. *Let $\{a_k\}$, $\{b_k\}$ be δ-frequency envelopes. Let C_{k,k_1,k_2,k_3} be as in Lemma 3.10. Then*

$$\sum_{(k_1,k_2,k_3)\in Z_2(k)\cup Z_3(k)} 2^{\max\{k,k_3\}-\max\{k_1,k_2\}}C_{k,k_1,k_2,k_3}a_{k_1}b_{k_2}c_{k_3} \lesssim a_kb_kc_k.$$

Proof. On $Z_3(k)$, $\max\{k_1,k_2\}\sim\max\{k,k_3\}$, and so the bound on $Z_3(k)$ follows from Corollary 3.11.

Note that $\max\{k_1, k_2\} > \max\{k, k_3\}$ on Z_2, where the sum is controlled by

$$\sum_{(k_1,k_2,k_3)\in Z_2(k)} 2^{\max\{k,k_3\}-\max\{k_1,k_2\}} 2^{-|k-k_3|/6} a_{k_1} b_{k_2} c_{k_3} \lesssim \sum_{k,k_3\le k_1-40} 2^{\max\{k,k_3\}-k_1} 2^{-|k-k_3|/6} a_{k_1} b_{k_1} c_{k_3},$$

Restricting the sum to $k_3 \le k$, we get

$$\sum_{k_3\le k\le k_1-40} 2^{-|k-k_1|} 2^{-|k-k_3|/6} a_{k_1} b_{k_1} c_{k_3} \lesssim a_k b_k c_k.$$

Over the complementary range $k \le k_3 \le k_1 - 40$, we have

$$\sum_{k\le k_3\le k_1-40} 2^{-|k_3-k_1|} 2^{-|k-k_3|/6} a_{k_1} b_{k_1} c_{k_3} \lesssim a_k b_k c_k \sum_{k\le k_3\le k_1-40} 2^{-|k_3-k_1|} 2^{-|k-k_3|/6} 2^{2\delta|k_1-k|} 2^{\delta|k-k-3|}.$$

Performing the change of variables $j := k_1 - k_3$, $l := k_3 - k$, we control the sum by

$$\sum_{j,l\ge 0} 2^{-j} 2^{-l/6} 2^{2\delta(j+l)} 2^{\delta l} \lesssim \sum_{j,l\ge 0} 2^{(2\delta-1)j} 2^{(3\delta-1/6)l} \lesssim 1. \qquad \square$$

Taking advantage of the local smoothing/maximal function spaces, we can obtain the following improvement.

Lemma 3.13 (main trilinear estimate improvement over Z_1). *The best constant C_{k,k_1,k_2,k_3} in (3-13) satisfies the improved estimate*

$$C_{k,k_1,k_2,k_3} \lesssim 2^{-|(k_1+k_2)/2-k|} \tag{3-15}$$

when $\{k_1, k_2, k_3\} \in Z_1(k)$.

4. Proof of Theorem 1.3

In this section we outline the proof of Theorem 1.3, taking as our starting point the local result stated in Theorem 1.1.

For technical reasons related to the function space definitions of the last section, it will be convenient to construct a solution φ on a time interval $(-2^{2\mathcal{K}}, 2^{2\mathcal{K}})$ for some given $\mathcal{K} \in \mathbb{Z}_+$ and proceed to prove bounds that are uniform in $\mathcal{K}$. We assume $1 \ll \mathcal{K} \in \mathbb{Z}_+$ is chosen and hereafter fixed. Invoking Theorem 1.1, we assume that we have a solution $\varphi \in C([-T, T] \to H^\infty_Q)$ of (1-1) on the time interval $[-T, T]$ for some $T \in (0, 2^{2\mathcal{K}})$. In order to extend φ to a solution on all of $(-2^{2\mathcal{K}}, 2^{2\mathcal{K}})$ with uniform bounds (uniform in $T, \mathcal{K}$), it suffices to prove uniform a priori estimates on

$$\sup_{t\in(-T,T)} \|\varphi(t)\|_{H^\sigma_Q}$$

for, say, σ in the interval $[1, \sigma_1]$, with $\sigma_1 \gg 1$ chosen sufficiently large ($\sigma_1 = 25$ will do).

The first step in our approach, carried out in Section 2, is to lift the Schrödinger map system (1-1) to the tangent bundle and view it with respect to the caloric gauge. Recall that the lift of (1-1) expressed in terms of the caloric gauge takes the form (2-7), or, equivalently,

$$(i\partial_t + \Delta)\psi_m = B_m + V_m, \tag{4-1}$$

with initial data $\psi_m(0)$. Here B_m and V_m respectively denote the magnetic and electric potentials (see (2-8) and (2-9) for definitions).

The goal then becomes proving a priori bounds on $\|\psi_m\|_{L_t^\infty H_x^\sigma}$. Herein lies the heart of the argument, and the purpose of this section is not only to give a high level description of the proof of Theorem 1.3, but also to outline the proof of the key a priori bounds. To establish these bounds, we in fact prove stronger frequency-localized estimates. The argument naturally splits into several components, and we consider each individually below.

Finally, to complete the proof of Theorem 1.3, we must transfer the a priori bounds on the derivative fields ψ_m back to bounds on the map φ, thereby allowing us to close a bootstrap argument. Once the derivative field bounds are established, this is, comparatively speaking, an easy task, and we take it up in the last subsection.

We return now to (4-1), projecting it to frequencies $\sim 2^k$ using the Littlewood–Paley multiplier P_k. Applying the linear estimate of Proposition 3.6 then yields

$$\|P_k\psi_m\|_{G_k(T)} \lesssim \|P_k\psi_m(0)\|_{L_x^2} + \|P_kV_m\|_{N_k(T)} + \|P_kB_m\|_{N_k(T)}. \tag{4-2}$$

In order to express control of the $G_k(T)$ norm of $P_k\psi_m$ in terms of the initial data, we introduce the following frequency envelopes. Let $\sigma_1 \in \mathbb{Z}_+$ be positive. For $\sigma \in [0, \sigma_1 - 1]$, set

$$b_k(\sigma) = \sup_{k' \in \mathbb{Z}} 2^{\sigma k'} 2^{-\delta|k-k'|} \|P_{k'}\psi_x\|_{G_k(T)}. \tag{4-3}$$

By (2-38), these envelopes are finite and in l^2. We abbreviate $b_k(0)$ by setting $b_k := b_k(0)$.

We now state the key result for solutions of the gauge field equation (4-1).

Theorem 4.1. *Assume $T \in (0, 2^{2\mathcal{K}})$ and $Q \in \mathbb{S}^2$. Choose $\sigma_1 \in \mathbb{Z}_+$ positive. Let $\varepsilon_1 > 0$ and let $\varphi \in H_Q^{\infty,\infty}(T)$ be a solution of the Schrödinger map system (1-1) whose initial data φ_0 has energy $E_0 := E(\varphi_0) < E_{\text{crit}}$ and satisfies the energy dispersion condition*

$$\sup_{k\in\mathbb{Z}} c_k \le \varepsilon_1. \tag{4-4}$$

Assume moreover that

$$\sum_{k\in\mathbb{Z}} \|P_k\psi_x\|^2_{L^4_{t,x}(I\times\mathbb{R}^2)} \le \varepsilon_1^2 \tag{4-5}$$

for any smooth extension φ on I, $[-T, T] \subset I \subset (-2^{2\mathcal{K}}, 2^{2\mathcal{K}})$. Suppose that the bootstrap hypothesis

$$b_k \le \varepsilon_1^{-1/10} c_k \tag{4-6}$$

is satisfied. Then, for ε_1 sufficiently small,

$$b_k(\sigma) \lesssim c_k(\sigma) \tag{4-7}$$

holds for all $\sigma \in [0, \sigma_1 - 1]$ and $k \in \mathbb{Z}$.

Proof. We use a continuity argument to prove Theorem 4.1. For $T' \in (0, T]$, let

$$\Psi(T') = \sup_{k\in\mathbb{Z}} c_k^{-1} \|P_k\psi_m(s=0)\|_{G_k(T')}.$$

Then $\psi : (0, T] \to [0, \infty)$ is well-defined, increasing, continuous, and satisfies

$$\lim_{T' \to 0} \psi(T') \lesssim 1.$$

The critical implication to establish is

$$\Psi(T') \le \varepsilon_1^{-1/10} \implies \Psi(T') \lesssim 1,$$

which in particular follows from

$$b_k \lesssim c_k. \tag{4-8}$$

We also must similarly establish

$$b_k(\sigma) \lesssim c_k(\sigma) \tag{4-9}$$

for $\sigma \in (0, \sigma_1 - 1]$. The next several subsections describe the main steps of the proof of (4-8) and (4-9), to which the bulk of the remainder of this paper is dedicated. In Section 4E we complete the high level argument used to prove (4-8) and (4-9). □

Corollary 4.2. *Given the conditions of Theorem 4.1,*

$$\| P_k |\partial_x|^\sigma \partial_m \varphi \|_{L_t^\infty L_x^2((-T,T)\times\mathbb{R}^2)} \lesssim c_k(\sigma) \tag{4-10}$$

holds for all $\sigma \in [0, \sigma_1 - 1]$.

The proof we defer to Section 4F.

Together Theorem 1.1, Theorem 4.1, and Corollary 4.2 are almost enough to establish Theorem 1.3. The next lemma provides the final piece. We also defer its proof to Section 4F.

Lemma 4.3. *We have*

$$\sum_{k\in\mathbb{Z}} \| P_k \psi_x \|_{L^4_{t,x}}^2 \sim \sum_{k\in\mathbb{Z}} \| P_k \partial_x \varphi \|_{L^4_{t,x}}^2 .$$

Note that this lemma affords us a condition equivalent to (4-5) whose advantage lies in the fact that it is not expressed in terms of gauges.

Proof of Theorem 1.3. Fix $\sigma_1 \in \mathbb{Z}_+$ positive and let $\varepsilon_1 = \varepsilon_1(\sigma_1) \ge 0$. It suffices to prove (1-7) on the time interval $[-T, T]$ provided the estimate is uniform in T. In view of Theorem 1.1 and mass-conservation, proving

$$\| \partial_x \varphi \|_{L_t^\infty \dot{H}_Q^\sigma((-T,T)\times\mathbb{R}^2)} \lesssim_\sigma \| \partial_x \varphi \|_{\dot{H}_Q^\sigma(\mathbb{R}^2)} \tag{4-11}$$

for $\sigma \in [0, \sigma_1 - 1]$ with $\sigma_1 = 25$ is enough to establish (1-6).

By virtue of Lemma 4.3, the assumptions of Theorem 1.3 are equivalent to those of Theorem 4.1. Therefore we have access to Corollary 4.2, which states that (4-10) holds for $\sigma \in [0, \sigma_1 - 1]$. Using (2-33) and the Littlewood–Paley square function completes the proof of (4-11).

Global existence and (1-7) then follow via a standard bootstrap argument from Theorem 1.1 and from the fact that the constants in (4-11) are uniform in T. □

The remainder of this section is organized as follows. In Section 4A we state the key lemmas of parabolic type that are used to control the electric and magnetic nonlinearities. In Section 4B we state bounds that rely principally upon local smoothing, including a bilinear Strichartz estimate; they find application in controlling the worst magnetic nonlinearity terms.

In Section 4C we piece together the parabolic estimates to control the electric potential. In Section 4D we decompose the magnetic potential into two main pieces and demonstrate how to control one of these pieces.

In Section 4E we close the bootstrap argument proving Theorem 4.1. Here the remaining piece of the magnetic potential is addressed using a certain nonlinear version of a bilinear Strichartz estimate.

Finally, in Section 4F, we prove Corollary 4.2 and Lemma 4.3.

4A. ***Parabolic estimates.*** By "parabolic estimates" we mean those that principally rely upon the smoothing effect of the harmonic map heat flow. We include here only those that play a direct role in controlling the nonlinearity $\mathcal{N}$. These are proved in Section 7, where a host of auxiliary parabolic estimates are included as well. As the proofs rely upon a bootstrap argument that takes advantage of energy dispersion (4-4), these bounds rely upon this smallness constraint implicitly. On the other hand, L^4 smallness (4-5) is not used in the proofs of these bounds, but rather only in their application in this paper.

Lemma 4.4. *For $\sigma \in [0, \sigma_1 - 1]$, the derivative fields ψ_m satisfy*

$$\|P_k \psi_m(s)\|_{F_k(T)} \lesssim (1 + s2^{2k})^{-4} 2^{-\sigma k} b_k(\sigma) \tag{4-12}$$

for $s \geq 0$.

This estimate is used in Section 4D in controlling the magnetic nonlinearity, which schematically looks like $A\partial_x \psi$. To recover the loss of derivative, it is important to take advantage of parabolic smoothing by invoking representation (2-23) of A. Within the integral we schematically have $\psi(s) D_x \psi(s)$, and hence (4-12) allows us to take advantage of (3-3)–(3-7) in bounding this term. We prove (4-12) in Section 7A.

Lemma 4.5. *For $\sigma \in [0, \sigma_1 - 1]$, the derivative fields ψ_l and connection coefficients A_m satisfy*

$$\|P_k(A_m(s)\psi_l(s))\|_{F_k(T)} \lesssim (s2^{2k})^{-3/8}(1 + s2^{2k})^{-2} 2^{-(\sigma-1)k} b_k(\sigma). \tag{4-13}$$

Like the previous estimate, this estimate is also used in Section 4D in controlling the magnetic nonlinearity. Its proof is given in Section 7B. The need for this estimate arises from the need to control $D_x\psi$ appearing in representation (2-23) of A.

The next several estimates are used in Section 4C to control the electric potential. In particular, they provide a source of smallness crucial here for closing the bootstrap argument. They are proved in Section 7B.

Lemma 4.6. *For $\sigma \in [2\delta, \sigma_1 - 1]$, the connection coefficient A_x satisfies*

$$\|A_x^2\|_{L^2_{t,x}} \lesssim \sup_{j \in \mathbb{Z}} b_j^2 \cdot \sum_{k \in \mathbb{Z}} b_k^2, \tag{4-14}$$

$$\|P_k A_x^2(0)\|_{L^2_{t,x}} \lesssim 2^{-\sigma k} b_k(\sigma) \cdot \sup_j b_j \cdot \sum_{l \in \mathbb{Z}} b_l^2. \tag{4-15}$$

Lemma 4.7. *For $\sigma \in [2\delta, \sigma_1 - 1]$, the connection coefficient A_t satisfies*

$$\|A_t\|_{L^2_{t,x}} \lesssim \Big(1 + \sum_{j\in\mathbb{Z}} b_j^2\Big)^2 \sum_{k\in\mathbb{Z}} \|P_k \psi_x(0)\|^2_{L^4_{t,x}}, \tag{4-16}$$

$$\|P_k A_t\|_{L^2_{t,x}} \lesssim \Big(1 + \sum_{p} b_p^2\Big) \tilde{b}_k 2^{-\sigma k} b_k(\sigma). \tag{4-17}$$

In subsequent estimates the following shorthand will be useful:

$$\epsilon := \Big(1 + \sum_{j\in\mathbb{Z}} b_j^2\Big)^2 \sum_{l\in\mathbb{Z}} \|P_l \psi_x(0)\|^2_{L^4_{t,x}} + \Big(1 + \sum_{l} b_l^2\Big) \sup_{k\in\mathbb{Z}} b_k^2. \tag{4-18}$$

Under the assumptions of Theorem 4.1, ϵ is a very small quantity, being at least as good as $O(\varepsilon_1^{1/2})$.

4B. *Smoothing and Strichartz.* The key result of Section 5 is the following frequency-localized bilinear Strichartz estimate.

Theorem 4.8. *Suppose that ψ_m satisfies (2-7) on $[-T, T]$. Assume $\sigma \in [0, \sigma_1 - 1]$. Let the frequency envelopes b_j and c_j be defined as in (4-3) and (2-32). Let ϵ be given by (4-18). Suppose also that $2^{j-k} \ll 1$. Then*

$$2^{k-j}(1 + s2^{2j})^8 \|P_j \psi_l(s) \cdot P_k \psi_m(0)\|^2_{L^2_{t,x}} \lesssim 2^{-2\sigma k} c_j^2 c_k^2(\sigma) + \epsilon^2 b_j^2 b_k^2(\sigma). \tag{4-19}$$

In Section 5B we split the proof into two cases: $s = 0$ and $s > 0$, the more involved being the $s = 0$ case. In either case, if instead we only were to appeal to the local smoothing-based estimate (3-11) and the frequency envelope definition (4-3), then we would get the bound

$$2^{k-j}(1 + s2^{2j})^8 \|P_j \psi_l(s) \cdot P_k \psi_m(0)\|^2_{L^2_{t,x}} \lesssim b_j^2 b_k^2.$$

In practice this sort of bound must needs be summed over $j \ll k$. When initial energy is assumed to be small, as is done in [Bejenaru et al. 2011c], the sum $\sum_j b_j^2 \ll 1$ is small, and consequently the resulting term perturbative. In our subthreshold energy setting this is no longer the case, as in fact the sum may be large. What (4-19) reveals, though, is that any b_j contributions come with a power of ε. In view of additional work which we present in due course, this turns out to be sufficient for establishing that $b_k \lesssim c_k$.

An interesting related bound is the following local smoothing estimate, also proved in Section 5B. It arises as an easy corollary of our proof of Theorem 4.8.

Theorem 4.9. *Suppose that ψ_m satisfies (2-7) on $[-T, T]$. Assume $\sigma \in [0, \sigma_1 - 1]$. Let the frequency envelopes $b_j(\sigma)$ and $c_j(\sigma)$ be defined as in (4-3) and (2-32). Also, let ϵ be given by (4-18). Then*

$$2^k \sup_{|j-k|\le 20} \sup_{\theta\in\mathbb{S}^1} \|P_{j,\theta} P_k \psi_m\|^2_{L^{\infty,2}_\theta} \lesssim 2^{-2\sigma k} c_k^2(\sigma) + \epsilon 2^{-2\sigma k} b_k^2(\sigma) \tag{4-20}$$

holds for each $k \in \mathbb{Z}$.

We note that (4-20) likely extends to $L^{\infty,2}_{\theta,\lambda}$ for λ satisfying $|\lambda| < 2^{k-40}$, though we do not prove this. For comparison, note that from the definition of (4-3) we have

$$2^k \sup_{|j-k|\le 20} \sup_{\theta\in\mathbb{S}^1} \sup_{|\lambda|<2^{k-40}} \|P_{j,\theta} P_k \psi_m\|^2_{L^{\infty,2}_{\theta,\lambda}} \lesssim 2^{-2\sigma k} b_k^2(\sigma). \tag{4-21}$$

On the other hand, while the right-hand side of (4-20) may indeed be large, it so happens thanks to our hypotheses of energy dispersion and L^4 smallness that the $b_k(\sigma)$ term is perturbative. For our purposes, this is a substantial improvement over (4-21). However, it can be seen from the argument in Section 4E that even an extension of (4-20) to $L^{\infty,2}_{\theta,\lambda}$ spaces is not sufficient for proving $b_k(\sigma) \lesssim c_k(\sigma)$: it is important that we can replace two "b_j" terms with corresponding "c_j" terms as in (4-19).

4C. *Controlling the electric potential V.*

Lemma 4.10. *Suppose that $\sigma < \frac{1}{6} - 2\delta$. Then the electric potential term V_m satisfies the estimate*

$$\|P_k V_m\|_{N_k(T)} \lesssim \big(\|A_x^2\|_{L^2_{t,x}} + \|A_t\|_{L^2_{t,x}} + \|\psi_x^2\|_{L^2_{t,x}}\big) 2^{-\sigma k} b_k(\sigma). \tag{4-22}$$

Proof. Letting $f \in \{A_t, A_x^2, \psi_x^2\}$, we bound $P_k(f\psi_x)$ in $N_k(T)$. Begin with the following Littlewood–Paley decomposition of $P_k(f\psi_x)$:

$$P_k(f\psi_x) = P_k(P_{<k-80} f\, P_{k-5<\cdot<k+5}\psi_x) + \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-80}} P_k(P_{k_1} f\, P_{k_2}\psi_x) + \sum_{\substack{|k_1-k_2|\le 90\\ k_1,k_2>k-80}} P_k(P_{k_1} f\, P_{k_2}\psi_x).$$

The first term is controlled using Hölder's inequality:

$$\begin{aligned}\|P_k(P_{<k-80} f\, P_{k-5<\cdot<k+5}\psi_x)\|_{N_k(T)} &\le \|P_k(P_{<k-80} f\, P_{k-5<\cdot<k+5}\psi_x)\|_{L^{4/3}_{t,x}} \\ &\le \|P_{<k-80} f\|_{L^2_{t,x}} \|P_{k-5<\cdot<k+5}\psi_x\|_{L^4_{t,x}}.\end{aligned}$$

To control the second term we apply (3-4):

$$\|P_k(P_{k_1} f\, P_{k_2}\psi_x)\|_{N_k(T)} \lesssim 2^{(k_2-k)/6} \|P_{k_1} f\|_{L^2_{t,x}} \|P_{k_2}\psi_x\|_{G_{k_2}(T)}.$$

Using (4-3), (2-30), and $\sigma < 1/6 - 2\sigma$, we conclude that

$$\Bigg\| \sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} P_k(P_{k_1} f\, P_{k_2}\psi_x)\Bigg\|_{N_k(T)} \lesssim 2^{-\sigma k} b_k(\sigma) \sum_{|k_1-k|\le 4} \|P_{k_1} f\|_{L^2_{t,x}}.$$

To control the high-high interaction, apply (3-5):

$$\|P_k(P_{k_1} f\, P_{k_2}\psi_x)\|_{N_k(T)} \lesssim 2^{(k-k_2)/6} \|P_{k_1} f\|_{L^2_{t,x}} \|P_{k_2}\psi_x\|_{G_{k_2}(T)}.$$

Therefore, by (4-3),

$$\sum_{\substack{|k_1-k_2|\le 90\\ k_1,k_2>k-80}} \|P_k(P_{k_1} f\, P_{k_2}\psi_x)\|_{N_k(T)} \lesssim \sum_{\substack{|k_1-k_2|\le 90\\ k_1,k_2>k-80}} 2^{(k-k_2)/6} \|P_{k_1} f\|_{L^2_{t,x}} 2^{-\sigma k_2} b_{k_2}(\sigma).$$

Using Cauchy–Schwarz and (2-31) yields

$$\sum_{\substack{|k_1-k_2|\le 90\\ k_1,k_2>k-80}} \|P_k(P_{k_1} f P_{k_2}\psi_x)\|_{N_k(T)} \lesssim 2^{-\sigma k} b_k(\sigma)\Bigg(\sum_{k_1\ge k-80} \|P_{k_1} f\|^2_{L^2_{t,x}}\Bigg)^{1/2},$$

and so, by switching the $L^2_{t,x}$ and l^2 norms, we get from the standard square function estimate that

$$\sum_{\substack{|k_1-k_2|\le 90\\ k_1,k_2>k-80}} \|P_k(P_{k_1} f P_{k_2}\psi_x)\|_{N_k(T)} \lesssim \|f\|_{L^2_{t,x}} 2^{-\sigma k} b_k(\sigma).$$ □

Corollary 4.11. *For $\sigma \in [0, \sigma_1 - 1]$ we have*

$$\|P_k V_m\|_{N_k(T)} \lesssim \epsilon 2^{-\sigma k} b_k(\sigma).$$

Proof. Given (4-22), this is a direct consequence of (4-14), (4-16), and the fact that

$$\|f\|^2_{L^4_{t,x}} \lesssim \sum_{k\in\mathbb{Z}} \|P_k f\|^2_{L^4_{t,x}}.$$

Therefore the result holds for $\sigma < 1/6 - 2\delta$.

To extend the proof to larger σ, we may mimic the proof of Lemma 4.10 by performing the same Littlewood–Paley decomposition and then, with regard to the first and third terms of the decomposition, proceeding as before in the proof of that lemma. The argument, however, must be modified in handling the term

$$\sum_{\substack{|k_1-k|\le 4\\ k_2\le k-80}} P_k(P_{k_1} f P_{k_2}\psi_x), \tag{4-23}$$

where $f \in \{A_t, A_x^2, \psi_x^2\}$. We take different approaches according to the choice of f.

When $f = A_x^2$, we apply (3-4) and invoke (4-15) to obtain

$$\begin{aligned}\Bigg\|\sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} P_k(P_{k_1} A_x^2 P_{k_2}\psi_x)\Bigg\|_{N_k(T)} &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} 2^{(k_2-k)/6} \|P_{k_1} A_x^2\|_{L^2_{t,x}} \|P_{k_2}\psi_x\|_{G_{k_2}(T)}\\ &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} 2^{(k_2-k)/6} 2^{-\sigma k_1} b_{k_1}(\sigma) b_{k_2} \cdot \sup_j b_j \cdot \sum_l b_l^2\\ &\lesssim 2^{-\sigma k} b_k(\sigma)\cdot b_k \cdot \sup_j b_j \cdot \sum_j b_j^2,\end{aligned}$$

In the case where $f = A_t$, we apply (3-4) and use (4-17) to conclude that

$$\Bigg\|\sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} P_k(P_{k_1} A_t P_{k_2}\psi_x)\Bigg\|_{N_k(T)} \lesssim 2^{-\sigma k} b_k(\sigma)\tilde{b}_k b_k\Bigg(1+\sum_p b_p^2\Bigg),$$

which suffices by Cauchy–Schwarz.

Finally we turn to $f = \psi_x^2$, which we further decompose as

$$f = 2\sum_{\substack{|j_1-k|\le 4\\ j_2<k-80}} P_{j_1}\psi_x P_{j_2}\psi_x + \sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k-80}} P_{j_1}\psi_x P_{j_2}\psi_x.$$

To control the high-low term, we apply estimate (3-7) and get

$$\sum_{\substack{|j_1-k|\le 4\\ j_2<k-80}} \|P_{j_1}\psi_x P_{j_2}\psi_x\|_{L^2} \lesssim \sum_{\substack{|j_1-k|\le 4\\ j_2<k-80}} 2^{(j_2-j_1)/6} b_{j_2} 2^{-\sigma j_1} b_{j_1}(\sigma) \lesssim 2^{-\sigma k} b_k b_k(\sigma).$$

We turn to the high-high case. The full trilinear expression is given by

$$\sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} P_k\biggl(P_{k_1}\biggl(\sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k_1-80}} P_{j_1}\psi_x P_{j_2}\psi_x\biggr)\cdot P_{k_2}\psi_x\biggr).$$

We can drop the P_{k_1} factor because of the summation ranges, obtaining

$$\sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} \sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k_1-80}} P_k(P_{j_1}\psi_x P_{j_2}\psi_x \cdot P_{k_2}\psi_x).$$

We apply estimate (3-4) with $h = P_{j_2}\psi_x P_{k_2}\psi_x$ to get

$$\begin{aligned}\sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} \sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k_1-80}} &\|P_k(P_{j_1}\psi_x P_{j_2}\psi_x \cdot P_{k_2}\psi_x)\|_{N_k(T)}\\ &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} \sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k_1-80}} 2^{-|j_1-k|/6}\|P_{j_1}\psi_x\|_{G_{j_1}(T)}\|P_{j_2}\psi_x P_{k_2}\psi_x\|_{L^2}.\end{aligned}$$

Next we use (3-7) to control the L^2 norm:

$$\begin{aligned}\sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} \sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k_1-80}} &2^{-|j_1-k|/6}\|P_{j_1}\psi_x\|_{G_{j_1}(T)}\|P_{j_2}\psi_x P_{k_2}\psi_x\|_{L^2}\\ &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2<k-80}} \sum_{\substack{|j_1-j_2|\le 8\\ j_1,j_2\ge k_1-80}} 2^{-|j_1-k|/6} 2^{-|j_2-k_2|/6} 2^{-\sigma j_1} b_{j_1}(\sigma) b_{j_2} b_{k_2}.\end{aligned}$$

In this sum we can replace the factor $2^{-|j_2-k_2|/6}$ by the larger factor $2^{-|k-k_2|/6}$, from which it is seen that the whole sum is controlled by

$$2^{-\sigma k} b_k(\sigma) b_k \sum_{k_2<k-80} 2^{-|k-k_2|/6} b_{k_2} \lesssim 2^{-\sigma k} b_k^2 b_k(\sigma). \qquad \square$$

4D. ***Decomposing the magnetic potential.*** We begin by introducing a paradifferential decomposition of the magnetic nonlinearity, splitting it into two pieces. This decomposition depends upon a frequency parameter $k \in \mathbb{Z}$, which we suppress in the notation; this same k will also be the output frequency whose behavior we are interested in controlling. The decomposition also depends upon the frequency gap parameter $\varpi \in \mathbb{Z}_+$. How ϖ is chosen and the exact role it plays are discussed in Section 5B. There it is shown that ϖ may be set equal to a sufficiently large universal constant (independent of ε, ε_1, k, etc.).

Define $A_{\mathrm{lo}\wedge\mathrm{lo}}$ as

$$A_{m,\mathrm{lo}\wedge\mathrm{lo}}(s) := -\sum_{k_1,k_2\le k-\varpi} \int_s^\infty \mathrm{Im}(\overline{P_{k_1}\psi_m} P_{k_2}\psi_s)(s')\,ds'$$

and $A_{\mathrm{hi}\vee\mathrm{hi}}$ as

$$A_{m,\mathrm{hi}\vee\mathrm{hi}}(s) := -\sum_{\max\{k_1,k_2\}>k-\varpi} \int_s^\infty \mathrm{Im}(\overline{P_{k_1}\psi_m} P_{k_2}\psi_s)(s')\,ds',$$

so that $A_m = A_{m,\mathrm{lo}\wedge\mathrm{lo}} + A_{m,\mathrm{hi}\vee\mathrm{hi}}$. Similarly define $B_{\mathrm{lo}\wedge\mathrm{lo}}$ as

$$B_{m,\mathrm{lo}\wedge\mathrm{lo}} := -i\sum_{k_3}\left(\partial_l(A_{l,\mathrm{lo}\wedge\mathrm{lo}}P_{k_3}\psi_m) + A_{l,\mathrm{lo}\wedge\mathrm{lo}}\partial_l P_{k_3}\psi_m\right)$$

and $B_{\mathrm{hi}\vee\mathrm{hi}}$ as

$$B_{m,\mathrm{hi}\vee\mathrm{hi}} := -i\sum_{k_3}\left(\partial_l(A_{l,\mathrm{hi}\vee\mathrm{hi}}P_{k_3}\psi_m) + A_{l,\mathrm{hi}\vee\mathrm{hi}}\partial_l P_{k_3}\psi_m\right),$$

so that $B_m = B_{m,\mathrm{lo}\wedge\mathrm{lo}} + B_{m,\mathrm{hi}\vee\mathrm{hi}}$.

Our goal is to control $P_k B_m$ in $N_k(T)$. We consider first $P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}$, performing a trilinear Littlewood–Paley decomposition. In order for frequencies k_1, k_2, k_3 to have an output in this expression at a frequency k, we must have $(k_1, k_2, k_3) \in Z_2(k) \cup Z_3(k) \cup Z_0(k)$, where

$$Z_0(k) := Z_1(k) \cap \{(k_1, k_2, k_3) \in \mathbb{Z}^3 : k_1, k_2 > k - \varpi\} \tag{4-24}$$

and the other $Z_j(k)$ are defined in (3-12). We apply Lemma 3.10 to bound $P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}$ in $N_k(T)$ by

$$\sum_{\substack{(k_1,k_2,k_3)\in \\ Z_2(k)\cup Z_3(k)\cup Z_0(k)}} \int_0^\infty 2^{\max\{k,k_3\}} C_{k,k_1,k_2,k_3} \|P_{k_1}\psi_x(s)\|_{F_{k_1}} \|P_{k_2}(D_l\psi_l(s))\|_{F_{k_2}} \|P_{k_3}\psi_m(0)\|_{G_{k_3}}\,ds,$$

which, thanks to (4-12) and (4-13), is controlled by

$$\sum_{\substack{(k_1,k_2,k_3)\in \\ Z_2(k)\cup Z_3(k)\cup Z_0(k)}} 2^{\max\{k,k_3\}} C_{k,k_1,k_2,k_3} b_{k_1}b_{k_2}b_{k_3} \int_0^\infty (1+s2^{2k_1})^{-4} 2^{k_2}(s2^{2k_2})^{-3/8}(1+s2^{2k_2})^{-2}\,ds.$$

As

$$\int_0^\infty (1+s2^{2k_1})^{-4} 2^{k_2}(s2^{2k_2})^{-3/8}(1+s2^{2k_2})^{-2}\,ds \lesssim 2^{-\max\{k_1,k_2\}}, \tag{4-25}$$

we reduce to

$$\sum_{\substack{(k_1,k_2,k_3)\text{ in} \\ Z_2(k)\cup Z_3(k)\cup Z_0(k)}} 2^{\max\{k,k_3\}-\max\{k_1,k_2\}} C_{k,k_1,k_2,k_3} b_{k_1}b_{k_2}b_{k_3}. \tag{4-26}$$

To estimate $P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}$ on $Z_2 \cup Z_3$, we apply Corollary 3.12 and use the energy dispersion hypothesis. As for $Z_0(k)$, we note that its cardinality $|Z_0(k)|$ satisfies $|Z_0(k)| \lesssim \varpi$ independently of k. Hence for fixed ϖ summing over this set is harmless given sufficient energy dispersion. We obtain a bound of

$$\|P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}\|_{N_k(T)} \lesssim b_k^2 b_k \lesssim \epsilon b_k. \tag{4-27}$$

Consider now the leading term $P_k B_{m,\mathrm{lo}\wedge\mathrm{lo}}$. Bounding this in N_k with any hope of summing requires the full strength of the decay that comes from the local smoothing/maximal function estimates. However, such bounds as are immediately at our disposal — (3-10) and (3-11) — do not bring $B_{m,\mathrm{lo}\wedge\mathrm{lo}}$ within the perturbative framework, instead yielding a bound of the form

$$\sum_{\substack{k_1,k_2\le k-\varpi\\ |k_3-k|\le 4}} b_{k_1} b_{k_2} b_{k_3},$$

which is problematic since even $\sum_{j\ll k} c_j^2 \sim E_0^2 = O(1)$ for k large enough. This stands in sharp contrast with the small energy setting.

In the next section, however, we are able to capture enough improvement in such estimates so as to barely bring $B_{m,\mathrm{lo}\wedge\mathrm{lo}}$ back within reach of our bootstrap approach.

Finally, we need for $\sigma > 0$ an estimate analogous to (4-27). Returning to the proof of (4-26), we remark that any b_{k_j} may be replaced by $2^{-\sigma k_j} b_{k_j}$; in order to obtain an analogue of (4-27), we must make replacements judiciously so as to retain summability. In particular, for any (k_1,k_2,k_3) in $Z_2(k)\cup Z_3(k)\cup Z_0(k)$, we replace $b_{k_{\max}}$ with $2^{-\sigma k_{\max}} b_{k_{\max}}(\sigma)$ so that (4-26) becomes

$$\sum_{\substack{(k_1,k_2,k_3)\in\\ Z_2(k)\cup Z_3(k)\cup Z_0(k)}} 2^{\max\{k,k_3\}-\max\{k_1,k_2\}} C_{k,k_1,k_2,k_3} b_{k_{\min}} b_{k_{\mathrm{mid}}} 2^{-\sigma k_{\max}} b_{k_{\max}}(\sigma),$$

where $k_{\min}, k_{\mathrm{mid}}, k_{\max}$ denote, respectively, the min, mid, and max of $\{k_1,k_2,k_3\}$. We have $k_{\max}\gtrsim k$ over the set $Z_2(k)\cup Z_3(k)\cup Z_0(k)$ (see (3-12) and (4-24) for definitions), which guarantees summability due to straightforward modifications of Corollaries 3.11 and 3.12. Therefore

$$\|P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}\|_{N_k(T)} \lesssim b_k^2 2^{-\sigma k} b_k(\sigma),$$

which, combined with (4-27) and the definition (4-18) of ϵ, implies this:

Corollary 4.12. *Assume $\sigma\in[0,\sigma_1-1]$. The term $B_{m,\mathrm{hi}\vee\mathrm{hi}}$ satisfies the estimate*

$$\|P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}\|_{N_k(T)} \lesssim \epsilon 2^{-\sigma k} b_k(\sigma). \tag{4-28}$$

4E. *Closing the gauge field bootstrap.* We turn first to the completion of the proof of Theorem 4.1, as we now have in place all of the estimates that we need to prove (4-8).

Using the main linear estimate of Proposition 3.6 and the decomposition introduced in Section 4D, we obtain

$$\|P_k\psi_m\|_{G_k(T)} \lesssim \|P_k\psi_m(0)\|_{L_x^2} + \|P_k V_m\|_{N_k(T)} + \|P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}\|_{N_k(T)} + \|P_k B_{m,\mathrm{lo}\wedge\mathrm{lo}}\|_{N_k(T)}. \tag{4-29}$$

In Sections 4C and 4D it is shown that $P_k V_m$ and $P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}$ are perturbative in the sense that

$$\|P_k V_m\|_{N_k(T)} + \|P_k B_{m,\mathrm{hi}\vee\mathrm{hi}}\|_{N_k(T)} \lesssim \epsilon 2^{-\sigma k} b_k(\sigma),$$

To handle $P_k B_{m,\mathrm{lo}\wedge\mathrm{lo}}$, we first write

$$P_k B_{m,\mathrm{lo}\wedge\mathrm{lo}} = -i\partial_l(A_{l,\mathrm{lo}\wedge\mathrm{lo}} P_k\psi_m) + R,$$

where R is a perturbative remainder (thanks to a slight modification of technical Lemma 5.11). Therefore

$$\|P_k\psi_m\|_{G_k(T)} \lesssim 2^{-\sigma k}c_k(\sigma) + \epsilon 2^{-\sigma k}b_k(\sigma) + \|\partial_l(A_{l,\mathrm{lo}\wedge\mathrm{lo}}P_k\psi_m)\|_{N_k(T)}. \tag{4-30}$$

Thus it remains to control $-i\partial_l(A_{l,\mathrm{lo}\wedge\mathrm{lo}}P_k\psi_m)$, which we expand as

$$-iP_k\partial_l \sum_{\substack{k_1,k_2\le k-\varpi \\ |k_3-k|\le 4}} \int_0^\infty \mathrm{Im}(\overline{P_{k_1}\psi_l}P_{k_2}\psi_s)(s')P_{k_3}\psi_m(0)\,ds', \tag{4-31}$$

and whose $N_k(T)$ norm we denote by N_{lo}. In the $\sigma=0$ case the key is to apply Theorem 4.8 to $\overline{P_{k_1}\psi_l}(s')$ and $P_{k_3}\psi_m(0)$, after first placing all of (4-31) in $N_k(T)$ using (3-10). We obtain

$$\begin{aligned} N_{\mathrm{lo}} &\lesssim 2^k \sum_{\substack{k_1,k_2\le k-\varpi \\ |k_3-k|\le 4}} 2^{-|k-k_2|/2}2^{-|k_1-k_3|/2}2^{-\max\{k_1,k_2\}}b_{k_2}\left(c_{k_1}c_{k_3}+\epsilon^{1/2}b_{k_1}b_{k_3}\right) \\ &\lesssim 2^k \sum_{k_1,k_2\le k-\varpi} 2^{(k_1+k_2)/2-k}2^{-\max\{k_1,k_2\}}b_{k_2}(c_{k_1}c_k+\epsilon^{1/2}b_{k_1}b_k). \end{aligned}$$

Without loss of generality we restrict the sum to $k_1\le k_2$:

$$\sum_{k_1\le k_2\le k-\varpi} 2^{(k_1-k_2)/2}b_{k_2}(c_{k_1}c_k+\epsilon^{1/2}b_{k_1}b_k).$$

Using the frequency envelope property to sum off the diagonal, we reduce to

$$N_{\mathrm{lo}} \lesssim \sum_{j\le k-\varpi}(b_jc_jc_k+\epsilon^{1/2}b_j^2b_k).$$

Combining this with (4-30) and the fact that R is perturbative, we obtain

$$b_k \lesssim c_k + \epsilon b_k + \sum_{j\le k-\varpi}(b_jc_jc_k+\epsilon^{1/2}b_j^2b_k), \tag{4-32}$$

which, in view of our choice of ϵ, reduces to

$$b_k \lesssim c_k + c_k\sum_{j\le k-\varpi} b_jc_j.$$

Squaring and applying Cauchy–Schwarz yields

$$b_k^2 \lesssim \left(1+\sum_{j\le k-\varpi} b_j^2\right)c_k^2. \tag{4-33}$$

Setting

$$B_k := 1+\sum_{j<k} b_j^2$$

in (4-33) leads to

$$B_{k+1} \le B_k(1+Cc_k^2)$$

with $C > 0$ independent of k. Therefore

$$B_{k+m} \le B_k \prod_{l=1}^{m}(1 + Cc_{k+l}^2) \le B_k \exp\biggl(C \sum_{l=1}^{m} c_{k+l}^2\biggr) \lesssim_{E_0} B_k.$$

Since $B_k \to 1$ as $k \to -\infty$, we conclude that

$$B_k \lesssim_{E_0} 1$$

uniformly in k, so that, in particular,

$$\sum_{j\in\mathbb{Z}} b_j^2 \lesssim 1, \tag{4-34}$$

which, joined with (4-33), implies (4-8).

The proof of (4-9) is almost an immediate consequence. Instead of (4-32), we obtain

$$b_k(\sigma) \lesssim c_k(\sigma) + \epsilon b_k(\sigma) + \sum_{j\le k-\varpi} (b_j c_j c_k(\sigma) + \epsilon^{1/2} b_j^2 b_k(\sigma)),$$

which suffices to prove (4-9) in view of (4-34).

4F. *De-gauging.* The previous subsections overcome the most significant obstacles encountered in proving conditional global regularity. All of the key estimates therein apply to the Schrödinger map system placed in the caloric gauge, and a bootstrap argument is in fact run and closed at that level. This final subsection justifies the whole approach, showing how to transfer these results obtained at the gauge level back to the underlying Schrödinger map itself.

Proof of (4-10). To gain control over the derivatives $\partial_m \varphi$ in $L_t^\infty L_x^2$, we utilize representation (2-2) and perform a Littlewood–Paley decomposition. We only indicate how to handle the term $v \cdot \mathrm{Re}\, \psi_m$, as the term $w \cdot \mathrm{Im}\, \psi_m$ may be handled similarly. Starting with

$$\begin{aligned} P_k(v \,\mathrm{Re}\, \psi_m) = \sum_{|k_2-k|\le 4} & P_k(P_{\le k-5} v \cdot P_{k_2} \mathrm{Re}\, \psi_m) \\ & + \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} P_k(P_{k_1} v \cdot P_{k_2} \mathrm{Re}\, \psi_m) + \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} P_k(P_{k_1} v \cdot P_{k_2} \mathrm{Re}\, \psi_m), \end{aligned} \tag{4-35}$$

we proceed to bound each term in $L_t^\infty L_x^2$.

In view of the fact that $|v| \equiv 1$, the low-high frequency interaction is controlled by

$$\begin{aligned} \sum_{|k_2-k|\le 4} \|P_k(P_{\le k-5} v \cdot P_{k_2} \mathrm{Re}\, \psi_m)\|_{L_t^\infty L_x^2} &\lesssim \|P_{\le k-5} v\|_{L_{t,x}^\infty} \|P_k \psi_m\|_{L_t^\infty L_x^2} \\ &\lesssim \|P_k \psi_m\|_{L_t^\infty L_x^2} \lesssim c_k. \end{aligned} \tag{4-36}$$

To control the high-low frequency interaction, we use Hölder's inequality, Bernstein's inequality, (2-33) and Bernstein's inequality again, and finally the bound (2-15) along with the summation rule (2-30):

$$\begin{aligned}\sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_k(P_{k_1}v\cdot P_{k_2}\operatorname{Re}\psi_m)\|_{L^\infty_t L^2_x} &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_{k_1}v\|_{L^\infty_t L^2_x}\|P_{k_2}\psi_m\|_{L^\infty_{t,x}}\\ &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_{k_1}v\|_{L^\infty_t L^2_x}\cdot 2^{k_2}\|P_{k_2}\psi_m\|_{L^\infty_t L^2_x}\\ &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_{k_1}\partial_x v\|_{L^\infty_t L^2_x}\cdot 2^{k_2-k}c_{k_2}\lesssim c_k. \qquad (4\text{-}37)\end{aligned}$$

To control the high-high frequency interaction, we use Bernstein's inequality, Cauchy–Schwarz, Bernstein again, (2-15), and finally (2-31):

$$\begin{aligned}\sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} \|P_k(P_{k_1}v\cdot P_{k_2}\operatorname{Re}\psi_m)\|_{L^\infty_t L^2_x} &\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^k\|P_{k_1}v\cdot P_{k_2}\operatorname{Re}\psi_m\|_{L^\infty_t L^1_x}\\ &\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^k\|P_{k_1}v\|_{L^\infty_t L^2_x}\|P_{k_2}\psi_m\|_{L^\infty_t L^2_x}\\ &\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^{k-k_1}\|P_{k_1}\partial_x v\|_{L^\infty_t L^2_x}\|P_{k_2}\psi_m\|_{L^\infty_t L^2_x}\\ &\lesssim \sum_{k_2\ge k-4} 2^{k-k_2}c_{k_2}\lesssim c_k. \qquad (4\text{-}38)\end{aligned}$$

Combining (4-36), (4-37), and (4-38) and applying them in (4-35), we obtain

$$\|P_k(v\operatorname{Re}\psi_m)\|_{L^\infty_t L^2_x}\lesssim c_k.$$

As the above calculation holds with w in place of v, we conclude (recalling (2-2)) that

$$\|P_k\partial_x\varphi\|_{L^\infty_t L^2_x}\lesssim c_k.$$

Hence (4-10) holds for $\sigma=0$.

Now we turn to the case $\sigma\in[0,\sigma_1-1]$. Using Bernstein's inequality in (4-36) and (4-38), we obtain

$$\sum_{|k_2-k|\le 4}\|P_k(P_{\le k-5}v\cdot P_{k_2}\operatorname{Re}\psi_m)\|_{L^\infty_t L^2_x}\lesssim 2^{-\sigma k}c_k(\sigma), \qquad (4\text{-}39)$$

$$\sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}}\|P_k(P_{k_1}v\cdot P_{k_2}\operatorname{Re}\psi_m)\|_{L^\infty_t L^2_x}\lesssim 2^{-\sigma k}c_k(\sigma), \qquad (4\text{-}40)$$

as well as analogous estimates with w in place of v. Such a direct argument, however, does not yield the analogue of (4-37). We circumvent this obstruction as follows. Let $\mathscr{C}\in(0,\infty)$ be the best constant for which

$$\|P_k\partial_x\varphi\|_{L^\infty_t L^2_x}\le \mathscr{C}2^{-\sigma k}c_k(\sigma) \qquad (4\text{-}41)$$

holds for $\sigma \in [0, \sigma_1 - 1]$. Such a constant exists by smoothness and the fact that the $c_k(\sigma)$ are frequency envelopes. In view of definition (2-34) and estimate (2-35), we similarly have

$$\|P_k \partial_x v(0)\|_{L_t^\infty L_x^2} \lesssim \mathscr{C} 2^{-\sigma k} c_k(\sigma). \tag{4-42}$$

Using (4-42) in (4-37), we obtain

$$\sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_k(P_{k_1} v \cdot P_{k_2} \operatorname{Re} \psi_m)\|_{L_t^\infty L_x^2} \lesssim \mathscr{C} 2^{-\sigma k} c_k c_k(\sigma). \tag{4-43}$$

From the representations (2-2) and (4-35), and from the estimates (4-39), (4-40), and (4-43), along with the analogous estimates for w, it follows that

$$\|P_k \partial_x \varphi\|_{L_t^\infty L_x^2} \lesssim (1 + c_k \mathscr{C}) 2^{-\sigma k} c_k(\sigma).$$

In view of energy dispersion ($c_k \le \varepsilon$) and the optimality of $\mathscr{C}$ in (4-41), we conclude that $\mathscr{C} \lesssim 1 + \varepsilon \mathscr{C}$, so that $\mathscr{C} \lesssim 1$. Therefore

$$\|P_k \partial_x^\sigma \partial_m \varphi\|_{L_t^\infty L_x^2} \sim 2^{\sigma k} \|P_k \partial_m \varphi\|_{L_t^\infty L_x^2} \lesssim c_k(\sigma),$$

which completes the proof of (4-10). □

It will be convenient in certain arguments to use the weaker frequency envelope defined by

$$\tilde{b}_k = \sup_{k' \in \mathbb{Z}} 2^{-\delta |k-k'|} \|P_{k'} \psi_x\|_{L^4_{t,x}}. \tag{4-44}$$

Proof of Lemma 4.3. Let us first establish

$$\sum_{k\in\mathbb{Z}} \|P_k \psi_x\|^2_{L^4_{t,x}} \lesssim \sum_{k\in\mathbb{Z}} \|P_k \partial_x \varphi\|^2_{L^4_{t,x}}.$$

We use (2-1), i.e., $\psi_m = v \cdot \partial_m \varphi + i w \cdot \partial_m \varphi$, but for the sake of exposition only treat $v \cdot \partial_m \varphi$. We start with the Littlewood–Paley decomposition

$$P_k \psi_m(0) = \sum_{|k_2-k|\le 4} P_k(P_{\le k-5} v \cdot P_{k_2} \partial_m \varphi) + \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} P_k(P_{k_1} v \cdot P_{k_2} \partial_m \varphi) + \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} P_k(P_{k_1} v \cdot P_{k_2} \partial_m \varphi).$$

In view of $|v| \equiv 1$, the $L^4_{t,x}$ norm of the low-high interaction is controlled by $\tilde{b}_k$ (see (4-44)). To control the high-low interaction, we use Hölder's and Bernstein's inequalities along with (2-15):

$$\begin{aligned}
\sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_k(P_{k_1} v \cdot P_{k_2} \partial_m \varphi)\|_{L^4_{t,x}} &\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} \|P_{k_1} v\|_{L_t^\infty L_x^4} \cdot \|P_{k_2} \partial_m \varphi\|_{L_t^4 L_x^\infty} \\
&\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} 2^{k_1/2} \|P_{k_1} v\|_{L_t^\infty L_x^2} 2^{k_2/2} \|P_{k_2} \partial_m \varphi\|_{L^4_{t,x}} \\
&\lesssim \sum_{\substack{|k_1-k|\le 4\\ k_2\le k-4}} 2^{k_1} \|P_{k_1} v\|_{L_t^\infty L_x^2} \tilde{b}_k \lesssim \tilde{b}_k.
\end{aligned}$$

To control the high-high interaction, we use Bernstein, Hölder, Bernstein again, and (2-15):

$$
\begin{aligned}
\sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} \|P_k(P_{k_1}v\cdot P_{k_2}\partial_m\varphi)\|_{L^4_{t,x}} &\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^{k/2}\|P_{k_1}v\cdot P_{k_2}\partial_m\varphi\|_{L^4_tL^2_x}\\
&\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^{k/2}\|P_{k_1}v\|_{L^\infty_tL^4_x}\|P_{k_2}\partial_m\varphi\|_{L^4_{t,x}}\\
&\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^{(k+k_1)/2}\|P_{k_1}v\|_{L^\infty_tL^2_x}\|P_k\partial_m\varphi\|_{L^4_{t,x}}\\
&\lesssim \sum_{\substack{|k_1-k_2|\le 8\\ k_1,k_2\ge k-4}} 2^{(k-k_1)/2}\|P_{k_1}\partial_x v\|_{L^\infty_tL^2_x}\|P_{k_2}\partial_m\varphi\|_{L^4_{t,x}} \lesssim \sum_{k_2\ge k-4} 2^{(k-k_2)/4}\tilde{b}_{k_2} \lesssim \tilde{b}_k.
\end{aligned}
$$

Therefore

$$\|P_k\psi_m(0)\|_{L^4_{t,x}} \lesssim \tilde{b}_k$$

and

$$\sum_{k\in\mathbb{Z}}\|P_k\psi_m(0)\|^2_{L^4_{t,x}} \lesssim \sum_{k\in\mathbb{Z}}\tilde{b}_k^2 \sim \sum_{k\in\mathbb{Z}}\|P_k\partial_m\varphi(0)\|^2_{L^4_{t,x}}.$$

By using (2-2), creating an L^4 frequency envelope for $P_k\partial_m\varphi(0)$, and reversing the roles of ψ_α and $\partial_\alpha\varphi$ in the preceding argument, we conclude the reverse inequality

$$\sum_{k\in\mathbb{Z}}\|P_k\partial_m\varphi(0)\|^2_{L^4_{t,x}} \lesssim \sum_{k\in\mathbb{Z}}\|P_k\psi_m(0)\|^2_{L^4_{t,x}}. \qquad \square$$

5. Local smoothing and bilinear Strichartz

The main goal of this section is to establish the improved bilinear Strichartz estimate of Theorem 4.8. As a by-product we also obtain the frequency-localized local smoothing estimate of Theorem 4.9.

Our approach is to first establish abstract local smoothing and bilinear Strichartz estimates for solutions to certain magnetic nonlinear Schrödinger equations. These are in the spirit of [Planchon and Vega 2009; 2012; Tao 2010]. We shall then apply these to Schrödinger maps, in particular to the paralinearized derivative field equations written with respect to the caloric gauge.

We introduce some notation. Let

$$I_k(\mathbb{R}^d) = \{\xi\in\mathbb{R}^d : |\xi|\in[-2^{k-1},2^{k+1}]\} \quad\text{and}\quad I_{(-\infty,k]} := \bigcup_{j\le k} I_j.$$

For a d-vector-valued function $B=(B_l)$ on $\mathbb{R}^d$ with real entries, define the magnetic Laplacian Δ_B, acting on complex-valued functions f, via

$$\Delta_B f := (\partial_x + iB)((\partial_x+iB)f) = \Delta f + i(\partial_l B_l)f + 2iB_l\partial_l f - B_l^2 f. \tag{5-1}$$

For a unit vector $\boldsymbol{e}\in\mathbb{S}^{d-1}$, denote by $\{x\cdot\boldsymbol{e}=0\}$ the orthogonal complement in $\mathbb{R}^d$ of the span of $\boldsymbol{e}$, equipped with the induced measure. Given $\boldsymbol{e}$, we can construct a positively oriented orthonormal basis

$\boldsymbol{e}, \boldsymbol{e}_1, \ldots, \boldsymbol{e}_{d-1}$ of $\mathbb{R}^d$ so that $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{d-1}$ form an orthonormal basis for $\{x \cdot \boldsymbol{e} = 0\}$. For complex-valued functions f on $\mathbb{R}^d$, define $E_{\boldsymbol{e}}(f) : \mathbb{R} \to \mathbb{R}$ as

$$E_{\boldsymbol{e}}(f)(x_0) := \int_{x \cdot \boldsymbol{e}=0} |f|^2 \, dx' = \int_{\mathbb{R}^{d-1}} |f(x_0 \boldsymbol{e} + x_j \boldsymbol{e}_j)|^2 \, dx', \tag{5-2}$$

where the implicit sum runs over $1, 2, \ldots, d-1$, and dx' is the standard $(d-1)$-dimensional Lebesgue measure. We also adopt the following notation for this section: for z, ζ complex,

$$z \wedge \zeta := z\bar{\zeta} - \bar{z}\zeta = 2i \operatorname{Im}(z\bar{\zeta}).$$

5A. *Key lemmas.*

Lemma 5.1 (abstract almost-conservation of energy). *Let $d \geq 1$ and $\boldsymbol{e} \in \mathbb{S}^{d-1}$. Let v be a $C_t^\infty(H_x^\infty)$ function on $\mathbb{R}^d \times [0, T]$ solving*

$$(i\partial_t + \Delta_{\mathcal{A}})v = \Lambda_v \tag{5-3}$$

with initial data v_0. Take $\mathcal{A}_l$ to be real-valued, smooth, and bounded, with $\Delta_{\mathcal{A}}$ defined via (5-1). Then

$$\|v\|_{L_t^\infty L_x^2}^2 \leq \|v_0\|_{L_x^2}^2 + \left| \int_0^T \int_{\mathbb{R}^d} v \wedge \Lambda_v \, dx dt \right|. \tag{5-4}$$

Proof. We begin with

$$\frac{1}{2}\partial_t \int |v|^2 \, dx = \int \operatorname{Im}(\bar{v}\partial_t v) \, dx,$$

which may equivalently be written as

$$i\partial_t \int |v|^2 \, dx = -\int v \wedge i\partial_t v \, dx.$$

Substituting from (5-3) yields

$$i\partial_t \int |v|^2 \, dx = \int v \wedge (\Delta_{\mathcal{A}} v - \Lambda_v) \, dx.$$

Expanding $\Delta_{\mathcal{A}}$ using (5-1) and using the straightforward relations

$$\partial_l(v \wedge i\mathcal{A}_l v) = v \wedge i(\partial_l \mathcal{A}_l)v + v \wedge 2i\mathcal{A}_l \partial_l v \quad \text{and} \quad \partial_l(v \wedge \partial_l v) = v \wedge \Delta v,$$

we get

$$i\partial_t \int |v|^2 \, dx = \int \partial_l(v \wedge \partial_l v) \, dx + \int \partial_l(v \wedge i\mathcal{A}_l v) \, dx - \int v \wedge \mathcal{A}_l^2 v \, dx - \int v \wedge \Lambda_v \, dx.$$

The first two terms on the right-hand side vanish upon integration in x; the third is equal to zero because $\mathcal{A}_l^2$ is real. Integrating in time and taking absolute values therefore yields

$$\left| \int_{\mathbb{R}^d} |v(T')|^2 - |v_0|^2 \, dx \right| = \left| \int_0^{T'} \int_{\mathbb{R}^d} v \wedge \Lambda_v \, dx dt \right|$$

for any time $T' \in (0, T]$. $\square$

Lemma 5.2 (local smoothing preparation). *Let $d \geq 1$ and $\boldsymbol{e} \in \mathbb{S}^{d-1}$. Let $j, k \in \mathbb{Z}$ and $j = k + O(1)$. Let $\varepsilon_m > 0$ be a small positive number such that $\varepsilon_m 2^{O(1)} \ll 1$. Let v be a $C_t^\infty(H_x^\infty)$ function on $\mathbb{R}^d \times [0, T]$ solving*

$$(i\partial_t + \Delta_{\mathcal{A}})v = \Lambda_v, \tag{5-5}$$

where $\mathcal{A}_l$ is real-valued, smooth, and satisfies the estimate

$$\|\mathcal{A}\|_{L^\infty_{t,x}} \leq \varepsilon_m 2^k. \tag{5-6}$$

The solution v is assumed to have (spatial) frequency support in I_k, with the additional constraint that $\boldsymbol{e} \cdot \xi \in [2^{j-1}, 2^{j+1}]$ for all ξ in the support of $\hat{v}$. Then

$$2^j \int_0^T E_{\boldsymbol{e}}(v)\, dt \lesssim \|v\|^2_{L^\infty_t L^2_x} + \left| \int_0^T \int_{x\cdot \boldsymbol{e} \geq 0} v \wedge \Lambda_v \, dx dt \right| + 2^j \int_0^T E_{\boldsymbol{e}}(v + i2^{-j}\partial_{\boldsymbol{e}} v)\, dt. \tag{5-7}$$

Proof. We begin by introducing

$$M_{\boldsymbol{e}}(t) := \int_{x\cdot \boldsymbol{e} \geq 0} |v(x,t)|^2 \, dx.$$

Then

$$0 \leq M_{\boldsymbol{e}}(t) \leq \|v(t)\|^2_{L^2_x(\mathbb{R}^d)} \leq \|v\|^2_{L^\infty_t L^2_x([-T,T]\times \mathbb{R}^d)}. \tag{5-8}$$

Differentiating in time yields

$$i\dot{M}_{\boldsymbol{e}}(t) = \int_{x\cdot \boldsymbol{e} \geq 0} v \wedge (i\partial_t v)\, dx = \int_{x\cdot \boldsymbol{e} \geq 0} v \wedge (\Delta_{\mathcal{A}} v - \Lambda_v)\, dx,$$

which may be rewritten as

$$i\dot{M}_{\boldsymbol{e}}(t) = \int_{x\cdot \boldsymbol{e} \geq 0} \partial_l (v \wedge (\partial_l + i\mathcal{A}_l)v)\, dx - \int_{x\cdot \boldsymbol{e} \geq 0} v \wedge \Lambda_v \, dx. \tag{5-9}$$

By integrating by parts,

$$\int_{x\cdot \boldsymbol{e} \geq 0} \partial_l (v \wedge (\partial_l + i\mathcal{A}_l)v)\, dx = -\int_{x\cdot \boldsymbol{e} = 0} v \wedge (\partial_{\boldsymbol{e}} v + i\boldsymbol{e}\cdot \mathcal{A} v)\, dx',$$

and therefore (5-9) may be rewritten as

$$-\int_{x\cdot \boldsymbol{e} = 0} v \wedge (\partial_{\boldsymbol{e}} v + i\boldsymbol{e}\cdot \mathcal{A} v)\, dx' = i\dot{M}_{\boldsymbol{e}}(t) + \int_{x\cdot \boldsymbol{e} \geq 0} v \wedge \Lambda_v \, dx. \tag{5-10}$$

On the one hand, we have the heuristic that $\partial_{\boldsymbol{e}} v \approx i2^j v$ since v has localized frequency support. On the other hand, since $\mathcal{A}$ is real-valued, we have

$$\int_0^T \int_{x\cdot \boldsymbol{e} = 0} v \wedge i\boldsymbol{e}\cdot \mathcal{A} v \, dx' dt = 2\int_0^T \int_{x\cdot \boldsymbol{e} = 0} \boldsymbol{e}\cdot \mathcal{A} |v|^2 \, dx' dt \tag{5-11}$$

and hence by assumption (5-6) also

$$\int_0^T \int_{x\cdot \boldsymbol{e} = 0} |\mathcal{A}|\, |v|^2 \, dx' dt \leq \varepsilon_m 2^k \int_0^T \int_{x\cdot \boldsymbol{e} = 0} |v|^2 \, dx' dt. \tag{5-12}$$

Together these facts motivate rewriting $v \wedge \partial_{\boldsymbol{e}} v$ as

$$v \wedge \partial_{\boldsymbol{e}} v = 2 \cdot i 2^j |v|^2 + v \wedge (\partial_{\boldsymbol{e}} v - i 2^j v). \tag{5-13}$$

Using (5-11), (5-13), and the bounds (5-12) and (5-8) in (5-10), we obtain by time-integration that

$$(1-\varepsilon_m 2^{k-j}) 2^j \int_0^T E_{\boldsymbol{e}}(v)\,dt \leq \|v\|^2_{L_t^\infty L_x^2} + \left| \int_0^T \int_{x\cdot\boldsymbol{e}\geq 0} v \wedge \Lambda_v \,dx dt \right| + 2\cdot 2^j \int_0^T \int_{x\cdot\boldsymbol{e}=0} |v + i2^{-j}\partial_{\boldsymbol{e}} v||v|\,dx' dt.$$

Applying Cauchy–Schwarz to the last term yields

$$2^j \int_0^T \int_{x\cdot\boldsymbol{e}=0} |v + i2^{-j}\partial_{\boldsymbol{e}} v||v|\,dx'dt \leq 8\cdot 2^j \int_0^T E_{\boldsymbol{e}}(v + i2^{-j}\partial_{\boldsymbol{e}} v)\,dt + \frac{1}{8}\cdot 2^j \int_0^T E_{\boldsymbol{e}}(v)\,dt.$$

Therefore (5-7). □

We now describe the constraints on the nonlinearity that we shall require in the abstract setting

Definition 5.3. Let $\mathcal{P}$ be a fixed finite subset of $\{1 < p < \infty\}$. A bilinear form $B(\cdot\,,\cdot)$ is said to be *adapted* to $\mathcal{P}$ provided it measures its arguments in Strichartz-type spaces, the estimate

$$\left| \int_0^T \int_{\mathbb{R}^d} f \wedge g \,dx dt \right| \lesssim B(f, g)$$

holds for all complex-valued functions f, g on $\mathbb{R}^d \times [0, T]$, Bernstein's inequalities hold in both arguments of B, and these arguments are measured in L_x^p only for $p \in \mathcal{P}$. Given $B(\cdot\,,\cdot)$ and $\boldsymbol{e} \in \mathbb{S}^{d-1}$, we define $B_{\boldsymbol{e}}(\cdot\,,\cdot)$ via

$$B_{\boldsymbol{e}}(f, g) := B(f, \chi_{\{x\cdot\boldsymbol{e}\geq 0\}} g).$$

Definition 5.4. Let $\boldsymbol{e} \in \mathbb{S}^{d-1}$ and let $\mathcal{A}_l$ be real-valued and smooth. Let v be a $C_t^\infty(H_x^\infty)$ function on $\mathbb{R}^d \times [0, T]$ solving

$$(i\partial_t + \Delta_{\mathcal{A}})v = \Lambda_v.$$

Assume v is (spatially) frequency-localized to I_k with the additional constraint that $\boldsymbol{e}\cdot\xi \in [2^{j-1}, 2^{j+1}]$ for all ξ in the support of $\hat{v}$. Define a sequence of functions $\{v^{(m)}\}_{m=1}^\infty$ by setting $v^{(1)} = v$ and

$$v^{(m+1)} := v^{(m)} + i2^{-j}\partial_{\boldsymbol{e}} v^{(m)}.$$

By (5-1) and the Leibniz rule,

$$(i\partial_t + \Delta_{\mathcal{A}})v^{(m)} = \Lambda_{v^{(m)}},$$

where

$$\Lambda_{v^{(m)}} := (1 + i2^{-j}\partial_{\boldsymbol{e}})\Lambda_{v^{(m-1)}} + i2^{-j}(i\partial_{\boldsymbol{e}}\partial_l A_l - \partial_{\boldsymbol{e}} A_l^2)v^{(m-1)} - 2^{-j+1}(\partial_{\boldsymbol{e}} A_l)\partial_l v^{(m-1)}.$$

The sequence $\{v^{(m)}\}_{m=1}^\infty$ is called the *derived sequence* corresponding to v.

Suppose we are given a form B adapted to $\mathcal{P}$. The derived sequence is said to be *controlled* with respect to $B_{\boldsymbol{e}}$ provided that $B_{\boldsymbol{e}}(v^{(m)}, \Lambda_{v^{(m)}}) < \infty$ for each $m \geq 1$.

We remark that if the derived sequence $\{v^{(m)}\}_{m=1}^{\infty}$ of v is controlled, then for all $l \geq 1$, the derived sequences $\{v^{(m)}\}_{m=l}^{\infty}$ are also controlled.

Theorem 5.5 (abstract local smoothing). *Let $d \geq 1$ and $\boldsymbol{e} \in \mathbb{S}^{d-1}$. Let $j, k \in \mathbb{Z}$ and $j = k + O(1)$. Let $\varepsilon_m > 0$ be a small positive number such that $\varepsilon_m 2^{O(1)} \ll 1$. Let $\eta > 0$. Let $\mathcal{P}$ be a fixed finite subset of $(1, \infty)$ with $2 \in \mathcal{P}$, and let B be a form adapted to $\mathcal{P}$. Let v be a $C_t^{\infty}(H_x^{\infty})$ function on $\mathbb{R}^d \times [0, T]$ solving*

$$(i\partial_t + \Delta_{\mathcal{A}})v = \Lambda_v, \tag{5-14}$$

where $\mathcal{A}_l$ is real-valued, smooth, has spatial Fourier support in $I_{(-\infty,k]}$, and satisfies the estimate

$$\|\mathcal{A}\|_{L^{\infty}_{t,x}} \leq \varepsilon_m 2^k. \tag{5-15}$$

The solution v is assumed to have (spatial) frequency support in I_k. We take Λ_v to be frequency-localized to $I_{(-\infty,k]}$. Assume moreover that

$$\boldsymbol{e} \cdot \xi \in [(1-\eta)2^j, (1+\eta)2^j] \tag{5-16}$$

for all ξ in the support of $\hat{v}$.

If the derived sequence of v is controlled with respect to $B_{\boldsymbol{e}}$, then there exists $\eta^* > 0$ such that, for all $0 \leq \eta < \eta^*$, the local smoothing estimate

$$2^j \int_0^T E_{\boldsymbol{e}}(v)\, dt \lesssim \|v\|^2_{L_t^{\infty} L_x^2} + B_{\boldsymbol{e}}(v, \Lambda_v) \tag{5-17}$$

holds uniformly in T and $j = k + O(1)$.

Proof. The foundation for proving (5-17) is (5-7), which for an adapted form $B_{\boldsymbol{e}}$ implies

$$2^j \int_0^T E_{\boldsymbol{e}}(v)\, dt \lesssim \|v\|^2_{L_t^{\infty} L_x^2} + B_{\boldsymbol{e}}(v, \Lambda_v) + 2^j \int_0^T E_{\boldsymbol{e}}(v + i2^{-j}\partial_{\boldsymbol{e}} v)\, dt. \tag{5-18}$$

Therefore our goal is control the last term in (5-18). This we do using a bootstrap argument that hinges upon the fact that $\tilde{v} := v + i2^{-j}\partial_{\boldsymbol{e}} v$ is the second term in the derived sequence of v, and that being "controlled" is an inherited property (in the sense of the comments following Definition 5.4).

By Bernstein's and Hölder's inequalities, we have

$$2^j \int_0^T E_{\boldsymbol{e}}(v)\, dt \lesssim 2^{2j} T \|v\|^2_{L_t^{\infty} L_x^2}.$$

for any v. For fixed $T > 0$ and $k \in \mathbb{Z}$, let $K_{T,k} \geq 1$ be the best constant for which the inequality

$$2^j \int_0^T E_{\boldsymbol{e}}(v)\, dt \leq K_{T,k}\big(\|v\|^2_{L_x^2} + B_{\boldsymbol{e}}(v, \Lambda_v)\big) \tag{5-19}$$

holds for all controlled sequences. Applying (5-19) to $\tilde{v}$ results in

$$2^j \int_0^T E_{\boldsymbol{e}}(\tilde{v})\, dt \leq K_{T,k}\big(\|\tilde{v}\|^2_{L_x^2} + B_{\boldsymbol{e}}(\tilde{v}, \Lambda_{\tilde{v}})\big), \tag{5-20}$$

and thus we seek to control norms of $\tilde{v}$ in terms of those of v.

Let $\widetilde{P}_k$, $\widetilde{P}_{j,\boldsymbol{e}}$ denote slight fattenings of the Fourier multipliers P_k, $P_{j,\boldsymbol{e}}$. On the one hand, Plancherel implies

$$\|(1+i2^{-j}\partial_{\boldsymbol{e}})\widetilde{P}_{j,\boldsymbol{e}}\widetilde{P}_k\|_{L^2_x\to L^2_x}\lesssim\eta. \tag{5-21}$$

On the other hand, Bernstein's inequalities imply

$$\|(1+i2^{-j}\partial_{\boldsymbol{e}})\widetilde{P}_{j,\boldsymbol{e}}\widetilde{P}_k\|_{L^p_x\to L^p_x}\lesssim 1,\qquad 1\le p\le\infty.$$

Therefore it follows from Riesz–Thorin interpolation that

$$\|(1+i2^{-j}\partial_{\boldsymbol{e}})\widetilde{P}_{j,\boldsymbol{e}}\widetilde{P}_k\|_{L^p_x\to L^p_x}\lesssim\begin{cases}\eta^{2/p} & 2\le p<\infty,\\ \eta^{2-2/p} & 1<p\le 2.\end{cases}$$

Restricting to $p\in\mathcal{P}$, we conclude that there exists a $q>0$ such that

$$\|(1+i2^{-j}\partial_{\boldsymbol{e}})\widetilde{P}_{j,\boldsymbol{e}}\widetilde{P}_k\|_{L^p_x\to L^p_x}\lesssim\eta^q \tag{5-22}$$

for all $p\in\mathcal{P}$ and all η small enough.

Applying (5-22) and Bernstein to $\tilde{v}$ yields

$$\|\tilde{v}\|_{L^2_x}\lesssim\eta^q\|v\|_{L^2_x},\qquad B_{\boldsymbol{e}}(\tilde{v},\Lambda_{\tilde{v}})\lesssim\eta^q B_{\boldsymbol{e}}(v,\Lambda_v),$$

which, combined with (5-20) and (5-18), leads to

$$2^j\int_0^T E_{\boldsymbol{e}}(v)\,dt\lesssim(1+\eta^q K_{T,k})\big(\|v\|^2_{L^\infty_t L^2_x}+B_{\boldsymbol{e}}(v,\Lambda_v)\big).$$

As $K_{T,k}$ is the best constant for which (5-19) holds, it follows that

$$K_{T,k}\lesssim 1+\eta^q K_{T,k}$$

and hence that $K_{T,k}\lesssim 1$ for η small enough. □

Corollary 5.6. *Given the assumptions of Theorem 5.5, we have*

$$2^j\int_0^T E_{\boldsymbol{e}}(v)\,dt\lesssim\|v_0\|^2_{L^2_x}+B(v,\Lambda_v)+B_{\boldsymbol{e}}(v,\Lambda_v).$$

Proof. This is an immediate consequence of Theorem 5.5 and Lemma 5.1. □

Corollary 5.7 (abstract bilinear Strichartz). *Let $d\ge 1$ and $\boldsymbol{e}\in\mathbb{S}^{d-1}$. Set $\tilde{\boldsymbol{e}}=(-\boldsymbol{e},\boldsymbol{e})/\sqrt{2}$. Let $j,k\in\mathbb{Z}$ and $j=k+O(1)$. Let $\varepsilon_m>0$ be a small positive number such that $\varepsilon_m 2^{O(1)}\ll 1$. Let $\eta>0$. Let $\mathcal{P}$ be a fixed finite subset of $(1,\infty)$ with $2\in\mathcal{P}$, and let $B_{\tilde{\boldsymbol{e}}}$ be a form that is adapted to $\mathcal{P}$.*

Let $w(x,y)$ be a $C^\infty_t(H^\infty_{x,y})$ function on $\mathbb{R}^{2d}\times[0,T]$, equal to w_0 at $t=0$ and solving

$$(i\partial_t+\Delta_{\mathcal{A}})w=\Lambda_w,$$

where $\mathcal{A}_{k'}$ is real-valued, smooth, has spatial Fourier support in $I_{(-\infty,k]}$, and satisfies the estimate

$$\|\mathcal{A}\|_{L^\infty_{t,x,y}}\le\varepsilon_m 2^k.$$

Assume w has (spatial) frequency support in I_k and that

$$\tilde{\boldsymbol{e}}\cdot\xi\in[(1-\eta)2^j,(1+\eta)2^j]$$

for all ξ in the support of $\widehat{w}$. Take Λ_w to be frequency-localized to $I_{(-\infty,k]}$.

Suppose that $w(x,y)$ admits a decomposition $w(x,y)=u(x)v(y)$, where u has frequency support in I_l, $l\ll k$. Use u_0, v_0 to denote $u(t=0)$, $v(t=0)$. If the derived sequence of w is controlled with respect to $B_{\tilde{\boldsymbol{e}}}$, then

$$\|uv\|^2_{L^2_{t,x}}\lesssim 2^{l(d-1)}2^{-j}\big(\|u_0\|^2_{L^2_x}\|v_0\|^2_{L^2_x}+B(w,\Lambda_w)+B_{\tilde{\boldsymbol{e}}}(w,\Lambda_w)\big) \tag{5-23}$$

uniformly in T and $j=k+O(1)$ provided η is small enough.

Proof. Taking into account that

$$\|w_0\|_{L^2_{x,y}}=\|u_0\|_{L^2_x}\|v_0\|_{L^2_x},$$

we apply Corollary 5.6 to w at $(x,y)=0$ and get

$$2^j\int_0^T E_{\tilde{\boldsymbol{e}}}(w)\,dt\lesssim\|u_0\|^2_{L^2_x}\|v_0\|^2_{L^2_x}+B(w,\Lambda_w)+B_{\tilde{\boldsymbol{e}}}(w,\Lambda_w). \tag{5-24}$$

We complete $(-\boldsymbol{e},\boldsymbol{e})/\sqrt{2}$ to a basis as follows:

$$(-\boldsymbol{e},\boldsymbol{e})/\sqrt{2},\ (0,\boldsymbol{e}_1),\ \dots,\ (0,\boldsymbol{e}_{d-1}),\ (\boldsymbol{e},\boldsymbol{e})/\sqrt{2},\ (\boldsymbol{e}_1,0),\ \dots,\ (\boldsymbol{e}_{d-1},0).$$

On the one hand, $E_{\tilde{\boldsymbol{e}}}(w)(0)$ is by definition (see (5-2)) equal to

$$\int_{\mathbb{R}}\int_{\mathbb{R}^{2d-2}}|u(0\cdot\boldsymbol{e}+r\boldsymbol{e}+x_j\boldsymbol{e}_j,t)v(0\cdot\boldsymbol{e}+r\boldsymbol{e}+y_j\boldsymbol{e}_j,t)|^2\,dx'dy'dr.$$

We rewrite it as

$$\int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}|v(r\boldsymbol{e}+y_j\boldsymbol{e}_j,t)|^2dy'\int_{\mathbb{R}^{d-1}}|u(r\boldsymbol{e}+x_j\boldsymbol{e}_j,t)|^2\,dx'dr. \tag{5-25}$$

On the other hand,

$$\|uv\|^2_{L^2_y}=\int_{\mathbb{R}^d}|u(y,t)|^2|v(y,t)|^2dy=\int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}|u(r\boldsymbol{e}+y_j\boldsymbol{e}_j)|^2|v(r\boldsymbol{e}+y_j\boldsymbol{e}_j)|^2dy'dr,$$

and by applying Bernstein to u in the y' variables, we obtain

$$\|uv\|^2_{L^2_y}\lesssim 2^{l(d-1)}\int_{\mathbb{R}}\int_{\mathbb{R}^{d-1}}|v(r\boldsymbol{e}+y_j\boldsymbol{e}_j)|^2dy'\int_{\mathbb{R}^{d-1}}|u(r\boldsymbol{e}+x_j\boldsymbol{e}_j)|^2\,dx'dr. \tag{5-26}$$

Together (5-26), (5-25), and (5-24) imply (5-23). □

5B. *Applying the abstract lemmas.* We would like to apply the abstract estimates just developed to the evolution equation (2-7). We work in the caloric gauge and adopt the magnetic potential decomposition introduced in Section 4D. Throughout we take ϵ as defined in (4-18).

Our starting point is the equation

$$(i\partial_t+\Delta)\psi_m=B_{m,\mathrm{lo}\wedge\mathrm{lo}}+B_{m,\mathrm{hi}\vee\mathrm{hi}}+V_m. \tag{5-27}$$

Applying Fourier multipliers P_k, $P_{j,\theta}P_k$, or variants thereof, we easily obtain corresponding evolution equations for $P_k\psi_m$, $P_{j,\theta}P_k$, etc. In rewriting a projection P of (5-27) in the form (5-3), evidently $\Delta_{\mathcal{A}}\psi_m$ should somehow come from $\Delta P\psi_m-PB_{m,\mathrm{lo}\wedge\mathrm{lo}}$, whereas $PB_{m,\mathrm{hi}\vee\mathrm{hi}}+PV_m$ ought to constitute the leading part of the nonlinearity Λ. Fourier multipliers P, however, do not commute with the connection coefficients A, and therefore in order to use the abstract machinery we must first track and control certain commutators. Toward this end we adopt some notation from [Tao 2001].

Following [Tao 2001, §1], we use $L_O(f_1, \dots, f_m)(s, x, t)$ to denote any multilinear expression of the form

$$L_O(f_1, \dots, f_m)(s, x, t) := \int K(y_1, \dots, y_{M(c)}) f_1(s, x - y_1, t) \dots f_m(s, x - y_{M(c)}, t)\, dy_1 \dots dy_{M(c)},$$

where the kernel K is a measure with bounded mass (and K may change from line to line). Moreover, the kernel of L_O does not depend upon the index α. Also, we extend this notation to vector or matrices by making K into an appropriate tensor. The expression $L_O(f_1, \dots, f_m)$ may be thought of as a variant of $O(f_1, \dots, f_m)$. It obeys two key properties. The first is a simple consequence of Minkowski's inequality; see, for example, [Tao 2001, Lemma 1].

Lemma 5.8. *Let $X_1, \dots, X_m, X$ be spatially translation-invariant Banach spaces such that the product estimate*

$$\|f_1 \cdots f_m\|_X \le C_0 \|f_1\|_{X_1} \cdots \|f_m\|_{X_m}$$

holds for all scalar-valued $f_i \in X_i$ and for some constant $C_0 > 0$. Then

$$\|L_O(f_1, \dots, f_m)\|_X \lesssim (Cd)^{Cm} C_0 \|f_1\|_{X_1} \cdots \|f_m\|_{X_m}$$

holds for all $f_i \in X_i$ that are scalars, d-dimensional vectors, or $d \times d$ matrices.

The next lemma is an adaptation of Lemma 2 in [Tao 2001].

Lemma 5.9 (Leibniz rule). *Let P_k' be a C^∞ Fourier multiplier whose frequency support lies in some compact subset of $I_k(\mathbb{R}^d)$. The commutator identity*

$$P_k'(fg) = f P_k' g + L_O(\partial_x f, 2^{-k} g)$$

holds.

Proof. Rescale so that $k = 0$ and let $m(\xi)$ denote the symbol of P_0' so that

$$\widehat{P_0' h}(\xi) := m(\xi) \hat{h}(\xi).$$

By the fundamental theorem of calculus, we have

$$\begin{aligned}\left(P_0'(fg) - f P_0' g\right)(s, x, t) &= \int_{\mathbb{R}^d} \check{m}(y)(f(s, x - y, t) - f(s, x, t)) g(s, x - y, t)\, dy \\ &= -\int_0^1 \int_{\mathbb{R}^d} \check{m}(y) y \cdot \partial_x f(s, x - ry, t) g(s, x - y, t)\, dy dr.\end{aligned}$$

The conclusion follows from the rapid decay of $\hat{m}$. $\square$

We are interested in controlling $P_{\theta,j}P_k\psi_m$ in $L^{\infty,2}_\theta$ over all $\theta\in\mathbb{S}^1$ and $|j-k|\le 20$. In the abstract framework, however, we assumed a much tighter localization than $P_{\theta,j}$ provides. Therefore we decompose $P_{\theta,j}$ as a sum

$$P=\sum_{l=1,\dots,O((\eta^*)^{-1})} P_{\theta,j,l}, \tag{5-28}$$

and it suffices by the triangle inequality to bound $P_{\theta,j,l}P_k\psi_m$. We note that this does not affect perturbative estimates since η^* is universal and in particular does not depend upon ε_1, ε.

For notational convenience set $P:=P_{\theta,j,l}P_k$. Applying P to (5-27) yields

$$(i\partial_t+\Delta)P\psi_m=P\left(B_{m,\mathrm{lo}\wedge\mathrm{lo}}+B_{m,\mathrm{hi}\vee\mathrm{hi}}+V_m\right).$$

Now

$$PB_{m,\mathrm{lo}\wedge\mathrm{lo}}=-iP\sum_{|k_3-k|\le 4}(\partial_l(A_{l,\mathrm{lo}\wedge\mathrm{lo}}P_{k_3}\psi_m)+A_{l,\mathrm{lo}\wedge\mathrm{lo}}\partial_l P_{k_3}\psi_m),$$

as P localizes to a region of the annulus I_k. Applying Lemma 5.9, we obtain

$$PB_{m,\mathrm{lo}\wedge\mathrm{lo}}=-i(\partial_l(A_{l,\mathrm{lo}\wedge\mathrm{lo}}P\psi_m)-iA_{l,\mathrm{lo}\wedge\mathrm{lo}}\partial_l P\psi_m)+R,$$

where

$$R:=\sum_{|k_3-k|\le 4}\left(L_O(\partial_x\partial_l A_{l,\mathrm{lo}\wedge\mathrm{lo}},2^{-k}P_{k_3}\psi_m)+L_O(\partial_x A_{l,\mathrm{lo}\wedge\mathrm{lo}},2^{-k}P_{k_3}\partial_l\psi_m)\right). \tag{5-29}$$

Set

$$\mathscr{A}_m:=A_{m,\mathrm{lo}\wedge\mathrm{lo}}.$$

Then

$$(i\partial_t+\Delta_{\mathscr{A}})P\psi_m=P(B_{m,\mathrm{hi}\vee\mathrm{hi}}+V_m)+\mathscr{A}_x^2P\psi_m+R. \tag{5-30}$$

It is this equation that we shall show fits within the abstract local smoothing framework.

First we check that Lemmas 5.1 and 5.2 apply. The main condition to check is (5-6). Key are the bound (2-14) and Bernstein, which together with the fact that $\mathscr{A}$ is frequency-localized to $I_{(-\infty,k]}$ provide the estimate

$$\|\mathscr{A}\|_{L^\infty_{t,x}}\lesssim 2^k.$$

To achieve the ε_m gain, we adjust ϖ, which forces a gap between I_k and the frequency support of $\mathscr{A}$, i.e., we localize $\mathscr{A}$ to $I_{(-\infty,k-\varpi]}$ instead. Thus it suffices to set $\varpi\in\mathbb{Z}_+$ equal to a sufficiently large universal constant.

There is more to check in showing that (5-30) falls within the purview of Theorem 5.5. Already we have $d=2$, $\boldsymbol{e}=\theta$, $\varepsilon_m\sim 2^{-\varpi}$, $\mathscr{A}_m:=A_{m,\mathrm{lo}\wedge\mathrm{lo}}$, $v=P_{\theta,j,l}P_k\psi_m$, and $\Lambda_v=P(B_{m,\mathrm{hi}\vee\mathrm{hi}}+V_m)+\mathscr{A}_x^2P\psi_m+R$.

Next we choose $\mathcal{P}$ based upon the norms used in N_k, with the exception of the local smoothing/maximal function estimates. To be precise, define the new norms $\widetilde{N}_k$ via

$$\|f\|_{\widetilde{N}_k(T)} := \inf_{f=f_1+f_2+f_3+f_4+f_5} \|f_1\|_{L^{4/3}_{t,x}} + 2^{k/6}\|f_2\|_{L^{3/2,6/5}_{\hat{\theta}_1}} + 2^{k/6}\|f_3\|_{L^{3/2,6/5}_{\hat{\theta}_2}} + 2^{-k/6}\|f_4\|_{L^{6/5,3/2}_{\hat{\theta}_1}} + 2^{-k/6}\|f_5\|_{L^{6/5,3/2}_{\hat{\theta}_2}}$$

and similarly $\widetilde{G}_k$ via

$$\|f\|_{\widetilde{G}_k(T)} := \|f\|_{L^\infty_t L^2_x} + \|f\|_{L^4_{t,x}} + 2^{-k/2}\|f\|_{L^4_x L^\infty_t} + 2^{-k/6}\sup_{\theta\in\mathbb{S}^1}\|f\|_{L^{3,6}_\theta} + 2^{k/6}\sup_{|j-k|\le 20}\sup_{\theta\in\mathbb{S}^1}\|P_{j,\theta} f\|_{L^{6,3}_\theta}.$$

Set $\mathcal{P} = \{2, 3, 3/2, 4, 4/3, 6, 5/6\}$. We define the form $B(\cdot\,,\cdot)$ via

$$B(f, g) := \|f\|_{\widetilde{G}_k(T)}\|g\|_{\widetilde{N}_k(T)} \tag{5-31}$$

and B_θ by

$$B_\theta(f, g) := B(f, \chi_{\{x\cdot\theta\ge 0\}} g) \tag{5-32}$$

as in Definition 5.3. That B_θ is adapted to $\mathcal{P}$ is a direct consequence of the definition.

Proposition 5.10. *Let $\eta > 0$ be a parameter to be specified later. Let also $d = 2$, $\boldsymbol{e} = \theta$, $\varepsilon_m \sim 2^{-\varpi}$, $\mathcal{A}_m := A_{m,\mathrm{lo}\wedge\mathrm{lo}}$, $v = P^{(\eta)}_{\theta,j,l} P_k\psi_m$, $\Lambda_v = P(B_{m,\mathrm{hi}\vee\mathrm{hi}} + V_m) + \mathcal{A}_x^2 P\psi_m + R$, and $\mathcal{P} = \{2, 3, 3/2, 4, 4/3, 6, 5/6\}$. Let B, B_θ be given by (5-31) and (5-32) respectively. Then the conditions of Theorem 5.5 are satisfied and the derived sequence of v is controlled with respect to B_θ so that conclusion (5-17) holds for $v = P^{(\eta)}_{\theta,j,l} P_k\psi_m$ given η sufficiently small.*

Proof. The only claim of Proposition 5.10 that remains to be verified is that the derived sequence of $v = P_{\theta,j,l}P_k\psi_m$ is controlled with respect to B_θ. In particular, we need to show that for each $q \ge 1$ we have

$$B_\theta(v^{(q)}, \Lambda_{v^{(q)}}) < \infty,$$

where $v^{(1)} := P_{\theta,j,l}P_k\psi_m$,

$$v^{(q+1)} := v^{(q)} + i2^{-j}\partial_\theta v^{(q)},$$

and

$$\Lambda_{v^{(q+1)}} := (1 + i2^{-j}\partial_\theta)\Lambda_{v^{(q)}} + i2^{-j}(i\partial_\theta\partial_l\mathcal{A}_l - \partial_\theta\mathcal{A}_l^2)v^{(q)} - 2^{-j+1}(\partial_\theta\mathcal{A}_l)\partial_l v^{(q)}.$$

We first prove the following lemma.

Lemma 5.11. *Let $\sigma \in [0, \sigma_1 - 1]$. The right-hand side of (5-30) satisfies*

$$\|P(B_{m,\mathrm{hi}\vee\mathrm{hi}} + V_m) + \mathcal{A}_x^2 P\psi_m + R\|_{\widetilde{N}_k(T)} \lesssim \epsilon 2^{-\sigma k} b_k(\sigma).$$

Proof. We will repeatedly use implicitly the fact that the multiplier $P_{\theta,j,l}$ is bounded on L^p, $1 \le p \le \infty$, so that in particular P obeys estimates that are at least as good as those obeyed by P_k.

From Corollaries 4.11 and 4.12 of Sections 4C and 4D it follows that $P_k(B_{m,\mathrm{hi}\vee\mathrm{hi}} + V_m)$ is perturbative and bounded in $\widetilde{N}_k(T)$ by $\epsilon 2^{-\sigma k} b_k(\sigma)$. The $\widetilde{N}_k(T)$ estimates on PV_m immediately imply the boundedness of $\mathcal{A}_x^2 P\psi_m$.

To estimate R, we apply Lemma 3.10 to bound $PB_{m,\mathrm{lo}\wedge\mathrm{lo}}$ by

$$\sum_{(k_1,k_2,k_3)\in Z_1(k)} \int_0^\infty 2^{\max\{k_1,k_2\}} 2^{k_3-k} C_{k,k_1,k_2,k_3} \|P_{k_1}\psi_x(s)\|_{F_{k_1}} \|P_{k_2}(D_l\psi_l(s))\|_{F_{k_2}} \|P_{k_3}\psi_m(0)\|_{G_{k_3}}\, ds,$$

which, in view of (4-12), (4-13), and (4-25), is controlled by

$$\sum_{(k_1,k_2,k_3)\in Z_1(k)} C_{k,k_1,k_2,k_3} b_{k_1} b_{k_2} 2^{-\sigma k_3} b_{k_3}(\sigma).$$

Summation is achieved thanks to Corollary 3.11. □

We return to the proof of the proposition, and in particular to showing that $B_\theta(v,\Lambda_v)<\infty$. With the important observation that the spatial multiplier $\chi_{x\cdot\theta\geq 0}$ is bounded on the spaces $\widetilde{N}_k(T)$, we may apply Lemma 5.11 to control $\chi_{x\cdot\theta\geq 0}\Lambda_v$ in $\widetilde{N}_k$. Since by assumption $P\psi_m$ is bounded in $\widetilde{G}_k(T)$ (even in $G_k(T)$), we conclude that $B_\theta(v,\Lambda_v)<\infty$.

Next we need to show $B_\theta(v^q,\Lambda_{v^q})<\infty$ for $q>1$. By Bernstein,

$$\|v^{(q)}\|_{\widetilde{G}_k(T)} \lesssim \|v^{(q-1)}\|_{\widetilde{G}_k(T)}.$$

Similarly,

$$\|(1+i2^{-j})\partial_\theta\Lambda_{v^{(q)}}\|_{\widetilde{N}_k(T)} \lesssim \|\Lambda_{v^{(q-1)}}\|_{\widetilde{N}_k(T)}.$$

Thus it remains to control $i2^{-j}(i\partial_\theta\partial_l\mathcal{A}_l-\partial_\theta\mathcal{A}_l^2)v^{(q)}$ and $2^{-j+1}(\partial_\theta\mathcal{A}_l)\partial_l v^{(q)}$ in $\widetilde{N}_k$ for each $q>1$. Both are consequences of arguments in Lemma 5.11: Boundedness of $2^{-j}(\partial_\theta\partial_l\mathcal{A}_l)v^{(q)}$ and $2^{-j+1}(\partial_\theta\mathcal{A}_l)\partial_l v^{(q)}$ follows directly from the argument used to control R and from Bernstein's inequality, whereas boundedness of $2^{-j}(\partial_\theta\mathcal{A}_l^2)v^{(q)}$ is a consequence of Bernstein and the estimates on $\mathcal{A}_x^2P\psi_m$ from Section 4C. □

Combining Lemma 5.11 and Proposition 5.10, we conclude that Corollary 5.6 applies to $v=P\psi_m$, with right-hand side bounded by $2^{-2\sigma k}c_k(\sigma)^2+\epsilon 2^{-2\sigma k}b_k(\sigma)^2$. In view of the decomposition (5-28), we conclude this:

Corollary 5.12. *Assume $\sigma\in[0,\sigma_1-1]$. The function $P_k\psi_m$ satisfies*

$$\sup_{|j-k|\leq 20}\sup_{\theta\in\mathbb{S}^1}\|P_{j,\theta}P_k\psi_m\|_{L_\theta^{\infty,2}} \lesssim 2^{-k/2}(2^{-\sigma k}c_k(\sigma)+\epsilon^{1/2}2^{-\sigma k}b_k(\sigma)).$$

This proves Theorem 4.9.

Our next objective is to apply Corollary 5.7 to the case where w splits as a product $u(x)v(y)$ where u, v are appropriate frequency localizations of ψ_m or $\overline{\psi_m}$. First we must find function spaces suitable for defining an adapted form. We start with $(i\partial_t+\Delta_{\mathcal{A}})w=\Lambda_w$ and observe how it behaves with respect to separation of variables. If $w(x,y)=u(x)v(y)$, then the left-hand side may be rewritten as $u\cdot(i\partial_t+\Delta_{\mathcal{A}_y})v+v\cdot(i\partial_t+\Delta_{\mathcal{A}_x})u$. Let $\Lambda_u:=(i\partial_t+\Delta_{\mathcal{A}_x})u$ and $\Lambda_v:=(i\partial_t+\Delta_{\mathcal{A}_y})v$. Then

$$(i\partial_t+\Delta_{\mathcal{A}})(uv)=u\Lambda_v+v\Lambda_u.$$

We control

$$\int_0^T\int_{\mathbb{R}^2\times\mathbb{R}^2} u(x)v(y)(\Lambda_u(x)v(y)+u(x)\Lambda_v(y))\,dxdydt$$

as follows: in the case of the first term $u(x)v(y)\Lambda_u(x)v(y)$ we place each $v(y)$ in $L_t^\infty L_y^2$; we bound $u(x)\Lambda_u(x)$ by placing $u(x)$ in G_j and $\Lambda_u(x)$ in $\widetilde{N}_j$. To control $u(x)v(y)u(x)\Lambda_v(y)$, we simply reverse the roles of u and v (and of x and y). This leads us to the spaces $\overline{N}_{k,l}$ defined by

$$\|f\|_{\overline{N}_{k,l}(T)} := \inf\Big\{ \|g_{2j-1}\|_{\widetilde{N}_l(T)} \|h_{2j-1}\|_{L_t^\infty L_y^2} + \|g_{2j}\|_{L_t^\infty L_x^2} \|h_{2j}\|_{\widetilde{N}_k(T)} : \\ J \in \mathbb{Z}_+ \text{ and} f(x,y) = \sum_{j=1}^{2J} \big(g_{2j-1}(x) h_{2j-1}(y) + g_{2j}(x) h_{2j}(y) \big) \Big\}, \tag{5-33}$$

and the spaces $\overline{G}_{k,l}$ defined via

$$\|f\|_{\overline{G}_{k,l}(T)} := \big\| \|f(x,y)\|_{\widetilde{G}_k(T)(y)} \big\|_{\widetilde{G}_l(T)(x)}. \tag{5-34}$$

We use these spaces to define the form $\overline{B}(\cdot\,,\cdot)$ by

$$\overline{B}(f,g) := \|f\|_{\overline{G}_{k,l}(T)} \|g\|_{\overline{N}_{k,l}(T)}, \tag{5-35}$$

and the form $\overline{B}_\Theta$ by

$$\overline{B}_\Theta(f,g) := \overline{B}(f, \chi_{\{(x,y)\cdot\Theta\ge 0\}} g), \tag{5-36}$$

where $\Theta := (-\theta, \theta)$.

Proposition 5.13. *Let $\eta > 0$ be a small parameter and $\varpi \in \mathbb{Z}_+$ a large parameter, both to be specified later. Let $j,k,l \in \mathbb{Z}$, $j = k + O(1)$, $l \ll k$. Let $d = 2$, $\boldsymbol{e} = \theta$, $\varepsilon_m \sim 2^{-\varpi}$, $\mathcal{A}_x := A_{m,\mathrm{lo}\wedge\mathrm{lo}}$, $v = P^{(\eta)}_{\theta,j,l} P_k \psi_m$, $\Lambda_v = P(B_{m,\mathrm{hi}\vee\mathrm{hi}} + V_m) + \mathcal{A}_x^2 P \psi_m + R$, and $\mathcal{P} = \{2, 3, 3/2, 4, 4/3, 6, 5/6\}$. Here R is given by (5-29). Also, let $u = \overline{P_l \psi_p}$, $p \in \{1,2\}$ and $\Lambda_u = P_l(B_{p,\mathrm{hi}\vee\mathrm{hi}} + V_p) + \mathcal{A}_x^2 P_l \psi_p + R'$, where R' is given by (5-29), but defined in terms of derivative field ψ_l and frequency l rather than ψ_m and k.*

Let $w(x,y) := u(x)v(y)$, $\mathcal{A} := (\mathcal{A}_x, \mathcal{A}_y)$, $\Lambda_w := \Lambda_u v + u \Lambda_v$. Then, for ϖ sufficiently large and η sufficiently small, the conditions of Corollary 5.7 are satisfied and (5-23) applies to $u(x)v(x)$.

Proof. The frequency support conditions on $\mathcal{A}$ and Λ_w are easily verified. That the L^∞ bound on $\mathcal{A}$ holds follows from (2-14) and Bernstein provided ϖ is large enough (see the discussion preceding Proposition 5.10). In order to guarantee the frequency support conditions on w, it is necessary to make the gap $l \ll k$ sufficiently large with respect to η.

That $\overline{B}_\Theta$ is adapted to $\mathcal{P}$ is a straightforward consequence of its definition. To see that the derived sequence of w is controllable, we look to the proof of Proposition 5.10 and the definitions of the $\overline{N}_{k,l}$, $\overline{G}_{k,l}$ spaces. □

In a spirit similar to that of the proof of Corollary 5.12, we may combine Lemma 5.11 and the proof of Proposition 5.10 to control $B(w,\Lambda_w) + B_\Theta(w,\Lambda_w)$; in fact, in measuring Λ_w in the $\overline{N}_{k,l}$ spaces, it suffices to take $J = 1$ (see (5-33)). Then we obtain $B(w,\Lambda_w) + B_\Theta(w,\Lambda_w) \lesssim \epsilon b_j 2^{-\sigma k} b_k(\sigma)$. Using decomposition (5-28) and the triangle inequality to bound $P_k \psi_m$ in terms of the bounds on $P^{(\eta)}_{\theta,j,l} P_k \psi_m$, we obtain the bilinear Strichartz analogue of Corollary 5.12. In our application, however, the lower-frequency term will not simply be $\overline{P_j \psi_l}$, but rather its heat flow evolution $\overline{P_j \psi_l}(s)$.

Corollary 5.14 (improved bilinear Strichartz). *Let $j, k \in \mathbb{Z}$, $j \ll k$, and let*

$$u \in \{P_j \psi_l, \overline{P_j \psi_l} : j \leq k - \varpi, l \in \{1, 2\}\}.$$

Then for $s \geq 0$, $\sigma \in [0, \sigma_1 - 1]$,

$$\|u(s) P_k \psi_m(0)\|_{L^2_{t,x}} \lesssim 2^{(j-k)/2}(1 + s2^{2j})^{-4} 2^{-\sigma k} \left(c_j c_k(\sigma) + \epsilon b_j b_k(\sigma)\right). \tag{5-37}$$

Proof. It only remains to prove (5-37) when $s > 0$. Let $v := P_k \psi_m$. Using the Duhamel formula, we write

$$u(s)v = \left(e^{s\Delta} u(0)\right) v(0) + \int_0^s e^{(s-s')\Delta} U(s')\, ds' \cdot v(0), \tag{5-38}$$

where U is defined by (2-21) in terms of u.

To control the nonlinear term $\int_0^s e^{(s-s')\Delta} U(s')\, ds' \cdot v(0)$ in L^2, we apply local smoothing estimate (3-11), which places the nonlinear evolution in $F_j(T)$ and $v(0)$ in $G_k(T)$. Using Lemma 7.11 to bound the $F_j(T)$ norm, we conclude that

$$\left\| \int_0^s e^{(s-s')\Delta} \widetilde{U}(s')\, ds' \cdot v(0) \right\|_{L^2_{t,x}} \lesssim \epsilon 2^{(j-k)/2} (1 + s2^{2j})^{-4} 2^{-\sigma k} b_j b_k(\sigma). \tag{5-39}$$

It remains to show that

$$\|(e^{s\Delta} u) v\|_{L^2_{t,x}} \lesssim (1 + s2^{2j})^{-4} 2^{(j-k)/2} 2^{-\sigma k} (c_j c_k(\sigma) + \epsilon b_j b_k(\sigma)), \tag{5-40}$$

which is not a direct consequence of the time $s = 0$ bound. Let $\mathcal{T}_a$ denote the spatial translation operator that acts on functions $f(x, t)$ according to $\mathcal{T}_a f(x, t) := f(x - a, t)$. If

$$\|(\mathcal{T}_{x_1} u)(\mathcal{T}_{x_2} v)\|_{L^2_{t,x}} \lesssim 2^{(j-k)/2} 2^{-\sigma k} (c_j c_k(\sigma) + \epsilon b_j b_k(\sigma)) \tag{5-41}$$

can be shown to hold for all $x_1, x_2 \in \mathbb{R}^2$, then (5-40) follows from Minkowski's and Young's inequalities.

Consider, then, a solution w to

$$(i\partial_t + \Delta_{\mathcal{A}}(x, t)) w(x, t) = \Lambda_w(x, t)$$

satisfying the conditions of Theorem 5.5. The translate $\mathcal{T}_{x_0} w(x, t)$ then satisfies

$$(i\partial_t + \Delta_{\mathcal{T}_{x_0}(\mathcal{A})(x,t)})(\mathcal{T}_{x_0} w)(x, t) = (\mathcal{T}_{x_0} \Lambda_w)(x, t).$$

The operator $\mathcal{T}_{x_0}$ clearly does not affect $L^\infty_{t,x}$ bounds or frequency support conditions. The only possible obstruction to concluding (5-17) is this: whereas the derived sequence of w is controlled with respect to $\boldsymbol{B_e}$, in the abstract setting it may no longer be the case that the derived sequence of $\mathcal{T}_{x_0} w$ is controlled. This is due to the presence of the spatial multiplier in the definition of $\boldsymbol{B_e}$. Fortunately, as already alluded to in the proof of Proposition 5.10, in our applications we do enjoy uniform boundedness with respect to *any* spatial multipliers appearing in the second argument of an adapted form $\boldsymbol{B_e}$. Therefore Proposition 5.13 holds for spatial translates of frequency projections of ψ_m, from which we conclude (5-41). □

This establishes Theorem 4.8.

6. The caloric gauge

In Section 6A we briefly recall from [Smith 2012a] the construction of the caloric gauge and some useful quantitative estimates. In Section 6B we prove the frequency-localized estimates stated in Section 2C.

6A. *Construction and basic results.* In brief, the basic caloric gauge construction goes as follows. Starting with H_Q^∞-class data $\varphi_0 : \mathbb{R}^2 \to \mathbb{S}^2$ with energy $E(\varphi_0) < E_{\text{crit}}$, evolve φ_0 in s via the heat flow equation (2-11). At $s = \infty$ the map trivializes. Place an arbitrary orthonormal frame $e(\infty)$ on $T_{\varphi(s=\infty)}\mathbb{S}^2$. Evolving this frame backward in time via parallel transport in the s direction yields a caloric gauge on $\varphi^* T_{\varphi(s=\infty)}\mathbb{S}^2$.

For energies $E(\varphi_0)$ sufficiently small, global existence and decay bounds may be proven directly using Duhamel's formula. In order to extend these results to all energies less than E_{crit}, we employ in [Smith 2012a] a concentration compactness argument that exploits the symmetries of (2-11) via concentration compactness.

In [Smith 2012a] the following energy densities play an important role in the quantitative arguments.

Definition 6.1. For each positive integer k, define the energy densities $\boldsymbol{e}_k$ of a heat flow φ by

$$\begin{aligned} \boldsymbol{e}_k &:= |(\varphi^*\nabla)_x^{k-1} \partial_x \varphi|^2 \\ &:= \langle (\varphi^*\nabla)_{j_1} \dots (\varphi^*\nabla)_{j_{k-1}} \partial_{j_k} \varphi, (\varphi^*\nabla)_{j_1} \dots (\varphi^*\nabla)_{j_{k-1}} \partial_{j_k} \varphi \rangle, \end{aligned} \tag{6-1}$$

where $j_1, \dots, j_k$ are summed over $1, 2$ and ∇ denotes the Riemannian connection on the sphere, i.e., for vector fields X, Y on the sphere $\nabla_X Y$ denotes the orthogonal projection of $\partial_X Y$ onto the sphere.

Theorem 6.2 [Smith 2012a]. *For any initial data $\varphi_0 \in H_Q^\infty$ with $E(\varphi_0) < E_{\text{crit}}$ there exists a unique global smooth heat flow φ with initial data φ_0. Moreover, φ satisfies the estimates*

$$\int_0^\infty \int_{\mathbb{R}^2} s^{k-1} \boldsymbol{e}_{k+1}(s, x)\, dx\, ds \lesssim_{E_0,k} 1, \tag{6-2}$$

$$\sup_{0<s<\infty} s^{k-1} \int_{\mathbb{R}^2} \boldsymbol{e}_k(s, x)\, dx \lesssim_{E_0,k} 1,$$

$$\sup_{\substack{0<s<\infty \\ x\in\mathbb{R}^2}} s^k \boldsymbol{e}_k(s, x) \lesssim_{E_0,k} 1,$$

$$\int_0^\infty s^{k-1} \sup_{x\in\mathbb{R}^2} \boldsymbol{e}_k(s, x)\, ds \lesssim_{E_0,k} 1, \tag{6-3}$$

for each $k \geq 1$, as well as the estimate

$$\int_0^\infty \int_{\mathbb{R}^2} \boldsymbol{e}_1^2(s, x)\, dx\, ds \lesssim_{E_0} 1. \tag{6-4}$$

We employ (6-2), (6-3), and (6-4) below.

6B. *Frequency-localized caloric gauge estimates.* The key estimate to establish is (2-35) for φ; most of the remaining estimates will be derived as corollaries of it. Our strategy is to exploit energy dispersion so that we can apply the Duhamel formula to a frequency localization of the heat flow equation (2-11), which for convenience we rewrite as

$$\partial_s \varphi = \Delta\varphi + \varphi \boldsymbol{e}_1. \tag{6-5}$$

Proof of (2-35) for φ. Let $\sigma_1 \in \mathbb{Z}_+$ be positive and let $S' \geq S \gg 0$. Let $\mathcal{K} \in \mathbb{Z}_+$, $T \in (0, 2^{2\mathcal{K}}]$ be fixed. Define for each $t \in (-T, T)$ the quantity

$$\mathscr{C}(S,t) := \sup_{\sigma\in[2\delta,\sigma_1]} \sup_{s\in[0,S]} \sup_{k\in\mathbb{Z}} (1+s2^{2k})^{\sigma_1} 2^{\sigma k} \gamma_k(\sigma)^{-1} \| P_k\varphi(s,\cdot,t)\|_{L^2_x(\mathbb{R}^2)}. \tag{6-6}$$

For fixed t the function $\mathscr{C}(S,t) : [0, S'] \to (0,\infty)$ is well-defined, continuous, and nondecreasing. Moreover, in view of the definition (2-34) of $\gamma_k(\sigma)$, it follows that $\lim_{S\to 0} \mathscr{C}(S,t) \lesssim 1$. A simple consequence of (6-6) is

$$\| P_k\varphi(s,\cdot,t)\|_{L^2_x(\mathbb{R}^2)} \leq \mathscr{C}(S,t)(1+s2^{2k})^{-\sigma_1} 2^{-\sigma k}\gamma_k(\sigma) \tag{6-7}$$

for $0 \leq s \leq S \leq S'$.

Our goal is to show $\mathscr{C}(S,t) \lesssim 1$ uniformly in S and t and our strategy is to apply Duhamel's formula to (6-5) and run a bootstrap argument. Beginning with the decomposition

$$P_k(\varphi\boldsymbol{e}_1) = \sum_{|k_2-k|\leq 4} P_k(P_{\leq k-5}\varphi \cdot P_{k_2}\boldsymbol{e}_1) + \sum_{|k_1-k|\leq 4} P_k(P_{k_1}\varphi \cdot P_{\leq k-5}\boldsymbol{e}_1) + \sum_{\substack{k_1,k_2\geq k-4\\ |k_1-k_2|\leq 8}} P_k(P_{k_1}\varphi\cdot P_{k_2}\boldsymbol{e}_1),$$

we proceed to place in L^2_x each of the three terms on the right-hand side; we then integrate in s and consider separately the low-high, high-low, and high-high frequency interactions.

Low-high interaction. By Duhamel and the triangle inequality it suffices to bound

$$\mathrm{LH}(s,t) := \int_0^s e^{-(s-s')2^{2k-2}} \sum_{|k_2-k|\leq 4} \| P_k(P_{\leq k-5}\varphi(s',\cdot,t)\cdot P_{k_2}\boldsymbol{e}_1(s',\cdot,t))\|_{L^2_x}\, ds'. \tag{6-8}$$

By Hölder's inequality, $|\varphi| \equiv 1$, and L^p-boundedness of the Littlewood–Paley multipliers,

$$\begin{aligned}\mathrm{LH}(s,t) &\lesssim \int_0^s e^{-(s-s')2^{2k-2}} \sum_{|k_2-k|\leq 4} \|P_{\leq k-5}\varphi\|_{L^\infty_x} \|P_{k_2}\boldsymbol{e}_1\|_{L^2_x}\, ds' \\ &\lesssim \int_0^s e^{-(s-s')2^{2k-2}} \sum_{|k_2-k|\leq 4} \|P_{k_2}\boldsymbol{e}_1(s',\cdot,t)\|_{L^2_x}\, ds'.\end{aligned}$$

To control the sum we further decompose $P_l\boldsymbol{e}_1 = P_l(\partial_x\varphi\cdot\partial_x\varphi)$ into low-high and high-high frequency interactions:

$$P_l\boldsymbol{e}_1 = 2\sum_{|l_1-l|\leq 4} P_l(P_{\leq l-5}\partial_x\varphi\cdot P_{l_1}\partial_x\varphi) + \sum_{\substack{l_1,l_2\geq l-4\\ |l_1-l_2|\leq 8}} P_l(P_{l_1}\partial_x\varphi\cdot P_{l_2}\partial_x\varphi). \tag{6-9}$$

Low-high interaction (i). We first attend to the low-high subcase. For convenience set Ξ_{lh} equal to the first term of the right-hand side of (6-9), i.e.,

$$\Xi_{lh}(s,x,t) := \sum_{|l_1-l|\leq 4} P_l(P_{\leq l-5}\partial_x\varphi(s,x,t)\cdot P_{l_1}\partial_x\varphi(s,x,t)).$$

By the triangle inequality, Hölder's inequality, Bernstein's inequality, the definition (6-1) for $\boldsymbol{e}_1(s,\cdot,t)$, and (6-7), it follows that

$$\begin{aligned}\|\Xi_{lh}(s,\cdot,t)\|_{L^2_x} &\lesssim \sum_{|l_1-l|\le 4}\|P_l(P_{\le l-5}\partial_x\varphi\cdot P_{l_1}\partial_x\varphi)\|_{L^2_x} \lesssim \sum_{|l_1-l|\le 4}\|P_{\le l-5}\partial_x\varphi\|_{L^\infty_x}\|P_{l_1}\partial_x\varphi\|_{L^2_x}\\ &\lesssim \sum_{|l_1-l|\le 4}\|P_{\le l-5}\partial_x\varphi\|_{L^\infty_x}2^{l_1}\|P_{l_1}\varphi\|_{L^2_x} \lesssim \|\sqrt{\boldsymbol{e}_1}\|_{L^\infty_x}2^l\sum_{|l_1-l|\le 4}\|P_{l_1}\varphi\|_{L^2_x}\\ &\lesssim \|\sqrt{\boldsymbol{e}_1}(s,\cdot,t)\|_{L^\infty_x}2^l2^{-\sigma l}\gamma_l(\sigma)\mathscr{C}(S,t)(1+s2^{2l})^{-\sigma_1}.\end{aligned}$$

As we apply this inequality in the case where $l=k_2$, $|k_2-k|\le 4$, we have

$$\begin{aligned}&\int_0^s e^{-(s-s')2^{2k-2}}\|\Xi_{lh}(s',\cdot,t)\|_{L^2_x}ds'\\ &\qquad\lesssim 2^k2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)\int_0^s e^{-(s-s')2^{2k-2}}\|\sqrt{\boldsymbol{e}_1}(s',\cdot,t)\|_{L^\infty_x}(1+s'2^{2k})^{-\sigma_1}ds'.\end{aligned}\tag{6-10}$$

Apply Cauchy–Schwarz. Clearly

$$\left(\int_0^s\|\sqrt{\boldsymbol{e}_1}(s',\cdot,t)\|^2_{L^\infty_x}\,ds'\right)^{1/2}\le\|\boldsymbol{e}_1(\cdot,\cdot,t)\|^{1/2}_{L^1_sL^\infty_x}.\tag{6-11}$$

We postpone applying (6-3) with $k=1$ to (6-11). As for the other factor, we have

$$\left(\int_0^s e^{-(s-s')2^{2k-1}}(1+s'2^{2k})^{-2\sigma_1}ds'\right)^{1/2}\lesssim\bigl(s(1+s2^{2k-1})^{-2\sigma_1}(1+s2^{2k})^{-1}\bigr)^{1/2}\tag{6-12}$$

since

$$\int_0^s e^{-(s-s')\lambda}(1+s'\lambda')^{-\alpha}ds'\lesssim s(1+\lambda s)^{-\alpha}(1+\lambda's)^{-1}$$

for $s\ge 0$, $0\le\lambda\le\lambda'$, and $\alpha>1$. Hence, applying Cauchy–Schwarz to (6-10) and using (6-11) and (6-12), we get

$$\begin{aligned}&\int_0^s e^{-(s-s')2^{2k-2}}\|\Xi_{lh}(s',\cdot,t)\|_{L^2_x}\,ds'\\ &\qquad\lesssim 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)2^ks^{1/2}(1+s2^{2k-1})^{-\sigma_1}(1+s2^{2k})^{-1/2}\|\boldsymbol{e}_1(t)\|^{1/2}_{L^1_sL^\infty_x([0,s]\times\mathbb{R}^2)}.\end{aligned}$$

Discarding $s^{1/2}2^k(1+s2^{2k})^{1/2}\le 1$, we conclude that

$$\int_0^s e^{-(s-s')2^{2k-2}}\|\Xi_{lh}(s',\cdot,t)\|_{L^2_x}\,ds'\lesssim 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s2^{2k-1})^{-\sigma_1}\|\boldsymbol{e}_1(t)\|^{1/2}_{L^1_sL^\infty_x([0,s]\times\mathbb{R}^2)}.\tag{6-13}$$

Low-high interaction (ii). We now move on to the high-high interaction subcase, setting Ξ_{hh} equal to the second term of the right-hand side of (6-9):

$$\Xi_{hh}(s,x,t):=\sum_{\substack{l_1,l_2\ge l-4\\|l_1-l_2|\le 8}}P_l(P_{l_1}\partial_x\varphi(s,x,t)\cdot P_{l_2}\partial_x\varphi(s,x,t)).$$

By the triangle inequality, Bernstein, and Cauchy–Schwarz,

$$\|\Xi_{hh}\|_{L^2_x} \lesssim \sum_{\substack{l_1,l_2\geq l-4\\ |l_1-l_2|\leq 8}} \|P_l(P_{l_1}\partial_x\varphi\cdot P_{l_2}\partial_x\varphi)\|_{L^2_x} \lesssim \sum_{\substack{l_1,l_2\geq l-4\\ |l_1-l_2|\leq 8}} 2^l \|P_{l_1}\partial_x\varphi\cdot P_{l_2}\partial_x\varphi\|_{L^1_x}$$
$$\lesssim \sum_{\substack{l_1,l_2\geq l-4\\ |l_1-l_2|\leq 8}} 2^l \|P_{l_1}\partial_x\varphi\|_{L^2_x}\|P_{l_2}\partial_x\varphi\|_{L^2_x}.$$

At this stage we apply Bernstein twice, exploiting $|l_1-l_2|\leq 8$, and get

$$\|P_{l_1}\partial_x\varphi\|_{L^2_x}\|P_{l_2}\partial_x\varphi\|_{L^2_x} \lesssim 2^{l_2}\|P_{l_1}\partial_x\varphi\|_{L^2_x}\|P_{l_2}\varphi\|_{L^2_x} \lesssim \|P_{l_1}|\partial_x|^2\varphi\|_{L^2_x}\|P_{l_2}\varphi\|_{L^2_x}.$$

So

$$\|\Xi_{hh}\|_{L^2_x} \lesssim 2^l \sum_{\substack{l_1,l_2\geq l-4\\ |l_1-l_2|\leq 8}} \|P_{l_1}|\partial_x|^2\varphi\|_{L^2_x}\|P_{l_2}\varphi\|_{L^2_x}.$$

Applying Cauchy–Schwarz yields

$$\|\Xi_{hh}\|_{L^2_x} \lesssim 2^l\Big(\sum_{l_1\geq l-4}\|P_{l_1}|\partial_x|^2\varphi\|^2_{L^2_x}\Big)^{1/2}\Big(\sum_{l_2\geq l-4}\|P_{l_2}\varphi\|^2_{L^2_x}\Big)^{1/2} \lesssim \||\partial_x|^2\varphi\|_{L^2_x}2^l\Big(\sum_{l_2\geq l-4}\|P_{l_2}\varphi\|^2_{L^2_x}\Big)^{1/2}. \tag{6-14}$$

As φ takes values in $\mathbb{S}^2$, which has constant curvature, we readily estimate ordinary derivatives by covariant ones:

$$|\partial_x^2\varphi| \lesssim \sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1. \tag{6-15}$$

Applying (6-15) in (6-14) and using (6-7), we arrive at

$$\|\Xi_{hh}(s,\cdot,t)\|_{L^2_x} \lesssim \|\sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1\|_{L^2_x}2^l\Big(\sum_{l_2\geq l-4}\|P_{l_2}\varphi\|^2_{L^2_x}\Big)^{1/2}$$
$$\lesssim \|(\sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1)(s,\cdot,t)\|_{L^2_x}2^l\mathscr{C}(S,t)\Big(\sum_{l_2\geq l-4}(1+s2^{2l_2})^{-2\sigma_1}2^{-2\sigma l_2}\gamma^2_{l_2}(\sigma)\Big)^{1/2}$$
$$\lesssim \|(\sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1)(s,\cdot,t)\|_{L^2_x}2^l\mathscr{C}(S,t)(1+s2^{2l})^{-\sigma_1}\Big(\sum_{l_2\geq l-4}2^{-2\sigma l_2}\gamma^2_{l_2}(\sigma)\Big)^{1/2}. \tag{6-16}$$

As $\sigma>\delta$ is bounded away from δ uniformly, we may apply summation rule (2-31) in (6-16). Recalling $l=k_2$ where $|k_2-k|\leq 4$, we conclude that

$$\|\Xi_{hh}(s,\cdot,t)\|_{L^2_x} \lesssim \|(\sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1)(s,\cdot,t)\|_{L^2_x}2^k2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s2^{2k})^{-\sigma_1}.$$

Integrating in s yields

$$\int_0^s e^{-(s-s')2^{2k-2}}\|\Xi_{hh}(s',\cdot,t)\|_{L^2_x}\,ds'$$
$$\lesssim 2^k2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)\int_0^s e^{-(s-s')2^{2k-2}}\big\|(\sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1)(s',\cdot,t)\big\|_{L^2_x}(1+s'2^{2k})^{-\sigma_1}\,ds'. \tag{6-17}$$

We use the triangle inequality to write $\|\sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1\|_{L^2_x} \le \|\sqrt{\boldsymbol{e}_2}\|_{L^2_x} + \|\boldsymbol{e}_1\|_{L^2_x}$ and split the integral in (6-17) into two pieces. By Cauchy–Schwarz and (6-12),

$$\begin{aligned}\int_0^s e^{-(s-s')2^{2k-2}} \|\boldsymbol{e}_1(s',\cdot,t)\|_{L^2_x}(1+s'2^{2k})^{-\sigma_1}\,ds' &\\ &\le \biggl(\int_0^s \|\boldsymbol{e}_1(s',\cdot,t)\|^2_{L^2_x}\,ds'\biggr)^{1/2}\biggl(\int_0^s e^{-(s-s')2^{2k-1}}(1+s'2^{2k})^{-2\sigma_1}\,ds'\biggr)^{1/2}\\ &\lesssim \|\boldsymbol{e}_1(t)\|_{L^2_{s,x}}\bigl(s(1+s2^{2k-1})^{-2\sigma_1}(1+s2^{2k})^{-1}\bigr)^{1/2}. \end{aligned}\tag{6-18}$$

To the remaining integral we also apply Cauchy–Schwarz and (6-12):

$$\begin{aligned}\int_0^s e^{-(s-s')2^{2k-2}} \|\sqrt{\boldsymbol{e}_2}(s',\cdot,t)\|_{L^2_x}(1+s'2^{2k})^{-\sigma_1}\,ds' &\\ &\le \biggl(\int_0^s \|\boldsymbol{e}_2(s',\cdot,t)\|_{L^1_x}\,ds'\biggr)^{1/2}\biggl(\int_0^s e^{-(s-s')2^{2k-1}}(1+s'2^{2k})^{-2\sigma_1}\,ds'\biggr)^{1/2}\\ &\lesssim \|\boldsymbol{e}_2(t)\|^{1/2}_{L^1_{s,x}}\bigl(s(1+s2^{2k-1})^{-2\sigma_1}(1+s2^{2k})^{-1}\bigr)^{1/2}. \end{aligned}\tag{6-19}$$

Hence, using Cauchy–Schwarz, (6-18), and (6-19) in (6-17), we conclude that

$$\int_0^s e^{-(s-s')2^{2k-2}}\|\Xi_{hh}(s',\cdot,t)\|_{L^2_x}\,ds' \lesssim 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s2^{2k-1})^{-\sigma_1}\bigl(\|\boldsymbol{e}_1(t)\|_{L^2_{s,x}}+\|\boldsymbol{e}_2(t)\|^{1/2}_{L^1_{s,x}}\bigr). \tag{6-20}$$

Low-high interaction: conclusion. Combining (6-13) and (6-20), we conclude in view of (6-8) and the decomposition (6-9) that

$$\mathrm{LH}(s,t) \lesssim 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s2^{2k-1})^{-\sigma_1}\bigl(\|\boldsymbol{e}_1(t)\|^{1/2}_{L^1_sL^\infty_x}+\|\boldsymbol{e}_1(t)\|_{L^2_{s,x}}+\|\boldsymbol{e}_2(t)\|^{1/2}_{L^1_{s,x}}\bigr). \tag{6-21}$$

High-low interaction. We now go on to bound the high-low interaction. By Duhamel and the triangle inequality it suffices to bound

$$\mathrm{HL}(s,t) := \int_0^s e^{-(s-s')2^{2k-2}}\sum_{|k_1-k|\le 4}\|P_k(P_{k_1}\varphi(s',\cdot,t)\cdot P_{\le k-5}\boldsymbol{e}_1(s',\cdot,t))\|_{L^2_x}\,ds'.$$

By Hölder's inequality, (6-7), and Bernstein's inequality, we have

$$\begin{aligned}\sum_{|k_1-k|\le 4}\|P_k(P_{k_1}\varphi(s,\cdot,t)\cdot P_{\le k-5}\boldsymbol{e}_1(s,\cdot,t))\|_{L^2_x} &\\ &\lesssim \sum_{|k_1-k|\le 4}\|P_{k_1}\varphi\|_{L^2_x}\|P_{\le k-5}\boldsymbol{e}_1\|_{L^\infty_x}\\ &\lesssim \|P_{\le k-5}\boldsymbol{e}_1(s,\cdot,t)\|_{L^\infty_x}\sum_{|k_1-k|\le 4}(1+s'2^{2k_1})^{-\sigma_1}2^{-\sigma k_1}\gamma_{k_1}(\sigma)\mathscr{C}(S,t)\\ &\lesssim 2^k\|P_{\le k-5}\boldsymbol{e}_1(s,\cdot,t)\|_{L^2_x}2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s'2^{2k})^{-\sigma_1}.\end{aligned}$$

Hence

$$\mathrm{HL}(s,t) \lesssim 2^k 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)\int_0^s e^{-(s-s')2^{2k-2}}(1+s'2^{2k})^{-\sigma_1}\|\boldsymbol{e}_1(s',\cdot,t)\|_{L^2_x}\,ds'.$$

Bounding the integral as in (6-18), we obtain

$$\mathrm{HL}(s,t) \lesssim 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s2^{2k-1})^{-\sigma_1}\|\boldsymbol{e}_1(t)\|_{L^2_{s,x}}. \tag{6-22}$$

High-high interaction. We conclude with the high-high interaction. Set

$$\mathrm{HH}(s,x,t) := \int_0^s e^{-(s-s')2^{2k-2}} \sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}} \|P_k(P_{k_1}\varphi(s,x,t)\cdot P_{k_2}\boldsymbol{e}_1(s,x,t))\|_{L^2_x}\,ds'.$$

By Bernstein, Cauchy–Schwarz, and (6-7),

$$\begin{aligned}
\sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}} \|P_k(P_{k_1}\varphi\cdot P_{k_2}\boldsymbol{e}_1)\|_{L^2_x} &\lesssim \sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}} 2^k\|P_{k_1}\varphi\|_{L^2_x}\|P_{k_2}\boldsymbol{e}_1\|_{L^2_x}\\
&\lesssim 2^k\Big(\sum_{k_1\ge k-4}\|P_{k_1}\varphi\|_{L^2_x}^2\Big)^{1/2}\Big(\sum_{k_2\ge k-4}\|P_{k_2}\boldsymbol{e}_1\|_{L^2_x}^2\Big)^{1/2}\\
&\lesssim 2^k\Big(\sum_{k_1\ge k-4}(1+s'2^{2k_1})^{-2\sigma_1}2^{-2\sigma k_1}\gamma_{k_1}(\sigma)^2\mathscr{C}(S,t)^2\Big)^{1/2}\|\boldsymbol{e}_1(s,\cdot,t)\|_{L^2_x}\\
&= \|\boldsymbol{e}_1(s,\cdot,t)\|_{L^2_x}2^k\mathscr{C}(S,t)\Big(\sum_{k_1\ge k-4}(1+s'2^{2k_1})^{-2\sigma_1}2^{-2\sigma k_1}\gamma_{k_1}(\sigma)^2\Big)^{1/2}.
\end{aligned}$$

We handle the sum as in (6-16), taking advantage of the frequency envelope summation rule (2-31), and conclude that

$$\mathrm{HH}(s,t) \lesssim 2^{-\sigma k}\gamma_k(\sigma)\mathscr{C}(S,t)(1+s2^{2k-1})^{-\sigma_1}\|\boldsymbol{e}_1(t)\|_{L^2_{s,x}}. \tag{6-23}$$

Wrapping up. For the linear term $e^{s\Delta}P_k\varphi$ we have

$$\|e^{s\Delta}P_k\varphi_0\|_{L^2_x} \le e^{-s2^{2k-2}}\|P_k\varphi_0\|_{L^2_x} \le e^{-s2^{2k-2}}2^{-\sigma k}\gamma_k(\sigma). \tag{6-24}$$

Using (6-21)–(6-24) in Duhamel's formula applied to the covariant heat equation (6-5), we have that for any $s\in[0,S]$, $t\in(-T,T)$,

$$\begin{aligned}
2^{\sigma k}\|P_k\varphi(s,\cdot,t)\|_{L^2_x}(1+s2^{2k})^{\sigma_1} &\lesssim \gamma_k(\sigma)+\mathrm{LL}(s,t)+\mathrm{LH}(s,t)+\mathrm{HH}(s,t)\\
&\lesssim \gamma_k(\sigma)+\gamma_k(\sigma)\mathscr{C}(S,t)\big(\|\boldsymbol{e}_1(t)\|^{1/2}_{L^1_sL^\infty_x}+\|\boldsymbol{e}_2(t)\|^{1/2}_{L^1_{s,x}}+\|\boldsymbol{e}_1(t)\|_{L^2_{s,x}}\big).
\end{aligned}$$

In view of (6-3) with $k=1$, (6-2) with $k=1$, and (6-4), we may split up the s-time interval $[0,\infty)$ into $O_{E_0}(1)$ intervals I_ρ on which

$$\|\boldsymbol{e}_1(t)\|^{1/2}_{L^1_sL^\infty_x(I_\rho\times\mathbb{R}^2)},\quad \|\boldsymbol{e}_2(t)\|^{1/2}_{L^1_sL^1_x(I_\rho\times\mathbb{R}^2)},\quad\text{and}\quad \|\boldsymbol{e}_1(t)\|_{L^2_sL^2_x(I_\rho\times\mathbb{R}^2)}$$

are all simultaneously small uniformly in t. By iterating a bootstrap argument $O_{E_0}(1)$ times beginning with interval I_1, we conclude that $\mathscr{C}(s,t)\lesssim 1$ for all $s>0$, uniformly in t. Therefore

$$\|P_k\varphi(s)\|_{L^\infty_tL^2_x} \lesssim (1+s2^{2k})^{-\sigma_1}2^{-\sigma k}\gamma_k(\sigma) \tag{6-25}$$

for $s\in[0,\infty)$ and $\sigma\ge 2\delta$. □

Remark 6.3. Having proven the quantitative bounds (2-35) for φ, one may establish as a corollary the qualitative bounds (2-36) for φ by using an inductive argument as in the proof of [Bejenaru et al. 2011c, Lemma 8.3]. We omit the proof, noting in particular that the argument deriving (2-36) from (2-35) does not require a small-energy hypothesis.

Proof of (2-35) *for* v, w. We begin by introducing the matrix-valued function

$$R(s,x,t) := \partial_s\varphi(s,x,t)\cdot\varphi(s,x,t)^\dagger - \varphi(s,x,t)\cdot\partial_s\varphi(s,x,t)^\dagger, \tag{6-26}$$

where here φ is thought of as a column vector. The dagger "†" denotes transpose. Using the heat flow equation (2-11) in (6-26), we rewrite R as

$$R = \Delta\varphi\cdot\varphi^\dagger - \varphi\cdot\Delta\varphi^\dagger \tag{6-27}$$

$$= \partial_m(\partial_m\varphi\cdot\varphi^\dagger - \varphi\cdot\partial_m\varphi^\dagger) \tag{6-28}$$

and proceed to bound its Littlewood–Paley projections $P_k R$ in L_x^2. Noting that by Bernstein we have

$$\|P_k(\partial_m(\partial_m\varphi\cdot\varphi^\dagger))\|_{L_x^2} \sim 2^k\|P_k(\partial_m\varphi\cdot\varphi^\dagger)\|_{L_x^2}, \tag{6-29}$$

we further decompose the nonlinearity $P_k(\partial_m\varphi\cdot\varphi^\dagger)$ as

$$P_k(\partial_m\varphi\cdot\varphi^\dagger) = \sum_{|k_2-k|\le 4} P_{\le k-4}\partial_m\varphi\cdot P_{k_2}\varphi^\dagger + \sum_{|k_1-k|\le 4} P_{k_1}\partial_m\varphi\cdot P_{\le k-4}\varphi^\dagger + \sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}} P_k(P_{k_1}\partial_m\varphi\cdot P_{k_2}\varphi^\dagger). \tag{6-30}$$

By Hölder's and Bernstein's inequalities, and by $|\varphi|\equiv 1$ and (6-25) with Bernstein,

$$\begin{aligned}\sum_{|k_2-k|\le 4}\|P_{\le k-4}\partial_m\varphi\cdot P_{k_2}\varphi\|_{L_x^2} &\lesssim \sum_{|k_2-k|\le 4} 2^k\|P_{\le k-4}\varphi\|_{L_x^\infty}\|P_{k_2}\varphi\|_{L_x^2}\\ &\lesssim 2^k(1+s2^{2k})^{-\sigma_1}2^{-\sigma k}\gamma_k(\sigma).\end{aligned} \tag{6-31}$$

Similarly,

$$\begin{aligned}\sum_{|k_1-k|\le 4}\|P_{k_1}\partial_m\varphi\cdot P_{\le k-4}\varphi\|_{L_x^2} &\lesssim \sum_{|k_1-k|\le 4}\|P_{k_1}\partial_m\varphi\|_{L_x^2}\|P_{\le k-4}\varphi\|_{L_x^\infty}\\ &\lesssim 2^k(1+s2^{2k})^{-\sigma_1}2^{-\sigma k}\gamma_k(\sigma).\end{aligned} \tag{6-32}$$

Finally, by Bernstein and Cauchy–Schwarz, energy decay, (6-25), and frequency envelope summation rule (2-31), we get

$$\begin{aligned}\sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}}\|P_k(P_{k_1}\partial_m\varphi\cdot P_{k_2}\varphi)\|_{L_x^2} &\lesssim \sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}} 2^k\|P_{k_1}\partial_m\varphi\|_{L_x^2}\|P_{k_2}\varphi\|_{L_x^2}\\ &\lesssim 2^k\sum_{k_2\ge k-4}\|P_{k_2}\varphi\|_{L_x^2}\\ &\lesssim 2^k\sum_{k_1\ge k-4}(1+s2^{2k_1})^{-\sigma_1}2^{-\sigma k_1}\gamma_{k_1}(\sigma)\\ &\lesssim 2^k(1+s2^{2k})^{-\sigma_1}2^{-\sigma k}\gamma_k(\sigma).\end{aligned} \tag{6-33}$$

Using the decomposition (6-30) and combining the cases (6-31), (6-32), and (6-33) to control (6-29), we conclude from the representation (6-28) of R that for fixed $t \in (-T, T)$,

$$2^{\sigma k} \| P_k R(s, \cdot, t) \|_{L^2_x} \lesssim 2^{2k} (1 + s 2^{2k})^{-\sigma_1} \gamma_k(\sigma).$$

As this estimate is uniform in T, it follows that

$$2^{\sigma k} \| P_k R(s) \|_{L^\infty_t L^2_x} \lesssim 2^{2k} (1 + s 2^{2k})^{-\sigma_1} \gamma_k(\sigma). \tag{6-34}$$

By arguing as in [Bejenaru et al. 2011c, Lemma 8.4], one may obtain the qualitative estimate

$$\sup_{s \geq 0} \big((1+s)^{(\sigma+2)/2} \| \partial_x^\sigma \partial_t^\rho R(s) \|_{L_t L^2_x} (1+s)^{(\sigma+3)/2} \| \partial_x^\sigma \partial_t^\rho R(s) \|_{L^\infty_{t,x}} \big) < \infty. \tag{6-35}$$

From the Duhamel representation of φ and the explicit formula for the heat kernel, one can easily show the qualitative bound[1]

$$\int_0^\infty \| R(s, \cdot, t) \|_{L^\infty_x} \, ds \lesssim_\varphi 1$$

as in [Smith 2012a, §7]. Hence we may define v as the unique solution of the ODE

$$\partial_s v = R(s) \cdot v \quad \text{and} \quad v(\infty) = Q', \tag{6-36}$$

where $Q' \in \mathbb{S}^2$ is chosen so that $Q \cdot Q' = 0$. This indeed coincides with the definition given in [Smith 2012a], since (6-36) is nothing other than the parallel transport condition $(\varphi^* \nabla)_s v = 0$ written explicitly in the setting $\mathbb{S}^2 \hookrightarrow \mathbb{R}^3$. Smoothness and basic convergence properties follow as in [Smith 2012a], to which we refer the reader for the precise results and proofs. Our goal here is to exploit (6-36) and (6-34) to prove (2-35) for v.

Using $\int_0^\infty \| \partial_x^\sigma \partial_t^\rho R(s) \|_{L^\infty_{t,x}} \, ds < \infty$ from (6-35), we conclude that

$$\sup_{s \geq 0} (1+s)^{(\sigma+1)/2} \| \partial_x^\sigma \partial_t^\rho (v(s) - Q') \|_{L^\infty_{t,x}} < \infty \tag{6-37}$$

for $\sigma, \rho \in \mathbb{Z}_+$. Integrating (6-36) in s from infinity, we get

$$v(s) - Q' + \int_s^\infty R(s') \cdot Q' \, ds' = - \int_s^\infty R(s') \cdot (v(s') - Q') \, ds', \tag{6-38}$$

which, combined with estimates (6-35) and (6-37), implies

$$\sup_{s \geq 0} \sup_{k \in \mathbb{Z}} (1+s)^{\sigma/2} 2^{\sigma k} \| P_k \partial_t^\rho v(s) \|_{L^\infty_t L^2_x} < \infty, \tag{6-39}$$

i.e., (2-36) for v. Projecting (6-36) to frequencies $\sim 2^k$ and integrating in s, we obtain

$$P_k(v(s)) = - \int_s^\infty P_k(R(s') \cdot v(s')) \, ds'. \tag{6-40}$$

[1] We may alternatively invoke (6-35) as in [Bejenaru et al. 2011c].

Set

$$\mathscr{C}_1(S,t) := \sup_{\sigma\in[2\delta,\sigma_1]} \sup_{s\in[S,\infty)} \sup_{k\in\mathbb{Z}} \gamma_k(\sigma)^{-1}(1+s2^{2k})^{\sigma_1-1}2^{\sigma k}\|P_k v(s,\cdot,t)\|_{L^2_x}.$$

That $\mathscr{C}_1(S,t)<\infty$ follows from (6-39) and $\sup_{k\in\mathbb{Z}}\gamma_k(\sigma)^{-1}2^{-\delta|k|}<\infty$. Consequently, for $s\in[S,\infty)$,

$$\|P_k v(s,\cdot,t)\|_{L^2_x} \lesssim \mathscr{C}_1(S,t)(1+s2^{2k})^{-\sigma_1+1}2^{-\sigma k}\gamma_k(\sigma). \tag{6-41}$$

We perform the Littlewood–Paley decomposition

$$\begin{aligned} P_k(R(s)v(s)) = \sum_{|k_2-k|\le 4} P_k(P_{\le k-4}R(s)P_{k_2}v(s)) \\ + \sum_{|k_1-k|\le 4} P_k(P_{k_1}R(s)P_{\le k-4}v(s)) + \sum_{k_2\ge k-4} P_k(P_{\ge k-4}R(s)P_{k_2}v(s)) \end{aligned} \tag{6-42}$$

and proceed to consider individually the various frequency interactions. By Hölder's inequality, Bernstein's inequality, and (6-41),

$$\begin{aligned} \sum_{|k_2-k|\le 4}\|P_k(P_{\le k-4}R(s)P_{k_2}v(s))\|_{L^2_x} &\lesssim \sum_{|k_2-k|\le 4}\|P_{\le k-4}R(s)\|_{L^2_x}\|P_{k_2}v(s)\|_{L^\infty_x} \\ &\lesssim \|R(s)\|_{L^2_x}\sum_{|k_2-k|\le 4}2^{k_2}\|P_{k_2}v(s)\|_{L^2_x} \\ &\lesssim \|R(s)\|_{L^2_x}2^k2^{-\sigma k}\gamma_k(\sigma)(1+s2^{2k})^{-\sigma_1+1}\mathscr{C}_1(S,t). \end{aligned} \tag{6-43}$$

By Hölder's inequality, $|v|\equiv 1$, and (6-34),

$$\begin{aligned} \sum_{|k_1-k|\le 4}\|P_k(P_{k_1}R(s)P_{\le k-4}v(s))\|_{L^2_x} &\lesssim \|P_{\le k-4}v(s)\|_{L^\infty_x}\sum_{|k_1-k|\le 4}\|P_{k_1}R(s)\|_{L^2_x} \\ &\lesssim 2^{2k}(1+s2^{2k})^{-\sigma_1}2^{-\sigma k}\gamma_k(\sigma). \end{aligned} \tag{6-44}$$

From Bernstein's inequality, Cauchy–Schwarz, (6-41), and $\sigma>2\delta$ with (2-31), it follows that

$$\begin{aligned} \sum_{k_2\ge k-4}\|P_k(P_{\ge k-4}R(s)P_{k_2}v(s))\|_{L^2_x} &\lesssim \sum_{k_2\ge k-4}2^k\|P_{\ge k-4}R(s)P_{k_2}v(s)\|_{L^1_x} \\ &\lesssim \|R(s)\|_{L^2_x}\,2^k\sum_{k_2\ge k-4}\|P_{k_2}v(s)\|_{L^2_x} \\ &\lesssim \|R(s)\|_{L^2_x}\,2^k\sum_{k_2\ge k-4}2^{-\sigma k_2}\gamma_{k_2}(\sigma)(1+s2^{2k_2})^{-\sigma_1+1}\mathscr{C}_1(S,t) \\ &\lesssim \|R(s)\|_{L^2_x}\,2^k2^{-\sigma k}\gamma_k(\sigma)(1+s2^{2k})^{-\sigma_1+1}\mathscr{C}_1(S,t). \end{aligned} \tag{6-45}$$

Using the decomposition (6-42) in (6-40) and combining the estimates (6-43), (6-44), and (6-45) gives

$$\begin{aligned} 2^{\sigma k}\|P_k v(s)\|_{L^2_x} &\le \int_s^\infty 2^{\sigma k}\|P_k(R(s')v(s'))\|_{L^2_x}\,ds' \\ &\lesssim \gamma_k(\sigma)\int_s^\infty 2^{2k}(1+s'2^{2k})^{-\sigma_1}\,ds' + \mathscr{C}_1(s,t)\gamma_k(\sigma)\int_s^\infty \|R(s')\|_{L^2_x}2^k(1+s'2^{2k})^{-\sigma_1+1}ds'. \end{aligned}$$

Applying Cauchy–Schwarz in s, we obtain

$$
\begin{aligned}
2^{\sigma k}\|P_k v(s)\|_{L^2_x} &\lesssim \gamma_k(\sigma)\int_s^\infty 2^{2k}(1+s'2^{2k})^{-\sigma_1}\,ds' \\
&\qquad + \mathscr{C}_1(s,t)\gamma_k(\sigma)\left(\int_s^\infty \|R(s')\|_{L^2_x}^2\,ds'\right)^{1/2}\left(\int_s^\infty 2^{2k}(1+s'2^{2k})^{-2\sigma_1+2}\,ds'\right)^{1/2} \\
&\lesssim \gamma_k(\sigma)+\mathscr{C}_1(s,t)\gamma_k(\sigma)\left(\int_s^\infty \|R(s')\|_{L^2_x}^2\,ds'\right)^{1/2}.
\end{aligned}
\tag{6-46}
$$

As noted in (6-15), we have $|\Delta\varphi| \le \sqrt{\boldsymbol{e}_2}+\boldsymbol{e}_1$, so it follows from the representation (6-27) of R that

$$|R(s,x,t)| \le |\boldsymbol{e}_1(s,x,t)| + |\sqrt{\boldsymbol{e}_2}(s,x,t)|. \tag{6-47}$$

As (6-47) implies

$$\int_0^\infty \|R(s)\|_{L^2_x}^2\,ds \lesssim \|\boldsymbol{e}_2\|_{L^1_{s,x}} + \|\boldsymbol{e}_1\|_{L^2_{s,x}}^2,$$

we therefore, in view of (6-2) with $k=1$ and (6-4), may choose S large so that the integral of the R term in (6-46) is small, say $\le \varepsilon$. Then

$$\mathscr{C}_1(S,t) \lesssim 1+\varepsilon\mathscr{C}_1(S,t),$$

so that $\mathscr{C}_1(S) \lesssim 1$ for such S. In fact, together (6-2) and (6-4) imply that we may divide the time interval $[0,\infty)$ into $O_{E_0}(1)$ subintervals I_ρ so that on each such subinterval

$$\int_{I_\rho} \|R(s)\|_{L^2_x}^2\,ds \le \varepsilon^2.$$

Hence by a simple iterative bootstrap argument we conclude that

$$\mathscr{C}_1(0,t) \lesssim 1. \tag{6-48}$$

As (6-48) is uniform in t, we have

$$\|P_k v(s,\cdot,t)\|_{L^2_x} \lesssim (1+s2^{2k})^{-\sigma_1+1}2^{-\sigma k}\gamma_k(\sigma). \tag{6-49}$$

By repeating the argument above with w in place of v (and appropriately modifying the boundary condition at ∞ in (6-36)), we get

$$\|P_k w(s,\cdot,t)\|_{L^2_x} \lesssim (1+s2^{2k})^{-\sigma_1+1}2^{-\sigma k}\gamma_k(\sigma) \tag{6-50}$$

and $\sup_{s\ge 0}\sup_{k\in\mathbb{Z}}(1+s)^{\sigma/2}2^{\sigma k}\|P_k\partial_t^\rho w(s)\|_{L^\infty_t L^2_x} < \infty$, and so (2-35) and (2-36) follow for w. □

Proof of (2-37). Recall that

$$\psi_m = v\cdot\partial_m\varphi + iw\cdot\partial_m\varphi = -\partial_m v\cdot\varphi - i\partial_m w\cdot\varphi. \tag{6-51}$$

Our first aim is to control $\|P_k\psi_x\|_{L^\infty_t L^2_x}$. We start with a Littlewood–Paley decomposition of $\partial_m v\cdot\varphi$:

$$
\begin{aligned}
&P_k(\partial_m v\cdot\varphi) \\
&\quad = \sum_{|k_2-k|\le 4} P_k(P_{\le k-5}\partial_m v\cdot P_{k_2}\varphi) + \sum_{|k_1-k|\le 4} P_k(P_{k_1}\partial_m v\cdot P_{\le k-5}\varphi) + \sum_{\substack{k_1,k_2\ge k-4\\ |k_1-k_2|\le 8}} P_k(P_{k_1}\partial_m v\cdot P_{k_2}\varphi).
\end{aligned}
\tag{6-52}
$$

To control the low-high frequency term we apply Hölder's inequality, energy decay, and (6-25) with Bernstein's inequality:

$$\sum_{|k_2-k|\le 4} \|P_k(P_{\le k-5}\partial_m v \cdot P_{k_2}\varphi)\|_{L^2_x} \lesssim \sum_{|k_2-k|\le 4} \|P_{\le k-5}\partial_m v\|_{L^2_x}\|P_{k_2}\varphi\|_{L^\infty_x} \lesssim (1+s2^{2k})^{-\sigma_1}2^k2^{-\sigma k}\gamma_k(\sigma). \tag{6-53}$$

We control the high-low frequency term by using Hölder's inequality, $|\varphi|\equiv 1$, and (6-49):

$$\sum_{|k_1-k|\le 4} \|P_k(P_{k_1}\partial_m v \cdot P_{\le k-5}\varphi)\|_{L^2_x} \lesssim \sum_{|k_1-k|\le 4} \|P_{k_1}\partial_m v\|_{L^2_x}\|P_{\le k-5}\varphi\|_{L^\infty_x} \lesssim (1+s2^{2k})^{-\sigma_1}2^k2^{-\sigma k}\gamma_k(\sigma). \tag{6-54}$$

To control the high-high frequency term, we use Bernstein's inequality and Cauchy–Schwarz, energy conservation and (6-25), and (2-31):

$$\begin{aligned}\sum_{\substack{k_1,k_2\ge k-4\\|k_1-k_2|\le 8}} \|P_k(P_{k_1}\partial_m v \cdot P_{k_2}\varphi)\|_{L^2_x} &\lesssim \sum_{\substack{k_1,k_2\ge k-4\\|k_1-k_2|\le 8}} 2^k\|P_{k_1}\partial_m v\|_{L^2_x}\|P_{k_2}\varphi\|_{L^2_x}\\ &\lesssim 2^k \sum_{k_2\ge k-4}(1+s2^{2k_2})^{-\sigma_1}2^{-\sigma k_2}\gamma_{k_2}(\sigma)\\ &\lesssim (1+s2^{2k})^{-\sigma_1}2^k2^{-\sigma k}\gamma_k(\sigma).\end{aligned} \tag{6-55}$$

We conclude using (6-53), (6-54), and (6-55) in representation (6-52) that

$$\|P_k(\partial_m v\cdot\varphi)\|_{L^2_x} \lesssim (1+s2^{2k})^{-\sigma_1}2^k2^{-\sigma k}\gamma_k(\sigma). \tag{6-56}$$

By repeating the argument with w in place of v, it follows that (6-56) also holds with w in place of v. Therefore, referring back to (6-51), we conclude that

$$\|P_k\psi_m\|_{L^2_x} \lesssim (1+s2^{2k})^{-\sigma_1}2^k2^{-\sigma k}\gamma_k(\sigma).$$

As this bound is uniform in t, (2-37) holds for ψ_m.

Recalling that

$$A_m = \partial_m v\cdot w,$$

and repeating the argument with w in place of φ and (6-50) in place of (6-25), we conclude that

$$\|P_k A_x(s)\|_{L^\infty_t L^2_x} \lesssim (1+s2^{2k})^{-\sigma_1+1}2^k2^{-\sigma k}\gamma_k(\sigma). \qquad \square$$

7. Proofs of parabolic estimates

The purpose of this section is to prove the parabolic heat-time estimates stated in Section 4A. Many of these estimates have counterparts in [Bejenaru et al. 2011c]. Nevertheless, our proofs are more involved since we only require energy dispersion, which is weaker than the small-energy assumption made in [Bejenaru et al. 2011c]. Some of the L^p estimates in Section 7B are new.

Throughout we assume ε_1 energy dispersion on the initial data as stated in (4-4) and we assume that the bootstrap hypothesis (4-6) holds. Let $\sigma_1 \in \mathbb{Z}_+$ be positive and fixed. We work exclusively with $\sigma \in [0, \sigma_1 - 1]$, even if this is not always explicitly stated. Set $\varepsilon = \varepsilon_1^{7/5}$ for short.

In this section we extensively use the spaces defined via (3-2). They provide a crucial gain in high-high frequency interactions, which is captured in Lemmas 7.2 and 7.14.

Lemma 7.1. *Let $f \in L^2_{k_1}(T)$, where $|k_1 - k| \le 20$, let $0 \le \omega' \le 1/2$, and let $h \in L^2_k(T)$. Then*

$$\|P_k(fg)\|_{F_k(T)} \lesssim \|f\|_{F_{k_1}(T)} \|g\|_{L^\infty_{t,x}},$$

$$\|P_k(fg)\|_{S_k^{\omega'}(T)} \lesssim \|f\|_{F_{k_1}(T)} 2^{k\omega'} \|g\|_{L_x^{2/\omega'} L_t^\infty},$$

$$\|h\|_{L^\infty_{t,x}} + 2^{k\omega'} \|h\|_{L_x^{2/\omega'} L_t^\infty} \lesssim 2^k \|h\|_{F_k(T)}.$$

Moreover, for f_{k_1}, g_{k_2} belonging to $L^2_{k_1}(T)$, $L^2_{k_2}(T)$ respectively, and with $|k_1 - k_2| \le 8$, we have

$$\|P_k(f_{k_1} g_{k_2})\|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim 2^k 2^{(k_2 - k)(1-\omega)} \|f_{k_1}\|_{S^\omega_{k_1}(T)} \|g_{k_2}\|_{S^0_{k_2}(T)}.$$

Proof. For the proofs, see [Bejenaru et al. 2011c, §3]. □

Lemma 7.2. *Assume that $T \in (0, 2^{2\mathcal{K}}]$, $f, g \in H^{\infty,\infty}(T)$, $P_k f \in F_k(T) \cap S^\omega_k(T)$, $P_k g \in F_k(T)$ for some $\omega \in [0, 1/2]$ and all $k \in \mathbb{Z}$, and*

$$\alpha_k = \sum_{|j-k| \le 20} \|P_j f\|_{F_j(T) \cap S^\omega_j(T)}, \quad \beta_k = \sum_{|j-k| \le 20} \|P_j g\|_{F_j(T)}.$$

Then, for any $k \in \mathbb{Z}$,

$$\|P_k(fg)\|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim \sum_{j \le k} 2^j (\beta_k \alpha_j + \alpha_k \beta_j) + 2^k \sum_{j \ge k} 2^{(j-k)(1-\omega)} \alpha_j \beta_j.$$

Proof. For the proof, see [Bejenaru et al. 2011c, §5]. □

7A. *Derivative field control.* The main purpose of this subsection is to establish the estimate (4-12), which states that

$$\|P_k \psi_m(s)\|_{F_k(T)} \lesssim (1 + s2^{2k})^{-4} 2^{-\sigma k} b_k(\sigma).$$

In the course of the proof we shall also establish auxiliary estimates useful elsewhere. Estimate (4-12) plays a key role in controlling the nonlinear paradifferential flow, allowing us to gain regularity by integrating in heat time. The proof uses a bootstrap argument and exploits the Duhamel formula.

Recall that the fields ψ_α, A_α, $\alpha = 1, 2, 3$, ($\psi_3 \equiv \psi_t$, $A_3 \equiv A_t$) satisfy (2-20), which states that

$$(\partial_s - \Delta)\psi_\alpha = U_\alpha.$$

We use representation (2-22) of the heat nonlinearity:

$$U_\alpha := 2i A_l \partial_l \psi_\alpha + i(\partial_l A_l)\psi_\alpha - A_x^2 \psi_\alpha + i\,\mathrm{Im}(\psi_\alpha \overline{\psi_l})\psi_l.$$

Hence ψ_α admits the representation

$$\psi_\alpha(s) = e^{s\Delta}\psi_\alpha(s_0) + \int_{s_0}^{s} e^{(s-s')\Delta} U_\alpha(s')\, ds' \tag{7-1}$$

for any $s \ge s_0 \ge 0$.

For each $k \in \mathbb{Z}$, set

$$a(k) := \sup_{s\in[0,\infty)} (1+s2^{2k})^4 \sum_{m=1,2} \|P_k \psi_m(s)\|_{F_k(T)},$$

and for $\sigma \in [0, \sigma_1 - 1]$ introduce the frequency envelopes

$$a_k(\sigma) = \sup_{j\in\mathbb{Z}} 2^{-\delta|k-j|} 2^{\sigma j} a(j). \tag{7-2}$$

The frequency envelopes $a_k(\sigma)$ are finite and in l^2 by (2-38) and (3-1).

Our goal is to show $a_k(\sigma) \lesssim b_k(\sigma)$, which in particular implies (4-12).

Lemma 7.3. *Suppose that ψ_x satisfies the bootstrap condition*

$$\|P_k \psi_x(s)\|_{F_k(T)\cap S_k^{1/2}(T)} \le \varepsilon_p^{-1/2} b_k (1+s2^{2k})^{-4}. \tag{7-3}$$

Then (4-12) holds.

We can take $\varepsilon_p = \varepsilon_1^{1/10}$, for instance. As in [Bejenaru et al. 2011c], this result may be strengthened:

Corollary 7.4. *The estimate (4-12) holds even when the bootstrap hypothesis (7-3) is dropped.*

Proof. Directly apply the argument of [Bejenaru et al. 2011c, Corollary 4.4], which we omit. □

The sequence of lemmas we prove in order to establish Lemma 7.3 culminates in Lemma 7.11, which controls the nonlinear term of the Duhamel formula (7-1) by $2^{-\sigma k} a_k(\sigma)$ along with suitable decay and an epsilon-gain arising from energy dispersion. Its immediate predecessor, Lemma 7.10, controls $P_k U_m$ in $F_k(T)$.

Referring back to (2-22) and seeing as how U_m contains the term $2i A_l \partial_l \psi_m$, we see that in order to apply the parabolic estimates of Lemma 7.1 toward controlling $P_k U_m$, it is necessary that we first control $P_k A_m$ in $F_k(T)$ in terms of the frequency envelopes $\{a_l(\sigma)\}$, and it is to this that we now turn.

For $k, k_0 \in \mathbb{Z}$ and $s \in [2^{2k_0-1}, 2^{2k_0+1})$, set

$$b_{k,s}(\sigma) = \begin{cases} \displaystyle\sum_{j=k}^{-k_0} a_j a_j(\sigma) & \text{if } k+k_0 \le 0, \\ 2^{k+k_0} a_{-k_0} a_k(\sigma) & \text{if } k+k_0 \ge 0. \end{cases}$$

Let $\mathscr{C}$ be the smallest number in $[1, \infty)$ such that

$$\|P_k A_m(s)\|_{F_k(T)\cap S_k^{1/2}(T)} \le \mathscr{C}(1+s2^{2k})^{-4} 2^{-\sigma k} b_{k,s}(\sigma) \tag{7-4}$$

for all $s \in [0, \infty)$, $k \in \mathbb{Z}$, $m = 1, 2$, and $\sigma \in [0, \sigma_1 - 1]$. While this constant is indeed finite, it is not a priori controlled by energy. To show that $\mathscr{C}$ is indeed controlled by energy, we use the integral representation

$$A_m(s) = -\sum_{l=1,2} \int_s^\infty \operatorname{Im}(\overline{\psi_m}(\partial_l \psi_l + i A_l \psi_l))(r)\, dr \tag{7-5}$$

and seek to control the Littlewood–Paley projection of the integrand in $F_k(T) \cap S_k^{1/2}(T)$. We treat differently the two types of terms in (7-5) that need to be controlled. In Lemma 7.5 we bound terms of the sort $P_k(\psi_x \overline{\psi_x})$ and $P_k(\psi_x \partial_x \overline{\psi_x})$ in $F_k(T) \cap S_k^{1/2}(T)$. In Lemma 7.6 we combine the estimate on $P_k(\psi_x \overline{\psi_x})$ with (7-4) to obtain control on $P_k(\psi_x \overline{\psi_x} A_x)$, gaining an epsilon from energy dispersion. Using (7-5) and exploiting the epsilon gain from energy dispersion will lead us to the conclusion of Lemma 7.7: $\mathscr{C} \lesssim 1$.

We use the following bracket notation in the sequel:

$$\langle f \rangle := (1 + f^2)^{1/2}.$$

Lemma 7.5. *For any $f, g \in \{\psi_m, \overline{\psi_m} : m = 1, 2\}$, $r \in [2^{2j-2}, 2^{2j+2}]$, $j \in \mathbb{Z}$, $i = 1, 2$, and $\sigma \in [0, \sigma_1 - 1]$, we have the bounds*

$$\| P_k(f(r)g(r)) \|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{-j} a_{-j} a_{\max(k,-j)}(\sigma) \tag{7-6}$$

and

$$\| P_k(f(r)\partial_i g(r)) \|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{-j} a_{-j} (2^k a_k(\sigma) + 2^{-j} a_{-j}(\sigma)). \tag{7-7}$$

Proof. By Lemma 7.2 with $\omega = 0$ we have

$$\| P_k(fg) \|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim \sum_{l \le k} 2^l \alpha_k \beta_l + \sum_{l \ge k} 2^l \alpha_l \beta_l, \tag{7-8}$$

where, due to the definition (7-2), α_k and β_k satisfy

$$\alpha_k \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} a_k(\sigma), \quad \beta_k \lesssim \langle 2^{j+k} \rangle^{-8} a_k. \tag{7-9}$$

Turning to the high-low frequency interaction first, we have using (7-9) and the frequency envelope property (2-29) that

$$\sum_{l \le k} 2^l \alpha_k \beta_l \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{-j} a_{-j} \sum_{l \le k} \langle 2^{j+l} \rangle^{-8} 2^{j+l} 2^{\delta |j+l|} a_k(\sigma). \tag{7-10}$$

Thus it remains to show that

$$\sum_{l \le k} \langle 2^{j+l} \rangle^{-8} 2^{j+l} 2^{\delta |j+l|} a_k(\sigma) \lesssim a_{\max(k,-j)}(\sigma), \tag{7-11}$$

which follows from pulling out a factor of $a_k(\sigma)$ or $a_{-j}(\sigma)$, according to whether $k + j \ge 0$ or $k + j < 0$, and then summing the remaining geometric series. In case $k + j < 0$ we pull out a factor of $a_{-j}(\sigma)$ via (2-29).

Turning to the high-high frequency interaction term, we have

$$\sum_{l \ge k} 2^l \alpha_l \beta_l \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{-j} a_{-j} \sum_{l \ge k} \langle 2^{j+l} \rangle^{-8} 2^{j+l} 2^{\delta |j+l|} a_l(\sigma), \tag{7-12}$$

and so it remains to show that

$$\sum_{l \ge k} \langle 2^{j+l} \rangle^{-8} 2^{j+l} 2^{\delta |j+l|} a_l(\sigma) \lesssim a_{\max(k,-j)}(\sigma). \tag{7-13}$$

When $k+j \geq 0$, we have, using (2-31),

$$\sum_{l \geq k} \langle 2^{j+l} \rangle^{-8} 2^{j+l} 2^{\delta|j+l|} a_l(\sigma) \lesssim a_k(\sigma) \sum_{l \geq k} 2^{(2\delta-1)(j+l)} \lesssim a_k(\sigma).$$

If $k+j \leq 0$, we control the sum with (2-30) if $l+j<0$ and with (2-31) if $l+j \geq 0$. Hence (7-13) holds.

Together (7-8)–(7-13) imply (7-6).

To establish (7-7) we follow a similar strategy. By Lemma 7.2 with $\omega = 0$ we have

$$\|P_k(f \partial_i g)\|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim \sum_{l \leq k} 2^l \alpha_l \beta_k + \sum_{l \geq k} 2^l \alpha_k \beta_l + \sum_{l \geq k} 2^l \alpha_l \beta_l, \tag{7-14}$$

where for any $\sigma \in [0, \sigma_1 - 1]$ we have

$$\alpha_k \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} a_k(\sigma) \quad \text{and} \quad \beta_k \lesssim \langle 2^{j+k} \rangle^{-8} 2^k 2^{-\sigma k} a_k(\sigma). \tag{7-15}$$

Beginning with the low-high frequency interaction, we have

$$\sum_{l \leq k} 2^l \alpha_l \beta_k \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^k a_k(\sigma) \sum_{l \leq k} \langle 2^{j+l} \rangle^{-8} 2^l a_l, \tag{7-16}$$

and so it remains to show that

$$\sum_{l \leq k} \langle 2^{j+l} \rangle^{-8} 2^l a_l \lesssim 2^{-j} a_{-j}. \tag{7-17}$$

If $k+j \leq 0$, then (7-17) holds due to (2-30). If $k+j \geq 0$, then we apply (2-30) and (2-31) according to whether $l+j \leq 0$ or $l+j > 0$.

Turning now to the high-low frequency interaction, we have

$$\sum_{l \leq k} 2^l \alpha_k \beta_l \lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{-j} a_{-j} 2^k a_k(\sigma) \sum_{l \leq k} \langle 2^{j+l} \rangle^{-8} 2^{l-k} 2^{l+j} 2^{\delta|l+j|}. \tag{7-18}$$

We need only check that

$$\sum_{l \leq k} \langle 2^{j+l} \rangle^{-8} 2^{l-k} 2^{l+j} 2^{\delta|l+j|} \lesssim 1,$$

which can be seen to hold by breaking into cases $k+j \leq 0$ and $k+j \geq 0$.

We conclude with the high-high frequency interaction:

$$\begin{aligned} \sum_{l \geq k} 2^l \alpha_l \beta_l &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} \sum_{l \geq k} \langle 2^{j+l} \rangle^{-8} 2^{2l} a_l(\sigma) a_l \\ &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{-2j} a_j a_j(\sigma) \sum_{l \geq k} \langle 2^{j+l} \rangle^{-8} 2^{2l+2j} 2^{2\delta|l+j|}. \end{aligned} \tag{7-19}$$

Here

$$\sum_{l \geq k} \langle 2^{j+l} \rangle^{-8} 2^{2l+2j} 2^{2\delta|l+j|} \lesssim 1, \tag{7-20}$$

which is seen to hold by considering separately the cases $k+j \geq 0$, $k+j < 0$.

Combining (7-16)–(7-20), we conclude (7-7). □

Lemma 7.6. *Let*

$$f(r) \in \{\overline{\psi_m}(r)\psi_l(r) : m, l = 1, 2\}, \quad g(r) \in \{A_m(r) : m = 1, 2\},$$

and $r \in [2^{2j-2}, 2^{2j+2}]$. *Then*

$$\|P_k(fg)(r)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \begin{cases} \varepsilon\mathscr{C}2^{-\sigma k}2^{-2j}a_{-j}a_{-j}(\sigma) & k+j \leq 0, \\ \varepsilon\mathscr{C}\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{-2j}b_{k,r}(\sigma) & k+j \geq 0. \end{cases}$$

Proof. We apply Lemma 7.2. By (7-6) and (7-4) we have

$$\alpha_k(r) \lesssim 2^{-\sigma k}\langle 2^{j+k}\rangle^{-8}2^{-j}a_{-j}a_{\max(k,-j)}(\sigma), \tag{7-21}$$

and

$$\beta_k(r) \lesssim \mathscr{C}2^{-\sigma k}\langle 2^{j+k}\rangle^{-8}b_{k,r}(\sigma), \tag{7-22}$$

for any $\sigma \in [0, \sigma_1 - 1]$.

We consider six cases, treating separately the low-high, high-low, and high-high frequency interactions, which we further divide according to whether $k+j \geq 0$ or $k+j \leq 0$.

Low-high frequency interaction with $k + j \geq 0$. Using (7-21) and (7-22), we have

$$\sum_{l\leq k} 2^l \alpha_l \beta_k \lesssim \mathscr{C}\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{-2j}b_{k,r}(\sigma)\sum_{l\leq k}2^l 2^{2j}\alpha_l, \tag{7-23}$$

and so it remains to verify that

$$\sum_{l\leq k} 2^l 2^{2j}\alpha_l \lesssim \varepsilon. \tag{7-24}$$

Taking $\sigma = 0$ in the bounds (7-21) for α_l and using (2-29), (2-31) yields

$$\begin{aligned}\sum_{l\leq k} 2^l 2^{2j}\alpha_l &\lesssim \sum_{l\leq k}\langle 2^{j+l}\rangle^{-8}2^l 2^{2j}2^{-j}a_{-j}a_{\max(l,-j)} \\ &= \sum_{l\leq -j}2^{l+j}a_{-j}^2 + \sum_{-j<l\leq k}\langle 2^{j+l}\rangle^{-8}2^{l+j}a_{-j}a_l \lesssim a_{-j}^2 + a_{-j}^2\sum_{-j<l\leq k}\langle 2^{j+l}\rangle^{-8}2^{(1+\delta)(l+j)} \lesssim \varepsilon,\end{aligned}$$

which proves (7-23).

High-low frequency interaction with $k + j \geq 0$. Taking $\sigma = 0$ in the bounds for $b_{l,r}$, we have

$$\sum_{l\leq k} 2^l \alpha_k \beta_l \lesssim \mathscr{C}\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{-2j}b_{k,r}(\sigma)\sum_{l\leq k}\langle 2^{j+l}\rangle^{-8}2^{l-k}b_{l,r}, \tag{7-25}$$

and so it remains to show that

$$\sum_{l\leq k}\langle 2^{j+l}\rangle^{-8}2^{l-k}b_{l,r} \lesssim \varepsilon. \tag{7-26}$$

We split the sum as follows:

$$\sum_{l\leq k}\langle 2^{j+l}\rangle^{-8}2^{l-k}b_{l,r} = \sum_{l\leq -j}\langle 2^{j+l}\rangle^{-8}2^{l-k}\sum_{q=l}^{-j}a_q^2 + \sum_{-j<l\leq k}\langle 2^{j+l}\rangle^{-8}2^{l-k}2^{l+j}a_{-j}a_l.$$

The first summand is controlled by

$$\sum_{l\le -j}\langle 2^{j+l}\rangle^{-8}2^{l-k}\sum_{q=l}^{-j}a_q^2\lesssim a_{-j}^2\sum_{l\le -j}2^{l-k}\sum_{q=l}^{-j}2^{-2\delta(j+q)}\lesssim a_{-j}^2\lesssim\varepsilon.$$

The second summand by may be handled similarly, thus proving (7-26).

High-high frequency interaction with $k+j\ge 0$. Taking $\sigma=0$ in the bound (7-22) for β_l, we have

$$\begin{aligned}\sum_{l\ge k}2^l\alpha_l\beta_l&\lesssim\langle 2^{j+l}\rangle^{-8}2^k\sum_{l\ge k}2^{l-k}2^{-\sigma l}2^{-j}a_{-j}a_l(\sigma)\mathscr{C}2^{l+j}a_{-j}a_l\\&\lesssim\mathscr{C}\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{-2j}b_{k,r}(\sigma)\sum_{l\ge k}\langle 2^{j+l}\rangle^{-8}2^{l-k}2^{\delta(l-k)}2^{l+j}a_{-j}a_l,\end{aligned}\tag{7-27}$$

and so it remains to show that

$$\sum_{l\ge k}\langle 2^{j+l}\rangle^{-8}2^{l-k}2^{\delta(l-k)}2^{l+j}a_{-j}a_l\lesssim\varepsilon,\tag{7-28}$$

which follows, for instance, from pulling out a_{-j}^2 via (2-29) and summing.

In view of (7-23)–(7-28), it follows from Lemma 7.2, with $\omega=0$ that

$$\|P_k(fg)(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim\varepsilon\mathscr{C}\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{-2j}b_{k,r}(\sigma)\quad\text{for } k+j\ge 0\tag{7-29}$$

as required.

Low-high frequency interaction with $k+j\le 0$. In this case it follows from (7-22) that

$$\beta_k\lesssim\mathscr{C}2^{-\sigma k}\sum_{p=k}^{-j}a_pa_p(\sigma),$$

so that

$$\begin{aligned}\sum_{l\le k}2^l\alpha_l\beta_k&\lesssim\mathscr{C}2^{-\sigma k}2^{-j}a_{-j}a_{-j}\sum_{p=k}^{-j}a_pa_p(\sigma)\sum_{l\le k}\langle 2^{j+l}\rangle^{-8}2^l\\&\lesssim\mathscr{C}2^{-\sigma k}2^{-2j}a_{-j}a_{-j}(\sigma)\cdot a_{-j}\sum_{p=k}^{-j}a_p2^{-\delta(j+p)}\sum_{l\le k}2^{l+j}.\end{aligned}\tag{7-30}$$

It remains to show that

$$a_{-j}\sum_{p=k}^{-j}a_p2^{-\delta(j+p)}\sum_{l\le k}2^{l+j}\lesssim\varepsilon,$$

which follows from pulling out a_p as an a_{-j} via (2-29) and summing.

High-low frequency interaction with $k+j\le 0$. In this case

$$\sum_{l\le k}2^l\alpha_k\beta_l\lesssim\mathscr{C}2^{-2j}a_{-j}a_{-j}(\sigma)\sum_{l\le k}2^{l+j}\sum_{p=l}^{-j}a_p^2,\tag{7-31}$$

and so we need to show that

$$\sum_{l\le k} 2^{l+j} \sum_{p=l}^{-j} a_p^2 \lesssim \varepsilon,$$

which follows by pulling out a_{-j}^2 and summing.

High-high frequency interaction with $k+j \le 0$. As a first step we write

$$2^k \sum_{l\ge k} 2^{(l-k)/2}\alpha_l\beta_l = 2^k \sum_{k\le l<-j} 2^{(l-k)/2}\alpha_l\beta_l + 2^k \sum_{l\ge -j} 2^{(l-k)/2}\alpha_l\beta_l. \tag{7-32}$$

The first summand is controlled by

$$2^k \sum_{k\le l<-j} 2^{(l-k)/2}\alpha_l\beta_l \lesssim \mathscr{C} 2^{-\sigma k}2^{-2j}a_{-j}a_{-j}(\sigma) \sum_{k\le l<-j} 2^{(l-k)/2}2^{k+j}2^{-\sigma(l-k)} \sum_{p=l}^{-j} a_p^2. \tag{7-33}$$

We have

$$\sum_{k\le l<-j} 2^{(l-k)/2}2^{k+j}2^{-\sigma(l-k)} \sum_{p=l}^{-j} a_p^2 \lesssim a_{-j}^2 2^{(k+j)/2} \sum_{k\le l<-j} 2^{-2\delta(j+l)} \lesssim \varepsilon,$$

which establishes the desired control on the first summand.

The second summand is controlled by

$$\begin{aligned}2^k \sum_{l\ge -j} 2^{(l-k)/2}\alpha_l\beta_l &\lesssim 2^k \sum_{l\ge -j} 2^{(l-k)/2}\langle 2^{j+l}\rangle^{-8}2^{-\sigma l}2^{-j}a_{-j}a_l(\sigma)\mathscr{C}\langle 2^{j+l}\rangle^{-8}2^{l+j}a_{-j}a_l \\ &\lesssim \mathscr{C}2^{-\sigma k}2^{-2j}a_{-j}a_{-j}(\sigma) \sum_{l\ge -j} 2^{(l-k)/2}2^{k+j}2^{(1+\delta)(l+j)}a_{-j}a_l,\end{aligned} \tag{7-34}$$

and so it remains to show that

$$\sum_{l\ge -j} 2^{(l-k)/2}2^{k+j}2^{(1+\delta)(l+j)}a_{-j}a_l \lesssim \varepsilon, \tag{7-35}$$

which follows from pulling out a_{-j}^2 and summing.

Combining (7-30)–(7-35), we conclude from applying Lemma 7.2 with $\omega = 1/2$ that

$$\|P_k(fg)(r)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon\mathscr{C}2^{-\sigma k}2^{-2j}a_{-j}a_{-j}(\sigma) \quad \text{for } k+j\le 0,$$

which, combined with (7-29) completes the proof of the lemma. □

Lemma 7.7. *For any $k\in\mathbb{Z}$ and $s\in[0,\infty)$ we have*

$$\|P_kA_m(s)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim (1+s2^{2k})^{-4}2^{-\sigma k}b_{k,s}(\sigma).$$

Proof. From the representation (7-5) for A_m it follows that

$$\begin{aligned}\|P_kA_m(s)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim &\int_s^\infty \|P_k(\overline{\psi_m}(r)\partial_l\psi_l(r))\|_{F_k(T)\cap S_k^{1/2}(T)}\,dr \\ &+ \int_s^\infty \|P_k(\overline{\psi_m}(r)\psi_lA_l(r))\|_{F_k(T)\cap S_k^{1/2}(T)}\,dr.\end{aligned} \tag{7-36}$$

Taking $k_0 \in \mathbb{Z}$ so that $s \in [2^{2k_0-1}, 2^{2k_0+1})$ and using (7-7), we see that the first term is dominated by

$$\sum_{j \geq k_0} \int_{2^{2j-1}}^{2^{2j+1}} \| P_k(\overline{\psi}_m(r)\partial_l \psi_l(r)) \|_{F_k(T)\cap S_k^{1/2}(T)} dr \lesssim 2^{-\sigma k} \sum_{j \geq k_0} \langle 2^{j+k} \rangle^{-8} (2^{j+k} a_{-j} a_k(\sigma) + a_{-j} a_{-j}(\sigma)). \tag{7-37}$$

We claim that

$$\sum_{j \geq k_0} \langle 2^{j+k} \rangle^{-8} (2^{j+k} a_{-j} a_k(\sigma) + a_{-j} a_{-j}(\sigma)) \lesssim (1 + s2^{2k})^{-4} b_{k,s}(\sigma). \tag{7-38}$$

When $k + k_0 \geq 0$, it follows from (2-29) that the left-hand side of (7-38) is bounded by

$$\begin{aligned} 2^{k_0+k} a_{-k_0} a_k(\sigma) \sum_{j \geq k_0} \langle 2^{j+k} \rangle^{-8} & \left(2^{j-k_0} 2^{\delta(j-k_0)} + 2^{-k_0-k} 2^{\delta(j-k_0)} 2^{\delta(k+j)} \right) \\ & \lesssim b_{k,s}(\sigma) \sum_{j \geq k_0} \langle 2^{j+k} \rangle^{-8} (2^{(1+\delta)(j-k_0)} + 2^{(\delta-1)(k_0+k)} 2^{2\delta(j-k_0)}), \end{aligned} \tag{7-39}$$

and so it suffices to show that

$$\sum_{j \geq k_0} \langle 2^{j+k} \rangle^{-8} 2^{2(j-k_0)} \lesssim \langle 2^{j+k_0} \rangle^{-8}, \tag{7-40}$$

which follows from series comparison, for instance.

Together (7-40) and (7-39), show that (7-38) holds for $k + k_0 \geq 0$.

If, on the other hand, $k + k_0 \leq 0$, then we split the sum in (7-38) according to whether $j + k \leq 0$ or $j + k > 0$. In the first case,

$$\sum_{k_0 \leq j \leq -k} \langle 2^{j+k} \rangle^{-8} (2^{j+k} a_{-j} a_k(\sigma) + a_{-j} a_{-j}(\sigma)) \lesssim \langle 2^{k_0+k} \rangle^{-8} b_{k,s}(\sigma) + \sum_{k_0 \leq j \leq -k} \langle 2^{j+k} \rangle^{-8} 2^{j+k} a_{-j} a_k(\sigma). \tag{7-41}$$

Then

$$\sum_{k_0 \leq j \leq -k} \langle 2^{j+k} \rangle^{-8} 2^{j+k} a_{-j} a_k(\sigma) \lesssim \sum_{k_0 \leq j \leq -k} \langle 2^{j+k} \rangle^{-8} 2^{j+k} a_{-j} a_{-j}(\sigma) 2^{-\delta(j+k)} \sim (1 + s2^{2k})^{-4} b_{k,s}(\sigma). \tag{7-42}$$

When $j + k > 0$ we have

$$\begin{aligned} \sum_{j > -k} \langle 2^{j+k} \rangle^{-8} (2^{j+k} a_{-j} a_k(\sigma) + a_{-j} a_{-j}(\sigma)) & \lesssim a_k a_k(\sigma) \sum_{j > -k} \langle 2^{j+k} \rangle^{-8} (2^{j+k} 2^{\delta(j+k)} + 2^{2\delta(j+k)}) \\ & \lesssim b_{k,s}(\sigma). \end{aligned} \tag{7-43}$$

Therefore (7-41) and (7-42) imply (7-38) holds when $k + k_0 \leq 0$ and $j + k \leq 0$ and (7-43) implies it holds when both $k + k_0 \leq 0$ and $j + k > 0$.

Having shown (7-38), we combine it with (7-37), concluding that

$$\int_s^\infty \| P_k(\overline{\psi}_m(r)\partial_l \psi_l(r)) \|_{F_k(T)\cap S_k^{1/2}(T)} dr \lesssim (1 + s2^{2k})^{-4} 2^{-\sigma k} b_{k,s}(\sigma). \tag{7-44}$$

We move on to control the second term in (7-36). By Lemma 7.6 and (7-38), this term is bounded by

$$\begin{aligned}\sum_{j\geq k_0}\int_{2^{2j-1}}^{2^{2j+1}}\|P_k(\overline{\psi_x}(r)\psi_x(r)A_x(r))\|_{F_k(T)\cap S_k^{1/2}(T)}\,dr\\ &\lesssim \mathscr{C}2^{-\sigma k}\varepsilon\sum_{j\geq k_0}\langle 2^{j+k}\rangle^{-8}(\mathbf{1}_-(k+j)a_{-j}a_{-j}(\sigma)+\mathbf{1}_+(k+j)b_{k,2^{2j}}(\sigma))\\ &\lesssim \mathscr{C}2^{-\sigma k}\varepsilon\langle 2^{k_0+k}\rangle^{-8}b_{k,2^{2k_0}}(\sigma).\end{aligned}\tag{7-45}$$

Together (7-36), (7-44), and (7-45) imply that

$$\|P_kA_m(s)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim 2^{-\sigma k}(1+s2^{2k})^{-4}b_{k,s}(\sigma)(1+\mathscr{C}\varepsilon),$$

from which it follows that $\mathscr{C}\lesssim 1+\mathscr{C}\varepsilon$ and hence $\mathscr{C}\lesssim 1$, proving the lemma. □

Lemma 7.8. *We have*

$$\|P_kA_l^2(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim\begin{cases}\varepsilon 2^{-\sigma k}2^{-j}a_{-j}a_{-j}(\sigma) & \textit{if } k+j\leq 0,\\ \varepsilon 2^{-\sigma k}2^{-j}b_{k,2^{2j}}(\sigma) & \textit{if } k+j\geq 0.\end{cases}$$

Proof. We apply Lemma 7.2 with $f=g=A_l$ and $\omega=0$ so that

$$\|P_k(A_l^2(r))\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim\sum_{l\leq k}2^l\alpha_k\beta_l+\sum_{l\geq k}2^l\alpha_l\beta_l,$$

where

$$\alpha_k\lesssim 2^{-\sigma k}\langle 2^{j+k}\rangle^{-8}b_{k,s}(\sigma),\quad \beta_k\lesssim\langle 2^{j+k}\rangle^{-8}b_{k,s}.$$

Case $k+j\leq 0$. We first consider the case $k+j\leq 0$ and proceed to control the high-low frequency interaction. We have

$$\begin{aligned}\sum_{l\leq k}2^l\alpha_k\beta_l &\lesssim 2^{-\sigma k}\sum_{l\leq k}2^lb_{k,2^{2j}}(\sigma)b_{l,2^{2j}}\lesssim 2^{-\sigma k}\sum_{p=k}^{-j}a_pa_p(\sigma)2^l\sum_{l\leq k}\sum_{q=l}^{-j}a_q^2\\ &\lesssim 2^{-\sigma k}a_{-j}a_{-j}(\sigma)\sum_{p=k}^{-j}2^{-2\delta(j+p)}\sum_{l\leq k}2^la_{-j}^2\sum_{q=l}^{-j}2^{-2\delta(j+q)}.\end{aligned}\tag{7-46}$$

It remains to show that

$$\sum_{p=k}^{-j}2^{-2\delta(j+p)}\sum_{l\leq k}2^la_{-j}^2\sum_{q=l}^{-j}2^{-2\delta(j+q)}\lesssim\varepsilon,\tag{7-47}$$

which follows from bounding a_{-j}^2 by ε and summing. To control the high-high interaction term we first split the sum as

$$\sum_{l\geq k}2^l\alpha_l\beta_l\lesssim\sum_{k\leq l<-j}2^l\alpha_l\beta_l+\sum_{l\geq -j}2^l\alpha_l\beta_l.\tag{7-48}$$

The first summand is controlled by

$$\sum_{k\le l<-j} 2^l\alpha_l\beta_l \lesssim 2^{-\sigma k}\sum_{k\le l<-j} 2^l b_{l,2^{2j}}(\sigma)b_{l,2^{2j}} \lesssim 2^{-\sigma k}2^{-j}\sum_{k\le l<-j} 2^{j+l}\sum_{p=l}^{-j} a_p a_p(\sigma)\sum_{q=l}^{-j} a_q^2.$$

Pulling out $a_{-j}^3 a_{-j}(\sigma)$ and summing implies

$$\sum_{k\le l<-j} 2^l\alpha_l\beta_l \lesssim \varepsilon 2^{-\sigma k}2^{-j}a_{-j}a_{-j}(\sigma). \tag{7-49}$$

The second summand is controlled by

$$\begin{aligned}\sum_{l\ge -j} 2^l\alpha_l\beta_l &\lesssim 2^{-\sigma k}\sum_{l\ge -j} 2^l\langle 2^{j+l}\rangle^{-8} b_{l,2^{2j}}(\sigma)b_{l,2^{2j}}\\ &\lesssim 2^{-\sigma k}\sum_{l\ge -j} 2^l\langle 2^{j+l}\rangle^{-8}2^{2(l+j)}a_{-j}^2 a_l a_l(\sigma) \lesssim \varepsilon 2^{-\sigma k}2^{-j}a_{-j}a_{-j}(\sigma).\end{aligned} \tag{7-50}$$

Combining (7-46)–(7-50), we conclude that

$$\|P_k A_l^2(r)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon 2^{-\sigma k}2^{-j}a_{-j}a_{-j}(\sigma) \quad \text{for } k+j\le 0. \tag{7-51}$$

Case $k+j\ge 0$. We now consider the case $k+j\ge 0$ and turn to the high-low frequency interaction, splitting it into two pieces:

$$\sum_{l\le k} 2^l\alpha_k\beta_l \le \sum_{l\le -j} 2^l\alpha_k\beta_l + \sum_{-j<l\le k} 2^l\alpha_k\beta_l. \tag{7-52}$$

The first summand is controlled by

$$\sum_{l\le -j} 2^l\alpha_k\beta_l \lesssim 2^{-\sigma k}2^{-j}b_{k,2^{2j}}(\sigma)\sum_{l\le -j} 2^{l+j}\langle 2^{j+k}\rangle^{-8} b_{l,2^{2j}}, \tag{7-53}$$

and so we need to show that

$$\sum_{l\le -j} 2^{l+j}\langle 2^{j+k}\rangle^{-8} b_{l,2^{2j}} \lesssim \varepsilon, \tag{7-54}$$

which follows from

$$\sum_{l\le -j} 2^{l+j}b_{l,2^{2j}} \lesssim \sum_{l\le -j} 2^{l+j}\sum_{p=l}^{-j} a_p^2 \lesssim a_{-j}^2\sum_{l\le -j} 2^{(1-2\delta)(l+j)} \lesssim \varepsilon.$$

The second summand in (7-52) is controlled by

$$\sum_{-j<l\le k} 2^l\alpha_k\beta_l \lesssim 2^{-\sigma k}2^{-j}b_{k,2^{2j}}(\sigma)\sum_{j<l\le k} \langle 2^{j+l}\rangle^{-8}2^{l+j}\langle 2^{j+k}\rangle^{-8}2^{l+j}a_{-j}a_l, \tag{7-55}$$

where we note that

$$\sum_{-j<l\le k} \langle 2^{j+l}\rangle^{-8}2^{2l+2j}a_{-j}a_l \lesssim a_{-j}^2\sum_{-j<l\le k} \langle 2^{j+l}\rangle^{-8}2^{(2+\delta)(l+j)} \lesssim \varepsilon. \tag{7-56}$$

We now turn to the high-high frequency interaction. We have

$$\begin{aligned}\sum_{l\geq k} 2^l\alpha_l\beta_l &\lesssim \sum_{l\geq k} 2^l 2^{-\sigma l}\langle 2^{j+l}\rangle^{-8}2^{2(l+j)}a_{-j}^2 a_l a_l(\sigma)\\ &\lesssim 2^{-\sigma k}2^{-j}2^{k+j}a_{-j}\sum_{l\geq k} 2^{l-k}2^{-\sigma(l-k)}\langle 2^{j+l}\rangle^{-8}2^{2(l+j)}a_{-j}a_l a_l(\sigma)\\ &\lesssim 2^{-\sigma k}2^{-j}b_{k,2^{2j}}(\sigma)\sum_{l\geq k}\langle 2^{j+l}\rangle^{-8}2^{(1+\delta)(l-k)}2^{2(l+j)}a_{-j}a_l.\end{aligned}\tag{7-57}$$

It remains to show that

$$\sum_{l\geq k}\langle 2^{j+l}\rangle^{-8}2^{(1+\delta)(l-k)}2^{2(l+j)}a_{-j}a_l \lesssim \varepsilon,\tag{7-58}$$

which follows from bounding $a_{-j}a_l$ by ε and summing.

Together (7-52)–(7-58) imply that

$$\|P_kA_l^2(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim \varepsilon 2^{-\sigma k}2^{-j}b_{k,2^{2j}}(\sigma)\quad \text{for } k+j\geq 0,$$

which combined with (7-51) implies the lemma. □

Set

$$c_{k,j}(\sigma)=\begin{cases}2^{-j}a_{-j}a_{-j}(\sigma) & \text{if } k+j\leq 0,\\ 2^{2k+j}a_{-j}a_k(\sigma) & \text{if } k+j\geq 0.\end{cases}\tag{7-59}$$

Lemma 7.9. *Let* $r\in[2^{2j-2},2^{2j+2}]$ *and let*

$$F\in\{A_l^2,\ \partial_lA_l,\ fg : l=1,2;\ f,g\in\{\psi_m,\overline{\psi_m} : m=1,2\}\}.$$

Then

$$\|P_kF(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim \langle 2^{j+k}\rangle^{-8}2^{-\sigma k}c_{k,j}(\sigma).\tag{7-60}$$

Proof. If $F=A_l^2$, then (7-60) is an immediate consequence of Lemma 7.8 when $k+j\leq 0$. If $k+j\geq 0$, then Lemma 7.8 implies

$$\|P_kA_l^2(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim \varepsilon 2^{-\sigma k}2^{-j}2^{k+j}a_{-j}a_{-j}(\sigma),$$

and multiplying the right-hand side by 2^{k+j} yields the desired estimate.

Consider now the case where $F=\partial_lA_l$. By Lemma 7.7, we have

$$\|P_k(\partial_lA_l)(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim 2^k\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}b_{k,2^{2j}}(\sigma).\tag{7-61}$$

When $k+j\geq 0$, we rewrite (7-61) as

$$\|P_k(\partial_lA_l)(r)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim \langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^k2^{k+j}a_{-j}a_k(\sigma),$$

which is the desired bound (7-60). If $k+j\le 0$, then (7-61) becomes

$$\|P_k(\partial_l A_l)(r)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^k \sum_{p=k}^{-j} a_p a_p(\sigma)$$
$$\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{-j} a_{-j} a_{-j}(\sigma) = \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} c_{k,j}(\sigma).$$

If $F = fg$, fg as in the statement of the lemma, then (7-60) follows directly from (7-6) when $k+j\le 0$. If $k+j\ge 0$, then to get (7-60) we multiply the right-hand side of (7-6) by 2^{2j+2k}. □

Set

$$d_{k,j} := \varepsilon \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k} (a_k(\sigma) + 2^{-3(k+j)/2} a_{-j}(\sigma)). \tag{7-62}$$

Lemma 7.10. *We have*

$$\|P_k U_m(r)\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k} (a_k(\sigma) + 2^{-3(k+j)/2} a_{-j}(\sigma)) =: d_{k,j}.$$

Proof. Using now (2-21) instead of (2-22), i.e., taking now

$$U_\alpha = i A_l \partial_l \psi_\alpha + i\partial_l(A_l\psi_\alpha) - A_x^2 \psi_\alpha + i\,\mathrm{Im}(\psi_\alpha \psi_l)\psi_l,$$

we have that it suffices to prove that

$$\|P_k(F(r)f(r))\|_{F_k(T)\cap S_k^{1/2}(T)} + 2^k \|P_k(A_l(r)f(r))\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim d_{k,j},$$

where

$$F \in \{A_l^2, \partial_l A_l, gh : l = 1,2;\ f,h \in \{\psi_m, \overline{\psi_m} : m = 1,2\}\}$$

and $f \in \{\psi_m, \overline{\psi_m} : m = 1,2\}$. We consider the terms $P_k(Ff)$, and $P_k(Af)$ separately.

Controlling $P_k(Ff)$. We apply Lemma 7.2 to $P_k(Ff)$, handling the different frequency interactions separately and according to cases. We record a consequence of (7-60):

$$\alpha_k \lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} c_{k,j}(\sigma),$$

Let us begin by assuming $k+j\le 0$. For the low-high frequency interaction, we have

$$\sum_{l\le k} 2^l \alpha_l \beta_k \lesssim 2^{-\sigma k} a_k(\sigma) \sum_{l\le k} 2^l c_{l,j} \lesssim 2^{-\sigma k} a_k(\sigma) \sum_{l\le k} 2^{l-j} a_{-j}^2 \lesssim \varepsilon 2^{-\sigma k} 2^{k-j} 2^{-\delta(k+j)} a_{-j}(\sigma). \tag{7-63}$$

In a similar manner we control the high-low frequency interaction by

$$\sum_{l\le k} 2^l \alpha_k \beta_l \lesssim 2^{-\sigma k} c_{k,j}(\sigma) \sum_{l\le k} 2^l a_l \lesssim 2^{-\sigma k} 2^{-j} a_{-j} a_{-j}(\sigma) \sum_{l\le k} 2^l a_l \lesssim \varepsilon 2^{-\sigma k} 2^{k-j} a_{-j}(\sigma). \tag{7-64}$$

The high-high frequency interaction we split into two sums:

$$2^k \sum_{l\ge k} 2^{(l-k)/2} \alpha_l \beta_l \lesssim 2^k \sum_{k\le l<-j} 2^{(l-k)/2} \alpha_l \beta_l + 2^k \sum_{l\ge -j} 2^{(l-k)/2} \alpha_l \beta_l. \tag{7-65}$$

We control the first summand using the definition (7-59) of $c_{k,j}(\sigma)$, the frequency envelope properties (2-29), (2-30), and energy dispersion:

$$\begin{aligned}2^k \sum_{k\le l<-j} 2^{(l-k)/2}\alpha_l\beta_l &\lesssim 2^k \sum_{k\le l<-j} 2^{(l-k)/2}2^{-\sigma l}c_{l,j}(\sigma)a_l\\ &\lesssim 2^{-\sigma k}2^{k-j}a_{-j}(\sigma)a_{-j}\sum_{k\le l<-j}2^{(l-k)/2}a_l\\ &\lesssim 2^{-\sigma k}2^{k-j}2^{-(k+j)/2}a_{-j}(\sigma)a_{-j}\sum_{k\le l<-j}2^{(l+j)/2}a_l\\ &\lesssim \varepsilon 2^{-\sigma k}2^{k-j}2^{-(k+j)/2}a_{-j}(\sigma).\end{aligned}\tag{7-66}$$

In like manner we control the second summand:

$$\begin{aligned}2^k \sum_{l\ge -j} 2^{(l-k)/2}\alpha_l\beta_l &\lesssim 2^k \sum_{l\ge -j}\langle 2^{j+l}\rangle^{-8}2^{(l-k)/2}2^{-\sigma l}c_{l,j}(\sigma)a_l\\ &\lesssim 2^k \sum_{l\ge -j}\langle 2^{j+l}\rangle^{-8}2^{(l-k)/2}2^{-\sigma l}2^{2l+j}a_{-j}a_l(\sigma)a_l\\ &\lesssim \varepsilon 2^{-\sigma k}2^{k-j}2^{-(k+j)/2}a_{-j}(\sigma).\end{aligned}\tag{7-67}$$

Combining (7-63)–(7-67), we conclude that

$$\|P_k(F(r)f(r))\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon 2^{-\sigma k}2^{k-j}2^{-(k+j)/2}a_{-j}(\sigma),\quad k+j\le 0.\tag{7-68}$$

We now turn to the case $k+j\ge 0$. In the low-high frequency interaction case, we have

$$\begin{aligned}\sum_{l\le k}2^l\alpha_l\beta_k &\lesssim \langle 2^{j+k}\rangle^{-8}2^{-\sigma k}a_k(\sigma)\sum_{l\le k}\langle 2^{j+l}\rangle^{-8}2^l c_{l,j}\\ &\lesssim \langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{2k}a_k(\sigma)\Bigg(\sum_{l\le -j}2^{l-2k}2^{-j}a_{-j}^2+\sum_{-j<l\le k}\langle 2^{j+l}\rangle^{-8}2^{l-2k}2^{l+j}a_{-j}a_l\Bigg).\end{aligned}\tag{7-69}$$

To estimate the first term we use

$$a_{-j}^2\sum_{l\le -j}2^{l-k}2^{-j-k}\lesssim \varepsilon 2^{-(j+k)}\cdot 2^{-(j+k)}\le\varepsilon,\tag{7-70}$$

and for the second

$$\begin{aligned}a_{-j}\sum_{-j<l\le k}\langle 2^{j+l}\rangle^{-8}2^{3l+j-2k}a_l &= a_{-j}\sum_{-j<l\le k}\langle 2^{j+l}\rangle^{-8}2^{l+j}2^{2l-2k}a_l\\ &\lesssim a_{-j}a_k\sum_{-j<l\le k}\langle 2^{j+l}\rangle^{-8}2^{l+j}2^{(2-\delta)(l-k)}\lesssim\varepsilon.\end{aligned}\tag{7-71}$$

In the high-low frequency interaction case, we have

$$\begin{aligned}\sum_{l\le k} 2^l \alpha_k \beta_l &\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} c_{k,j}(\sigma) \sum_{l\le k} \langle 2^{j+l}\rangle^{-8} 2^l a_l \\ &\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k+j} a_{-j} a_k(\sigma) \sum_{l\le k} \langle 2^{j+l}\rangle^{-8} 2^l a_l \\ &\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k} a_k(\sigma) a_{-j}^2. \end{aligned} \tag{7-72}$$

In the high-high frequency interaction case we have

$$\begin{aligned}\sum_{l\ge k} 2^l \alpha_l \beta_l &\lesssim \sum_{l\ge k} \langle 2^{j+l}\rangle^{-8} 2^l 2^{-\sigma l} a_l(\sigma) c_{l,j} \\ &\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} \sum_{l\ge k} \langle 2^{j+l}\rangle^{-8} 2^l a_l(\sigma) 2^{2l+j} a_{-j} a_l \\ &\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k} a_k(\sigma) a_{-j}^2. \end{aligned} \tag{7-73}$$

From (7-69)–(7-73) we conclude that

$$\| P_k(F(r) f(r)) \|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k} a_k(\sigma), \quad k+j \ge 0. \tag{7-74}$$

Controlling $2^k P_k(Af)$. We now apply Lemma 7.2 to $P_k(A_l f)$. Note that

$$\alpha_k \lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} b_{k,r}(\sigma)$$

because of Lemma 7.7, and that

$$\beta_k \lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} a_k(\sigma).$$

We begin by assuming $k+j\le 0$. The low-high frequency interaction is controlled by

$$\begin{aligned}\sum_{l\le k} 2^l \alpha_l \beta_k &\lesssim 2^{-\sigma k} a_k(\sigma) \sum_{l\le k} 2^l \sum_{p=l}^{-j} a_l^2 \\ &\lesssim 2^{-\sigma k} 2^{-\delta(k+j)} a_{-j}^2 a_{-j}(\sigma) \sum_{l\le k} 2^l \sum_{p=l}^{-j} 2^{-2(j+p)}. \end{aligned}$$

Summing yields

$$2^k \sum_{l\le k} 2^l \alpha_l \beta_k \lesssim 2^{2k} 2^{-\sigma k} 2^{-(k+j)/2} a_{-j}^2 a_{-j}(\sigma). \tag{7-75}$$

Control over the high-low frequency interaction follows from

$$\begin{aligned}\sum_{l\le k} 2^l \alpha_k \beta_l &\lesssim 2^{-\sigma k} \sum_{p=k}^{-j} a_p a_p(\sigma) \sum_{l\le k} 2^l a_l \\ &\lesssim 2^k 2^{-\sigma k} 2^{-2\delta(k+j)} a_{-j} a_k a_{-j}(\sigma). \end{aligned} \tag{7-76}$$

We now turn to the high-high frequency interaction. We begin by splitting the sum:

$$2^k \sum_{l \geq k} 2^{(l-k)/2} \alpha_l \beta_l \lesssim 2^k \sum_{k \leq l < -j} 2^{(l-k)/2} \alpha_l \beta_l + 2^k \sum_{l \geq -j} 2^{(l-k)/2} \alpha_l \beta_l. \tag{7-77}$$

Then

$$\begin{aligned} 2^k \sum_{k \leq l < -j} 2^{(l-k)/2} \alpha_l \beta_l &\lesssim 2^k 2^{-\sigma k} a_{-j}(\sigma) \sum_{k \leq l < -j} 2^{(l-k)/2} 2^{-\delta(j+l)} \sum_{p=l}^{-j} a_p^2 \\ &\lesssim 2^k 2^{-\sigma k} 2^{-(k+j)/2} a_{-j}^2 a_{-j}(\sigma). \end{aligned} \tag{7-78}$$

As for the second summand, we have

$$\begin{aligned} 2^k \sum_{l \geq -j} 2^{(l-k)/2} \alpha_l \beta_l &\lesssim 2^k \sum_{l \geq -j} \langle 2^{j+l} \rangle^{-8} 2^{(l-k)/2} 2^{l+j} a_{-j} a_l(\sigma) 2^{-\sigma l} a_l \\ &\lesssim 2^k 2^{-\sigma k} 2^{-(k+j)/2} a_{-j}^2 a_{-j}(\sigma). \end{aligned} \tag{7-79}$$

Combining (7-75)– (7-79) yields

$$2^k \| P_k(A_l(r) f(r)) \|_{F_k(T) \cap S_k^{1/2}(T)} \lesssim \varepsilon 2^{2k} 2^{-\sigma k} 2^{-(k+j)/2} a_{-j}(\sigma), \quad k + j \leq 0. \tag{7-80}$$

Now let us assume that $k + j \geq 0$. The low-high frequency interaction we first split into two pieces:

$$\sum_{l \leq k} 2^l \alpha_l \beta_k \lesssim \sum_{l \leq -j} 2^l \alpha_l \beta_k + \sum_{-j < l \leq k} 2^l \alpha_l \beta_k. \tag{7-81}$$

For the first term, we have

$$\begin{aligned} \sum_{l \leq -j} 2^l \alpha_l \beta_k &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} a_k(\sigma) \sum_{l \leq -j} \sum_{p=l}^{-j} a_p^2 \\ &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} a_{-j}^2 a_k(\sigma) \sum_{l \leq -j} 2^l \sum_{p=l}^{-j} 2^{-2\delta(j+p)}. \end{aligned} \tag{7-82}$$

Then

$$\sum_{l \leq -j} 2^l \sum_{p=l}^{-j} 2^{-2\delta(j+p)} \lesssim \sum_{l \leq -j} 2^l 2^{-2\delta(j+l)} \lesssim 2^{-j} \leq 2^k. \tag{7-83}$$

As for the second summand,

$$\begin{aligned} \sum_{-j < l \leq k} 2^l \alpha_l \beta_k &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} a_k(\sigma) \sum_{-j < l \leq k} \langle 2^{j+l} \rangle^{-8} 2^l 2^{l+j} a_{-j} a_l \\ &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^k a_{-j}^2 a_k(\sigma). \end{aligned} \tag{7-84}$$

The high-low frequency interaction is controlled by

$$\begin{aligned} \sum_{l \leq k} 2^l \alpha_k \beta_l &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^{k+j} a_{-j} a_k(\sigma) \sum_{l \leq k} \langle 2^{j+l} \rangle^{-8} 2^l a_l \\ &\lesssim \langle 2^{j+k} \rangle^{-8} 2^{-\sigma k} 2^k a_{-j}^2 a_k(\sigma). \end{aligned} \tag{7-85}$$

Finally, the high-high frequency interaction is controlled by

$$\begin{aligned}\sum_{l\geq k} 2^l \alpha_l \beta_l &\lesssim \sum_{l\geq k} \langle 2^{j+l}\rangle^{-8} 2^l 2^{l+j} a_{-j} a_l 2^{-\sigma l} a_l(\sigma) \\ &\lesssim \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^k a_{-j} a_k a_k(\sigma). \end{aligned} \tag{7-86}$$

Thus, in view of (7-81)–(7-86), we have shown that

$$2^k \| P_k(A_l(r) f(r)) \|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon \langle 2^{j+k}\rangle^{-8} 2^{-\sigma k} 2^{2k} a_k(\sigma), \quad k+j\geq 0. \tag{7-87}$$

Combining (7-68), (7-74), (7-80), and (7-87) proves the lemma. □

Lemma 7.11. *We have*

$$\left\| \int_0^s e^{(s-s')\Delta} P_k U_m(s')\, ds' \right\|_{F_k(T)\cap S_k^{1/2}(T)} \lesssim \varepsilon (1+s2^{2k})^{-4} 2^{-\sigma k} a_k(\sigma).$$

Proof. Let $k_0 \in \mathbb{Z}$ be such that $s\in[2^{2k_0-1}, 2^{2k_0+1})$. If $k+k_0\leq 0$, then it follows from Lemma 7.10 that

$$\begin{aligned}\left\| \int_0^s e^{(s-r)\Delta} P_k U_m(r)\, dr \right\|_{F_k(T)\cap S_k^{1/2}(T)} &\lesssim \sum_{j\leq k_0} \int_{2^{2j-1}}^{2^{2j+1}} \| P_k U_m(r) \|_{F_k(T)\cap S_k^{1/2}(T)}\, dr \\ &\lesssim \sum_{j\leq k_0} 2^{2j} \varepsilon 2^{-\sigma k} 2^{2k} (a_k(\sigma) + 2^{-3(k+j)/2} a_{-j}(\sigma)) \\ &\lesssim \varepsilon 2^{-\sigma k} a_k(\sigma) \sum_{j\leq k_0} 2^{2k+2j} (1 + 2^{-3(k+j)/2} 2^{-\delta(k+j)}) \\ &\lesssim \varepsilon 2^{-\sigma k} a_k(\sigma). \end{aligned}$$

On the other hand, if $k+k_0>0$, then

$$\begin{aligned}\left\| \int_0^s e^{(s-r)\Delta} P_k U_m(r)\, dr \right\|_{F_k(T)\cap S_k^{1/2}(T)} &\lesssim \int_0^{s/2} \| e^{(s-r)\Delta} P_k U_m(r) \|_{F_k(T)\cap S_k^{1/2}(T)}\, dr \\ &\qquad + \int_{s/2}^s \| e^{(s-r)\Delta} P_k U_m(r) \|_{F_k(T)\cap S_k^{1/2}(T)}\, dr \\ &\lesssim \sum_{j\leq k_0} 2^{-20(k+k_0)} 2^{2j} d_{k,j} + 2^{2k_0} d_{k,k_0} \\ &\lesssim 2^{-20(k_0+k)} \sum_{j\leq k_0} 2^{2j} d_{k,j} + 2^{-2k} d_{k,k_0}. \end{aligned} \tag{7-88}$$

By Lemma 7.10 and the fact that $k+k_0>0$, we have

$$2^{-2k} d_{k,k_0} \lesssim \varepsilon \langle 2^{k_0+k}\rangle^{-8} 2^{-\sigma k} a_k(\sigma)$$

and

$$\begin{aligned}2^{-20(k_0+k)}\sum_{j\le k_0}2^{2j}d_{k,j}&\lesssim 2^{-20(k_0+k)}\sum_{j\le k_0}\varepsilon\langle 2^{j+k}\rangle^{-8}2^{-\sigma k}2^{2k}\left(2^{2j}a_k(\sigma)+2^{j/2}2^{-3k/2}a_{-j}(\sigma)\right)\\&\lesssim \varepsilon 2^{-\sigma k}a_k(\sigma)2^{-20(k_0+k)}\sum_{j\le k_0}\langle 2^{j+k}\rangle^{-8}\left(2^{2j+2k}+2^{(j+k)/2}2^{\delta|j+k|}\right)\\&\lesssim \varepsilon\langle 2^{k_0+k}\rangle^{-8}2^{-\sigma k}a_k(\sigma),\end{aligned}$$

which, combined with (7-88), completes the proof of the lemma. ☐

Lemma 7.12. *The following bound from* (4-12) *holds*:

$$\|P_k\psi_m(s)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim (1+s2^{2k})^{-4}2^{-\sigma k}b_k(\sigma).$$

Proof. In view of (7-1), we have

$$P_k\psi_m(s)=e^{s\Delta}P_k\psi_m(0)+\int_0^s e^{(s-r)\Delta}P_kU_m(r)\,dr.$$

Then it follows from Lemma 7.11 that

$$\|P_k\psi_m(s)\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim 2^{-\sigma k}(1+s2^{2k})^{-4}(b_k(\sigma)+\varepsilon a_k(\sigma)),\quad 0\le\sigma\le\sigma_1-1.$$

Therefore $a_k(\sigma)\lesssim b_k(\sigma)+\varepsilon a_k(\sigma)$ and hence

$$a_k(\sigma)\lesssim b_k(\sigma), \tag{7-89}$$

as required. ☐

7B. ***Connection coefficient control.*** The main results of this subsection are the $L^2_{t,x}$ bounds (4-14) and (4-16), respectively proven in Corollary 7.19 and Lemma 7.21, and the frequency-localized $L^2_{t,x}$ bounds (4-15) and (4-17), respectively proven in Corollaries 7.20 and 7.22.

Lemma 7.13. *Let* $s\in[2^{2j-2},2^{2j+2}]$. *Then*

$$\|P_k(A_l(s)\psi_m(s))\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim \varepsilon(1+s2^{2k})^{-3}(s2^{2k})^{-3/8}2^k2^{-\sigma k}b_k(\sigma).$$

Proof. Using (7-80) and (2-29), we have

$$2^k\|P_k(A_l(s)\psi_m(s))\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim \varepsilon 2^{2k}2^{-\sigma k}2^{-(1/2+\delta)(k+j)}a_k(\sigma). \tag{7-90}$$

Combining (7-90), (7-87), and (7-89) then yields

$$\|P_k(A_l(s)\psi_m(s))\|_{F_k(T)\cap S_k^{1/2}(T)}\lesssim\begin{cases}\varepsilon(s2^{2k})^{-3/8}2^k2^{-\sigma k}b_k(\sigma) & \text{if } k+j\le 0,\\ \varepsilon(1+s2^{2k})^{-4}2^k2^{-\sigma k}b_k(\sigma) & \text{if } k+j\ge 0,\end{cases}$$

which proves the lemma. ☐

Lemma 7.14 [Bejenaru et al. 2011c, §5]. *Assume that $T \in (0, 2^{2\mathcal{K}}]$, $f, g \in H^{\infty,\infty}(T)$, $P_k f \in S_k^\omega(T)$, and $P_k g \in L^4_{t,x}$ for some $\omega \in [0, 1/2]$ and all $k \in \mathbb{Z}$. Set*

$$\mu_k := \sum_{|j-k|\le 20} \|P_j f\|_{S_{k'}^\omega(T)}, \qquad \nu_k := \sum_{|j-k|\le 20} \|P_j g\|_{L^4_{t,x}}.$$

Then, for any $k \in \mathbb{Z}$,

$$\|P_k(fg)\|_{L^4_{t,x}} \lesssim \sum_{j\le k} 2^j \mu_j \nu_k + \sum_{j \le k} 2^{(k+j)/2} \mu_k \nu_j + 2^k \sum_{j\ge k} 2^{-\omega(j-k)} \mu_j \nu_j.$$

Lemma 7.15. *We have*

$$\|P_k \psi_s(0)\|_{L^4_{t,x}} + \|P_k \psi_t(0)\|_{L^4_{t,x}} \lesssim 2^k \tilde{b}_k \Big(1 + \sum_j b_j^2\Big).$$

Proof. We only treat $\psi_t(0)$ since $\psi_s(0)$ and $\psi_t(0)$ differ only by a factor of i. As $\psi_t(0) = i D_l(0)\psi_l(0)$, we have

$$\psi_t(0) = i\partial_l \psi_l(0) - A_l(0)\psi_l(0).$$

Clearly

$$\|P_k \partial_l \psi_l(0)\|_{L^4_{t,x}} \lesssim 2^k \|P_k \psi_x(0)\|_{L^4_{t,x}} \lesssim 2^k \tilde{b}_k.$$

For the remaining term, we apply Lemma 7.14, bounding $P_j A_l(0)$ in $S_j^{1/2}$ by $\sum_p b_p^2$, which follows from Lemma 7.7. We get

$$\|P_k(A_l(0)\psi_l(0))\|_{L^4_{t,x}} \lesssim \sum_{j\le k} 2^j \Big(\sum_p b_p^2\Big)\tilde{b}_k + \sum_{j\le k} 2^{(k+j)/2}\Big(\sum_p b_p^2\Big)\tilde{b}_j + 2^k \sum_{j \ge k} 2^{-(j-k)/2}\Big(\sum_p b_p^2\Big)\tilde{b}_j.$$

Therefore

$$\|P_k(A_l \psi_l(0))\|_{L^4_{t,x}} \lesssim 2^k \tilde{b}_k \Big(\sum_j b_j^2\Big). \qquad \square$$

Corollary 7.16. *We have*

$$\|P_k \psi_s(0)\|_{L^4_{t,x}} + \|P_k \psi_t(0)\|_{L^4_{t,x}} \lesssim 2^k 2^{-\sigma k} b_k(\sigma)\Big(1 + \sum_j b_j^2\Big).$$

Proof. Without loss of generality, we prove the bound only for ψ_t. We have

$$\|P_k \partial_l \psi_l(0)\|_{L^4_{t,x}} \lesssim 2^k \|P_k \psi_x(0)\|_{L^4_{t,x}} \lesssim 2^k 2^{-\sigma k} b_k(\sigma).$$

It remains to control $P_k(A_l(0)\psi_l(0))$ in $L^4_{t,x}$. The obstruction to applying Lemma 7.14 as we did in Lemma 7.15 is the high-low interaction, for which summation can be achieved only for small σ. If we restrict the range of σ to $\sigma < 1/2 - 2\delta$, then we ensure the constant remains bounded and can apply Lemma 7.14 as in Lemma 7.15.

For $\sigma \ge 1/2 - 2\delta$, we can still apply the bounds of Lemma 7.14 to the low-high and high-high interactions. For the remaining high-low interaction, we bound $A_l(0)$ in $L^4_{t,x}$ and $\psi_l(0)$ in $L^\infty_{t,x}$. In particular, we have, thanks to (7-95) and Bernstein, that

$$\begin{aligned}\sum_{\substack{|j_1-k|\le 4\\ j_2\le k+4}} \|P_k(P_{j_1}A_l(0)P_{j_2}\psi_l(0))\|_{L^4_{t,x}} &\lesssim \sum_{\substack{|j_1-k|\le 4\\ j_2\le k+4}} \|P_{j_1}A_l(0)\|_{L^4_{t,x}}\|P_{j_2}\psi_l(0)\|_{L^\infty_{t,x}}\\ &\lesssim \sum_{\substack{|j_1-k|\le 4\\ j_2\le k+4}} 2^{-\sigma j_1}b_{j_1}b_{j_1}(\sigma)2^{j_2}\|P_{j_2}\psi_l(0)\|_{L^\infty_t L^2_x}\\ &\lesssim \sum_{j_2\le k+4} 2^{-\sigma k}b_k b_k(\sigma)2^{j_2}b_{j_2}\\ &\lesssim 2^{-\sigma k}b_k^2 b_k(\sigma)\sum_{j_2\le k+4} 2^k 2^{(j_2-k)+(k-j_2)\delta} \lesssim 2^{-\sigma k}2^k b_k^2 b_k(\sigma). \qquad \square\end{aligned}$$

Lemma 7.17. *We have*

$$\|P_k\psi_s(s)\|_{L^4_{t,x}} + \|P_k\psi_t(s)\|_{L^4_{t,x}} \lesssim (1+s2^{2k})^{-2}2^k\tilde{b}_k\Big(1+\sum_j b_j^2\Big).$$

Proof. We treat only $\psi_t(s)$ since the proof for $\psi_s(s)$ is analogous. From (7-1) we have

$$\psi_t(s) = e^{s\Delta}\psi_t(0) + \int_0^s e^{(s-r)\Delta}U_t(r)\,dr.$$

We claim that

$$\Big\|\int_0^s e^{(s-r)\Delta}P_kU_t(r)\,dr\Big\|_{L^4_{t,x}} \lesssim \varepsilon(1+s2^{2k})^{-2}2^k\tilde{b}_k\Big(1+\sum_j b_j^2\Big), \tag{7-91}$$

which combined with Lemma 7.15 and a standard iteration argument proves the lemma.

As in the proof of Lemma 7.11, we take

$$F \in \{A_l^2,\ \partial_l A_l,\ fg : l=1,2;\ f,g\in\{\psi_m,\overline{\psi_m} : m=1,2\}\}.$$

By (7-60) and (7-89) we have

$$\|P_kF(r)\|_{S_k^{1/2}(T)} \lesssim \varepsilon^{1/2}(1+s2^{2k})^{-2}(s2^{2k})^{-5/8}2^k b_k. \tag{7-92}$$

Moreover, by Lemma 7.7,

$$\|P_kA_l(r)\|_{S_k^{1/2}(T)} \lesssim \varepsilon^{1/2}(1+s2^{2k})^{-3}(s2^{2k})^{-1/8}b_k. \tag{7-93}$$

Applying Lemma 7.14 with $\omega = 1/2$ yields

$$\|P_k(F(r)\psi_t(r))\|_{L^4_{t,x}} + 2^k\|P_k(A_l(r)\psi_t(r))\|_{L^4_{t,x}} \lesssim \varepsilon(1+s2^{2k})^{-2}(s2^{2k})^{-7/8}2^k\tilde{b}_k\Big(1+\sum_j b_j^2\Big). \tag{7-94}$$

Integrating with respect to s yields

$$\int_0^s (1+(s-r)2^{2k})^{-N}(1+r2^{2k})^{-2}(r2^{2k})^{-7/8}\,dr \lesssim 2^{-2k}(1+s2^{2k})^{-2},$$

which, together with (7-94), implies (7-91). $\square$

Lemma 7.18. *We have*

$$\|P_k A_m(0)\|_{L^4_{t,x}} \lesssim 2^{-\sigma k} b_k b_k(\sigma). \tag{7-95}$$

Proof. We have

$$\|P_k \psi_m(s)\|_{S^0_k} \lesssim (1+s2^{2k})^{-4} 2^{-\sigma k} b_k(\sigma)$$

and

$$\|P_k(D_l\psi_l)(s)\|_{L^4_{t,x}} \lesssim (1+s2^{2k})^{-3}(s2^{2k})^{-3/8} 2^k 2^{-\sigma k} b_k(\sigma).$$

Applying Lemma 7.14 with $\omega = 0$, we get

$$\begin{aligned}
\|P_k A_m(0)\|_{L^4_{t,x}} &\lesssim \sum_{l=1,2} \int_0^\infty \|P_k(\overline{\psi_m(s)} D_l \psi_l(s))\|_{L^4_{t,x}}\, ds \\
&\lesssim 2^{-\sigma k} \sum_{j\le k} b_j b_k(\sigma) 2^{j+k} \int_0^\infty (1+s2^{2k})^{-3}(s2^{2k})^{-3/8}\, ds \\
&\qquad + 2^{-\sigma k} \sum_{j\le k} b_k(\sigma) b_j 2^{(k+j)/2} 2^j \int_0^\infty (1+s2^{2k})^{-4}(s2^{2j})^{-3/8}\, ds \\
&\qquad + \sum_{j\ge k} 2^{-\sigma j} b_j(\sigma) b_j 2^{k-j} 2^{2j} \int_0^\infty (1+s2^{2j})^{-7}(s2^{2j})^{-3/8}\, ds.
\end{aligned}$$

Call the integrals I_1, I_2, and I_3, respectively. Clearly I_1 and I_3 satisfy $I_1 \lesssim 2^{-2k}$ and $I_3 \lesssim 2^{-2j}$. By Cauchy–Schwarz, I_2 satisfies

$$I_2 \lesssim \left(\int_0^\infty (1+s2^{2k})^{-8}(1+s2^{2j})^4\, ds\right)^{1/2} \left(\int_0^\infty (1+s2^{2j})^{-4}(s2^{2j})^{-3/8}\, ds\right)^{1/2} \lesssim 2^{-j-k}.$$

Therefore

$$\|P_k A_m(0)\|_{L^4_{t,x}} \lesssim 2^{-\sigma k} b_k(\sigma) \sum_{j\le k} \bigl(b_j 2^{j-k} + b_j 2^{(j-k)/2}\bigr) + 2^{-\sigma k} \sum_{j\ge k} b_j(\sigma) b_j 2^{k-j} \lesssim 2^{-\sigma k} b_k b_k(\sigma). \qquad \square$$

Corollary 7.19. *We have*

$$\|A_x^2(0)\|_{L^2_{t,x}} \lesssim \sup_{j\in\mathbb{Z}} b_j^2 \cdot \sum_{k\in\mathbb{Z}} b_k^2.$$

Proof.

$$\|A_x^2(0)\|_{L^2_{t,x}} \lesssim \|A_x(0)\|^2_{L^4_{t,x}} \lesssim \sum_{k\in\mathbb{Z}} \|P_k A_x(0)\|^2_{L^4_{t,x}} \lesssim \sup_{j\in\mathbb{Z}} b_j^2 \cdot \sum_{k\in\mathbb{Z}} b_k^2. \qquad \square$$

Corollary 7.20. *Let $\sigma \ge 2\delta$. Then*

$$\|P_k A_x^2(0)\|_{L^2_{t,x}} \lesssim 2^{-\sigma k} b_k(\sigma) \cdot \sup_j b_j \cdot \sum_{l\in\mathbb{Z}} b_l^2.$$

Proof. We perform a Littlewood–Paley decomposition and invoke Corollary 7.19.

Consider first the high-low interactions:

$$\sum_{\substack{|j_2-k|\le 4\\ j_1\le k-5}} \|P_k(P_{j_1}A_x P_{j_2}A_x)\|_{L^2} \lesssim \sum_{\substack{|j_2-k|\le 4\\ j_1\le k-5}} \|P_{j_1}A_x\|_{L^4} \|P_{j_2}A_x\|_{L^4} \lesssim 2^{-\sigma k} b_k b_k(\sigma) \sum_{j_1\le k-5} b_{j_1}^2.$$

Next consider the high-high interactions:

$$\sum_{\substack{j_1,j_2\geq k-4\\|j_1-j_2|\leq 8}} \|P_k(P_{j_1}A_xP_{j_2}A_x)\|_{L^2} \lesssim \sum_{\substack{j_1,j_2\geq k-4\\|j_1-j_2|\leq 8}} \|P_{j_1}A_x\|_{L^4}\|P_{j_2}A_x\|_{L^4} \lesssim \sum_{j\geq k-4} 2^{-\sigma j} b_j(\sigma) b_j^3.$$

Using the frequency envelope property, we bound this last sum by

$$\sum_{j\geq k-4} 2^{-\sigma j} b_j(\sigma) b_j^3 \lesssim 2^{-\sigma k} b_k(\sigma) \sum_{j\geq k-4} 2^{-\sigma(j-k)} 2^{\delta(j-k)} b_j^3 \lesssim 2^{-\sigma k} b_k(\sigma) \sup_{j\geq k-4} b_j \cdot \sum_{j\geq k-4} b_j^2.$$

It is in controlling this last sum that we use $\sigma > \delta+$. □

Lemma 7.21. *We have*

$$\|A_t(0)\|_{L^2_{t,x}} \lesssim \Big(1+\sum_j b_j^2\Big)^2 \sum_k \|P_k\psi_x(0)\|^2_{L^4_{t,x}}.$$

Proof. We begin with

$$\|A_t(0)\|_{L^2_{t,x}} \lesssim \int_0^\infty \|(\overline{\psi_t}\cdot D_l\psi_l)(s)\|_{L^2_{t,x}}\, ds. \tag{7-96}$$

If we define

$$\mu_k(s) := \sup_{k'\in\mathbb{Z}} 2^{-\delta|k-k'|}\|P_k\psi_t(s)\|_{L^4_{t,x}} \quad\text{and}\quad \nu_k(s) := \sup_{k'\in\mathbb{Z}} 2^{-\delta|k-k'|}\|P_k(D_l\psi_l)(s)\|_{L^4_{t,x}}, \tag{7-97}$$

then

$$\|(\overline{\psi_t}\cdot D_l\psi_l)(s)\|_{L^2_{t,x}} \lesssim \sum_k \mu_k(s)\sum_{j\leq k}\nu_j(s) + \sum_k \nu_k(s)\sum_{j\leq k}\mu_j(s). \tag{7-98}$$

From Lemmas 7.15, 7.12, and 7.13, it follows that

$$\mu_k(s), \nu_k(s) \lesssim (1+s2^{2k})^{-2} 2^k \tilde{b}_k \Big(1+\sum_p b_p^2\Big). \tag{7-99}$$

Combining (7-96), (7-98), and (7-99), we have

$$\begin{aligned}
\|A_t(0)\|_{L^2_{t,x}} &\lesssim \sum_k \mu_k(s)\sum_{j\leq k}\nu_j(s)\\
&\lesssim \Big(1+\sum_p b_p^2\Big)^2 \sum_k 2^k\tilde{b}_k \sum_{j\leq k} 2^j\tilde{b}_j \int_0^\infty (1+s2^{2j})^{-2}(1+s2^{2k})^{-2}\, ds\\
&\lesssim \Big(1+\sum_p b_p^2\Big)^2 \sum_k 2^k\tilde{b}_k \sum_{j\leq k} 2^j\tilde{b}_j \int_0^\infty (1+s2^{2k})^{-2}\, ds\\
&\lesssim \Big(1+\sum_p b_p^2\Big)^2 \sum_k 2^{2k}\tilde{b}_k^2 \int_0^\infty (1+s2^{2k})^{-2}\, ds\\
&\lesssim \Big(1+\sum_p b_p^2\Big)^2 \sum_k \tilde{b}_k^2.
\end{aligned}$$

□

As a corollary of the proof, we also obtain this:

Corollary 7.22. *Let $\sigma \geq 2\delta$. Then*

$$\|P_k A_t\|_{L^2_{t,x}} \lesssim \Big(1+\sum_p b_p^2\Big)\tilde{b}_k 2^{-\sigma k} b_k(\sigma).$$

Proof. We start by modifying the proof of Lemma 7.21, taking μ_k and ν_k as in (7-97). Then

$$\begin{aligned}\|P_k A_t\|_{L^2} &\lesssim \int_0^\infty \|P_k(\overline{\psi_t}\cdot D_l\psi_l)(s)\|_{L^2_{t,x}}\,ds\\ &\lesssim \int_0^\infty \Big(\mu_k(s)\sum_{j\leq k}\nu_j(s) + \nu_k\sum_{j\leq k}\mu_j(s) + \sum_{j\geq k}\mu_j(s)\nu_j(s)\Big)ds.\end{aligned}$$

Combining Lemmas 7.12 and 7.13 gives a bound on ν_k of

$$\|\nu_k(s)\|_{L^4} \lesssim (1+s2^{2k})^{-3}(s2^{2k})^{-3/8}2^k 2^{-\sigma k} b_k(\sigma), \tag{7-100}$$

which leads to

$$\int_0^\infty \nu_k \sum_{j\leq k}\mu_j(s)\,ds \lesssim \Big(1+\sum_p b_p^2\Big)\tilde{b}_k 2^{-\sigma k} b_k(\sigma).$$

Also, by using (7-99) for μ_k and (7-100) for ν_k yields

$$\begin{aligned}\int_0^\infty \sum_{j\geq k}\mu_j(s)\nu_j(s)\,ds &\lesssim \Big(1+\sum_p b_p^2\Big)\sum_{j\geq k}2^{2j}2^{-\sigma j}b_j(\sigma)\tilde{b}_j\int_0^\infty (1+s2^{2j})^{-3}(s2^{2j})^{-3/8}\,ds\\ &\lesssim \Big(1+\sum_p b_p^2\Big)\sum_{j\geq k}2^{-\sigma j}b_j(\sigma)\tilde{b}_j\\ &\lesssim \Big(1+\sum_p b_p^2\Big)2^{-\sigma k}b_k(\sigma)\sum_{j\geq k}2^{(\delta-\sigma)(j-k)}\tilde{b}_j \lesssim \Big(1+\sum_p b_p^2\Big)2^{-\sigma k}\tilde{b}_k b_k(\sigma).\end{aligned}$$

Here we have used $\sigma \geq 2\delta$. It remains to consider

$$\int_0^\infty \mu_k(s)\sum_{j\leq k}\nu_j(s)ds.$$

Suppose that

$$\mu_k(s) \lesssim (1+s2^{2k})^{-2}2^k 2^{-\sigma k} b_k(\sigma)\Big(1+\sum_p b_p^2\Big). \tag{7-101}$$

Then

$$\begin{aligned}\int_0^\infty \mu_k(s)\sum_{j\leq k}\nu_j(s)\,ds &\lesssim \Big(1+\sum_p b_p^2\Big)^2 2^{-\sigma k}b_k(\sigma)2^k\sum_{j\leq k}\int_0^\infty (1+s2^{2k})^{-2}(1+s2^{2j})^{-2}2^j\tilde{b}_j\,ds\\ &\lesssim \Big(1+\sum_p b_p^2\Big)^2 2^{-\sigma k}b_k(\sigma)2^k\sum_{j\leq k}2^j\tilde{b}_j\int_0^\infty (1+s2^{2k})^{-2}\,ds\\ &\lesssim \Big(1+\sum_p b_p^2\Big)^2 2^{-\sigma k}b_k(\sigma)2^{2k}\tilde{b}_k\cdot 2^{-2k} = \Big(1+\sum_p b_p^2\Big)^2 2^{-\sigma k}b_k(\sigma)\tilde{b}_k.\end{aligned}$$

Hence it remains to establish (7-101).

By Corollary 7.16, (7-101) holds when $s = 0$. To extend this estimate to $s > 0$, we proceed as in the proof of Lemma 7.17, replacing bounds (7-92) and (7-93) with their $\sigma > 0$ analogues as needed; that these analogues hold follows from the bounds referenced in establishing (7-92) and (7-93). To obtain the analogue of (7-94), we apply Lemma 7.14, choosing to use $\sigma > 0$ bounds only over the high frequency ranges. □

Acknowledgments

The author sincerely thanks Terence Tao for all of his guidance, support, and encouragement, and thanks Daniel Tataru and Ioan Bejenaru for their encouragement and helpful talks and discussions on wave maps and Schrödinger maps. The author would also like to thank the anonymous referee for pointing out that the proof of Lemma 4.10 is only valid for $\sigma < \frac{1}{2} - \delta^+$, for supplying a suggestion as to how that lemma might be extended to larger σ, and for helpful suggestions pertaining to overall exposition.

References

[Bejenaru 2008a] I. Bejenaru, "Global results for Schrödinger maps in dimensions $n \geq 3$", *Comm. Partial Differential Equations* **33**:1-3 (2008), 451–477. MR 2009h:35392 Zbl 1148.35083

[Bejenaru 2008b] I. Bejenaru, "On Schrödinger maps", *Amer. J. Math.* **130** (2008), 1033–1065. MR 2009g:35296 Zbl 1159.35065

[Bejenaru and Tataru 2010] I. Bejenaru and D. Tataru, "Near soliton evolution for equivariant Schroedinger Maps in two spatial dimensions", preprint, 2010. arXiv 1009.1608

[Bejenaru et al. 2007] I. Bejenaru, A. D. Ionescu, and C. E. Kenig, "Global existence and uniqueness of Schrödinger maps in dimensions $d \geq 4$", *Adv. Math.* **215**:1 (2007), 263–291. MR 2008i:35124 Zbl 1152.35049

[Bejenaru et al. 2011a] I. Bejenaru, A. Ionescu, C. E. Kenig, and D. Tataru, "Equivariant Schrödinger maps in two spatial dimensions", preprint, 2011. arXiv 1112.6122

[Bejenaru et al. 2011b] I. Bejenaru, A. D. Ionescu, and C. E. Kenig, "On the stability of certain spin models in $2+1$ dimensions", *J. Geom. Anal.* **21**:1 (2011), 1–39. MR 2012b:35318 Zbl 1215.35145

[Bejenaru et al. 2011c] I. Bejenaru, A. D. Ionescu, C. E. Kenig, and D. Tataru, "Global Schrödinger maps in dimensions $d \geq 2$: small data in the critical Sobolev spaces", *Ann. of Math.* (2) **173**:3 (2011), 1443–1506. MR 2012g:58048 Zbl 1233.35112

[Bejenaru et al. 2012] I. Bejenaru, A. Ionescu, C. E. Kenig, and D. Tataru, "Equivariant Schrödinger Maps in two spatial dimensions: the $\mathbb{H}^2$ target", *ArXiv e-prints: 1212.2566* (December 2012).

[Bourgain 1998] J. Bourgain, "Refinements of Strichartz' inequality and applications to 2D-NLS with critical nonlinearity", *Int. Math. Res. Not.* **1998**:5 (1998), 253–283. MR 99f:35184 Zbl 0917.35126

[Chang et al. 2000] N.-H. Chang, J. Shatah, and K. Uhlenbeck, "Schrödinger maps", *Comm. Pure Appl. Math.* **53**:5 (2000), 590–602. MR 2001a:35161 Zbl 1028.35134

[D'Ancona and Fanelli 2008] P. D'Ancona and L. Fanelli, "Strichartz and smoothing estimates of dispersive equations with magnetic potentials", *Comm. Partial Differential Equations* **33**:4-6 (2008), 1082–1112. MR 2009i:35257 Zbl 1160.35363

[D'Ancona et al. 2010] P. D'Ancona, L. Fanelli, L. Vega, and N. Visciglia, "Endpoint Strichartz estimates for the magnetic Schrödinger equation", *J. Funct. Anal.* **258**:10 (2010), 3227–3240. MR 2011e:35325 Zbl 1188.81061

[Ding and Wang 2001] W. Ding and Y. Wang, "Local Schrödinger flow into Kähler manifolds", *Sci. China Ser. A* **44**:11 (2001), 1446–1464. MR 2002k:53129 Zbl 1019.53032

[Dodson and Smith 2013] B. Dodson and P. Smith, "A controlling norm for energy-critical Schrödinger maps", *ArXiv e-prints: 1302:3879* (2013).

[Erdoğan et al. 2008] M. B. Erdoğan, M. Goldberg, and W. Schlag, "Strichartz and smoothing estimates for Schrödinger operators with large magnetic potentials in $\mathbb{R}^3$", *J. Eur. Math. Soc.* (*JEMS*) **10**:2 (2008), 507–531. MR 2009h:35069 Zbl 1152.35021

[Erdoğan et al. 2009] M. B. Erdoğan, M. Goldberg, and W. Schlag, "Strichartz and smoothing estimates for Schrödinger operators with almost critical magnetic potentials in three and higher dimensions", *Forum Math.* **21**:4 (2009), 687–722. MR 2010j:35091 Zbl 1181.35208

[Fanelli and Vega 2009] L. Fanelli and L. Vega, "Magnetic virial identities, weak dispersion and Strichartz inequalities", *Math. Ann.* **344**:2 (2009), 249–278. MR 2010j:35427 Zbl 1163.35005

[Gustafson and Koo 2011] S. Gustafson and E. Koo, "Global well-posedness for 2D radial Schrödinger maps into the sphere", preprint, 2011. arXiv 1105.5659

[Gustafson et al. 2008] S. Gustafson, K. Kang, and T.-P. Tsai, "Asymptotic stability of harmonic maps under the Schrödinger flow", *Duke Math. J.* **145**:3 (2008), 537–583. MR 2009k:58030 Zbl 1170.35091

[Gustafson et al. 2010] S. Gustafson, K. Nakanishi, and T.-P. Tsai, "Asymptotic stability, concentration, and oscillation in harmonic map heat-flow, Landau–Lifshitz, and Schrödinger maps on $\mathbf{R}^2$", *Comm. Math. Phys.* **300**:1 (2010), 205–242. MR 2011j:58020 Zbl 1205.35294

[Ionescu and Kenig 2005] A. D. Ionescu and C. E. Kenig, "Well-posedness and local smoothing of solutions of Schrödinger equations", *Math. Res. Lett.* **12**:2-3 (2005), 193–205. MR 2006h:35224 Zbl 1077.35108

[Ionescu and Kenig 2006] A. D. Ionescu and C. E. Kenig, "Low-regularity Schrödinger maps", *Differential Integral Equations* **19**:11 (2006), 1271–1300. MR 2009m:35485 Zbl 1212.35449

[Ionescu and Kenig 2007a] A. D. Ionescu and C. E. Kenig, "Global well-posedness of the Benjamin–Ono equation in low-regularity spaces", *J. Amer. Math. Soc.* **20**:3 (2007), 753–798. MR 2008f:35350 Zbl 1123.35055

[Ionescu and Kenig 2007b] A. D. Ionescu and C. E. Kenig, "Low-regularity Schrödinger maps. II. Global well-posedness in dimensions $d \geq 3$", *Comm. Math. Phys.* **271**:2 (2007), 523–559. MR 2009m:35486 Zbl 1137.35068

[Kato 2005] J. Kato, "Existence and uniqueness of the solution to the modified Schrödinger map", *Math. Res. Lett.* **12**:2-3 (2005), 171–186. MR 2006b:35302 Zbl 1082.35140

[Kato and Koch 2007] J. Kato and H. Koch, "Uniqueness of the modified Schrödinger map in $H^{3/4+\epsilon}(\mathbb{R}^2)$", *Comm. Partial Differential Equations* **32**:1-3 (2007), 415–429. MR 2008h:58028 Zbl 05150098

[Kenig and Nahmod 2005] C. E. Kenig and A. R. Nahmod, "The Cauchy problem for the hyperbolic-elliptic Ishimori system and Schrödinger maps", *Nonlinearity* **18**:5 (2005), 1987–2009. MR 2006d:35227 Zbl 1213.35358

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Small solutions to nonlinear Schrödinger equations", *Ann. Inst. H. Poincaré Anal. Non Linéaire* **10**:3 (1993), 255–288. MR 94h:35238 Zbl 0786.35121

[Kenig et al. 1998] C. E. Kenig, G. Ponce, and L. Vega, "Smoothing effects and local existence theory for the generalized nonlinear Schrödinger equations", *Invent. Math.* **134**:3 (1998), 489–545. MR 99k:35166 Zbl 0928.35158

[Kenig et al. 2000] C. E. Kenig, G. Ponce, and L. Vega, "On the initial value problem for the Ishimori system", *Ann. Henri Poincaré* **1**:2 (2000), 341–384. MR 2001g:35237 Zbl 1021.35105

[Kenig et al. 2004] C. E. Kenig, G. Ponce, and L. Vega, "The Cauchy problem for quasi-linear Schrödinger equations", *Invent. Math.* **158**:2 (2004), 343–388. MR 2005f:35283 Zbl 1177.35221

[Krieger and Schlag 2012] J. Krieger and W. Schlag, *Concentration compactness for critical wave maps*, EMS Monographs in Mathematics **5**, European Mathematical Society, Zürich, 2012. MR 2895939 Zbl 06004782

[Krieger et al. 2008] J. Krieger, W. Schlag, and D. Tataru, "Renormalization and blow up for charge one equivariant critical wave maps", *Invent. Math.* **171**:3 (2008), 543–615. MR 2009b:58061 Zbl 1139.35021

[Landau 1967] L. D. Landau, *Collected papers of L. D. Landau*, edited by D. ter Haar, Gordon and Breach Science Publishers, New York, 1967. MR 38 #5577

[Lin and Wang 2008] F. Lin and C. Wang, *The analysis of harmonic maps and their heat flows*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008. MR 2011a:58030 Zbl 1203.58004

[McGahagan 2007] H. McGahagan, "An approximation scheme for Schrödinger maps", *Comm. Partial Differential Equations* **32**:1-3 (2007), 375–400. MR 2009e:35265 Zbl 1122.35138

[Merle et al. 2011a] F. Merle, P. Raphaël, and I. Rodnianski, "Blow up dynamics for smooth equivariant solutions to the energy critical Schrödinger map", *C. R. Math. Acad. Sci. Paris* **349**:5-6 (2011), 279–283. MR 2012e:58029 Zbl 1213.35139

[Merle et al. 2011b] F. Merle, P. Raphaël, and I. Rodnianski, "Blow up dynamics for smooth equivariant solutions to the energy critical Schrödinger map", *C. R. Math. Acad. Sci. Paris* **349**:5-6 (2011), 279–283. MR 2012e:58029 Zbl 1213.35139

[Nahmod et al. 2003] A. Nahmod, A. Stefanov, and K. Uhlenbeck, "On Schrödinger maps", *Comm. Pure Appl. Math.* **56**:1 (2003), 114–151. MR 2003i:58030 Zbl 1028.58018

[Nahmod et al. 2007] A. Nahmod, J. Shatah, L. Vega, and C. Zeng, "Schrödinger maps and their associated frame systems", *Int. Math. Res. Not.* **2007**:21 (2007), Art. ID rnm088, 29. MR 2009b:58063 Zbl 1142.35087

[Papanicolaou and Tomaras 1991] N. Papanicolaou and T. N. Tomaras, "Dynamics of magnetic vortices", *Nuclear Phys. B* **360**:2-3 (1991), 425–462. MR 92c:82114

[Perelman 2012] G. Perelman, "Blow up dynamics for equivariant critical Schrödinger maps", preprint, 2012. arXiv 1212.6709

[Planchon 2012] F. Planchon, "Existence globale et scattering pour les solutions de masse finie de l'équation de Schrödinger cubique en dimension deux, d'après Benjamin Dodson, Rowan Killip, Terence Tao, Monica Visan et Xiaoyi Zhang", exposé 1042 in *Séminaire Bourbaki, 2010/2011*, Astérisque **348**, Soc. math. de France, Paris, 2012.

[Planchon and Vega 2009] F. Planchon and L. Vega, "Bilinear virial identities and applications", *Ann. Sci. Éc. Norm. Supér.* (4) **42**:2 (2009), 261–290. MR 2010b:35441 Zbl 1192.35166

[Rodnianski and Sterbenz 2010] I. Rodnianski and J. Sterbenz, "On the formation of singularities in the critical O(3) σ-model", *Ann. of Math.* (2) **172**:1 (2010), 187–242. MR 2011i:58023 Zbl 1213.35392

[Rodnianski et al. 2009] I. Rodnianski, Y. A. Rubinstein, and G. Staffilani, "On the global well-posedness of the one-dimensional Schrödinger map flow", *Anal. PDE* **2**:2 (2009), 187–209. MR 2010i:53128 Zbl 1191.35258

[Smith 2012a] P. Smith, "Geometric renormalization below the ground state", *Int. Math. Res. Not.* **2012**:16 (2012), 3800–3844. Zbl 06078264

[Smith 2012b] P. Smith, "Global regularity of critical Schrödinger maps: subthreshold dispersed energy", preprint, 2012. arXiv 1112.0251

[Stefanov 2007] A. Stefanov, "Strichartz estimates for the magnetic Schrödinger equation", *Adv. Math.* **210**:1 (2007), 246–303. MR 2009e:81071 Zbl 1121.35023

[Sterbenz and Tataru 2010a] J. Sterbenz and D. Tataru, "Energy dispersed large data wave maps in $2+1$ dimensions", *Comm. Math. Phys.* **298**:1 (2010), 139–230. MR 2011g:58045 Zbl 1218.35129

[Sterbenz and Tataru 2010b] J. Sterbenz and D. Tataru, "Regularity of wave-maps in dimension $2+1$", *Comm. Math. Phys.* **298**:1 (2010), 231–264. MR 2011h:58026 Zbl 1218.35057

[Strichartz 1977] R. S. Strichartz, "Restrictions of Fourier transforms to quadratic surfaces and decay of solutions of wave equations", *Duke Math. J.* **44**:3 (1977), 705–714. MR 58 #23577 Zbl 0372.35001

[Sulem et al. 1986] P.-L. Sulem, C. Sulem, and C. Bardos, "On the continuous limit for a system of classical spins", *Comm. Math. Phys.* **107**:3 (1986), 431–454. MR 87k:82024 Zbl 0614.35087

[Tao 2001] T. Tao, "Global regularity of wave maps. II. Small energy in two dimensions", *Comm. Math. Phys.* **224**:2 (2001), 443–544. MR 2002h:58052 Zbl 1020.35046

[Tao 2004] T. Tao, "Geometric renormalization of large energy wave maps", pp. Exp. No. XI, 32 in *Journées "Équations aux Dérivées Partielles"*, École Polytech., Palaiseau, 2004. MR 2006i:58044 Zbl 1087.58019

[Tao 2006a] T. Tao, "Gauges for the Schrödinger map", 2006, Available at http://tinyurl.com/d6g499k.

[Tao 2006b] T. Tao, *Nonlinear dispersive equations: local and global analysis*, CBMS Regional Conference Series in Mathematics **106**, Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2006. MR 2008i:35211 Zbl 1106.35001

[Tao 2008a] T. Tao, "Global regularity of wave maps, III: Large energy from $\mathbf{R}^{1+2}$ to hyperbolic spaces", preprint, 2008. arXiv 0805.4666

[Tao 2008b] T. Tao, "Global regularity of wave maps, IV: Absence of stationary or self-similar solutions in the energy class", preprint, 2008. arXiv 0806.3592

[Tao 2008c] T. Tao, "Global regularity of wave maps, V: Large data local wellposedness and perturbation theory in the energy class", preprint, 2008. arXiv 0808.0368

[Tao 2009a] T. Tao, "Global regularity of wave maps, VI: Abstract theory of minimal-energy blowup solutions", preprint, 2009. arXiv 0906.2833

[Tao 2009b] T. Tao, "Global regularity of wave maps, VII: Control of delocalised or dispersed solutions", preprint, 2009. arXiv 0908.0776

[Tao 2010] T. Tao, "A physical space proof of the bilinear Strichartz and local smoothing estimates for the Schrödinger equation", preprint, 2010, Available at http://terrytao.files.wordpress.com/2013/07/bilinear.pdf.

[Tataru 2001] D. Tataru, "On global existence and scattering for the wave maps equation", *Amer. J. Math.* **123**:1 (2001), 37–77. MR 2002c:58045 Zbl 0979.35100

Received 19 Apr 2011. Revised 8 Jun 2012. Accepted 6 Jul 2012.

PAUL SMITH: smith@math.berkeley.edu
Department of Mathematics, University of California, Berkeley, 970 Evans Hall, Berkeley, CA 94720-3840, United States

Analysis & PDE

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2013 is US $160/year for the electronic version, and $310/year (+$35, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 6 No. 3 2013