SOLITON RESOLUTION FOR ENERGY-CRITICAL WAVE MAPS IN THE EQUIVARIANT CASE

JACEK JENDREJ AND ANDREW LAWRIE

ABSTRACT. We consider the equivariant wave maps equation $\mathbb{R}^{1+2} \to \mathbb{S}^2$, in all equivariance classes $k \in \mathbb{N}$. We prove that every finite energy solution resolves, continuously in time, into a superposition of asymptotically decoupling harmonic maps and free radiation.

CONTENTS

1. INTRODUCTION

1.1. **Setting of the problem.** We study wave maps from the Minkowski space $\mathbb{R}^{1+2}_{t,x}$ into the two-sphere $\mathbb{S}^2$, under k-equivariant symmetry. These are formal critical points of the Lagrangian action,

$$\mathscr{L}(\Psi) = \frac{1}{2} \iint_{\mathbb{R}^{1+2}_{t,x}} \left(-|\partial_t \Psi(t,x)|^2 + |\nabla \Psi(t,x)|^2 \right) \mathrm{d}x \mathrm{d}t,$$

restricted to the class of maps $\Psi : \mathbb{R}^{1+2}_{t,x} \to \mathbb{S}^2 \subset \mathbb{R}^3$ that take the form,

$$\Psi(t, re^{i\theta}) = (\sin u(t,r) \cos k\theta, \sin u(t,r) \sin k\theta, \cos u(t,r)) \in \mathbb{S}^2 \subset \mathbb{R}^3,$$

for some fixed $k \in \{1, 2, \ldots\}$. Above u is the colatitude measured from the north pole, the metric on $\mathbb{S}^2$ is $\mathrm{d}s^2 = \mathrm{d}u^2 + \sin^2 u \, \mathrm{d}\omega^2$, and (r, θ) are polar coordinates on $\mathbb{R}^2$.

The general $\mathbb{S}^2$-valued wave maps equation in two space dimensions is called the $O(3)$ sigma model in high energy physics literature. It is a canonical example of a geometric wave equation as it generalizes the free scalar wave equation to the setting of manifold-valued maps. The static solutions given by finite energy harmonic maps are amongst the simplest examples of topological solitons as they admit Bogomol'nyi structure [2]; other examples include kinks in scalar field theories on the line, vortices in Ginzburg-Landau equations, magnetic monopoles, Skyrmions, and Yang-Mills instantons; see [50] for an extensive treatment of field theories admitting topological solitons from the point of view of mathematical physics.

Our interest in k-equivariant wave maps stems from the richness of their nonlinear dynamics in the relatively simple setting of the geometrically natural scalar semilinear wave equation,

$$\partial_t^2 u(t,r) - \Delta u(t,r) + \frac{k^2}{r^2} \frac{\sin 2u(t,r)}{2} = 0, \quad (t,r) \in \mathbb{R} \times (0, \infty), \tag{1.1}$$

which is the Euler-Lagrange equation associated to $\mathscr{L}(\Psi)$ under the k-equivariant symmetry reduction. Here $\Delta := \partial_r^2 + r^{-1}\partial_r$ is the radial Laplacian in 2-dimensions. The conserved energy for (1.1) is given by

$$E(u, \partial_t u)(t) := 2\pi \int_0^\infty \frac{1}{2} \left((\partial_t u(t,r))^2 + (\partial_r u(t,r))^2 + k^2 \frac{\sin^2 u(t,r)}{r^2} \right) r \mathrm{d}r. \tag{1.2}$$

We will often write pairs of functions using boldface, $\boldsymbol{v} = (v, \dot{v})$, noting that the notation $\dot{v}$ will not, in general, refer to a time derivative of v but rather just to the second component of $\boldsymbol{v}$. With this notation the Cauchy problem for (1.1) can be rephrased as the Hamiltonian system

$$\partial_t \boldsymbol{u}(t) = J \circ \mathrm{D}E(\boldsymbol{u}(t)), \qquad \boldsymbol{u}(T_0) = \boldsymbol{u}_0, \tag{1.3}$$

where

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathrm{D}E(\boldsymbol{u}(t)) = \begin{pmatrix} -\Delta u(t) + k^2 r^{-2} 2^{-1} \sin(2u(t)) \\ \dot{u}(t) \end{pmatrix}.$$

Both (1.2) and (1.3) are invariant under the scaling

$$(u(t,r), \partial_t u(t,r)) \mapsto \left(u(t/\lambda, r/\lambda), \lambda^{-1} \partial_t u(t/\lambda, r/\lambda) \right), \qquad \lambda > 0,$$

and thus (1.1) is called energy-critical.

The natural setting in which to consider the Cauchy problem for (1.1) is the space of initial data $\boldsymbol{u}_0$ with finite energy, $E(\boldsymbol{u}_0) < \infty$. The set of finite energy data is split into disjoint sectors, $\mathcal{E}_{\ell,m}$, which for $\ell, m \in \mathbb{Z}$, are defined by

$$\mathcal{E}_{\ell,m} := \left\{ (u_0, \dot{u}_0) \mid E(u_0, \dot{u}_0) < \infty, \quad \lim_{r \to 0} u_0(r) = \ell\pi, \quad \lim_{r \to \infty} u_0(r) = m\pi \right\}.$$

These sectors, which are preserved by the flow, are related to the topological degree of the full map $\Psi_0 : \mathbb{R}^2 \to \mathbb{S}^2$: if $m - \ell$ is even and $(u_0, 0) \in \mathcal{E}_{\ell,m}$, then the corresponding map Ψ with polar angle u_0 is topologically trivial, whereas for odd $m - \ell$ the map has degree k.

The sets $\mathcal{E}_{\ell,m}$ are affine spaces, parallel to the linear space $\mathcal{E} := \mathcal{E}_{0,0} = H \times L^2$, which we endow with the norm,

$$\|\boldsymbol{u}_0\|_{\mathcal{E}}^2 := \|\dot{u}_0\|_{L^2}^2 + \|u_0\|_H^2 := \int_0^\infty \left((\dot{u}_0(r))^2 + (\partial_r u_0(r))^2 + k^2 \frac{(u_0(r))^2}{r^2} \right) r \mathrm{d}r.$$

The linearization of (1.1) about the zero solution is given by

$$\partial_t^2 v - \Delta v + \frac{k^2}{r^2} v = 0, \tag{1.4}$$

and the flow for (1.4) preserves the $\mathcal{E}$ norm.

The unique k-equivariant harmonic map is given explicitly by

$$Q(r) := 2 \arctan(r^k).$$

Here uniqueness means up to scaling, sign change, and adding a multiple of π, i.e., every finite energy stationary solution to (1.1) takes the form $Q_{\mu,\sigma,m}(r) = m\pi + \sigma Q(r/\mu)$ for some $\mu \in (0, \infty), \sigma \in \{0, -1, 1\}$ and $m \in \mathbb{Z}$. The pair $\boldsymbol{Q}_\lambda := (Q_\lambda, 0)$ and its rescaled versions $\boldsymbol{Q}_\lambda(r) := (Q_\lambda(r), 0) := Q(\lambda^{-1}r)$ for $\lambda > 0$, are minimizers of the energy E within the class $\mathcal{E}_{0,1}$; in fact, $E(\boldsymbol{Q}_\lambda) = 4\pi k$. We denote, $\boldsymbol{\pi} := (\pi, 0)$.

1.2. Statement of the results. Our main result is formulated as follows.

Theorem 1 (Soliton Resolution). *Let $k \in \mathbb{N}$, let $\ell, m \in \mathbb{Z}$, and let $\boldsymbol{u}(t)$ be a finite energy solution to (1.1) with initial data $\boldsymbol{u}(0) = \boldsymbol{u}_0 \in \mathcal{E}_{\ell,m}$, defined on its maximal forward interval of existence $[0, T_+)$.*

(Global solution) If $T_+ = \infty$, there exist a time $T_0 > 0$, a solution $\boldsymbol{u}_{\mathrm{L}}^*(t) \in \mathcal{E}$ to the linear wave equation (1.4), an integer $N \geq 0$, continuous functions $\lambda_1(t), \ldots, \lambda_N(t) \in C^0([T_0, \infty))$, signs $\iota_1, \ldots, \iota_N \in \{-1, 1\}$, and $\boldsymbol{g}(t) \in \mathcal{E}$ defined by

$$\boldsymbol{u}(t) = m\boldsymbol{\pi} + \sum_{j=1}^N \iota_j (\boldsymbol{Q}_{\lambda_j(t)} - \boldsymbol{\pi}) + \boldsymbol{u}_{\mathrm{L}}^*(t) + \boldsymbol{g}(t),$$

such that

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{j=1}^N \frac{\lambda_j(t)}{\lambda_{j+1}(t)} \to 0 \quad as \quad t \to \infty,$$

where above we use the convention that $\lambda_{N+1}(t) = t$.

(Blow-up solution) If $T_+ < \infty$, there exist a time $T_0 < T_+$, integers m_∞, m_Δ, a mapping $\boldsymbol{u}_0^* \in \mathcal{E}_{0,m_\infty}$, an integer $N \geq 1$, continuous functions $\lambda_1(t), \ldots, \lambda_N(t) \in C^0([T_0, T_+))$, signs $\iota_1, \ldots, \iota_N \in \{-1, 1\}$, and $\boldsymbol{g}(t) \in \mathcal{E}$ defined by

$$\boldsymbol{u}(t) = m_\Delta \boldsymbol{\pi} + \sum_{j=1}^N \iota_j (\boldsymbol{Q}_{\lambda_j(t)} - \boldsymbol{\pi}) + \boldsymbol{u}_0^* + \boldsymbol{g}(t),$$

such that

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{j=1}^N \frac{\lambda_j(t)}{\lambda_{j+1}(t)} \to 0 \quad as \quad t \to T_+,$$

where above we use the convention that $\lambda_{N+1}(t) = T_+ - t$.

Analogous statements hold for the backwards-in-time evolution.

Remark 1.1. This type of behavior is referred to as soliton resolution. A recent preprint by Duyckaerts, Kenig, Martel, and Merle proved Theorem 1 in the case $k = 1$ using the method of energy channels; see [18]. Roughly, energy channels refer to measurements of the portion of energy that a linear or nonlinear wave radiates outside fattened light cones. Such exterior energy estimates were introduced by Duyckaerts, Kenig, and Merle [24] in their proof of the soliton resolution conjecture for the radial energy critical NLW in 3 space dimensions; see also [25–27] for the treatment of all odd dimensions. The approach we take to prove Theorem 1 is independent of the method of energy channels.

Remark 1.2. The soliton resolution problem is inspired by the theory of completely integrable systems, e.g., [30,67,68], motivated by numerical simulations, [31,80], and by the bubbling theory of harmonic maps in the elliptic and parabolic settings [60,61,72,78,79]; see also [16,18,27] for discussions on the history of the problem.

Remark 1.3. Our method establishes the exact analog of Theorem 1 in the case of the equivariant Yang-Mills equation, by making the usual analogy between equivariant Yang-Mills and $k = 2$-equivariant wave maps; see Cazenave, Shatah, and Tahvildar-Zadeh [4] for the formulation. There, the harmonic map is replaced by the first instanton.

Remark 1.4. The Cauchy problem for energy-critical wave maps restricted to radially symmetric initial data (which is distinct from the k-equivariant symmetry reduction considered here as there are no radially symmetric nontrivial harmonic maps) was understood in classic works by Christodoulou and Tahvildar-Zadeh [7,8], Struwe [74,75], extended by Nahas in [56], and settled definitively by Chiodaroli, Krieger, and Lührmann [5], who developed a concentration compactness theory in that setting.

Remark 1.5. Theorem 1 is a qualitative description of the dynamics of all finite energy solutions to (1.1). A natural, challenging question is to ask which types of configurations of solitons and radiation are realized in solutions. The first results of this nature were constructions of solutions blowing up in finite time by bubbling off a single harmonic map by Krieger, Schlag, and Tataru [46], Rodnianski and Sterbenz [64], and Raphaël and Rodnianski [62]. In [36], the first author constructed a solution exhibiting more than one bubble in the decomposition, showing the existence of a solution that forms a 2-bubble in infinite time with zero radiation in equivariance classes $k \geq 2$. In [65] Rodriguez showed that no such 2-bubble occurs in the case $k = 1$, proving that the only non-scattering solution with energy $= 2E(\boldsymbol{Q})$ blows up by bubbling of a single harmonic map in finite time, while radiating $\boldsymbol{u}_0^* = -\boldsymbol{Q}$. It is not known if there are any solutions with more than one bubble in the decomposition when $k = 1$. As a consequence of our analysis, we prove that such a solution would necessarily have non-zero radiation, thus extending the result of Rodriguez to any number of bubbles (see Proposition 1.8 below).

It is natural to ask about the fate of solutions with more than one bubble in the decomposition in the opposite time direction. An answer to this question was given by the authors in [41] for the 2-bubble solution $\boldsymbol{u}_{(2)}(t)$ constructed by the first author in [36]. We showed that any 2-bubble in forward time must scatter freely in backwards time. When the scales of the bubbles become comparable, this collision completely annihilates the 2-bubble structure and the entire solution becomes free radiation, i.e., the collision is inelastic. Viewing the evolution of $\boldsymbol{u}_{(2)}(t)$ in forward time, this means that the 2-soliton structure emerges from pure radiation, and constitutes an orbit connecting two different dynamical behaviors. We later showed in [38,40] that $\boldsymbol{u}_{(2)}(t)$ is the unique 2-bubble solution up to sign, translation, and scaling in equivariance classes $k \geq 4$.

Crucial to the proof of scattering after the collision in the case of two bubbles is the fact that the 2-bubble configurations considered in [41] are minimal in the sense that any solution in $\mathcal{E}_{0,0}$

with energy $< 2E(\boldsymbol{Q})$ must scatter (see [11]). While inelasticity of collisions is still expected in the case of solutions with more than two bubbles in one time direction, such a solution can still exhibit bubbling behavior even after a collision that produces radiation – for example a solution in $\mathcal{E}_{0,1}$ with three bubbles and no radiation in one direction could have one bubble with non-zero radiation in the other direction. While we do not consider such refined two-directional analysis here, a relatively straightforward corollary of the proof of Theorem 1 is that there can be *no elastic collisions of pure multi-bubbles*, which we formulate as a proposition below.

Definition 1.6. With the notations from the statement of Theorem 1, we say that $\boldsymbol{u}$ is a *pure multi-bubble* in the forward time direction if $\boldsymbol{u}_{\mathrm{L}}^* = 0$ in the case $T_+ = +\infty$, and $\boldsymbol{u}_0^* = 0$ in the case $T_+ < +\infty$.

We say that $\boldsymbol{u}$ is a pure multi-bubble in the backward time direction if $t \mapsto \boldsymbol{u}(-t)$ is a pure multi-bubble in the forward time direction.

Proposition 1.7. *Stationary solutions are the only pure multi-bubbles in both time directions.*

In the special case of equivariance class $k = 1$, the following much stronger result holds.

Proposition 1.8. *If $k = 1$, then stationary solutions are the only pure multi-bubbles in forward (or backward) time.*

Remark 1.9. We note that Proposition 1.7 was also proved in the case $k = 1$ for (1.1) in the recent preprint [18], as well as for the energy critical focusing NLW under radial symmetry and in odd space dimensions in [24, 27], all via a different approach based on energy channels. As mentioned above, the case of $N = 2$ bubbles was already considered in [41]. See [51–53] for more regarding the inelastic soliton collision problem for non-integrable PDEs.

1.3. **History of progress on the problem.** Our proof of Theorem 1 is built on top of two significant partial results, namely (1) that the radiation term, $\boldsymbol{u}_{\mathrm{L}}^*$ in the global case and $\boldsymbol{u}_0^*$ in the blow-up setting, can be identified continuously in time, and (2) that the resolution is known to hold along a well-chosen sequence of times. The result (1) was established in [10–12, 42] as a consequence of the classical work of Shatah and Tahvildar-Zadeh [70], and we make explicit use of this fact. The latter result (2) was proved by Côte [10] and Jia and Kenig [42] using Struwe's classical bubbling analysis [73], many ideas from Duyckaerts, Kenig, and Merle's seminal works [21–23], and several new insights particular to (1.1). While the sequential resolution certainly inspires part of our argument, we cannot use it simply as a black box, but rather we revisit the proof and derive more precise information from the analysis of Côte, and Jia and Kenig as we explain in the next section.

We discuss these prior results in more detail. To unify the blow-up and global-in-time settings we make the following conventions. Consider a finite energy wave map $\boldsymbol{u}(t) \in \mathcal{E}_{\ell,m}$. We assume that either $\boldsymbol{u}(t)$ blows up in backwards time at $T_- = 0$ and is defined on an interval $I_* := (0, T_0]$, or $\boldsymbol{u}(t)$ is global in forward time and defined on the interval $I_* := [T_0, \infty)$ where in both cases $T_0 > 0$. We let $T_* := 0$ in the blow-up case and $T_* := \infty$ in the global case.

Extraction of the radiation. Below we will use the notation $\mathcal{E}(r_1, r_2)$ to denote the local energy norm

$$\|\boldsymbol{g}\|_{\mathcal{E}(r_1,r_2)}^2 := \int_{r_1}^{r_2} \left((\dot{g})^2 + (\partial_r g)^2 + \frac{k^2}{r^2} g^2 \right) r \mathrm{d}r,$$

By convention, $\mathcal{E}(r_0) := \mathcal{E}(r_0, \infty)$ for $r_0 > 0$. The local nonlinear energy is denoted $E(\boldsymbol{u}_0; r_1, r_2)$. We adopt similar conventions as for $\mathcal{E}$ regarding the omission of r_2, or both r_1 and r_2.

Theorem 1.10 (Identification of the radiation). [10, Propositions 5.1, 5.2] *Let $\boldsymbol{u}(t) \in \mathcal{E}_{\ell,m}$ be a finite energy wave map on an interval I_* as above. Then, the limit $\pi\mathbb{Z} \ni m_\Delta \pi := \lim_{t \to T_*} u(t, \frac{1}{2}t)$*

exists, and there is an integer $m_\infty \in \mathbb{Z}$, a finite energy wave map $\boldsymbol{u}^*(t) \in \mathcal{E}_{0,m_\infty}$ called the radiation, and a function $\rho : I_* \to (0,\infty)$ that satisfies,

$$\lim_{t \to T_*} \left((\rho(t)/t)^k + \|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - m_\Delta \boldsymbol{\pi}\|^2_{\mathcal{E}(\rho(t))} \right) = 0. \tag{1.5}$$

Moroever, for any $\alpha \in (0,1)$,

$$E(\boldsymbol{u}^*(t); 0, \alpha t) \to 0 \quad as \quad t \to T_*. \tag{1.6}$$

Remark 1.11. In the global setting, i.e., $I_* = [T_0, \infty)$ we must have $m_\infty = 0$ and the linear wave $\boldsymbol{u}^*_{\mathrm{L}}(t) \in \mathcal{E}$ that appears in Theorem 1 is the unique solution to the linear equation (1.4) satisfying,

$$\|\boldsymbol{u}^*(t) - \boldsymbol{u}^*_{\mathrm{L}}(t)\|_{\mathcal{E}} \to 0 \quad as \quad t \to \infty,$$

which one obtains via the existence of wave operators; see Lemma 2.8. In the finite time blow-up setting the radiation $\boldsymbol{u}^*_0 \in \mathcal{E}_{0,m_\infty}$ that appears in Theorem 1 is initial data for $\boldsymbol{u}^*(t)$, i.e., the radiation $\boldsymbol{u}^*(t)$ in Theorem 1.10 satisfies $\boldsymbol{u}(t,r) = m_\Delta \boldsymbol{\pi} + \boldsymbol{u}^*(t,r)$ for $r > t$. With this definition and energy conservation, Theorem 1.15 implies the energy identity,

$$E(\boldsymbol{u}) = NE(\boldsymbol{Q}) + E(\boldsymbol{u}^*). \tag{1.7}$$

We remark that (1.6) in the case $T_* = \infty$ uses the estimates for the even dimensional free scalar wave equation proved by Côte, Kenig, and Schlag in [15].

The identification of $\boldsymbol{u}^*(t)$ and the vanishing (1.5) uses fundamental technique of Shatah and Tahvildar-Zadeh [70] (see also Christodoulou and Tahvildar-Zadeh [7] for the case of spherically symmetric wave maps); in [70] it is proved that every singular wave map has asymptotically no energy in the self-similar region of the cone, i.e.,

$$E(\boldsymbol{u}(t); \alpha t, t) \to 0 \quad as \quad t \to T_*$$

for each $\alpha \in (0,1)$ in the case $T_* = 0$, and

$$\lim_{A \to \infty} \limsup_{t \to T_*} E(\boldsymbol{u}(t), \alpha t, t - A) = 0$$

in the case $T_* = \infty$. Note that the latter refined estimate for globally defined wave maps was proved in [12] using methods from [7, 70].

Remark 1.12. The radiation field can be identified in several other contexts and by different means. For example, Tao accomplished this in [77] for certain high dimensional NLS. For critical nonlinear waves with power-type nonlinearities, the radiation field can be identified even outside radial symmetry; see the work of Duyckaerts, Kenig, and Merle [20].

Sequential soliton resolution. The first result in this direction was Struwe's bubbling theorem [73], which showed that any smooth solution to (1.1) that develops a singularity in finite time must do so by bubbling off at least one harmonic map, locally in space, along some sequence of times.

A deep insight of Duyckaerts, Kenig, and Merle, proved in [22] for the energy critical NLW, is that once the linear radiation is subtracted from the solution, the entire remainder should exhibit strong sequential compactness – it decomposes into a finite sum of asymptotically decoupled elliptic objects, in our case these are stationary harmonic maps, along at least one time sequence, up to an error that vanishes in the energy space. A crucial tool in proving such a compactness statement is the remarkable theory of profile decompositions for dispersive equations developed by Bahouri and Gérard [1]. However, after finding the profiles and their space-time concentration properties (in our case their scales) via the main result in [1], one must identify them as elliptic

objects (solitons) by some means, and then prove that the error vanishes in the sense of energy, rather than the weaker form of compactness (vanishing in the sense of a Strichartz norm) given by [1]. In the wave map case, this program was carried out by Côte, Kenig, the second author, and Schlag [11,12] (using the even dimensional exterior energy estimates proved by Côte, Kenig, and Schlag in [15]) for solutions to (1.1) with $k = 1$ in $\mathcal{E}_{0,1}$ with $E < 3E(\boldsymbol{Q})$. The latter condition restricted the number of possible configurations to those with a single bubble, and in this special case the sequential resolution could easily be upgraded to a continuous one using the variational characterization of $\boldsymbol{Q}$ and the coercivity of the energy functional.

In our setting, the sequential resolution was proved by Côte [10] in the case $k = 1$, and Jia and Kenig [42] in the case $k = 2$, namely that Theorem 1 holds along a well-chosen sequence of times. These works used the bubbling theory of Struwe [73] to identify the profiles as harmonic maps, and in the latter paper the authors used a novel nonlinear multiplier identity to obtain the convergence of the error in the energy space – in fact, we make use of this same identity in this work, see Section 3. A minor technical observation, which we explain in Remark 3.2, yields their result in all equivariance classes $k \in \mathbb{N}$. Before stating it, we introduce some notation.

Definition 1.13 (Multi-bubble configuration). Given $M \in \{0, 1, \ldots\}$, $m \in \mathbb{Z}$, $\vec{\iota} = (\iota_1, \ldots, \iota_M) \in \{-1, 1\}^M$ and an increasing sequence $\vec{\lambda} = (\lambda_1, \ldots, \lambda_M) \in (0, \infty)^M$, a *multi-bubble configuration* is defined by the formula

$$\boldsymbol{Q}(m, \vec{\iota}, \vec{\lambda}; r) := m\boldsymbol{\pi} + \sum_{j=1}^{M} \iota_j \big(\boldsymbol{Q}_{\lambda_j}(r) - \boldsymbol{\pi}\big).$$

Remark 1.14. If $M = 0$, it should be understood that $\boldsymbol{Q}(m, \vec{\iota}, \vec{\lambda}; r) = m\boldsymbol{\pi}$ for all $r \in (0, \infty)$, where $\vec{\iota}$ and $\vec{\lambda}$ are 0-element sequences, that is the unique functions $\emptyset \to \{-1, 1\}$ and $\emptyset \to (0, \infty)$, respectively.

We state the main theorems from Côte [10] and Jia, Kenig [42] using this notation.

Theorem 1.15 (Sequential soliton resolution). [10, Theorem 1.1], [42, Theorem 1.2] *Let $k \in \mathbb{N}$, $\ell, m \in \mathbb{Z}$, and let $\boldsymbol{u}(t) \in \mathcal{E}_{\ell,m}$ be a finite energy wave map on an interval I_* as above. Let $m_\Delta, m_\infty \in \mathbb{Z}$, and the radiation $\boldsymbol{u}^*(t) \in \mathcal{E}_{0,m_\infty}$ be as in Theorem 1.10. Then, there exists an integer $N \geq 0$, a sequence of times $t_n \to T_*$, signs $\vec{\iota}_n \in \{-1, 1\}^N$, and scales $\vec{\lambda}_n \in (0, \infty)^N$ such that,*

$$\lim_{n \to \infty} \left(\|\boldsymbol{u}(t_n) - \boldsymbol{u}^*(t_n) - \boldsymbol{Q}(m, \vec{\iota}_n, \vec{\lambda}_n)\|_{\mathcal{E}} + \sum_{j=1}^{N} \frac{\lambda_{n,j}}{\lambda_{n,j+1}} \right) = 0,$$

where above we use the convention $\lambda_{n,N+1} := t_n$.

Remark 1.16. The Duyckaerts, Kenig, and Merle approach from [22] to sequential soliton resolution has been successful in other settings. The same authors with Jia proved the sequential decomposition for the full energy critical NLW (i.e., not assuming radial symmetry) in [16] and for wave maps outside equivariant symmetry for data with energy slightly above the ground state [17], where the perturbative regularity theory of Tao [76] could be used; see also the bubbling theory of Grinis [34]. See also [13] for the radially symmetric energy critical NLW in four space dimensions, and [66] for the same equation in odd space dimensions.

1.4. **Summary of the proof: collision intervals and no-return analysis.** The challenging nature of bridging the gap between Theorem 1.15, which is the resolution along one sequence of times, and Theorem 1 is apparent from the following consideration. The sequence $t_n \to T_*$ in Theorem 1.15 gives no relationship between the lengths of the time intervals $[t_n, t_{n+1}]$ and

the concentration scales $\vec{\lambda}_n$ of the various harmonic maps in the decomposition. One immediate enemy is then the possibility of *elastic collisions*. If colliding solitons could recover their shape after a collision, then one could potentially encounter the following scenario: the solution approaches a multi-soliton configuration for a sequence of times, but in between infinitely many collisions take place, so that there is no soliton resolution in continuous time.

We describe our approach. Fix $\boldsymbol{u}(t) \in \mathcal{E}_{\ell,m}$, a finite energy solution to (1.1) on the time interval I_* as defined above. Let $N \geq 0$, $m_\infty, m_\Delta \in \mathbb{Z}$, and the radiation $\boldsymbol{u}^*(t) \in \mathcal{E}_{0,m_\infty}$ be as in Theorem 1.15. We define a *multi-bubble proximity function* at each $t \in I_*$ by

$$\mathbf{d}(t) := \inf_{\vec{\iota},\vec{\lambda}} \left(\|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\iota}, \vec{\lambda})\|_{\mathcal{E}}^2 + \sum_{j=1}^{N} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \right)^{\frac{1}{2}}, \tag{1.8}$$

where $\vec{\iota} := (\iota_1, \ldots, \iota_N) \in \{-1, 1\}^N$, $\vec{\lambda} := (\lambda_1, \ldots, \lambda_N) \in (0, \infty)^N$, and $\lambda_{N+1} := t$. We note that $\mathbf{d}(t)$ is a continuous function on I_*.

With this notation, we see that Theorem 1.15 gives a monotone sequence of times $t_n \to T_*$ such that,

$$\lim_{n \to \infty} \mathbf{d}(t_n) = 0.$$

Theorem 1 is an immediate consequence of showing that $\lim_{t \to T_*} \mathbf{d}(t) = 0$. We argue by contradiction, assuming that $\limsup_{t \to T_*} \mathbf{d}(t) > 0$. This means that there is some sequence of times where $\boldsymbol{u}(t) - \boldsymbol{u}^*(t)$ approaches an N-bubble and another sequence of times for which it stays bounded away from N-bubble configurations. It is natural to rule out this behavior by proving what is called a *no-return* lemma. In this generality, our approach is inspired by no-return results for one soliton by Duyckaerts and Merle [28, 29], Nakanishi and Schlag [57, 58], and Krieger, Nakanishi and Schlag [44, 45]. The exponential instability considered in those works is absent here, but is replaced by attractive nonlinear interactions between the solitons. This latter consideration, and indeed the overall scheme of the proof is based on our previous work [41], where modulation analysis of bubble interactions was used for the first time in the context of the soliton resolution problem (in fact, we recently showed that the collision analysis in [41] yielded a quick proof of Theorem 1 in the special cases when at most two bubbles appear in the decomposition; see [39]).

The basic tool we use is the standard virial functional

$$\mathfrak{v}(t) := \int_0^\infty \partial_t u(t) r \partial_r u(t) \chi_{\rho(t)} \, r \mathrm{d}r,$$

where the cut-off χ is placed along a Lipschitz curve $r = \rho(t)$ that will be carefully chosen (note that a time-dependent cut-off of the virial functional was also used in [57, 58]). Differentiating $\mathfrak{v}(t)$ in time we have,

$$\mathfrak{v}'(t) = -\int_0^\infty |\partial_t u(t, r)|^2 \chi_{\rho(t)}(r) \, r\mathrm{d}r + \Omega_{\rho(t)}(\boldsymbol{u}(t)), \tag{1.9}$$

where $\Omega_{\rho(t)}(\boldsymbol{u}(t))$ is the error created by the cut-off. Importantly, this error has structure, see Lemmas 2.4 and 4.18, and satisfies the estimates,

$$\Omega_{\rho(t)}(\boldsymbol{u}(t)) \lesssim (1 + |\rho'(t)|) \min\{E(\boldsymbol{u}(t); \rho(t), 2\rho(t)), \mathbf{d}(t)\}.$$

Roughly, this allows us to think of $\mathfrak{v}(t)$ as a Lyapunov functional for our problem, localized to scale $\rho(t)$, with "almost" critical points given by multi-bubbles $\boldsymbol{\mathcal{Q}}(m, \vec{\iota}, \vec{\lambda})$. Indeed, if $\boldsymbol{u}(t)$ is close to a multi-bubble up to scale $\rho(t)$, and $|\rho'(t)| \lesssim 1$, then $|\mathfrak{v}'(t)| \lesssim \mathbf{d}(t)$.

Our first result is a localized compactness lemma. In Section 3 we prove the following: given a sequence of wave maps $\boldsymbol{u}_n(t) \in \mathcal{E}_{\ell,m}$ on time intervals $[0, \tau_n]$ with bounded energy, and a sequence $R_n \to \infty$ such that

$$\lim_{n \to \infty} \frac{1}{\tau_n} \int_0^{\tau_n} \int_0^{R_n \tau_n} |\partial_t u_n(t,r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t = 0,$$

one can find a new sequence $1 \ll r_n \ll R_n$ and a sequence of times $s_n \in [0, \tau_n]$, so that up to passing to a subsequence of the $\boldsymbol{u}_n$, we have $\lim_{n \to \infty} \boldsymbol{\delta}_{r_n \tau_n}(\boldsymbol{u}_n(s_n)) = 0$. Here $\boldsymbol{\delta}_R(\boldsymbol{u})$ is a local (up to scale R) version of the distance function $\mathbf{d}$. We note that the sequential decomposition Theorem 1.15 is an almost immediate consequence of the localized compactness lemma along with the Shatah and Tahvildar-Zadeh theory; see Remark 3.2. The proof of the compactness lemma is very similar in spirit to the analysis of Côte [10] and Jia and Kenig [42].

We give a caricature of the no-return analysis, pointing the reader to the technical arguments in Sections 4, 5 for the actual arguments. We would like to integrate (1.9) over intervals $[a_n, b_n]$ with $a_n, b_n \to T_*$ such that $\mathbf{d}(a_n), \mathbf{d}(b_n) \ll 1$ but contain some subinterval $[c_n, d_n] \subset [a_n, b_n]$ on which $\mathbf{d}(t) \simeq 1$; such intervals exist under the contradiction hypothesis. From (1.9) we obtain,

$$\int_{a_n}^{b_n} \int_0^{\rho(t)} |\partial_t u(t,r)|^2 \, r \mathrm{d}r \, \mathrm{d}t \lesssim \rho(a_n)\mathbf{d}(a_n) + \rho(b_n)\mathbf{d}(b_n) + \int_{a_n}^{b_n} \left| \Omega_{\rho(t)}(\boldsymbol{u}(t)) \right| \, \mathrm{d}t. \quad (1.10)$$

We consider the choice of $\rho(t)$. One can use the sequential compactness lemma so that choosing $\rho(t)/(d_n - c_n) \gg 1$ we have,

$$\int_{c_n}^{d_n} \int_0^{\rho(t)} |\partial_t u(t,r)|^2 \, r \mathrm{d}r \, \mathrm{d}t \gtrsim d_n - c_n, \quad (1.11)$$

and one can expect that the integral of the error $\int_{c_n}^{d_n} \left| \Omega_{\rho(t)}(\boldsymbol{u}(t)) \right| \, \mathrm{d}t \ll |d_n - c_n|$ is absorbed into the left-hand side by choosing $\rho(t)$ to lie in a region where $\boldsymbol{u}(t)$ has negligible energy.

To complete the proof one would need to show that the error generated on the intervals $[a_n, c_n]$ and $[d_n, b_n]$ can also be absorbed into the left-hand side, and moreover that the terms $\rho(a_n)\mathbf{d}(a_n), \rho(b_n)\mathbf{d}(b_n) \ll d_n - c_n$. To accomplish this, we require a more careful choice of the intervals $[a_n, b_n]$ and placement of the cut-off $\rho(t)$, which motivates the notion of *collision intervals* introduced in Section 4.1. These allow us to distinguish between "interior" bubbles that come into collision, and "exterior" bubbles, which stay coherent throughout the intervals $[a_n, b_n]$, and to ensure we place the cutoff in the region between the interior and exterior bubbles.

Given $K \in \{1, \ldots, N\}$, we say that an interval $[a, b]$ is a collision interval with parameters $0 < \epsilon < \eta$ and $N - K$ exterior bubbles for some $1 \leq K \leq N$, if $\mathbf{d}(a), \mathbf{d}(b) \leq \epsilon$, there exists a $c \in [a, b]$ with $\mathbf{d}(c) \geq \eta$, and a curve $r = \rho_K(t)$ outside of which $\boldsymbol{u}(t) - \boldsymbol{u}^*(t)$ is within ϵ of an $N - K$-bubble in the sense of (1.8) (a localized version of $\mathbf{d}(t)$); see Definition 4.4. We now define K to be the *smallest* non-negative integer for which there exists $\eta > 0$, a sequence $\epsilon_n \to 0$, and sequences $a_n, b_n \to T_*$, so that $[a_n, b_n]$ are collision intervals with parameters ϵ_n, η and $N - K$ exterior bubbles, and we write $[a_n, b_n] \in \mathcal{C}_K(\epsilon_n, \eta)$; see Section 4.1 for the proof that K is well-defined and ≥ 1, under the contradiction hypothesis.

We revisit (1.10) on a sequence of collision intervals $[a_n, b_n] \in \mathcal{C}_K(\epsilon_n, \eta)$. Near the endpoints a_n, b_n, $\boldsymbol{u}(t) - \boldsymbol{u}^*(t)$ is close to an N-bubble configuration and we denote the interior scales, which will come into collision, by $\vec{\lambda} = (\lambda_1, \ldots, \lambda_K)$ and the exterior scales, which stay coherent, by $\vec{\mu} = (\vec{\mu}_{K+1}, \ldots, \vec{\mu}_N)$. We assume for simplicity in this discussion that the collision intervals have only a single subinterval $[c_n, d_n]$ as above, and that $\mathbf{d}(t)$ is sufficiently small on the intervals $[a_n, c_n]$ and $[d_n, b_n]$ so that the interior scales are well defined (via modulation theory) there. We call $[a_n, c_n], [d_n, b_n]$ *modulation intervals* and $[c_n, d_n]$ *compactness intervals*.

The scale of the Kth bubble $\lambda_K(t)$ plays an important role and must be carefully tracked. We will need to also make sense of this scale on the compactness intervals, where the bubble itself may lose its shape from time to time. We do this by energy considerations; see Definition 5.1. Crucially, the minimality of K can be used to ensure that the intervals $[c_n, d_n]$ as above satisfy $d_n - c_n \simeq \max\{\lambda_K(c_n), \lambda_K(d_n)\}$; see Lemma 5.4. Thus the first terms on the right-hand-side of (1.10) can be absorbed using (1.11) by ensuring $\rho(a_n) = o(\epsilon_n^{-1})\lambda_K(a_n), \rho(b_n) = o(\epsilon_n^{-1})\lambda_K(b_n)$ if we can additionally prove that the scale $\lambda_K(t)$ does not change much on the modulation intervals. Note that our choice of cut-off will satisfy $\lambda_K(t) \ll \rho(t) \ll \mu_{K+1}(t)$.

We must also absorb the errors $(\int_{a_n}^{c_n} + \int_{d_n}^{b_n})|\Omega_{\rho(t)}(\boldsymbol{u}(t))|\,\mathrm{d}t \lesssim (\int_{a_n}^{c_n} + \int_{d_n}^{b_n})\mathbf{d}(t)\,\mathrm{d}t$ on the modulation intervals. Here we perform a refined modulation analysis on the interior bubbles, which allows us to track the growth of $\mathbf{d}(t)$ through a collision of (possibly) many bubbles. Roughly, up to scale $\rho(t)$, $\boldsymbol{u}(t)$ looks like a K-bubble, and using the implicit function theorem we define modulation parameters $\vec{\iota}, \vec{\lambda}(t)$, and error $\boldsymbol{g}(t)$ with

$$\boldsymbol{u}(t,r) = \boldsymbol{\mathcal{Q}}(m_n, \vec{\iota}, \vec{\lambda}(t); r) + \boldsymbol{g}(t,r), \quad \text{if} \ \ r \le \rho(t), \quad \langle \Lambda Q_{\lambda_j(t)} \mid g(t) \rangle = 0, \ \ \text{for} \ \ j = 1, \ldots, K,$$

where $\Lambda := r\partial_r$ is the generator of the H-invariant scaling (note that for $k = 1, 2$ the decomposition is slightly different due to the slow decay of ΛQ) and

$$\langle \phi \mid g \rangle := \int_0^\infty \phi(r)g(r)\,r\mathrm{d}r, \qquad \text{for} \ \phi, g : (0, \infty) \to \mathbb{R}. \tag{1.12}$$

The orthogonality conditions and an expansion of the nonlinear energy of $\boldsymbol{u}(t)$ up to scale $\rho(t)$ lead to the coercivity estimate,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{j \notin \mathcal{A}} \left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^{\frac{k}{2}} \lesssim \max_{i \in \mathcal{A}} \left(\frac{\lambda_i(t)}{\lambda_{i+1}(t)}\right)^{\frac{k}{2}} + o_n(1) \simeq \mathbf{d}(t) + o_n(1),$$

where $\mathcal{A} = \{j \in 1, \ldots, K-1 : \iota_j \ne \iota_{j+1}\}$ captures the alternating bubbles (which experience an attractive interaction force) and the $o_n(1)$ term comes from errors due to the presence of the radiation $\boldsymbol{u}^*$ in the region $r \lesssim \rho(t) \ll t$. In fact, since $\mathbf{d}(t)$ grows out of the modulation intervals we can absorb these errors into $\mathbf{d}(t)$ by enlarging the parameter ϵ_n and requiring the lower bound $\mathbf{d}(t) \ge \epsilon_n$ on the modulation intervals.

The growth of $\mathbf{d}(t)$ is then captured by the dynamics of the alternating bubbles, which, since (1.1) is second order, enter at the level of $\lambda_j''(t)$. However, it is not clear how to derive useful estimates from the equation for $\lambda''(t)$ obtained by twice differentiating the orthogonality conditions. To cancel terms with critical size, but indeterminate sign, we introduce a localized virial correction to $\lambda_j' \simeq -\iota_j\|\Lambda Q\|_{L^2}^{-2}\lambda_j^{-1}\langle \Lambda Q_{\lambda_j} \mid \dot{g} \rangle$, defining

$$\beta_j(t) = -\iota_j\|\Lambda Q\|_{L^2}^{-2}\langle \Lambda Q_{\lambda_j(t)} \mid \dot{g}(t) \rangle - \|\Lambda Q\|_{L^2}^{-2}\left\langle \underline{A}(\lambda_j(t))g(t) \mid \dot{g}(t) \right\rangle,$$

where $\underline{A}(\lambda)$ is a truncated (to scale λ) version of $\underline{\Lambda} = \Lambda + 1$, the generator of L^2 scaling. Roughly, we show in Sections 4.3 and 5.2, that if the distance $\mathbf{d}(t)$ is dominated at a local minimal point t_0 by the ratio between the j-th bubble and its larger neighbor with opposite sign, then we can control dynamics of $\beta_j(t)$ near t_0, showing that $\mathbf{d}(t)$ grows in a controlled way until some other bubble ratio becomes dominant, and so on, until we exit the modulation interval. All the while we can ensure that the Kth scale does not move much, and we obtain bounds of the form $(\int_{a_n}^{c_n} + \int_{d_n}^{b_n})\mathbf{d}(t)\,\mathrm{d}t \lesssim \mathbf{d}(c_n)^{\frac{2}{k}}\lambda_K(a_n) + \mathbf{d}(d_n)^{\frac{2}{k}}\lambda_K(b_n)$ (see the "ejection" Lemma 5.8). Thus the errors can be absorbed into the left-hand side of (1.10) and we obtain a contradiction.

A similar, but simpler refined modulation analysis was performed in [41]. The use of such refinements to modulation parameters to obtain dynamical control was introduced by the first author in the context of a two-bubble construction for NLS in [35]. The notion of localized

virial corrections in the context of energy/Morawetz-type estimates was developed by Raphaël and Szeftel in [63].

1.5. **Notational conventions.** The energy is denoted E, $\mathcal{E}$ is the energy space, $\mathcal{E}_{\ell,m}$ are the finite energy sectors.

Given a function $\phi(r)$ and $\lambda > 0$, we denote by $\phi_\lambda(r) = \phi(r/\lambda)$, the H-invariant re-scaling, and by $\phi_{\underline{\lambda}}(r) = \lambda^{-1}\phi(r/\lambda)$ the L^2-invariant re-scaling. We denote by $\Lambda := r\partial_r$ and $\underline{\Lambda} := r\partial r + 1$ the infinitesimal generators of these scalings. We denote $\langle \cdot \mid \cdot \rangle$ the radial $L^2(\mathbb{R}^2)$ inner product given by (1.12).

We denote k the equivariance degree and $f(u) := \frac{1}{2}\sin 2u$ the nonlinearity in (1.1). We let χ be a smooth cut-off function, supported in $r \leq 2$ and equal 1 for $r \leq 1$.

The general rules we follow giving names to various objects are:
- index of an infinite sequence: n
- sequences of small numbers: $\gamma, \delta, \epsilon, \zeta, \eta, \theta$
- scales of bubbles and quantities describing the spatial scales: $\lambda, \mu, \nu, \xi, \rho$; in general we call λ the scale of the interior bubbles and μ the exterior ones (once these notions are defined)
- moment in time: $t, s, \tau, a, b, c, d, e, f$
- indices in summations: i, j, ℓ
- time intervals: I, J
- number of bubbles: K, M, N
- signs are denoted ι and σ
- boldface is used for pairs of elements related to the Hamiltonian structure; an arrow is used for vectors (finite sequences) in other contexts.

We call a "constant" a number which depends only on the equivariance degree k and the number of bubbles N. Constants are denoted C, C_0, C_1, c, c_0, c_1. We write $A \lesssim B$ if $A \leq CB$ and $A \gtrsim B$ if $A \geq cB$. We write $A \ll B$ if $\lim_{n\to\infty} A/B = 0$.

For any sets X, Y, Z we identify $Z^{X \times Y}$ with $(Z^Y)^X$, which means that if $\phi : X \times Y \to Z$ is a function, then for any $x \in X$ we can view $\phi(x)$ as a function $Y \to Z$ given by $(\phi(x))(y) := \phi(x, y)$.

2. Preliminaries

2.1. **Basic properties of finite energy maps.** We aggregate here several well known results.

Lemma 2.1. *Fix integers ℓ, m. For every $\epsilon > 0$ and $R_0 > 1$, there exists a $\delta > 0$ with the following property. Let $0 \leq R_1 < R_2 \leq \infty$ with $R_2/R_1 \geq R_0$, and $\boldsymbol{u} \in \mathcal{E}_{\ell,m}$ be such that $E((u, 0); R_1, R_2) < \delta$. Then, there exists $\ell_0 \in \mathbb{Z}$ such that $|u(r) - \ell_0\pi| < \epsilon$ for almost all $r \in (R_1, R_2)$.*

Moreover, there exist constants $C = C(R_0), \alpha = \alpha(R_0) > 0$ such that if $E((u, 0); R_1, R_2) < \alpha$, then

$$\|\boldsymbol{u} - \ell_0\boldsymbol{\pi}\|_{\mathcal{E}(R_1, R_2)} \leq CE(\boldsymbol{u}; R_1, R_2). \tag{2.1}$$

Proof. By an approximation argument we can assume $(u, 0) \in \mathcal{E}_{\ell,m}$ is smooth. First, we show that for any $\epsilon_0 > 0$, there exists $r_0 \in [R_1, R_2]$ such that $|u(r_0) - \ell_0\pi| < \epsilon_0$ for some $\ell_0 \in \mathbb{Z}$ as long as $E((u, 0); R_1, R_2)$ is sufficiently small. If not, one could find $\epsilon_1 > 0$, $0 < R_1 < R_2$, and a sequence $(u_n, 0) \in \mathcal{E}_{\ell,m}$ so that $E((u_n, 0); R_1; R_2) \to 0$ as $n \to \infty$ but such that $\inf_{r\in[R_1, R_2], \ell\in\mathbb{Z}} |u_n(r) - \ell\pi| \geq \epsilon_1$. The latter condition gives a constant $c(\epsilon_1) > 0$ such that $\inf_{r\in[R_1, R_2]} |\sin(u_n(r))| \geq c(\epsilon_1)$. But then

$$E((u_n, 0); R_1; R_2) \geq \frac{k^2}{2} \int_{R_1}^{R_2} \sin^2(u_n(r)) \frac{\mathrm{d}r}{r} \geq \frac{k^2}{2} c(\epsilon_1)^2 \log(R_2/R_1),$$

which is a contradiction. Next define the function, $G(u) = \int_0^u |\sin \rho| \, d\rho$, and for $r_1 \in (R_1, R_2)$ note the inequality,

$$|G(u(r_0)) - G(u(r_1))| = \left| \int_{u(r_1)}^{u(r_0)} |\sin \rho| \, d\rho \right| = \left| \int_{r_1}^{r_0} |\sin u(r)| \, |\partial_r u(r)| \, dr \right| \lesssim E((u,0); R_1, R_2).$$

We conclude using that G is continuous and increasing that $|u(r) - \ell_0 \pi| < \epsilon$ for all $r \in (R_1, R_2)$. As long as $\epsilon > 0$ is small enough we see that in fact, $\sin^2(u(r)) \geq \frac{1}{2} |u(r) - \ell_0 \pi|^2$ for all $r \in (R_1, R_2)$ and (2.1) follows. $\square$

We have the following version of the principle of finite speed of propagation.

Lemma 2.2. *Let $\boldsymbol{u}(t)$ be a solution to (1.1) on the time interval $[0, T]$. Then*

$$E(\boldsymbol{u}(T); 0, R - T) \leq E(\boldsymbol{u}(0); 0, R), \qquad \text{for all } R \geq T.$$

Proof. It suffices to consider the case of a smooth solution and then approximate a finite energy solution by smooth ones. For a proof in the smooth case, see [70, Section 2]. $\square$

Remark 2.3. The energy conservation yields the following equivalent formulation:

$$E(\boldsymbol{u}(T); R + T) \leq E(\boldsymbol{u}(0); R), \qquad \text{for all } R \geq 0. \tag{2.2}$$

We have the following virial identity.

Lemma 2.4 (Virial identity)**.** *Let $\boldsymbol{u}(t)$ be a solution to (1.1) on an open time interval I and $\rho : I \to (0, \infty)$ a Lipschitz function. Then for almost all $t \in I$,*

$$\frac{d}{dt} \left\langle \partial_t u(t) \mid \chi_{\rho(t)}^2 \, r \partial_r u(t) \right\rangle = -\int_0^\infty (\partial_t u(t,r) \chi_{\rho(t)}(r))^2 \, r dr + \Omega_{\rho(t)}(\boldsymbol{u}(t)),$$

where

$$\Omega_{\rho(t)}(\boldsymbol{u}(t)) := -2 \frac{\rho'(t)}{\rho(t)} \int_0^\infty \partial_t u(t,r) r \partial_r u(t,r) \chi_{\rho(t)}(r) \Lambda \chi_{\rho(t)}(r) \, r dr$$

$$- \int_0^\infty \left((\partial_t u(t,r))^2 + (\partial_r u(t,r))^2 - k^2 \frac{\sin^2 u(t,r)}{r^2} \right) \chi_{\rho(t)}(r) \Lambda \chi_{\rho(t)}(r) \, r dr.$$

Proof. The proof is a direct computation along with an approximation argument for fixed $t \in I$, assuming ρ is differentiable at t. $\square$

2.2. **Local Cauchy theory.** The following theorem was proved by Shatah and Tahvildar-Zadeh in [70, 71].

Lemma 2.5 (Local well-posedness)**.** [71, Theorem 1.1], [69, Theorem 8.1] [70] *Let $\ell, m \in \mathbb{Z}$ and let $\boldsymbol{u}_0 \in \mathcal{E}_{\ell,m}$. Then, there exist a maximal time interval of existence $(T_-, T_+) = I_{\max}(\boldsymbol{u}_0) \ni 0$ on which (1.1) admits a unique solution $\boldsymbol{u}(t)$ in the space $C^0(I_{\max}; \mathcal{E}_{\ell,m})$ with $\boldsymbol{u}(0) = \boldsymbol{u}_0$.*

In fact, there exists $\epsilon_0 > 0$ with the following property. Let $\boldsymbol{u}_0 \in \mathcal{E}_{\ell,m}$, $\tau > 0$ and suppose the solution $\boldsymbol{u}(t)$ to (1.1) with data $\boldsymbol{u}(0) = \boldsymbol{u}_0$ is defined on the interval $[0, \tau)$, i.e., in $C^0([0, \tau); \mathcal{E}_{\ell,m})$. Suppose that there exists a time $0 \leq t < \tau$ and a number $R > \tau - t$ such that,

$$E(\boldsymbol{u}(t), 0; R) < \epsilon_0.$$

Then, $T_+(\boldsymbol{u}) > \tau$.

See Struwe [73, p. 817] for the continuation criterion in the second paragraph of Lemma 2.5 in the case of smooth initial data, and see [69, Theorem 8.1] for the global well-posedness theorem for energy class equivariant wave maps with sufficiently small energy. Key to the proof are Strichartz estimates for the wave equation (see, e.g., Lindblad, Sogge [47], and Ginibre, Velo [33]), after noticing that the linearization of (1.1) about the zero solution is equivalent, in the energy space, to the free scalar wave equation in dimension $d = 2k + 2$. Indeed, the linearization of (1.1) about the zero solution is given by the linear wave equation,

$$\partial_t^2 v - \partial_r^2 v - \frac{1}{r}\partial_r v + \frac{k^2}{r^2}v = 0. \tag{2.3}$$

We will sometimes use the notation $v_{\mathrm{L}}(t) = S_{\mathrm{L}}(t)\boldsymbol{v}_0$ as the unique solution to (2.3) with initial data $\boldsymbol{v}_{\mathrm{L}}(0) = \boldsymbol{v}_0 \in \mathcal{E}$. The mapping $\mathcal{E} \ni \boldsymbol{v}(t) \mapsto \boldsymbol{W}(t) \in (\dot{H}^1 \times L^2)_{\mathrm{rad}}(\mathbb{R}^{2k+2})$ defined by

$$\boldsymbol{W}(t,r) := (r^{-k}v(t,r), r^{-k}\partial_t v(t,r))$$

satisfies $\|\boldsymbol{v}(t)\|_{\mathcal{E}} \simeq \|\boldsymbol{W}(t)\|_{(\dot{H}^1 \times L^2)_{\mathrm{rad}}(\mathbb{R}^{2k+2})}$ and $\boldsymbol{v}(t) \in \mathcal{E}$ solves (2.3) if and only if $\boldsymbol{W}(t) \in (\dot{H}^1 \times L^2)_{\mathrm{rad}}$ solves

$$\partial_t^2 W - \Delta_{2k+2} W = 0, \tag{2.4}$$

where $\Delta_{2k+2} = \partial_r^2 + \frac{2k+1}{r}\partial_r$ is the radial Laplacian in dimension $d = 2k + 2$.

For equivariance classes $k > 2$, this leads to a spatial dimension $d > 6$ and inconvenient technical complications. However, we observed in [41] that one may give a unified local Cauchy theory for (1.1) for all equivariance classes $k \in \mathbb{N}$ based on Strichartz estimates for linear waves with a critical repulsive potential proved by Planchon, Stalker, Tahvildar-Zadeh [59]. For this purpose, consider the mapping,

$$v_0(r) \mapsto V_0(r) := r^{-1}v_0(r), \quad \dot{v}_0(r) \mapsto \dot{V}_0(r) := r^{-1}\dot{v}_0(r). \tag{2.5}$$

We see that $\boldsymbol{v}(t) = (v(t), \partial_t v(t))$ solves (2.3) if and only if $\boldsymbol{V}(t) = (v(t), \partial_t v(t))$ solves

$$\partial_t^2 V - \partial_r^2 V - \frac{3}{r}\partial_r V + \frac{k^2-1}{r^2}V = 0. \tag{2.6}$$

For each $k \geq 1$, define the norm H_k for radially symmetric functions V on $\mathbb{R}^4$ by

$$\|V\|_{H_k(\mathbb{R}^4)}^2 := \int_0^\infty \left[(\partial_r V)^2 + \frac{(k^2-1)}{r^2}V^2 \right] r^3 \, \mathrm{d}r.$$

Solutions $\boldsymbol{V}(t)$ to (2.6) conserve the $H_k \times L^2$ norm and by Hardy's inequality we have

$$\|V\|_{H_k(\mathbb{R}^4)} \simeq \|V\|_{\dot{H}^1(\mathbb{R}^4)}.$$

Thus the mapping (2.5) satisfies

$$\|(V_0, \dot{V}_0)\|_{\dot{H}^1 \times L^2(\mathbb{R}^4)} \simeq \|(V_0, \dot{V}_0)\|_{H_k \times L^2(\mathbb{R}^4)} = \|(v_0, \dot{v}_0)\|_{H \times L^2(\mathbb{R}^2)}.$$

We conclude that the Cauchy problem for (2.6) with initial data in $\dot{H}^1 \times L^2(\mathbb{R}^4)$ is equivalent to the Cauchy problem for (2.3) for initial data $(v_0, \dot{v}_0) \in H \times L^2$. As a consequence, Strichartz estimates for solutions to (2.3) are inherited from Strichartz estimates for (2.6) proved by Planchon, Stalker, and Tahvildar-Zadeh [59].

Lemma 2.6 (Strichartz estimates for (2.6)). [59, Corollary 3.9] *Fix $k \geq 1$ and let $\boldsymbol{V}(t)$ be a radial solution to the linear equation*

$$\partial_t^2 V - \partial_r^2 V - \frac{3}{r}\partial_r V + \frac{k^2-1}{r^2}V = F(t,r), \quad \boldsymbol{V}(0) = (V_0, \dot{V}_0) \in \dot{H}^1 \times L^2(\mathbb{R}^4).$$

Then, for any time interval $0 \in I \subset \mathbb{R}$ we have

$$\|V\|_{(L_t^3 L_x^6 \cap L_{t,x}^5)(I \times \mathbb{R}^4)} + \sup_{t \in I} \|\boldsymbol{V}(t)\|_{\dot{H}^1 \times L^2(\mathbb{R}^4)} \lesssim \|\boldsymbol{V}(0)\|_{\dot{H}^1 \times L^2(\mathbb{R}^4)} + \|F\|_{L_t^1 L_x^2(I \times \mathbb{R}^4)},$$

where the implicit constant above is independent of I.

We define the Strichartz norm,

$$\|v\|_{\mathcal{S}(I)} := \|r^{-\frac{3}{5}} v\|_{L_{t,r}^5(I)} + \|r^{-\frac{2}{3}} v\|_{L_t^3 L_r^6(I)}$$

and recall that the notation L_r^p refers to the Lebesgue space on $(0, \infty)$ with respect to the measure $r \, dr$.

Corollary 2.7 (Strichartz estimates for (2.3)). *Fix $k \geq 1$ and let $\boldsymbol{v}(t)$ be a radial solution to the linear equation*

$$\partial_t^2 v - \partial_r^2 v - \frac{1}{r} \partial_r v + \frac{k^2}{r^2} v = F(t, r), \quad \boldsymbol{v}(0) = (v_0, \dot{v}_0) \in \mathcal{E} = H \times L^2.$$

Then, for any time interval $0 \in I \subset \mathbb{R}$ we have

$$\|v\|_{\mathcal{S}(I)} + \|\boldsymbol{v}(t)\|_{L_t^\infty (H \times L^2)(I)} \lesssim \|\boldsymbol{v}(0)\|_{H \times L^2} + \|F\|_{L_t^1 L_r^2(I)},$$

where the implicit constant above is independent of I.

Writing the Cauchy problem for (1.1) in the class $\mathcal{E} = \mathcal{E}_{0,0}$ as

$$\partial_t^2 u - \Delta u + \frac{k^2}{r^2} u = \frac{k^2}{2r^2}(2u - \sin 2u) \tag{2.7}$$
$$\boldsymbol{u}(0) = (u_0, u_1) \in \mathcal{E} = H \times L^2.$$

a standard argument based on the contraction mapping principle yields the following result; see for example [14].

Lemma 2.8 (Cauchy theory in $\mathcal{E}_{0,0}$). *There exist $\delta_0 > 0$ and a functions $C : [0, \infty) \to (0, \infty)$ with the following properties. Let $A \geq 0$ and let $\boldsymbol{u}_0 = (u_0, u_1) \in \mathcal{E}$ with $\|\boldsymbol{u}_0\|_{\mathcal{E}} \leq A$. Let $I \ni 0$ be an open interval such that*

$$\|S_{\mathrm{L}}(t)\boldsymbol{u}_0\|_{\mathcal{S}(I)} = \delta \leq \delta_0.$$

Then there exists a unique solution $\boldsymbol{u}(t)$ to (2.7) in the space $C^0(I; \mathcal{E}) \cap \mathcal{S}(I)$ with initial data $\boldsymbol{u}(0) = \boldsymbol{u}_0$. Moreover, $\boldsymbol{u}(t)$ satisfies the bounds $\|u\|_{\mathcal{S}(I)} \leq C(A)\delta$, and $\|\boldsymbol{u}\|_{L_t^\infty(I;\mathcal{E})} \leq C(A)$. To each solution $\boldsymbol{u}(t)$ to (2.7) we can associate a maximal interval of existence $I_{\max}(\boldsymbol{u})$ such that for each compact subinterval $I' \subset I_{\max}$ we have $\|u\|_{\mathcal{S}(I')} < \infty$.

Moreover, small data scattering holds: there exists ϵ_0 small enough so that if $\boldsymbol{u}_0 \in \mathcal{E}$ satisfies $E(\boldsymbol{u}_0) < \epsilon_0$, the solution $\boldsymbol{u}(t)$ given above is defined globally in time, satisfies the bound,

$$\sup_{t \in \mathbb{R}} \|\boldsymbol{u}(t)\|_{\mathcal{E}} + \|u\|_{\mathcal{S}(\mathbb{R})} \lesssim \|\boldsymbol{u}_0\|_{\mathcal{E}}$$

and scatters in the following sense: there exist solutions $\boldsymbol{u}_{\mathrm{L}}^\pm(t) \in \mathcal{E}$ to (2.3) such that

$$\|\boldsymbol{u} - \boldsymbol{u}_{\mathrm{L}}^\pm(t)\|_{\mathcal{E}} \to 0 \quad as \quad t \to \pm\infty \tag{2.8}$$

Conversely, the existence of wave operators holds, i.e., for any solution $\boldsymbol{v}_{\mathrm{L}}(t) \in \mathcal{E}$ to the linear equation (2.3), there exists a unique, global-in-forward time solution $\boldsymbol{u}(t) \in \mathcal{E}$ to (2.7) such that (2.8) holds as $t \to \infty$. An analogous statement holds for negative times.

We make note of the following estimate proved in [10], which is relevant for the vanishing of the error in the linear profile decomposition stated in the next section.

Lemma 2.9. [10, Lemma 2.11] *There exists a uniform constant $C > 0$ such that every solution $\boldsymbol{v}(t) \in \mathcal{E}$ to (2.3) satisfies,*

$$\|v\|_{L^\infty_{t,r}(\mathbb{R})} \le C\|\boldsymbol{v}(0)\|_{\mathcal{E}}^{\frac{3}{8}}\|r^{-\frac{3}{5}}v\|_{L^5_{t,r}(\mathbb{R})}^{\frac{5}{8}}.$$

2.3. Profile decomposition.

Bahouri-Gérard-type linear profile decompositions [1] are an essential ingredient in the study of solutions to (1.1); see also [3, 32, 48, 49, 54]. We make explicit use of a version adapted to sequences of functions in the affine spaces $\mathcal{E}_{\ell,m}$ proved by Jia and Kenig in [42], which synthesized Côte's analysis in [10]; see also [11] which treats sequences in $\mathcal{E}_{0,0}$.

Lemma 2.10 (Linear profile decomposition). [42, Lemma 5.5] [1] *Let $\ell, m \in \mathbb{Z}$ and let $\boldsymbol{u}_n$ be a sequence in $\mathcal{E}_{\ell,m}$ with $\limsup_{n\to\infty} E(\boldsymbol{u}_n) < \infty$. Then, there exists $K_0 \in \{0, 1, 2, \dots\}$, sequences $\lambda_{n,j} \in (0, \infty)$ for $j \in \{1, \dots, K_0\}$, $\sigma_{n,i} \in (0, \infty)$, and $t_{n,i} \in \mathbb{R}$, as well as mappings $\boldsymbol{\psi}^j \in \mathcal{E}_{\ell_j, m_j}$ with $E(\boldsymbol{\psi}^j) < \infty$ and finite energy solutions $\boldsymbol{v}^i_{\mathrm{L}}$ to (2.3) such that for each $J \ge 1$,*

$$\boldsymbol{u}_n = m\boldsymbol{\pi} + \sum_{j=1}^{K_0}\left(\left(\psi^j\left(\frac{\cdot}{\lambda_{n,j}}\right), \frac{1}{\lambda_{n,j}}\dot\psi^j\left(\frac{\cdot}{\lambda_{n,j}}\right)\right) - m_j\boldsymbol{\pi}\right)$$

$$+ \sum_{i=1}^{J}\left(v^i_{\mathrm{L}}\left(\frac{-t_{n,i}}{\sigma_{n,i}}, \frac{\cdot}{\sigma_{n,i}}\right), \frac{1}{\sigma_{n,i}}\partial_t v^i_{\mathrm{L}}\left(\frac{-t_{n,i}}{\sigma_{n,i}}, \frac{\cdot}{\sigma_{n,i}}\right)\right) + \boldsymbol{w}^J_{n,0}(\cdot)$$

where, denoting by $\boldsymbol{w}^J_{n,\mathrm{L}}(t)$ the solution to the linear wave equation (2.3) with initial data $\boldsymbol{w}^J_{n,0}$, the following hold:

- *the parameters $\lambda_{n,j}$ satisfy*

$$\lambda_{n,1} \ll \lambda_{n,2} \ll \cdots \ll \lambda_{n,K_0} \quad as \quad n \to \infty;$$

  and for each j one of $\lambda_{n,j} \to 0$, $\lambda_{n,j} = 1$ for all n, or $\lambda_{n,j} \to \infty$ as $n \to \infty$, holds;

- *for each i, either $t_{n,i} = 0$ for all n or $\lim_{n\to\infty}\frac{-t_{n,i}}{\sigma_{n,i}} = \pm\infty$. Either $\sigma_{n,i} \to 0$, $\sigma_{n,i} = 1$ for all n, or $\sigma_{n,i} \to \infty$ as $n \to \infty$;*

- *for each $i \in \mathbb{N}$,*

$$\frac{\lambda_{n,j}}{\sigma_{n,i}} + \frac{\sigma_{n,i}}{\lambda_{n,j}} + \frac{|t_{n,i}|}{\lambda_{n,j}} \to \infty \quad as \quad n \to \infty \quad \forall j = 1, \dots, K_0;$$

- *the scales $\sigma_{n,i}$ and times $t_{n,i}$ satisfy,*

$$\frac{\sigma_{n,i}}{\sigma_{n,i'}} + \frac{\sigma_{n,i'}}{\sigma_{n,i}} + \frac{|t_{n,i} - t_{n,i'}|}{\sigma_{n,i}} \to \infty \quad as \quad n \to \infty;$$

- *the integers ℓ_j and m_j satisfy, $|\ell_j - m_j| \ge 1$, and,*

$$\ell = m + \sum_{j=1}^{K_0}(\ell_j - m_j);$$

- *the error term $\boldsymbol{w}^J_n$ satisfies,*

$$(w^J_{n,0}(\lambda_{n,j}\cdot), \lambda_{n,j}\dot w^J_{n,0}(\lambda_{n,j}\cdot)) \rightharpoonup 0 \in \mathcal{E} \quad as \quad n \to \infty$$

$$(w^J_{n,\mathrm{L}}(t_{n,i}, \sigma_{n,i}\cdot), \sigma_{n,i}\partial_t w^J_{n,\mathrm{L}}(t_{n,i}, \sigma_{n,i}\cdot)) \rightharpoonup 0 \in \mathcal{E} \quad as \quad n \to \infty$$

  for each $J \ge 1$, each $j = 1, \dots, K_0$, and $i \in \mathbb{N}$, and vanishes strongly in the sense that

$$\lim_{J\to\infty}\limsup_{n\to\infty}\left(\|w^J_{n,\mathrm{L}}\|_{L^\infty_{t,r}(\mathbb{R})} + \|w_{n,\mathrm{L}}\|_{\mathcal{S}(\mathbb{R})}\right) = 0;$$

- *the following pythagorean decomposition of the nonlinear energy holds: for each $J \geq 1$,*

$$E(\boldsymbol{u}_n) = \sum_{j=1}^{K_0} E(\boldsymbol{\psi}^j) + \sum_{i=1}^{J} E\big((v_{\mathrm{L}}^j(-t_{n,i}/\sigma_{n,i}), \partial_t v_{\mathrm{L}}^j(-t_{n,i}/\sigma_{n,i}))\big) + E(\boldsymbol{w}_n^J) + o_n(1) \quad (2.9)$$

as $n \to \infty$.

Remark 2.11. The pythagorean expansion of the nonlinear energy in the case $K_0 = 0$ was treated in [11, Lemma 2.16]. The case with $K_0 \geq 1$ was treated in the recent preprint [18, Appendix B.2].

Remark 2.12. We call the pairs $(\boldsymbol{\psi}^j, \lambda_{n,j})$ and the triplets $(\boldsymbol{v}_{\mathrm{L}}^i, \sigma_{n,i}, t_{n,i})$ *profiles*. Following Bahouri and Gérard [1] we refer to the profiles $(\boldsymbol{\psi}^j, \lambda_{n,j})$ and the profiles $(\boldsymbol{v}_{\mathrm{L}}^i, \sigma_{n,i}, 0)$ as *centered*, to the profiles $(\boldsymbol{v}_{\mathrm{L}}^i, \sigma_{n,i}, t_{n,i})$ with $-t_{n,i}/\sigma_{n,i} \to \infty$ as $n \to \infty$ as *outgoing*, and those with $-t_{n,i}/\sigma_{n,i} \to -\infty$ as *incoming*.

In Section 3 we will need to evolve the linear profiles via the flow for (1.1) in the special case when all of the centered profiles are given by harmonic maps. In this setting we define *nonlinear profiles* as follows. Given a profile $(\boldsymbol{v}_{\mathrm{L}}^i, \sigma_{n,i}, t_{n,i})$ as in Lemma 2.10 we define the corresponding nonlinear profile, $(\boldsymbol{v}_{\mathrm{NL}}^i, \sigma_{n,i}, t_{n,i})$ as the unique solution to (1.1) such that for all $-t_{n,i}/\sigma_{n,i} \in I_{\max}(\boldsymbol{v}_{\mathrm{NL}})$ we have,

$$\lim_{n \to \infty} \|\boldsymbol{v}_{\mathrm{NL}}^i(-\frac{t_{n,i}}{\sigma_{n,i}}) - \boldsymbol{v}_{\mathrm{L}}^i(-\frac{t_{n,i}}{\sigma_{n,i}})\|_{\mathcal{E}} = 0$$

The existence of nonlinear profiles follows from the local Cauchy theory in Lemma 2.8 in the case of a centered linear profile, i.e., $t_{n,i} = 0$, and from the existence of wave operators statement in Lemma 2.8 in the case of outgoing/incoming profiles, i.e., $-t_{n,i}/\sigma_{n,i} \to \pm\infty$.

Lemma 2.13 (Nonlinear profile decomposition). *Let $\ell, m \in \mathbb{Z}$ and let $\boldsymbol{u}_n$ be a sequence in $\mathcal{E}_{\ell,m}$ with $\limsup_{n \to \infty} E(\boldsymbol{u}_n) < \infty$. Assume the linear profile decomposition for $\boldsymbol{u}_n$ given by the Lemma 2.10 takes the form*

$$\boldsymbol{u}_n = m\boldsymbol{\pi} + \sum_{j=1}^{K_0} \big((Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}\big) + \sum_{i=1}^{J} \Big(v_{\mathrm{L}}^i\big(\tfrac{-t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big), \tfrac{1}{\sigma_{n,i}}\partial_t v_{\mathrm{L}}^i\big(\tfrac{-t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big)\Big) + \boldsymbol{w}_{n,0}^J(\cdot),$$

that is, all of the profiles $(\boldsymbol{\psi}^j, \lambda_{n,j})$ for $1 \leq j \leq K_0$ as in Lemma 2.10 are given by harmonic maps $(Q, \lambda_{n,j})$. There exists a constant $\delta_0 > 0$ sufficiently small with the following properties. Let $i_0 \in \mathbb{N}$, $\tau_0 > 0$ and assume that for each $i \in \mathbb{N}$, and for each $1 \leq j \leq K_0$,

$$\frac{\tau_0 \sigma_{n,i_0} - t_{n,i}}{\sigma_{n,i}} < T_{+,i}(\boldsymbol{v}_{\mathrm{NL}}^i), \quad \limsup_{n \to \infty} \|v_{\mathrm{NL}}^i\|_{\mathcal{S}((-\frac{t_{n,i}}{\sigma_{n,i}}, \frac{\tau_0 \sigma_{n,i_0} - t_{n,i}}{\sigma_{n,i}}))} < \infty,$$

$$\text{and} \quad \tau_0 \frac{\sigma_{n,i_0}}{\lambda_{n,j}} \leq \delta_0, \tag{2.10}$$

for all n. Then for each n sufficiently large, the wave map evolution $\boldsymbol{u}_n(t)$ of the data $\boldsymbol{u}_n(0) = \boldsymbol{u}_n$ is defined on the interval $[0, \tau_0 \sigma_{n,i_0}]$ and the following nonlinear profile decomposition holds: for each $t \in [0, \tau_0 \sigma_{n,i_0}]$ the sequence $\boldsymbol{z}_n^J(t)$ defined by

$$\boldsymbol{u}_n(t) = m\boldsymbol{\pi} + \sum_{j=1}^{K_0} \big((Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}\big) + \sum_{i=1}^{J} \Big(v_{\mathrm{NL}}^i\big(\tfrac{t - t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big), \tfrac{1}{\sigma_{n,i}}\partial_t v_{\mathrm{NL}}^i\big(\tfrac{t - t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big)\Big)$$

$$+ \boldsymbol{w}_{n,\mathrm{L}}^J(t) + \boldsymbol{z}_n^J(t),$$

satisfies,

$$\lim_{J\to\infty}\limsup_{n\to\infty}\Big(\sup_{t\in[0,\tau_0\sigma_{n,i_0}]}\|\boldsymbol{z}_n^J(t)\|_{\mathcal{E}}+\|z_n^J\|_{\mathcal{S}([0,\tau_0\sigma_{n,i_0}])}\Big)=0.$$

Remark 2.14. We note that the harmonic maps in the nonlinear profile decomposition are static solutions to (1.1), but their presence in a linear profile decomposition may lead to eventual singularities in the nonlinear flow. This leads us to the hypothesis in the second line of (2.10), which ensures that we are only considering the nonlinear evolution of $\boldsymbol{u}_n(t)$ on time intervals shorter than the length scales of the harmonic maps, thus avoiding the possibility of a singularity.

The key ingredient in the proof of Lemma 2.13 is the following modification of the now standard nonlinear perturbation lemma [43, Theorem 2.20]; see also [1, Section IV].

Lemma 2.15 (Nonlinear perturbation lemma). *Fix integers ℓ, m. There are continuous functions $\epsilon_0, C_0 : (0,\infty) \to (0,\infty)$ with the following properties. Let I be an open interval and let $\boldsymbol{u}, \boldsymbol{v} \in C^0(I; \mathcal{E}_{\ell,m})$ such that for some $A \geq 0$,*

$$\|\boldsymbol{u}-\boldsymbol{v}\|_{L_t^\infty(I;\mathcal{E})}+\|r^{-\frac{2}{3}}\sin v\|_{L_t^3(I;L_r^6)}\leq A$$

and

$$\|\mathrm{eq}(u)\|_{L_t^1(I;L_r^2)}+\|\mathrm{eq}(v)\|_{L_t^1(I;L_r^2)}+\|\boldsymbol{w}_0\|_{\mathcal{S}(I)}\leq\epsilon\leq\epsilon_0(A)$$

where $\mathrm{eq}(u) := \partial_t^2 u - \Delta u + k^2 r^{-2} f(u)$ in the sense of distributions, and $\boldsymbol{w}_0(t) := \vec{S}_{\mathrm{L}}(t-t_0)(\boldsymbol{u}(t_0)-\boldsymbol{v}(t_0))$ is the linear evolution of the difference, i.e., the solution to (2.3), where $t_0 \in I$ is arbitrary, but fixed, and where $\vec{S}_{\mathrm{L}}(t)\boldsymbol{v} := (S_{\mathrm{L}}(t)\boldsymbol{v}, \partial_t S_{\mathrm{L}}(t)\boldsymbol{v})$. Then,

$$\|\boldsymbol{u}(t)-\boldsymbol{v}(t)-\boldsymbol{w}_0(t)\|_{L^\infty(I;\mathcal{E})}+\|u-v\|_{\mathcal{S}(I)}\leq C_0(A)\epsilon.$$

Proof of Lemma 2.15. Let $X(I)$ denote the space $L_t^3(I; L_r^6)$ in this proof. Define $\boldsymbol{w}(t) := \boldsymbol{u}(t) - \boldsymbol{v}(t)$ and let $e := \partial_t^2 u - \Delta u + k^2 r^{-2}\sin(u)\cos(u) - (\partial_t^2 v - \Delta v + k^2 r^{-2}\sin(v)\cos(v)) = \mathrm{eq}(u) - \mathrm{eq}(v)$. Let $t_0 \in I$, fix a small constant δ_0 to be determined below and partition the right-half of I as follows,

$$t_0 < t_1 < t_2 < \cdots < t_n \leq \infty, \quad I_j := (t_j, t_{j+1}), \quad I\cap(t_0,\infty)=(t_0,t_n),$$

$$\|r^{-\frac{2}{3}}\sin v\|_{L_t^3(I_j;L_r^6)}\leq\delta_0\ \text{ for }\ j=0,\ldots,n-1,\ \text{ and }\ n\leq C(A;\delta_0).$$

We omit the estimate on $I \cap (-\infty, t_0)$ since it is the same by symmetry. Let $w_j(t) := S_{\mathrm{L}}(t-t_j)\boldsymbol{w}(t_j)$ for all $0 \leq j < n$, where S_{L} is the linear propagator for (2.3). Then

$$w(t)=w_0(t)+\int_{t_0}^t S_{\mathrm{L}}(t-s)\big(0, e-k^2r^{-2}(f(v+w)-f(v)-w)\big)(s)\,\mathrm{d}s,$$

which implies that, for some absolute constant $C_1 \geq 1$,

$$\begin{aligned}\|w-w_0\|_{X(I_0)}&\lesssim\|e-k^2r^{-2}(f(v+w)-f(v)-w)\|_{L_t^1 L_r^2(I_0)}\\&\leq C_1\epsilon+C_1(\delta_0^2+\|w\|_{X(I_0)}^2)\|w\|_{X(I_0)}.\end{aligned}\qquad(2.11)$$

In the second estimate above we have used the expansion,

$$\begin{aligned}f(v+w)-f(v)-w&=\frac{1}{2}(\sin(2v+2w)-\sin 2v-2w)\\&=-2w\sin^2 v-2\sin v\cos v\sin^2 w+O(|w|^3),\end{aligned}$$

to estimate the terms on the right. Note that in (2.11) we are using in an essential way the divisibility of the $X(I)$ norm. Note that $\|w\|_{X(I_0)} < \infty$ provided I_0 is a finite interval. If I_0 is half-infinite, then we first need to replace it with an interval of the form $[t_0, N)$, and let $N \to \infty$

after performing estimates which are uniform in N. Now assume that $C_1\delta_0^2 \le \frac{1}{4}$ and fix δ_0 in this fashion. By means of the continuity method, (2.11) implies that $\|w\|_{X(I_0)} \le 8C_1\epsilon$. Next, Duhamel's formula gives

$$w_1(t) - w_0(t) = \int_{t_0}^{t_1} S_{\mathrm{L}}(t-s)\big(0, e - k^2 r^{-2}(f(v+w) - f(v) - w)\big)(s)\,\mathrm{d}s$$

from which we obtain

$$\|w_1 - w_0\|_{X(\mathbb{R})} \lesssim \int_{t_0}^{t_1} \big\|\big(e - k^2 r^{-2}(f(v+w) - f(v) - w)\big)(s)\big\|_2 \,\mathrm{d}s$$

which is estimated as in (2.11). We conclude that $\|w_1\|_{X(\mathbb{R})} \le 8C_1\epsilon$. In a similar fashion one verifies that for all $0 \le j < n$

$$\|w - w_j\|_{X(I_j)} + \|w_{j+1} - w_j\|_{X(\mathbb{R})} \lesssim \|e - k^2 r^{-2}(f(v+w) - f(v) - w)\|_{L_t^1 L_r^2(I_j)}$$
$$\le C_1\epsilon + C_1(\delta_0^2 + \|w\|_{X(I_j)}^2)\|w\|_{X(I_j)} \qquad (2.12)$$

where $C_1 \ge 1$ is as above. By induction in j we have

$$\|w\|_{X(I_j)} + \|w_j\|_{X(\mathbb{R})} \le C(j)\,\epsilon \quad \forall\, 1 \le j < n.$$

This requires that $\epsilon < \epsilon_0(n)$ which can be achieved as long as $\epsilon_0(A)$ is chosen small enough. Repeating the estimate (2.12), but with the full $\mathcal{S}(I)$ norm and the energy piece $L_t^\infty \mathcal{E}$ included on the left-hand side completes the proof. □

Sketch of the proof of Lemma 2.13. The proof is very similar to [21, Proof of Proposition 2.8] or [11, Proof of Proposition 2.17] and we give a brief sketch below, mainly to address how the nonlinear profiles given by harmonic maps are handled.

Let $I_n = [0, \tau_n) \subset [0, \tau_0 \sigma_{n,i_0}]$ be any half-open subinterval on which the wave map evolution $\boldsymbol{u}_n(t)$ is defined. By (2.10), the sequence

$$\boldsymbol{v}_n^J(t) := m\boldsymbol{\pi} + \sum_{j=1}^{K_0}\big((Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}\big) + \sum_{i=1}^J \Big(v_{\mathrm{NL}}^i\big(\tfrac{t - t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big), \tfrac{1}{\sigma_{n,i}}\partial_t v_{\mathrm{NL}}^i\big(\tfrac{t - t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big)\Big)$$

is well defined on the time intervals $[0, \tau_0 \sigma_{n,i_0}]$. The idea is to apply Lemma 2.15 to the sequences $\boldsymbol{u}_n$ and $\boldsymbol{v}_n^J$ on I_n for large n and so we need to check that the hypothesis of Lemma 2.15 are satisfied. First, $\boldsymbol{u}_n(t)$ solves (1.1) so $\mathrm{eq}(u_n) = 0$. Next we claim that

$$\lim_{n\to\infty} \|\mathrm{eq}(v_n^J)\|_{L_t^1 L_r^2([0,\tau_0\sigma_{n,i_0}])} = 0 \qquad (2.13)$$

for any *fixed* J. Denoting $v_{\mathrm{NL},n}^i(t) := v_{\mathrm{NL}}^i\big(\tfrac{t-t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big)$ we have,

$$\big|\mathrm{eq}(u_n^J)(t)\big| = \frac{k^2}{2r^2}\Big| \sin(2v_n^J(t)) - \sum_{j=1}^{K_0} \sin 2Q_{\lambda_{n,j}} - \sum_{i=1}^J \sin 2(v_{\mathrm{NL},n}^i(t))\Big|.$$

And hence (2.13) follows from an argument based the pseudo-orthogonality of the parameters, the hypothesis (2.10), and repeated use of the identity,

$$\sin(A+B) - \sin A - \sin B = -2\sin A \sin^2(B/2) - 2\sin B \sin^2(A/2).$$

Next, note that the last condition in (2.10) implies that

$$\limsup_{n\to\infty} \|r^{-\frac{2}{3}}\sin Q_{\lambda_{n,j}}\|_{L_t^3([0,\tau_0\sigma_{n,i_0}];L_r^6)} \lesssim 1,$$

for each $j \in 1, \ldots, K_0$. In fact, is crucial that,

$$\limsup_{n \to \infty} \|r^{-\frac{2}{3}} \sin(v_n^J)\|_{L_t^3 L_r^6(I_n)} \lesssim 1$$

uniformly in J. This is possible thanks to the small data theory from Lemma 2.8 together with the pythagorean expansion of the energy (2.9). Indeed, there exists J_1 such that for each $i \geq J_1$, we must have $\limsup_{n \to \infty} E(\boldsymbol{v}_{\mathrm{L},n}^j(0)) < \epsilon_0$ where ϵ_0 is as in Lemma 2.8 and $v_{\mathrm{L},n}^i(t) := v_{\mathrm{L}}^i(\frac{t - t_{n,i}}{\sigma_{n,i}}, \frac{\cdot}{\sigma_{n,i}})$. Using again the pseudo-orthogonality of the parameters and (2.10) along with Lemma 2.8 we obtain,

$$\limsup_{n \to \infty} \|r^{-\frac{2}{3}} \sin(\sum_{i \geq J_1} v_{\mathrm{NL},n}^i)\|_{L_t^3 L_r^6(I_n)}^3 \lesssim \limsup_{n \to \infty} \sum_{i \geq J_1} \|r^{-\frac{2}{3}} v_{\mathrm{NL},n}^i\|_{L_t^3 L_r^6(I_n)}^3$$

$$\lesssim \limsup_{n \to \infty} \sum_{i \geq J_1} \|v_{\mathrm{L},n}^i\|_{\mathcal{E}}^3 < \infty,$$

where the last inequality implicitly uses the fact that for all $\boldsymbol{v} \in \mathcal{E}$ with $E(\boldsymbol{v}) \leq \epsilon$ sufficiently small we have $\|\boldsymbol{v}\|_{\mathcal{E}} \simeq E(\boldsymbol{v})$. One may now apply Lemma 2.15 and conclude, for instance, that

$$\limsup_{n \to \infty} \sup_{t \in I_n} \|\boldsymbol{u}_n(t) - \boldsymbol{v}_n^J(t) - \boldsymbol{w}_{\mathrm{L},n}^J(t)\|_{\mathcal{E}} = 0, \tag{2.14}$$

for each interval $I_n \subset [0, \tau_0 \sigma_{n,i_0}]$ on which $\boldsymbol{u}_n(t)$ is defined. In fact, by Lemma 2.5 this is sufficient to deduce that $T_+(\boldsymbol{u}_n) > \tau_0 \sigma_{n,i_0}$ for all sufficiently large n as long as δ_0 as in (2.10) is chosen small enough. To see this, suppose for contradiction there is some subsequence $\boldsymbol{u}_n(t)$ and a sequence $\tau_n \to 0$ for which $\boldsymbol{u}_n(t)$ has maximal forward interval existence given by $I_n = [0, \tau_n \sigma_{n,i_0})$. Fix $J > i_0$ and let $\epsilon > 0$ be a constant to be determined below. Since each of the profiles is well-defined up till time $\tau_0 \sigma_{n,i_0}$, and using crucially the second line in (2.10) (in particular that $\lambda_{n,j} \gtrsim \sigma_{n,i_0}$ for each j), we can find $A_n = A_n(\epsilon) > 0$ such that

$$\sum_{j=1}^{K_0} E(Q_{\lambda_{n,j}}; 0, A_n) + \sum_{i=1}^{J} E(\boldsymbol{v}_{\mathrm{NL},n}^i(\tau_n \sigma_{n,i_0}); 0, A_n) + E(\boldsymbol{w}_{\mathrm{L},n}^J(\tau_n \sigma_{n,i_0}); 0, A_n) < \epsilon$$

and such that $s_n := \tau_n \sigma_{n,i_0} - \frac{A_n}{4} > 0$. By finite speed of propagation and the above we have

$$\sum_{j=1}^{K_0} E(Q_{\lambda_{n,j}}; 0, A_n/2) + \sum_{i=1}^{J} E(\boldsymbol{v}_{\mathrm{NL},n}^i(s_n); 0, A_n/2) + E(\boldsymbol{w}_{\mathrm{L},n}^J(s_n); 0, A_n/2) < \epsilon.$$

Combing the above with (2.14), we obtain,

$$E(\boldsymbol{u}_n(s_n); 0, A_n/2) \lesssim \epsilon$$

as long as n is taken sufficiently large. Since $\tau_n \sigma_{n,i_0} - s_n = A_n/4 < A_n/2$, we see by Lemma 2.5 that $\tau_n \sigma_{n,i_0}$ cannot be a maximal time for $\boldsymbol{u}_n$ as long as $\epsilon > 0$ is chosen small enough, a contradiction. This completes the proof. □

In Section 3 we need an additional fact about profile decompositions satisfying additional hypothesis proved in [19]. First, a preliminary lemma.

Lemma 2.16. [19, Claim 2] *Let $(f_n, g_n) \in \mathcal{E}$ be a sequence of functions, bounded in $\mathcal{E}$ and assume that there exists a sequence $\alpha_n > 0$ of positive numbers such that*

$$\|g_n\|_{L^2(r \geq \alpha_n)} \to 0 \quad as \quad n \to \infty.$$

Let $\{s_n\} \subset \mathbb{R}$ be any sequence such that $\lim_{n\to\infty} |s_n|/\alpha_n = \infty$, and denote by $\boldsymbol{v}_0 = (v_0, \dot{v}_0) \in \mathcal{E}$ the following weak limit,

$$\vec{S}_{\mathrm{L}}(-s_n)(f_n, g_n) \rightharpoonup \boldsymbol{v}_0 \in \mathcal{E}.$$

Then, $\vec{S}_{\mathrm{L}}(s_n)(f_n, g_n) \rightharpoonup (v_0, -\dot{v}_0) \in \mathcal{E}$, where $\vec{S}_{\mathrm{L}}(t)(f, g) = (S_{\mathrm{L}}(t)(f, g), \partial_t S_{\mathrm{L}}(t)(f, g))$.

As a consequence one has the following lemma.

Lemma 2.17. [19, Claim 3] *Let* $\ell, m \in \mathbb{Z}$ *and let* $\boldsymbol{u}_n$ *be a sequence in* $\mathcal{E}_{\ell,m}$ *with* $\limsup_{n\to\infty} E(\boldsymbol{u}_n) < \infty$. *Assume the sequence* $\boldsymbol{u}_n$ *admits a profile decomposition of the form,*

$$\boldsymbol{u}_n = m\boldsymbol{\pi} + \sum_{j=1}^{K_0}(Q\big(\tfrac{\cdot}{\lambda_{n,j}}\big),0) - \boldsymbol{\pi}) + \sum_{i=1}^{J}\Big(v_{\mathrm{L}}^i\big(\tfrac{-t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big), \tfrac{1}{\sigma_{n,i}}\partial_t v_{\mathrm{L}}^i\big(\tfrac{-t_{n,i}}{\sigma_{n,i}}, \tfrac{\cdot}{\sigma_{n,i}}\big)\Big) + \boldsymbol{w}_{n,0}^J(\cdot),$$

that is, all of the profiles $(\boldsymbol{\psi}^j, \lambda_{n,j})$ *for* $1 \le j \le K_0$ *as in Lemma 2.10 are given by harmonic maps* $(\boldsymbol{Q}, \lambda_{n,j})$. *Assume in addition that,*

$$\|\dot{u}_n\|_{L^2} \to 0 \quad as \quad n \to \infty.$$

Then, after passing to a subsequence, for each profile $(\boldsymbol{v}_{\mathrm{L}}^i, \sigma_{n,i}, t_{n,i})$ *we can ensure that either,*

$$t_{n,i} = 0 \; \forall \, n \quad and \quad \dot{v}_{\mathrm{L}}^i(0) = 0$$

or,

$$-\frac{t_{n,i}}{\sigma_{n,i}} \to \pm\infty \quad and \quad \exists i' \ne i \text{ such that } v_{\mathrm{L}}^i(t) = v_{\mathrm{L}}^{i'}(-t) \; \forall t, \; t_{n,i} = -t_{n,i'} \; \sigma_{n,i} = -\sigma_{n,i'} \; \forall n.$$

2.4. Multi-bubble configurations.
In this section we study properties of finite energy maps near a multi-bubble configuration.

The operator $\mathcal{L}_{\mathcal{Q}}$ obtained by linearization of (1.1) about an M-bubble configuration $\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$ is given by,

$$\mathcal{L}_{\mathcal{Q}}\, g := \mathrm{D}^2 E_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))g = -\Delta g + \frac{k^2}{r^2}f'(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))g,$$

where $f'(z) = \cos 2z$ and $E_{\mathbf{p}}$ is the potential energy,

$$E_{\mathbf{p}}(\psi) := \int_0^\infty \frac{1}{2}\Big((\partial_r \psi(r))^2 + k^2\frac{\sin^2 \psi(r)}{r^2}\Big)r\,\mathrm{d}r.$$

Given $\boldsymbol{g} = (g, \dot{g}) \in \mathcal{E}$,

$$\big\langle \mathrm{D}^2 E(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))\boldsymbol{g} \mid \boldsymbol{g} \big\rangle = \int_0^\infty \Big(\dot{g}(r)^2 + (\partial_r g(r))^2 + \frac{k^2}{r^2}f'(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))g(r)^2 \Big)r\mathrm{d}r.$$

An important instance of the operator $\mathcal{L}_{\mathcal{Q}}$ is given by linearizing (1.1) about a single harmonic map $\mathcal{Q}(m, M, \vec{\iota}, \vec{\lambda}) = Q_\lambda$. In this case we use the short-hand notation,

$$\mathcal{L}_\lambda := (-\Delta + \frac{k^2}{r^2}) + \frac{k^2}{r^2}(f'(Q_\lambda) - 1).$$

We write $\mathcal{L} := \mathcal{L}_1$. For each $k \ge 1$,

$$\Lambda Q(r) := r\partial_r Q(r) = k \sin Q = 2k\frac{r^k}{1 + r^{2k}}.$$

When $k \ge 2$, ΛQ is a zero energy eigenfunction for $\mathcal{L}$, i.e.,

$$\mathcal{L}\Lambda Q = 0, \quad and \quad \Lambda Q \in L^2_{\mathrm{rad}}(\mathbb{R}^2).$$

When $k = 1$, $\mathcal{L}\Lambda Q = 0$ holds but $\Lambda Q \notin L^2$ due to slow decay as $r \to \infty$ and 0 is called a threshold resonance. Indeed, for $R > 0$,

$$\int_0^R (\Lambda Q(r))^2 \, r \, dr = -\frac{2R^2}{1+R^2} + 2\log(1+R^2) = 4\log R + O(1) \quad \text{as} \quad R \to \infty. \quad (2.15)$$

On the other hand when $k = 1$, $\underline{\Lambda}\Lambda Q$ has an important cancellation which leads to improved decay,

$$\underline{\Lambda}\Lambda Q = \frac{4r}{(1+r^2)^2}, \quad (2.16)$$

so $\underline{\Lambda}\Lambda Q \in L^1 \cap L^\infty$ and $\langle \underline{\Lambda}\Lambda Q \mid \Lambda Q \rangle = 2$, whereas for $k \geq 2$, $\langle \underline{\Lambda}\Lambda Q \mid \Lambda Q \rangle = 0$.

We define a smooth non-negative function $\mathcal{Z} \in C^\infty(0,\infty) \cap L^1((0,\infty), r\,dr)$ by

$$\mathcal{Z}(r) := \begin{cases} \chi(r)\Lambda Q(r) & \text{if } k = 1, 2 \\ \Lambda Q(r) & \text{if } k \geq 3 \end{cases} \quad (2.17)$$

and note that

$$\langle \mathcal{Z} \mid \Lambda Q \rangle > 0. \quad (2.18)$$

In fact the precise form of $\mathcal{Z}$ is not so important, rather only that it is not perpendicular to ΛQ and has sufficient decay and regularity. We fix it as above because of the convenience of setting $\mathcal{Z} = \Lambda Q$ if $k \geq 3$. We record the following localized coercivity lemma proved in [36].

Lemma 2.18 (Localized coercivity for $\mathcal{L}$). [36, Lemma 5.4] *Fix $k \geq 1$. There exist uniform constants $c < 1/2, C > 0$ with the following properties. Let $g \in H$. Then,*

$$\langle \mathcal{L}g \mid g \rangle \geq c\|g\|_H^2 - C \langle \mathcal{Z} \mid g \rangle^2.$$

If $R > 0$ is large enough then,

$$(1 - 2c)\int_0^R \left((\partial_r g)^2 + k^2\frac{g^2}{r^2}\right) r dr + c\int_R^\infty \left((\partial_r g)^2 + k^2\frac{g^2}{r^2}\right) r dr + \left\langle \frac{k^2}{r^2}(f'(Q) - 1)g \mid g \right\rangle$$
$$\geq -C \langle \mathcal{Z} \mid g \rangle^2.$$

If $r > 0$ is small enough, then

$$(1 - 2c)\int_r^\infty \left((\partial_r g)^2 + k^2\frac{g^2}{r^2}\right) r dr + c\int_0^r \left((\partial_r g)^2 + k^2\frac{g^2}{r^2}\right) r dr + \left\langle \frac{k^2}{r^2}(f'(Q) - 1)g \mid g \right\rangle$$
$$\geq -C \langle \mathcal{Z} \mid g \rangle^2.$$

As a consequence, (see for example [37, Proof of Lemma 2.4] for an analogous argument) one obtains the following coercivity property of the operator $\mathcal{L}_\mathcal{Q}$.

Lemma 2.19. *Fix $k \geq 1$, $M \in \mathbb{N}$. There exist $\eta, c_0 > 0$ with the following properties. Consider the subset of M-bubble configurations $\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$ for $\vec{\iota} \in \{-1, 1\}^M$, $\vec{\lambda} \in (0, \infty)^M$ such that,*

$$\sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k \leq \eta^2. \quad (2.19)$$

Let $g \in H$ be such that

$$0 = \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid g \right\rangle \quad \text{for} \quad j = 1, \ldots M.$$

for some $\vec{\lambda}$ as in (2.19). Then,

$$\left\langle D^2 E_\mathbf{p}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))g \mid g \right\rangle \geq c_0\|g\|_H^2.$$

The following technical lemma is useful when computing interactions between bubbles at different scales.

Lemma 2.20. *For any $\lambda \leq \mu$ and $\alpha, \beta > 0$ with $\alpha \neq \beta$ the following bound holds:*

$$\int_0^\infty \max\left(1, \frac{r}{\lambda}\right)^{-\alpha} \max\left(1, \frac{\mu}{r}\right)^{-\beta} \frac{\mathrm{d}r}{r} \lesssim_{\alpha,\beta} \left(\frac{\lambda}{\mu}\right)^{\min(\alpha,\beta)}.$$

For any $\alpha > 0$ the following bound holds:

$$\int_0^\infty \max\left(1, \frac{r}{\lambda}\right)^{-\alpha} \max\left(1, \frac{\mu}{r}\right)^{-\alpha} \frac{\mathrm{d}r}{r} \lesssim_\alpha \left(\frac{\lambda}{\mu}\right)^\alpha \log\left(\frac{\mu}{\lambda}\right).$$

Proof. This is a straightforward computation, considering separately the regions $0 < r \leq \lambda$, $\lambda \leq r \leq \mu$, and $r \geq \mu$. $\qquad\square$

Using the above, along with the formula for $\mathcal{Z}$ in (2.17) we obtain the following.

Corollary 2.21. *Let $\mathcal{Z}$ be as in (2.17) and suppose that $\lambda, \mu > 0$ satisfy $\lambda/\mu \leq 1$. Then,*

$$\left\langle \mathcal{Z}_{\underline{\lambda}} \mid \Lambda Q_{\underline{\mu}} \right\rangle \lesssim \begin{cases} (\lambda/\mu)^{k+1} & \text{if } k = 1,2 \\ (\lambda/\mu)^{k-1} & \text{if } k \geq 3 \end{cases} \quad , \quad \left\langle \mathcal{Z}_{\underline{\mu}} \mid \Lambda Q_{\underline{\lambda}} \right\rangle \lesssim \begin{cases} 1 & \text{if } k = 1 \\ (\lambda/\mu)^{k-1} & \text{if } k \geq 2 \end{cases} .$$

Lemma 2.20 is also used to prove the following two lemmas. The first gives leading order terms in an expansion of the nonlinear energy functional about an M-bubble configuration.

Lemma 2.22. *Fix $k \geq 1$, $M \in \mathbb{N}$. For any $\theta > 0$, there exists $\eta > 0$ with the following property. Consider the subset of M-bubble $\mathcal{Q}(m, \iota, \vec{\lambda})$ configurations such that*

$$\sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k \leq \eta.$$

Then,

$$\left| E(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) - ME(\boldsymbol{Q}) - 16k\pi \sum_{j=1}^{M-1} \iota_j \iota_{j+1} \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k \right| \leq \theta \sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k. \tag{2.20}$$

Moreover, there exists a uniform constant $C > 0$ such that for any $g \in H$,

$$\left| \left\langle DE_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) \mid g \right\rangle \right| \leq C \|g\|_H \sum_{j=1}^M \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k. \tag{2.21}$$

The next lemma gives the nonlinear interaction force between bubbles, which we introduce notation for below. Given an M-bubble configuration, $\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$ we set

$$f_{\mathbf{i}}(m, \vec{\iota}, \vec{\lambda}) := -DE_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) = -\frac{k^2}{r^2}\left(f(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) - \sum_{j=1}^M \iota_j f(Q_{\lambda_j}) \right). \tag{2.22}$$

Lemma 2.23. *Let $k \geq 1$, $M \in \mathbb{N}$. For any $\theta > 0$ there exists $\eta > 0$ with the following property. Let $\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$ be an M-bubble configuration with*

$$\sum_{j=0}^M \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k \leq \eta,$$

under the convention that $\lambda_0 = 0$, $\lambda_{M+1} = \infty$. Then, we have,

$$\left| \left\langle \Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m, \vec{\iota}, \vec{\lambda}) \right\rangle + \iota_{j-1} 8k^2 \left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k - \iota_{j+1} 8k^2 \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k \right| \leq \theta\left(\left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k + \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k \right)$$

where here $f_{\mathbf{i}}(m, \vec{\iota}, \vec{\lambda})$ is defined in (2.22).

Proof of Lemma 2.22. To prove (2.20) we proceed iteratively, singling out first the contributions to the leading order from the first bubble $\iota_1 Q_{\lambda_1}$, writing,

$$
\begin{aligned}
&\frac{1}{2\pi}\Big(E(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) - M E(\mathbf{Q}) \Big) \\
={}&\frac{1}{2} \int_0^\infty \Big((\partial_r \mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))^2 + k^2 \frac{\sin^2 \mathcal{Q}(m, \vec{\iota}, \vec{\lambda})}{r^2} \Big) r \, dr - \sum_{j=1}^M \frac{1}{2} \int_0^\infty \Big((\partial_r Q_{\lambda_j})^2 + k^2 \frac{\sin^2 Q_{\lambda_j}}{r^2} \Big) r \, dr \\
={}&\sum_{1 \le i < j \le M} \int_0^\infty \iota_i \iota_j \partial_r Q_{\lambda_i} \partial_r Q_{\lambda_j} r \, dr + \frac{k^2}{2} \int_0^\infty \frac{\sin^2(\sum_{i=1}^M \iota_i Q_{\lambda_i}) - \sum_{i=1}^M \sin^2 Q_{\lambda_i}}{r^2} r \, dr \\
={}&\int_0^\infty \iota_1 \partial_r Q_{\lambda_1} \sum_{j=2}^M \iota_j \partial_r Q_{\lambda_j} r \, dr + \frac{k^2}{2} \int_0^\infty \frac{\sin^2(\sum_{i=1}^M \iota_i Q_{\lambda_i}) - \sin^2 Q_{\lambda_1} - \sin^2(\sum_{i=2}^M \iota_i Q_{\lambda_i})}{r^2} r \, dr \\
&+ \sum_{2 \le i < j \le M} \int_0^\infty \iota_i \iota_j \partial_r Q_{\lambda_i} \partial_r Q_{\lambda_j} r \, dr + \frac{k^2}{2} \int_0^\infty \frac{\sin^2(\sum_{i=2}^M \iota_i Q_{\lambda_i}) - \sum_{i=2}^M \sin^2 Q_{\lambda_i}}{r^2} r \, dr.
\end{aligned}
\tag{2.23}
$$

We will call a term in the expansion negligible if it can be made small relative to $\sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k$ by taking η small. We expand further the second term on the right above using the identity, $\sin^2(A + B) - \sin^2 A - \sin^2 B = \frac{1}{2} \sin 2A \sin 2B - 2 \sin^2 A \sin^2 B$, obtaining,

$$
\begin{aligned}
\sin^2(\sum_{i=1}^M \iota_i Q_{\lambda_i}) - \sin^2 Q_{\lambda_1} - \sin^2(\sum_{i=2}^M \iota_i Q_{\lambda_i}) ={}& \iota_1 \sin 2Q_{\lambda_1} \sum_{j=2}^M \iota_j Q_{\lambda_j} \\
&- \iota_1 \sin 2Q_{\lambda_1} \Big(\sum_{j=2}^M \iota_j Q_{\lambda_j} - \frac{1}{2} \sin(2 \sum_{j=2}^M \iota_j Q_{\lambda_j}) \Big) - 2 \sin^2(\sum_{j=2}^M \iota_j Q_{\lambda_j}) \sin^2 Q_{\lambda_1}.
\end{aligned}
\tag{2.24}
$$

The last term above satisfies the bound

$$
\Big| \sin^2(\sum_{j=2}^M \iota_j Q_{\lambda_j}) \sin^2 Q_{\lambda_1} \Big| \lesssim (\Lambda Q_{\lambda_1})^2 \sum_{j=2}^M (\Lambda Q_{\lambda_j})^2,
$$

and using Lemma 2.20 we see that its contribution is negligible. The first term on the right is integrated by parts, using the identity $r^{-1} \partial_r (r \partial_r Q) = k^2 r^{-2} 2^{-1} \sin 2Q$ yielding,

$$
\frac{k^2}{2} \int_0^\infty \frac{\iota_1 \sin 2Q_{\lambda_1} \sum_{j=2}^M \iota_j Q_{\lambda_j}}{r^2} r \, dr = -\int_0^\infty \iota_1 \partial_r Q_{\lambda_1} \sum_{j=2}^M \iota_j \partial_r Q_{\lambda_j} r \, dr,
$$

which exactly cancels the first term on the right of the last equality in (2.23). The second term on the right in (2.24) contributes the leading order term involving λ_1. We integrate by parts

and use the identity $\Lambda Q = k \sin Q$ to obtain,

$$-\frac{k^2}{2} \int_0^\infty \frac{\iota_1 \sin 2Q_{\lambda_1} \left(\sum_{j=2}^M \iota_j Q_{\lambda_j} - \frac{1}{2} \sin(2 \sum_{j=2}^M \iota_j Q_{\lambda_j}) \right)}{r^2} r \, \mathrm{d}r$$

$$= \iota_1 \int_0^\infty \Lambda Q_{\lambda_1} \sum_{j=2}^M \iota_j \Lambda Q_{\lambda_j} (1 - \cos 2 \sum_{j=2}^M \iota_j Q_{\lambda_j}) \frac{\mathrm{d}r}{r}$$

$$= \iota_1 \iota_2 \frac{2}{k^2} \int_0^\infty \Lambda Q_{\lambda_1} (\Lambda Q_{\lambda_2})^3 \frac{\mathrm{d}r}{r} + 2 \iota_1 \iota_2 \int_0^\infty \Lambda Q_{\lambda_1} \Lambda Q_{\lambda_2} \left(\sin^2(\sum_{j=2}^M \iota_j Q_{\lambda_j}) - \sin^2 Q_{\lambda_2} \right) \frac{\mathrm{d}r}{r}$$

$$+ 2 \iota_1 \int_0^\infty \Lambda Q_{\lambda_1} \sum_{j=3}^M \iota_j \Lambda Q_{\lambda_j} \sin^2(\sum_{j=2}^M \iota_j Q_{\lambda_j}) \frac{\mathrm{d}r}{r}.$$

The last two terms above are negligible due to Lemma 2.20 with smallness coming from the fact that every term involves the bubble at scale λ_1 and at least one bubble at scale λ_j with $j \geq 3$. The first term on the right is of leading order and satisfies,

$$\iota_1 \iota_2 \frac{2}{k^2} \int_0^\infty \Lambda Q_{\lambda_1} (\Lambda Q_{\lambda_2})^3 \frac{\mathrm{d}r}{r} = \iota_1 \iota_2 8k \left(\frac{\lambda_1}{\lambda_2} \right)^k (1 + O((\lambda_1/\lambda_2)^{2k})),$$

and we refer the reader to [41, Proof of (3.18) p. 1285] where the above is proved.

To complete the proof of (2.20) we iterate the exact analysis as above with the last line in (2.23), noting that it is of the exact same form as the expression in the previous equality of (2.23), only with the bubbles from $2 \leq j \leq M$.

To prove (2.21) we use (2.22) and note that for $\ell = m - \sum_{i=1}^M \iota_i$ we have

$$f(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) = \frac{1}{2} \sin(2\ell\pi + 2 \sum_{i=1}^M \iota_i Q_{\lambda_i}) = \frac{1}{2} \sin(2 \sum_{i=1}^M \iota_i Q_{\lambda_i})$$

Expanding, we obtain the expression,

$$\frac{r^2}{k^2} DE_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) = \frac{1}{2} \sin(2 \sum_{i=2}^M \iota_i Q_{\lambda_i} + 2\iota_1 Q_{\lambda_1}) - \frac{1}{2} \sum_{i=1}^M \iota_i \sin 2Q_{\lambda_i}$$

$$= -\sin \left(2 \sum_{i=2}^M \iota_i Q_{\lambda_i} \right) \sin^2 Q_{\lambda_1} - \iota_1 \sin^2 \left(\sum_{i=2}^M \iota_i Q_{\lambda_i} \right) \sin 2Q_{\lambda_1}$$

$$+ \frac{1}{2} \sin(2 \sum_{i=2}^M \iota_i Q_{\lambda_i}) - \frac{1}{2} \sum_{i=2}^M \iota_i \sin 2Q_{\lambda_i}$$

Iterating this expansion in the last line above and using the identity $k \sin Q = \Lambda Q$ we obtain the pointwise estimates,

$$|DE_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}))| \lesssim \frac{1}{r^2} \sum_{i,j,\ell \text{ not all equal}} \Lambda Q_{\lambda_i} \Lambda Q_{\lambda_j} \Lambda Q_{\lambda_\ell}$$

Using Lemma (2.20) we then have,

$$\left| \left\langle DE_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) \mid g \right\rangle \right| \lesssim \|g\|_{L^\infty} \sum_{j=1}^M \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \lesssim \|g\|_H \sum_{j=1}^M \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k,$$

which completes the proof. $\square$

Proof of Lemma 2.23. Fixing $j \in \{1, \ldots, M\}$, we expand, isolating the terms that are quadratic in $\sin Q_{\lambda_j} = k^{-1} \Lambda Q_{\lambda_j}$, and interacting with only its nearest neighboring bubbles $Q_{\lambda_{j-1}}$ and $Q_{\lambda_{j+1}}$,

$$\frac{1}{2}\sin(2\sum_{i\neq j}\iota_i Q_{\lambda_i} + 2\iota_j Q_{\lambda_j}) - \frac{1}{2}\sum_{i=1}^{M}\iota_i \sin 2Q_{\lambda_i}$$

$$= -\sin\left(2\sum_{i\neq j}\iota_i Q_{\lambda_i}\right)\sin^2 Q_{\lambda_j} - \iota_j \sin^2\left(\sum_{i\neq j}\iota_i Q_{\lambda_i}\right)\sin 2Q_{\lambda_j}$$

$$+ \frac{1}{2}\sin(2\sum_{i\neq j}\iota_i Q_{\lambda_i}) - \frac{1}{2}\sum_{i\neq j}\iota_i \sin 2Q_{\lambda_i}$$

$$= -\iota_{j+1}\sin 2Q_{\lambda_{j+1}}\sin^2 Q_{\lambda_j} - \iota_{j-1}\sin 2Q_{\lambda_{j-1}}\sin^2 Q_{\lambda_j}$$

$$+ 2\sin(2\iota_{j-1}Q_{\lambda_{j-1}})\sin^2\left(\sum_{i\neq j-1,j}\iota_i Q_{\lambda_i}\right)\sin^2 Q_{\lambda_j} \qquad (2.25)$$

$$+ 2\sin(2\iota_{j+1}Q_{\lambda_{j+1}})\sin^2\left(\sum_{i\neq j-1,j,j+1}\iota_i Q_{\lambda_i}\right)\sin^2 Q_{\lambda_j}$$

$$+ 2\sin 2\left(\sum_{i\neq j-1,j}\iota_i Q_{\lambda_i}\right)\sin^2 Q_{\lambda_{j-1}}\sin^2 Q_{\lambda_j}$$

$$- \iota_j \sin^2\left(\sum_{i\neq j}\iota_i Q_{\lambda_i}\right)\sin 2Q_{\lambda_j} + \frac{1}{2}\left(\sin(2\sum_{i\neq j}\iota_i Q_{\lambda_i}) - \sum_{i\neq j}\iota_i \sin 2Q_{\lambda_i}\right)$$

$$= -\iota_{j+1}\sin 2Q_{\lambda_{j+1}}\sin^2 Q_{\lambda_j} - \iota_{j-1}\sin 2Q_{\lambda_{j-1}}\sin^2 Q_{\lambda_j} + \Psi_j(\vec{\iota}, \vec{\lambda}_j)$$

where the error term $\Psi_j(\vec{\iota}, \vec{\lambda}_j)$ satisfies,

$$\left|\Psi_j(\vec{\iota}, \vec{\lambda}_j)\right| \lesssim (\Lambda Q_{\lambda_j})^2\Big(\Lambda Q_{\lambda_{j-1}}\sum_{i\neq j-1,j}(\Lambda Q_{\lambda_i})^2 + (\Lambda Q_{\lambda_{j-1}})^2\sum_{i\neq j-1,j}\Lambda Q_{\lambda_i}\Big)$$

$$+ (\Lambda Q_{\lambda_j})^2\Lambda Q_{\lambda_{j+1}}\sum_{i\neq j-1,j,j+1}(\Lambda Q_{\lambda_i})^2 + \Lambda Q_{\lambda_j}\sum_{i\neq j}(\Lambda Q_{\lambda_i})^2.$$

Using the pointwise bound above along with Lemma 2.20 the function $\Psi_j(\vec{\iota}, \vec{\lambda}_j)$ satisfies,

$$\left|\left\langle \Lambda Q_{\lambda_j} \mid r^{-2}\Psi_j(\vec{\iota}, \vec{\lambda}_j)\right\rangle\right| \le \theta(\eta)\left(\left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k + \left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k\right),$$

where $\theta(\eta) > 0$ is a function that tends to zero as $\eta \to 0$. It follows that

$$\left\langle \Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m, \vec{\iota}, \vec{\lambda})\right\rangle \simeq +\iota_{j+1}\left\langle r^{-2}(\Lambda Q_{\lambda_j})^3 \mid \sin 2Q_{\lambda_{j+1}}\right\rangle + \iota_{j-1}\left\langle r^{-2}(\Lambda Q_{\lambda_j})^3 \mid \sin 2Q_{\lambda_{j-1}}\right\rangle$$

$$= \iota_{j+1}\left\langle r^{-2}(\Lambda Q_{\lambda_j/\lambda_{j+1}})^3 \mid \sin 2Q\right\rangle + \iota_{j-1}\left\langle r^{-2}(\Lambda Q_{\lambda_j/\lambda_{j-1}})^3 \mid \sin 2Q\right\rangle$$

where "$\simeq$" above means up to negligible terms. Note that,

$$\sin 2Q = 4r^k\frac{1 - r^{2k}}{(1 + r^{2k})^2} = 4r^k + O(r^{3k}) \quad \text{if} \ \ r \ll 1$$

$$= -4r^{-k} + O(r^{-3k}) \quad \text{if} \ \ r \gg 1.$$

Via residue calculus we compute,

$$\int_0^\infty (\Lambda Q(r))^3 4 r^k \frac{\mathrm{d}r}{r} = 32k^3 \int_0^\infty \frac{r^k}{(r^k + r^{-k})^3} \frac{\mathrm{d}r}{r} = 8k^2$$

$$\int_0^\infty (\Lambda Q(r))^3 4 r^{-k} \frac{\mathrm{d}r}{r} = 32k^3 \int_0^\infty \frac{r^{-k}}{(r^k + r^{-k})^3} \frac{\mathrm{d}r}{r} = 8k^2$$

And thus,

$$\left| \iota_{j+1} \left\langle r^{-2} (\Lambda Q_{\lambda_j/\lambda_{j+1}})^3 \mid \sin 2Q \right\rangle - \iota_{j+1} 8 k^2 \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \right| \le \theta(\eta) \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k$$

$$\left| \iota_{j-1} \left\langle r^{-2} (\Lambda Q_{\lambda_j/\lambda_{j-1}})^3 \mid \sin 2Q \right\rangle + \iota_{j-1} 8 k^2 \left(\frac{\lambda_{j-1}}{\lambda_j} \right)^k \right| \le \theta(\eta) \left(\frac{\lambda_{j-1}}{\lambda_j} \right)^k$$

where $\theta(\eta) \to 0$ as $\eta \to 0$, which completes the proof; see [41, Proof of Claim 3.14] for more details of this computation. $\square$

The following modulation lemma plays an important role in our analysis. Before stating it, we define a proximity function to M-bubble configurations. Fixing m, M we observe that $\mathcal{Q}(m, \vec{\iota}, \vec{\lambda}; r)$ is an element of $\mathcal{E}_{\ell,m}$, where

$$\ell = \ell(m, M, \vec{\iota}) := m - \sum_{j=1}^{M} \iota_j \tag{2.26}$$

Definition 2.24. Fix m, M as in Definition 1.13 and let $\boldsymbol{v} \in \mathcal{E}_{\ell,m}$ for some $\ell \in \mathbb{Z}$. Define,

$$\mathbf{d}(\boldsymbol{v}) = \mathbf{d}_{m,M}(\boldsymbol{v}) := \inf_{\vec{\iota}, \vec{\lambda}} \left(\| \boldsymbol{v} - \mathcal{Q}(m, \vec{\iota}, \vec{\lambda}) \|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \right)^{\frac{1}{2}}.$$

where the infimum is taken over all vectors $\vec{\lambda} = (\lambda_1, \dots, \lambda_M) \in (0, \infty)^M$ and all $\vec{\iota} = \{\iota_1, \dots, \iota_M\} \in \{-1, 1\}^M$ satisfying (2.26).

Lemma 2.25 (Static modulation lemma). *Fix $k \ge 1$ and $M \in \mathbb{N}$. There exists $\eta, C > 0$ with the following properties. Let m be as in Definition 1.13 and $\mathbf{d}_{m,M}$ as in Definition 2.24. Let $\theta > 0$, $\ell \in \mathbb{Z}$, and let $\boldsymbol{v} \in \mathcal{E}_{\ell,m}$ be such that*

$$\mathbf{d}_{m,M}(\boldsymbol{v}) \le \eta, \quad and \quad E(\boldsymbol{v}) \le M E(\boldsymbol{Q}) + \theta^2, \tag{2.27}$$

Then, there exists a unique choice of $\vec{\lambda} = (\lambda_1, \dots, \lambda_M) \in (0, \infty)^M$, $\vec{\iota} \in \{-1, 1\}^M$, and $g \in H$, such that setting $\boldsymbol{g} = (g, \dot{v})$, we have

$$\boldsymbol{v} = \mathcal{Q}(m, \vec{\iota}, \vec{\lambda}) + \boldsymbol{g},$$
$$0 = \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid g \right\rangle, \quad \forall j = 1, \dots, M, \tag{2.28}$$

along with the estimates,

$$\mathbf{d}_{m,M}(\boldsymbol{v})^2 \le \| \boldsymbol{g} \|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \le C \mathbf{d}_{m,M}(\boldsymbol{v})^2, \tag{2.29}$$

and,

$$\| \boldsymbol{g} \|_{\mathcal{E}}^2 + \sum_{j \notin \mathcal{A}} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \le C \max_{j \in \mathcal{A}} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k + \theta^2, \tag{2.30}$$

where $\mathcal{A} := \{ j \in \{1, \dots, M-1\} : \iota_j \ne \iota_{j+1} \}$.

Remark 2.26. We use the following, less standard, version of the implicit function theorem in the proof of Lemma 2.25.

Let X, Y, Z be Banach spaces, $(x_0, y_0) \in X \times Y$, and $\delta_1, \delta_2 > 0$. Consider a mapping $G : B(x_0, \delta_1) \times B(y_0, \delta_2) \to Z$, continuous in x and C^1 in y. Assume $G(x_0, y_0) = 0$, $(D_y G)(x_0, y_0) =: L_0$ has bounded inverse L_0^{-1}, and

$$\|L_0 - D_y G(x, y)\|_{\mathcal{L}(Y,Z)} \leq \frac{1}{3\|L_0^{-1}\|_{\mathcal{L}(Z,Y)}}$$

$$\|G(x, y_0)\|_Z \leq \frac{\delta_2}{3\|L_0^{-1}\|_{\mathcal{L}(Z,Y)}}, \tag{2.31}$$

for all $\|x - x_0\|_X < \delta_1$ and $\|y - y_0\|_Y < \delta_2$. Then, there exists a continuous function $\varsigma : B(x_0, \delta_1) \to B(y_0, \delta_2)$ such that for all $x \in B(x_0, \delta_1)$, $y = \varsigma(x)$ is the unique solution of $G(x, \varsigma(x)) = 0$ in $B(y_0, \delta_2)$.

This is proved in the same way as the usual implicit function theorem, see, e.g., [6, Section 2.2]. The essential point is that the bounds (2.31) give uniform control on the size of the open set where the Banach contraction mapping theorem is applied.

Proof of Lemma 2.25. The argument is very similar to [41, Proof of Lemma 3.1] and we only give a brief sketch. Let $\eta_0 := \mathbf{d}_{m,M}(\boldsymbol{v})$. By (2.27) there exists some choice of $\vec{\iota} \in \{-1, 1\}^M$ and $\widetilde{\vec{\lambda}} \in (0, \infty)^M$ such that

$$\widetilde{g} := v - \mathcal{Q}(m, \vec{\iota}, \widetilde{\vec{\lambda}}) \quad \text{satisfies} \quad \eta_0^2 \leq \|\widetilde{g}\|_H^2 + \sum_{j=1}^{M-1} \left(\frac{\widetilde{\lambda}_j}{\widetilde{\lambda}_{j+1}}\right)^k \leq 4\eta_0^2$$

Define $F : H \times (0, \infty)^M \to H$, by

$$F(g, \vec{\lambda}) := g + \mathcal{Q}(m, \vec{\iota}, \widetilde{\vec{\lambda}}) - \mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$$

Note that, $F(0, \widetilde{\vec{\lambda}}) = 0$ and

$$\|F(g, \vec{\lambda})\|_H \lesssim \|g\|_H + \sum_{j=1}^{M} \left|\frac{\lambda_j}{\widetilde{\lambda}_j} - 1\right|. \tag{2.32}$$

Next, define $G : H \times (0, \infty)^M \to \mathbb{R}^M$ by,

$$G(g, \vec{\lambda}) := \left(\frac{1}{\lambda_1} \left\langle \mathcal{Z}_{\underline{\lambda_1}} \mid F(g, \vec{\lambda}) \right\rangle, \ldots, \frac{1}{\lambda_M} \left\langle \mathcal{Z}_{\underline{\lambda_M}} \mid F(g, \vec{\lambda}) \right\rangle \right)$$

note that $G(0, \widetilde{\vec{\lambda}}) = \vec{0}$, and we record the computation,

$$\lambda_j \partial_{\lambda_j} G_j(g, \vec{\lambda}) = -\frac{1}{\lambda_j} \left\langle [(\underline{\Lambda} + 1)\mathcal{Z}]_{\underline{\lambda_j}} \mid F(g, \vec{\lambda}) \right\rangle - \iota_j \left\langle \mathcal{Z} \mid \Lambda Q \right\rangle$$

$$\lambda_i \partial_{\lambda_i} G_j(g, \vec{\lambda}) = -\iota_i \frac{\lambda_i}{\lambda_j} \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}} \right\rangle \quad \text{if} \quad i \neq j \tag{2.33}$$

At this point, it is convenient to change variables, letting $\ell_j := \log \lambda_j$ and $\widetilde{G}(g, \vec{\ell}) = G(g, \vec{\lambda})$. Note that $\partial_{\ell_j} = \lambda_j \partial_{\lambda_j}$. From (2.32) we see that $\widetilde{G}(\cdot, \cdot)$ is continuous near $0 \in H$ in the first slot and is C^1 near $\widetilde{\vec{\ell}} = (\log \widetilde{\lambda}_1, \ldots, \log \widetilde{\lambda}_M)$ in the last M variables. We compute,

$$L_0 := D_{\ell_1, \ldots \ell_M} \widetilde{G}(g, \vec{\ell}) \big|_{g=0, \vec{\ell}=\widetilde{\vec{\ell}}} = (A_{ij})_{1 \leq i,j \leq M}$$

where (A_{ij}) is the $M \times M$ matrix with entries,

$$A_{jj} = -\iota_j \langle \mathcal{Z} \mid \Lambda Q \rangle, \quad A_{ij} = -\iota_j \frac{\lambda_j}{\lambda_i} \left\langle \mathcal{Z}_{\underline{\lambda_i}} \mid \Lambda Q_{\underline{\lambda_j}} \right\rangle \quad \text{if} \ \ i \neq j \tag{2.34}$$

which one may check, using (2.18) and Corollary 2.21 is invertible and $\|L_0\|^{-1} = O(1)$. The conditions in (2.31) are readily verified, and one may take $\delta_1 = C_1 \eta$ and $\delta_2 = C_2 \eta$ in the notation of Remark 2.26 in that case for uniform constants C_1, C_2. Indeed,

$$|G(g, \vec{\lambda})| \lesssim \|g\|_H$$

and thus the second condition in (2.31) is verified. One may verify the first condition in (2.31) using (2.33) and (2.34).

An application of Remark 2.26 yields a continuous mapping $\varsigma : B_H(0; \delta_1) \to B_{\mathbb{R}^M}(0; \delta_2)$ such that

$$\widetilde{G}(g_0, \vec{\ell}) = \vec{0} \Longleftrightarrow \vec{\ell} = \varsigma(g_0).$$

We define

$$g := F(\widetilde{g}, \varsigma(\widetilde{g})), \quad \vec{\ell} := \varsigma(\widetilde{g}).$$

Setting $\lambda_j = e^{\ell_j}$, and $\boldsymbol{g} = (g, \dot{v})$, by construction we then have,

$$\boldsymbol{v} = \boldsymbol{\mathcal{Q}}(m, \vec{\iota}, \vec{\lambda}) + \boldsymbol{g},$$

and $\boldsymbol{g}$ satisfies (2.28) and (2.29).

To prove the remaining estimates we expand the nonlinear energy of $\boldsymbol{v}$,

$$ME(\boldsymbol{Q}) + \theta^2 \geq E(\boldsymbol{v}) = E(\boldsymbol{\mathcal{Q}}(m, \vec{\iota}, \vec{\lambda}) + \boldsymbol{g})$$

$$= E(\boldsymbol{\mathcal{Q}}(m, \vec{\iota}, \vec{\lambda})) + \left\langle \mathrm{D}\, E(\boldsymbol{\mathcal{Q}}(m, \vec{\iota}, \vec{\lambda})) \mid \boldsymbol{g} \right\rangle + \frac{1}{2} \left\langle \mathrm{D}^2\, E(\boldsymbol{\mathcal{Q}}(m, \vec{\iota}, \vec{\lambda}))\boldsymbol{g} \mid \boldsymbol{g} \right\rangle + O(\|\boldsymbol{g}\|_{\mathcal{E}}^3)$$

and apply the conclusions of Lemma 2.19 and Lemma 2.22. This completes the proof. $\qquad \square$

Lemma 2.27. *Let $k \geq 1$. There exists $\eta > 0$ sufficiently small with the following property. Let $m, \ell \in \mathbb{Z}$, $M, L \in \mathbb{N}$, $\vec{\iota} \in \{-1, 1\}^M, \vec{\sigma} \in \{-1, 1\}^L$, $\vec{\lambda} \in (0, \infty)^M, \vec{\mu} \in (0, \infty)^L$, and w be such that $E_{\mathbf{p}}(w) < \infty$ and,*

$$\|w - \mathcal{Q}(m, \vec{\iota}, \vec{\lambda})\|_H^2 + \sum_{j=1}^{M-1} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \leq \eta, \tag{2.35}$$

$$\|w - \mathcal{Q}(\ell, \vec{\sigma}, \vec{\mu})\|_H^2 + \sum_{j=1}^{L-1} \left(\frac{\mu_j}{\mu_{j+1}} \right)^k \leq \eta. \tag{2.36}$$

Then, $m = \ell$, $M = L$, $\vec{\iota} = \vec{\sigma}$. Moreover, for every $\theta > 0$ the number $\eta > 0$ above can be chosen small enough so that

$$\max_{j=1,\dots M} |\frac{\lambda_j}{\mu_j} - 1| \leq \theta. \tag{2.37}$$

Proof of Lemma 2.27. From (2.35) we see that $\lim_{r \to \infty} w(r) = m\pi$, and from (2.36) we see that $\lim_{r \to \infty} w(r) = \ell\pi$. Hence, $m = \ell$.

Next, let $g_\lambda := w - \mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$ and $g_\mu := w - \mathcal{Q}(\ell, \vec{\sigma}, \vec{\mu})$. By expanding the nonlinear potential energy we have,

$$E_{\mathbf{p}}(w) = E_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) + \left\langle DE_{\mathbf{p}}(\mathcal{Q}(m, \vec{\iota}, \vec{\lambda})) \mid g_\lambda \right\rangle + O(\|g_\lambda\|_H^2).$$

Choosing $\eta > 0$ small enough so that Lemma 2.22 applies, we see that

$$ME(\boldsymbol{Q}) - C\eta \le E_{\mathbf{p}}(w) \le ME(\boldsymbol{Q}) + C\eta,$$

for some $C > 0$. By an identical argument,

$$LE(\boldsymbol{Q}) - C\eta \le E_{\mathbf{p}}(w) \le LE(\boldsymbol{Q}) + C\eta.$$

It follows that $M = L$. Next, we prove that $\eta > 0$ can be chosen small enough to ensure that $\vec{\iota} = \vec{\sigma}$. Suppose not, then we can find a sequence w_n with $E_{\mathbf{p}}(w_n) \le C$, and sequences $\vec{\iota}_n, \vec{\sigma}_n, \vec{\lambda}_n, \vec{\mu}_n$ so that,

$$\|w_n - \mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n)\|_H^2 + \sum_{j=1}^{M-1} \left(\frac{\lambda_{n,j}}{\lambda_{n,j+1}}\right)^k = o_n(1) \quad \text{as} \quad n \to \infty,$$

$$\|w_n - \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n)\|_H^2 + \sum_{j=1}^{M-1} \left(\frac{\mu_{n,j}}{\mu_{n,j+1}}\right)^k = o_n(1) \quad \text{as} \quad n \to \infty,$$

but with $\vec{\iota}_n \ne \vec{\sigma}_n$ for every n. We may assume without loss of generality that

$$0 = \lim_{r \to 0} w_n(r) = \lim_{r \to 0} \mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n; r) = \lim_{r \to 0} \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n; r)$$

and we note that above limits agree mean that we must have $\sum_{j=1}^{M} \iota_{n,j} = \sum_{j=1}^{M} \sigma_{n,j}$ for each n. Passing to a subsequence we may assume that there exists an index $j_0 \ge 1$ such that $\iota_{j,n} = \sigma_{j,n}$ for every $j < j_0$ and every n and $\iota_{j_0,n} \ne \sigma_{j_0,n}$ for every n. We have,

$$\begin{aligned} \|\mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n) - \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n)\|_H & \\ \le \|w_n - \mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n)\|_H &+ \|w_n - \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n)\|_H = o_n(1). \end{aligned} \tag{2.38}$$

First we show that $j_0 > 1$ Assume for contradiction that $j_0 = 1$. Then, we may assume that $\iota_{n,1} = 1$, $\sigma_{n,1} = -1$ and $\lambda_{n,1} < \mu_{n,1}$ for all n. It follows that

$$\mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n) - \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n) \ge \frac{\pi}{4} \quad \forall r \in [\lambda_{n,1}, 2\lambda_{n,1}],$$

for all n large enough. But then,

$$\|\mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n) - \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n)\|_H^2 \ge \int_{\lambda_{n,1}}^{2\lambda_{n,1}} (\pi/4)^2 \frac{\mathrm{d}r}{r} \ge (\pi/4)^2 \log 2,$$

for all sufficiently large n, which contradicts (2.38). So $\iota_{1,n} = \sigma_{n,1}$ for all n. Thus $j_0 > 1$. But then by a nearly identical argument we can show that we must have $\lambda_{n,j} \simeq \mu_{n,j}$ uniformly in n for all $j < j_0$. Again we may assume (after passing to a subsequence) that $\lambda_{n,j_0} < \mu_{n,j_0}$. It follows again that for all sufficiently large n we have,

$$\left|\mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n) - \mathcal{Q}(m, \vec{\sigma}_n, \vec{\mu}_n)\right| \ge \frac{\pi}{4} \quad \forall r \in [\lambda_{n,j_0}, 2\lambda_{n,j_0}],$$

which again yields a contradiction. Hence we must have $\vec{\iota} = \vec{\sigma}$.

Finally, we prove (2.37). Suppose (2.37) fails. Then there exists $\theta_0 > 0$ and sequences $\vec{\lambda}_n, \vec{\mu}_n$ such that

$$\|\mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n) - \mathcal{Q}(m, \vec{\iota}_n, \vec{\mu}_n)\|_H = o_n(1),$$

but

$$\sup_{j=1,\dots,M} |\lambda_{n,j}/\mu_{n,j} - 1| \ge \theta_0, \tag{2.39}$$

for all n. Following the same logic as before we note that we must have $\lambda_{n,j} \simeq \mu_{n,j}$ uniformly in n. But then we have,

$$\|\mathcal{Q}(m, \vec{\iota}_n, \vec{\lambda}_n) - \mathcal{Q}(m, \vec{\iota}_n, \vec{\mu}_n)\|_H^2 = \sum_{j=1}^M \|Q_{\lambda_{n,j}} - Q_{\mu_{n,j}}\|_H^2 + o_n(1),$$

which implies that $\|Q_{\lambda_{n,j}} - Q_{\mu_{n,j}}\|_H = o_n(1)$ for every j, yielding a contradiction with (2.39). This completes the proof. $\qquad\square$

3. Localized sequential bubbling

The goal of this section is to prove a localized sequential bubbling lemma for sequences of wave maps with vanishing averaged kinetic energy on an expanding region of space. The main result, and the arguments used to prove it are in the spirit of the main theorems in Côte [10] and Jia and Kenig [42], and also use many ideas from Struwe [73] and Duyckaerts, Kenig, and Merle [22].

To state the compactness lemma, we define a localized distance function,

$$\boldsymbol{\delta}_R(\boldsymbol{u}) := \inf_{m,M,\vec{\iota},\vec{\lambda}} \left(\|u - \mathcal{Q}(m, \vec{\iota}, \vec{\lambda})\|_{H(r \le R)}^2 + \|\dot{u}\|_{L^2(r \le R)}^2 + \sum_{j=1}^M \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \right)^{\frac{1}{2}}. \qquad (3.1)$$

where the infimum above is taken over all $m \in \mathbb{Z}$, $M \in \{0, 1, 2, \dots\}$, and all vectors $\vec{\iota} \in \{-1, 1\}^M, \vec{\lambda} \in (0, \infty)^M$, and here we use the convention that the last scale $\lambda_{M+1} = R$.

Lemma 3.1 (Compactness Lemma). *Let* $\ell, m \in \mathbb{Z}$. *Let* $\rho_n > 0$ *be a sequence of positive numbers and let* $\boldsymbol{u}_n(t) \in \mathcal{E}_{\ell,m}$ *be a sequence of wave maps on the time intervals* $[0, \rho_n]$ *such that* $\limsup_{n \to \infty} E(\boldsymbol{u}_n) < \infty$.

Suppose there exists a sequence $R_n \to \infty$ *such that,*

$$\lim_{n \to \infty} \frac{1}{\rho_n} \int_0^{\rho_n} \int_0^{\rho_n R_n} |\partial_t u_n(t, r)|^2 \, r \, dr \, dt = 0.$$

Then, up to passing to a subsequence of the $\boldsymbol{u}_n$, *there exists a time sequence* $t_n \in [0, \rho_n]$ *and a sequence* $r_n \le R_n$ *with* $r_n \to \infty$ *such that*

$$\lim_{n \to \infty} \boldsymbol{\delta}_{r_n \rho_n}(\boldsymbol{u}_n(t_n)) = 0.$$

Remark 3.2. We note that Theorem 1.15 in the blow-up case is a quick consequence of Lemma 3.1 together with the fundamental result of Shatah and Tahvildar-Zadeh [70], that for wave map developing a singularity at $T_- = 0$ one has,

$$\lim_{t \to 0} \frac{1}{t} \int_0^t \int_0^\tau |\partial_t u(\tau, r)|^2 \, r \, dr \, dt = 0.$$

In the global case $T_+ = \infty$ one uses,

$$\lim_{A \to \infty} \limsup_{T \to \infty} \frac{1}{T} \int_A^T \int_0^{t-A} |\partial_t u(t, r)|^2 \, r \, dr \, dt = 0,$$

proved in [12] using the analysis of [70].

3.1. **Prior results on bubbling.** The proof of Lemma 3.1 requires several preliminary lemmas, including two Real Analysis results, which we address first.

Lemma 3.3. *If $a_{k,n}$ are positive numbers such that $\lim_{n\to\infty} a_{k,n} = \infty$ for all $k \in \mathbb{N}$, then there exists a sequence of positive numbers b_n such that $\lim_{n\to\infty} b_n = \infty$ and $\lim_{n\to\infty} a_{k,n}/b_n = \infty$ for all $k \in \mathbb{N}$.*

Proof. For each k and each n define $\widetilde{a}_{k,n} = \min\{a_{1,n}, \ldots, a_{k,n}\}$. Then the sequences $\widetilde{a}_{k,n} \to \infty$ as $n \to \infty$ for each k, but also satisfy $\widetilde{a}_{k,n} \le a_{k,n}$ for each k, n, as well as $\widetilde{a}_{j,n} \le \widetilde{a}_{k,n}$ if $j > k$. Next, choose a strictly increasing sequence $\{n_k\}_k \subset \mathbb{N}$ such that $\widetilde{a}_{k,n} \ge k^2$ as long as $n \ge n_k$. For n large enough, let $b_n \in \mathbb{N}$ be determined by the condition $n_{b_n} \le n < n_{b_n+1}$. Observe that $b_n \to \infty$ as $n \to \infty$. Now fix any $\ell \in \mathbb{N}$ and let n be such that $b_n > \ell$. We then have

$$a_{\ell,n} \ge \widetilde{a}_{\ell,n} \ge \widetilde{a}_{b_n,n} \ge b_n^2 \gg b_n.$$

Thus the sequence b_n has the desired properties. $\qquad\square$

If $f : [0,1] \to [0, +\infty]$ is a measurable function, we denote by

$$Mf(\tau) := \sup_{I \ni \tau; I \subset [0,1]} \frac{1}{|I|} \int_I f(t)\, \mathrm{d}t$$

its Hardy-Littlewood maximal function. Recall the weak-L^1 boundedness estimate

$$|\{\tau \in [0,1] : Mf(\tau) > \alpha\}| \le \frac{3}{\alpha} \int_0^1 f(t)\mathrm{d}t, \qquad \text{for all } \alpha > 0, \tag{3.2}$$

see [55, Section 2.3].

Lemma 3.4. *Let f_n be a sequence of continuous positive functions defined on $[0,1]$ such that $\lim_{n\to\infty} \int_0^1 f_n(t)\mathrm{d}t = 0$ and let g_n be a uniformly bounded sequence of real-valued continuous functions on $[0,1]$ such that $\limsup_{n\to\infty} \int_0^1 g_n(t)\mathrm{d}t \le 0$. Then there exists a sequence $t_n \in [0,1]$ such that*

$$\lim_{n\to\infty} Mf_n(t_n) = 0, \quad \lim_{n\to\infty} f_n(t_n) = 0, \quad and \quad \limsup_{n\to\infty} g_n(t_n) \le 0.$$

Proof. Let α_n be a sequence such that $\int_0^1 f_n(t)\mathrm{d}t \ll \alpha_n \ll 1$. Let $A_n := \{t \in [0,1] : Mf_n(t) + f_n(t) \le \alpha_n\}$. By (3.2) and Tchebychev's inequality $\lim_{n\to\infty} |A_n| = 1$. Since g_n is uniformly bounded, we have

$$\int_{[0,1]\setminus A_n} |g_n(t)|\mathrm{d}t \lesssim |[0,1] \setminus A_n| \to 0,$$

which implies

$$\limsup_{n\to\infty} \int_{A_n} g_n(t)\mathrm{d}t \le 0.$$

It suffices to take $t_n \in A_n$ such that $g_n(t_n) \le |A_n|^{-1} \int_{A_n} g_n(t)\mathrm{d}t$. $\qquad\square$

A key ingredient of the proof of Lemma 3.1 is a Struwe-type bubbling lemma [73]. We require the version proved in [10, 42].

Lemma 3.5 (Bubbling). *[73], [10, Proposition 3.1], [42, Lemma 5.6] Let $\sigma > 0$ and let $\alpha_n \to 0$ and $\beta_n \to \infty$ be two sequences. Let $\boldsymbol{v}_n$ be a sequence of wave maps, i.e., solutions to (1.1), on the time interval $[0, \sigma]$ such that $\limsup_{n\to\infty} E(\boldsymbol{v}_n) < \infty$. Suppose that*

$$\lim_{n\to\infty} \frac{1}{\sigma} \int_0^\sigma \int_{\alpha_n}^{\beta_n} |\partial_t v_n(t,r)|^2\, r\, \mathrm{d}r\, \mathrm{d}t = 0.$$

Then, there exists an integer m_0, $\iota_0 \in \{-1, 0, +1\}$, and a scale $\lambda_0 > 0$ such that, up to passing to a subsequence, we have

$$\boldsymbol{v}_n \to m_0\boldsymbol{\pi} + \iota_0\boldsymbol{Q}_{\lambda_0}$$

in the space $(L_t^2(\mathcal{E}))_{\mathrm{loc}}([0, \sigma] \times (0, \infty))$. In addition $\boldsymbol{v}_n \to m_0\boldsymbol{\pi} + \iota_0\boldsymbol{Q}_{\lambda_0}$ locally uniformly in $[0, \sigma] \times (0, \infty)$. And finally, $\boldsymbol{v}_n(0) \to m_0\boldsymbol{\pi} + \iota_0\boldsymbol{Q}_{\lambda_0}$ in the space $\mathcal{E}_{\mathrm{loc}}((0, \infty))$.

The lengthy proof of the Lemma 3.1 will consist of several steps, which are designed to reduce the proof to the exact scenarios already considered by Côte in [10, Proof of Lemma 3.5] and then by Jia-Kenig in [42, Proof of Theorem 3.2]. In particular, we will seek to apply the following result from [42].

Lemma 3.6. [42, Theorem 3.2] *Let $\boldsymbol{v}_n$ be a sequence of wave maps, i.e., solutions to (1.1), on the time interval $[0, 1]$ such that $\limsup_{n\to\infty} E(\boldsymbol{v}_n) < \infty$. Suppose that there exists a sequence $t_n \in [0, 1]$, and integer $K_0 \geq 0$, and scales $\lambda_{n,1} \ll \cdots \ll \lambda_{n,K_0} \lesssim 1$ such that*

$$\boldsymbol{v}_n(t_n) = m_1\boldsymbol{\pi} + \sum_{j=1}^{K_0} \iota_j((Q(\frac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}) + \boldsymbol{w}_{n,0},$$

where $\|\boldsymbol{w}_{n,0}\|_{L^\infty \times L^2} \to 0$ and $\|\boldsymbol{w}_{n,0}\|_{\mathcal{E}(r \geq r_n^{-1})} \to 0$ as $n \to \infty$ for some sequence $r_n \to \infty$. Suppose in addition that, $\|\boldsymbol{w}_{n,0}\|_{\mathcal{E}(A^{-1}\lambda_n \leq r \leq A\lambda_n)} \to 0$ as $n \to \infty$ for any sequence $\lambda_n \lesssim 1$ and any $A > 1$, and finally, that

$$\limsup_{n\to\infty} \int_0^\infty \left(k^2\frac{\sin^2(2v_n(t_n))}{2r^2} + (\partial_r v_n(t_n))^2 2\cos(2v_n(t_n))\right) r\,\mathrm{d}r \leq 0. \tag{3.3}$$

Then,

$$\|\boldsymbol{w}_{n,0}\|_{\mathcal{E}} \to 0 \quad as \quad n \to \infty.$$

Remark 3.7. Lemma 3.6 is not stated in [42] exactly as given above. However, an examination of [42, Proof of Theorem 3.2] shows that this is precisely what is established. The heart of the matter lies in the fact that the Jia-Kenig virial functional (3.3) vanishes at Q, i.e.,

$$\int_0^\infty \left(k^2\frac{\sin^2(2Q)}{2r^2} + (\partial_r Q)^2 2\cos(2Q)\right) r\,\mathrm{d}r = 0,$$

but gives coercive control of the energy in regions where $v_n(t, r)$ is near integer multiples of π.

3.2. Proof of the compactness lemma.

Proof of Lemma 3.1. Rescaling we may assume that $\rho_n = 1$ for each n.

Step 1. We claim that there exist $\sigma_n \in [0, \frac{1}{3}]$, $\tau_n \in [\frac{2}{3}, 1]$, such that

$$\lim_{n\to\infty} \int_{\sigma_n}^{\tau_n} \int_0^\infty \left(k^2\frac{\sin^2(2u_n)}{2r^2}\chi - (\partial_r^2 u_n + \frac{1}{r}\partial_r u_n)\sin(2u_n)\chi\right) r\mathrm{d}r\mathrm{d}t = 0, \tag{3.4}$$

where χ is a smooth cut-off function equal 1 on $[0, \frac{1}{2}]$, with support in $[0, 1]$. Here and later in the argument the second term in the integrand in (3.4) is to be interpreted as the expression obtained after integration by parts, which is well defined due to the finiteness of the energy.

Since

$$\lim_{n\to\infty} \int_0^{\frac{1}{3}} \int_0^{R_n} (\partial_t u_n)^2 \, r\mathrm{d}r\mathrm{d}t = 0 \quad \text{and} \quad \lim_{n\to\infty} \int_{\frac{2}{3}}^1 \int_0^{R_n} (\partial_t u_n)^2 \, r\mathrm{d}r\mathrm{d}t = 0$$

there exist $\sigma_n \in [0, \frac{1}{3}]$, $\tau_n \in [\frac{2}{3}, 1]$ such that,

$$\lim_{n \to \infty} \int_0^{R_n} (\partial_t u_n(\sigma_n))^2 \, r \mathrm{d}r = 0 \quad \text{and} \quad \lim_{n \to \infty} \int_0^{R_n} (\partial_t u_n(\tau_n))^2 \, r \mathrm{d}r = 0. \tag{3.5}$$

For $t \in [\sigma_n, \tau_n]$, we have the following Jia-Kenig virial identity; see [42, Lemma 3.5 and Lemma 3.10].

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \int_0^\infty \partial_t u_n \sin(2u_n) \chi \, r \mathrm{d}r &= \int_0^\infty 2 \cos(2u_n)(\partial_t u_n)^2 \chi \, r \mathrm{d}r \\
&+ \int_0^\infty (\partial_r^2 u_n + \frac{1}{r} \partial_r u_n - k^2 \frac{\sin(2u_n)}{2r^2}) \sin(2u_n) \chi \, r \mathrm{d}r.
\end{aligned} \tag{3.6}$$

By the Cauchy-Schwarz inequality, the boundedness of the nonlinear energy and (3.5), we see that

$$\lim_{n \to \infty} \int_0^\infty \left(|\partial_t u_n(\sigma_n)| |\sin(2u_n(\sigma_n))| + |\partial_t u_n(\tau_n)| |\sin(2u_n(\tau_n))| \right) \chi \, r \mathrm{d}r = 0.$$

Integrating (3.6) between σ_n and τ_n, and using the above, we obtain (3.4).

Step 2. We rescale again so that $[\sigma_n, \tau_n]$ becomes $[0, 1]$. We apply Lemma 3.4, to

$$f_n(t) := \int_0^{R_n} |\partial_t u_n(t, r)|^2 \, r \, \mathrm{d}r,$$

$$g_n(t) := \int_0^\infty \left(k^2 \frac{\sin^2(2u_n)}{2r^2} - (\partial_r^2 u_n + \frac{1}{r} \partial_r u_n) \sin(2u_n) \right) \chi \, r \mathrm{d}r$$

(integrating by parts the second term in g_n, we see that this is a uniformly bounded sequence of continuous functions) and we find a sequence $\{t_n\} \in [0, 1]$ such that we have vanishing of the maximal function of the local kinetic energy,

$$\lim_{n \to \infty} \sup_{I \ni t_n; I \subset [0,1]} \frac{1}{|I|} \int_I \int_0^{R_n} |\partial_t u_n(t, r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t = 0,$$

$$\text{and} \quad \lim_{n \to \infty} \int_0^{R_n} |\partial_t u_n(t_n, r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t = 0, \tag{3.7}$$

and also pointwise vanishing of a localized Jia-Kenig virial functional,

$$\limsup_{n \to \infty} \int_0^\infty \left(k^2 \frac{\sin^2(2u_n(t_n))}{2r^2} - (\partial_r^2 u_n(t_n) + \frac{1}{r} \partial_r u_n(t_n)) \sin(2u_n(t_n)) \right) \chi \, r \mathrm{d}r \leq 0. \tag{3.8}$$

We emphasize the conclusion from the first steps is the existence of the sequence t_n such that (3.7) and (3.8) hold.

Step 3. Now that we have chosen the sequence $t_n \in [0, 1]$, we may, after passing to a subsequence, assume that $t_n \to t_0 \in [0, 1]$.

We apply Lemma 2.10 to the sequence $\boldsymbol{u}_n(t_n)$, obtaining profiles $(\boldsymbol{\psi}^j, \lambda_{n,j})$ and $(\boldsymbol{v}_L^i, t_{n,i}, \sigma_{n,i})$, and $\boldsymbol{w}_{n,0}^J$, so that, using the notation,

$$\boldsymbol{\psi}_n^j := \left(\psi^j(\cdot/\lambda_{n,j}), \lambda_{n,j}^{-1} \dot{\psi}^j(\cdot/\lambda_{n,j}) \right), \quad \boldsymbol{v}_{L,n}^i(0) := \left(v_L^i(\frac{-t_{n,i}}{\sigma_{n,i}}, \frac{\cdot}{\sigma_{n,i}}), \sigma_{n,i}^{-1} \partial_t v_L^i(\frac{-t_{n,i}}{\sigma_{n,i}}, \frac{\cdot}{\sigma_{n,i}}) \right),$$

we have

$$\boldsymbol{u}_n(t_n) = m_0 \boldsymbol{\pi} + \sum_{j=1}^{K_0} (\boldsymbol{\psi}_n^j - m_j \boldsymbol{\pi}) + \sum_{i=1}^{J} \boldsymbol{v}_{L,n}^i(0) + \boldsymbol{w}_{n,0}^J \tag{3.9}$$

satisfying the conclusions of Lemma 2.10. We refer to the profiles $(\boldsymbol{\psi}^j, \lambda_{n,j})$ as well as the profiles $(\boldsymbol{v}_{\mathrm{L}}^i(0), t_{n,i}, \sigma_{n,i})$ with $t_{n,i} = 0$ for all n as *centered* profiles (here the subscript $_{\mathrm{L}}$ on $\boldsymbol{v}_{\mathrm{L}}^i$ is superfluous). We refer to the profiles $(\boldsymbol{v}_{\mathrm{L}}^i(0), t_{n,i}, \sigma_{n,i})$ with $-t_{n,i}/\sigma_{n,i} \to \pm\infty$ as *outgoing/incoming* profiles.

Step 4.(Centered profiles at large scales) At each step, we will impose conditions on the choice of the ultimate choice of sequence $r_n \to \infty$. Consider the sets of indices

$$\mathcal{J}_{c,\infty,} := \{j \in \{1, \ldots, K_0\} \mid \lim_{n\to\infty} \lambda_{n,j} = \infty\}$$
$$\mathcal{I}_{c,\infty} := \{i \in \mathbb{N} \mid t_{n,i} = 0 \ \forall n, \ \text{ and } \ \lim_{n\to\infty} \sigma_{n,i} = \infty\}.$$

Using Lemma 3.3 we choose a sequence $r_{0,n} \to \infty$ so that $r_{0,n} \ll R_n, \lambda_{n,j}, \sigma_{n,i}$ for each $\lambda_{n,j}$ with $j \in \mathcal{J}_{c,\infty}$ and each $\sigma_{n,i}$ with $i \in \mathcal{I}_{c,\infty}$. We note that by construction we have,

$$E(\boldsymbol{\psi}_n^j; 0, r_{0,n}) \to 0 \ \text{ as } \ n \to \infty,$$
$$E((v_{\mathrm{L}}^i(0, \cdot/\sigma_{n,i}), \sigma_{n,i}^{-1} \dot{v}_{\mathrm{L}}^i(0, \cdot/\sigma_{n,i})); 0, r_{0,n}) \to 0 \ \text{ as } \ n \to \infty, \tag{3.10}$$

for any of the indices $j \in \mathcal{J}_{c,\infty}, i \in \mathcal{I}_{c,\infty}$.

Step 5.(Centered profiles at bounded scales) Consider the sets of indices

$$\mathcal{J}_{c,0} := \{j \in \{1, \ldots, K_0\} \mid \lim_{n\to\infty} \lambda_{n,j} < \infty\}$$
$$\mathcal{I}_{c,0} := \{i \in \mathbb{N} \mid t_{n,i} = 0 \ \forall n, \ \text{ and } \ \lim_{n\to\infty} \sigma_{n,i} < \infty\}$$

We use Lemma 3.5 to show that each of the associated profiles must be a harmonic map.

Consider first the case of a profile $(\boldsymbol{\psi}^j, \lambda_{n,j})$ with $j \in \mathcal{J}_{c,0}$. Define,

$$\boldsymbol{u}_n^j(t,r) = (u_n^j(t,r), \partial_t u_n^j(t,r)) := (u_n(t_n + \lambda_{n,j}t, \lambda_{n,j}r), \lambda_{n,j}\partial_t u_n(t_n + \lambda_{n,j}t, \lambda_{n,j}r))$$

and note that $\boldsymbol{u}_n^j$ is a wave map on the interval $t \in [-t_n/\lambda_{n,j}, (\rho - t_n)/\lambda_{n,j}]$. Consider the case $t_n \to t_0 < 1$, (the other possible limits have nearly identical arguments). Recall that we have the weak convergence $\boldsymbol{u}_n^j(0) \rightharpoonup \boldsymbol{\psi}^j$. Moreover,

$$\frac{1}{\sigma} \int_0^\sigma \int_0^{\frac{R_n}{\lambda_{n,j}}} |\partial_t u_n^j(t,r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t = \frac{1}{\sigma} \int_0^\sigma \int_0^{\frac{R_n}{\lambda_{n,j}}} |\lambda_{n,j}\partial_t u_n(t_n + \lambda_{n,j}t, \lambda_{n,j}r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t$$
$$= \frac{1}{\sigma\lambda_{n,j}} \int_{t_n}^{t_n+\lambda_{n,j}\sigma} \int_0^{R_n} |\partial_t u_n(s,y)|^2 \, y \, \mathrm{d}y \, \mathrm{d}s \to 0 \ \text{ as } \ n \to \infty$$

where the last line follows from (3.7) after fixing $\sigma > 0$ small enough so that $t_n + \lambda_{n,j}\sigma \leq 1$ for all n large enough. Thus by Lemma 3.5 we conclude that there exists $\widetilde{m}_j, \iota_j, \lambda_{0,j}$ so that $\boldsymbol{\psi}^j = \widetilde{m}_j\boldsymbol{\pi} + \iota_j\boldsymbol{Q}_{\lambda_{0,j}}$.

The cases of profiles $(\boldsymbol{v}_{\mathrm{L}}^i(0), t_{n,i}, \sigma_{n,i})$ with $i \in \mathcal{I}_{c,0}$ are completely analogous. And we conclude that each of these profiles must satisfy

$$\boldsymbol{v}_{\mathrm{L}}^i(0, r) = (0, 0).$$

since each $\boldsymbol{v}_{\mathrm{L}}^i(0) \in \mathcal{E}$ and thus can only be a constant harmonic map.

Step 6.(Incoming/outgoing profiles with $\lim_{n\to\infty} |t_{n,i}| = \infty$) We next treat profiles $(\boldsymbol{v}_{\mathrm{L}}^i, t_{n,i}, \sigma_{n,i})$ that satisfy,

$$-\frac{t_{n,i}}{\sigma_{n,i}} \to \pm\infty.$$

Up to passing to a subsequence of $\boldsymbol{u}_n(t_n)$ we may assume that $-t_{n,i} \to t_\infty \in [-\infty, \infty]$. Consider the set of indices,

$$\mathcal{I}_{\mathrm{L},\infty} := \{i \in \mathbb{N} \mid -\frac{t_{n,i}}{\sigma_{n,i}} \to \pm\infty \text{ and } |t_{n,i}| \to \infty\}.$$

We impose additional restrictions on the sequence r_n. We require that $r_n \leq \frac{1}{2}|t_{n,i}|$ for sufficiently large n, for each sequence $t_{n,i}$ in $\mathcal{I}_{\mathrm{L},\infty}$. So at this stage, we again use Lemma 3.3 to choose a sequence $r_{1,n} \to \infty$ such that $r_{1,n} \leq r_{0,n}$ and $r_{1,n} \leq \frac{1}{2}|t_{n,i}|$ for sufficiently large n, for each sequence $t_{n,i}$ in $\mathcal{J}_{\mathrm{L},\infty}$.

Since $\boldsymbol{v}_{\mathrm{L}}^i$ is a solution to (2.3) we know that it asymptotically concentrates all of its energy near the light-cone. In fact, a direct consequence of [15, Theorem 4] is that

$$\lim_{s \to \pm\infty} \|\boldsymbol{v}_{\mathrm{L}}^i(s)\|_{\mathcal{E}(r \leq \frac{1}{2}|s|)} = 0.$$

Thus, if $i \in \mathcal{I}_{\mathrm{L},\infty}$ and as long as $r_{1,n} \leq \frac{1}{2}|t_{n,i}|$ for n large enough, we see that $\sigma_{n,i}^{-1}r_{1,n} \leq \frac{1}{2}\sigma_{n,i}^{-1}|t_{n,i}|$ and thus

$$\|\boldsymbol{v}_{\mathrm{L}}^i(-t_{n,i}/\sigma_{n,i})\|_{\mathcal{E}(r \leq r_{1,n}\sigma_{n,i}^{-1})} \to 0 \quad \text{as} \quad n \to \infty. \tag{3.11}$$

by the above and we conclude that any such profile does not contribute to the asymptotic size of $\boldsymbol{\delta}_{r_{1,n}}(\boldsymbol{u}_n(t_n))$.

Step 7. (Incoming/outgoing profiles with $\lim_{n\to\infty}|t_{n,i}| < \infty$) Next, we consider profiles $(\boldsymbol{v}_{\mathrm{L}}^i, t_{n,i}, \sigma_{n,i})$ such that

$$-\frac{t_{n,i}}{\sigma_{n,i}} \to \pm\infty \quad \text{and} \quad -t_{n,i} \to t_{\infty,i} \in \mathbb{R}$$

and we denote by $\mathcal{I}_{\mathrm{L},0}$ the set indices labeling all such profiles, and note that $\sigma_{n,i} \to 0$ as $n \to \infty$ for each $i \in \mathcal{I}_{\mathrm{L},0}$. We claim that any such profile must satisfy $\boldsymbol{v}_{\mathrm{L}}^i \equiv 0$. The argument we use follows closely the argument given in [19, Erratum]. As there are few technical changes due to setting of the current problem, we reproduce the argument here.

We claim that there exists a new sequence $\sqrt{r_{1,n}} \leq r_{2,n} \leq r_{1,n}$ such that

$$\lim_{n \to \infty} \sup_{t \in [0,1]} E(\boldsymbol{u}_n(t); A_n^{-1}r_{2,n}, A_n r_{2,n}) = 0 \tag{3.12}$$

for some $1 \ll A_n \ll r_{2,n}$. By the finite speed of propagation, it suffices to have

$$\lim_{n \to \infty} E(\boldsymbol{u}_n(0); A_n^{-1}r_{2,n}, A_n r_{2,n}) = 0,$$

and then replace A_n by its half, for example.

Let A_n be the largest integer such that $A_n^{2A_n} \leq \sqrt{r_{1,n}}$. Obviously, $1 \ll A_n \ll \sqrt{r_{1,n}}$. For $l \in \{0, 1, \ldots, A_n - 1\}$, set $R_n^{(l)} := A_n^{2l}\sqrt{r_{1,n}}$, so that $A_n^{-1}R_n^{(l+1)} = A_n R_n^{(l)}$, thus

$$\sum_{l=0}^{A_n-1} E(\boldsymbol{u}_n(0); A_n^{-1}R_n^{(l)}, A_n R_n^{(l)}) \leq E(\boldsymbol{u}_n(0)).$$

Since all the terms of the sum are positive, there exists $l_0 \in \{0, 1, \ldots, A_n - 1\}$ such that $r_{2,n} := R_n^{(l_0)}$ satisfies

$$E(\boldsymbol{u}_n(0); A_n^{-1}r_{2,n}, A_n r_{2,n}) \leq A_n^{-1}E(\boldsymbol{u}_n(0)) \to 0,$$

proving (3.12).

Next, using the finite speed of propagation along with (3.12), we pass to a new sequence of maps $\widetilde{\boldsymbol{u}}_n$ with vanishing average kinetic energy on the whole space. To see this, first we use Lemma 2.1 to find a sequence $y_n \in [2r_{2,n}, 4r_{2,n}]$ and integers $m_n \in \mathbb{Z}$ such that

$$|u_n(0, y_n) - m_n\pi| \to 0 \quad \text{as} \quad n \to \infty.$$

Since ℓ, m are fixed and $\limsup_{n\to\infty} E(\boldsymbol{u}_n) < \infty$, the integers $m_n \in [-L, L]$ for all n for some $L > 0$. Hence, after passing to a subsequence, we may assume that $m_n = m_1$ is a fixed integer for each n. We define a sequence of truncated initial data $\widetilde{\boldsymbol{u}}_n$ as follows,

$$\widetilde{\boldsymbol{u}}_n(t_n, r) = \chi_{2r_{2,n}}(r)\boldsymbol{u}_n(t_n, r) + (1 - \chi_{2r_{2,n}}(r))m_1\boldsymbol{\pi}. \tag{3.13}$$

Using (3.12), we have $E(\widetilde{\boldsymbol{u}}_n(t_n); \frac{1}{8}r_{2,n}, 8r_{2,n}) \to 0$ as $n \to \infty$. Let $\widetilde{\boldsymbol{u}}_n(t)$ denote the wave map evolution of the data $\widetilde{\boldsymbol{u}}_n(t_n)$, which we observe, using the vanishing of the energy of the data on the region $[r_{2,n}/8, 8r_{2,n}]$ is well defined on the interval $[0, 1]$ for large n. In fact, using the finite speed of propagation and the monotonicity of the energy on truncated cones, we see that $\widetilde{\boldsymbol{u}}_n(t)$ satisfies,

$$\widetilde{\boldsymbol{u}}_n(t, r) = \boldsymbol{u}_n(t, r) \ \text{ if } \ r \le r_{2,n}, \quad \text{and} \quad \sup_{t \in [0,1]} E(\widetilde{\boldsymbol{u}}_n(t); r_{2,n}, \infty) \to 0 \ \text{ as } \ n \to \infty. \tag{3.14}$$

Next, from the decomposition (3.9) we have,

$$\begin{aligned}
\widetilde{\boldsymbol{u}}_n(t_n) = {} & m_1\boldsymbol{\pi} + \sum_{j \in \mathcal{J}_{c,0}} (\iota_j Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}) + \sum_{i \le J, \, i \in \mathcal{I}_{L,0}} \boldsymbol{v}_{L,n}^i(0) + \chi_{2r_{2,n}} \boldsymbol{w}_{n,0}^J \\
& - \chi_{2r_{2,n}} m_1\boldsymbol{\pi} + \chi_{2r_{2,n}} m_0\boldsymbol{\pi} + (\chi_{2r_{2,n}} - 1) \sum_{j \in \mathcal{J}_{c,0}} (\iota_j Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}) \\
& + (\chi_{2r_{2,n}} - 1) \sum_{i \le J, \, i \in \mathcal{I}_{L,0}} \boldsymbol{v}_{L,n}^i(0) + \chi_{2r_{2,n}} \sum_{j \in \mathcal{J}_{c,\infty}} \boldsymbol{\psi}_n^j
\end{aligned} \tag{3.15}$$

where above we have allowed the abuse of notation, $\lambda_{n,j} \leftrightarrow \lambda_{n,j}\lambda_{0,j}$, for the profiles with indices in $\mathcal{J}_{c,0}$. Using the same logic used to deduce (3.10) and (3.11) we have,

$$E(Q_{\lambda_{n,j}}; r_{2,n}, \infty) \to 0 \ \text{ as } \ n \to \infty, \quad \|\boldsymbol{v}_L^i(-t_{n,i}/\sigma_{n,i})\|_{\mathcal{E}(r \ge r_{2,n}\sigma_{n,i}^{-1})} \to 0 \ \text{ as } \ n \to \infty$$

for any fixed $j \in \mathcal{J}_{c,0}$ or $i \in \mathcal{I}_{L,0}$. Thus, using (2.9) and the above along with (3.10) and (3.11) we see that the last three lines in (3.15) can effectively be absorbed into the error and writing

$$\widetilde{\boldsymbol{w}}_{n,0}^J(r) := \chi_{2r_{2,n}}(r)\boldsymbol{w}_{n,0}^J(r) + o_n(1)$$

we obtain the decomposition,

$$\widetilde{\boldsymbol{u}}_n(t_n) = m_1\boldsymbol{\pi} + \sum_{j \in \mathcal{J}_{c,0}} \iota_j((Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}) + \sum_{i \le J, \, i \in \mathcal{I}_{L,0}} \boldsymbol{v}_{L,n}^i(0) + \widetilde{\boldsymbol{w}}_{n,0}^J. \tag{3.16}$$

The above is a profile decomposition for $\widetilde{\boldsymbol{u}}_n(t_n)$ in the sense of Lemma 2.10 – the vanishing properties of the error $\widetilde{\boldsymbol{w}}_{n,0}^J$ are inherited from the corresposnding vanishing properties of $\boldsymbol{w}_{n,0}^J$ using e.g., [15, Lemmas 10 and 11] after noting the correspondence between the linear wave equation (2.6) and the $2k + 2$-dimensional radially symmetric free wave equation (see also [21, Claim A.1 and Claim 2.11] for the treatment of the wave equation in odd dimensions).

Assume for the sake of contradiction that there exists a nonzero profile $(\boldsymbol{v}_L^{i_0}, \sigma_{n,i_0}, t_{n,i_0})$ with index $i_0 \in \mathcal{I}_{L,0}$, and assume without loss of generality that

$$\frac{-t_{n,i_0}}{\sigma_{n,i_0}} \to +\infty \quad \text{as} \quad n \to \infty.$$

Using (3.7) and (3.14) we have

$$
\lim_{n\to\infty} \sup_{I \ni t_n; I \subset [0,1]} \frac{1}{|I|} \int_I \int_0^\infty |\partial_t \widetilde{u}_n(t,r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t = 0,
$$

$$
\|\partial_t \widetilde{u}_n(t_n)\|_{L^2} \to 0 \quad \text{as} \quad n \to \infty,
$$

(3.17)

and we can apply (after passing to a subsequence) Lemma 2.17 to deduce the existence of a matching profile $(\boldsymbol{v}_{\mathrm{L}}^{i_1}, \sigma_{n,i_1}, t_{n,i_1})$ such that for all $s \in \mathbb{R}$,

$$
v_{\mathrm{L}}^{i_1}(s) = v_{\mathrm{L}}^{i_0}(-s), \quad \sigma_{n,i_1} = \sigma_{n,i_0}, \quad \text{and} \quad t_{n,i_1} = -t_{n,i_0} \; \forall n.
$$

After relabeling we may assume that $i_1 = i_0 + 1$.

We claim that there exists $\tau_0 > 0$ so that, in addition to (3.17), we also have,

$$
\lim_{n\to\infty} \|\partial_t \widetilde{u}_n(t_n + \tau_0 \sigma_{n,i_0})\|_{L^2} = 0.
$$

(3.18)

To see this, assume for simplicity that $t_n \to t_0 < 1$ (the other possible scenarios are similar). Passing to a subsequence, we may assume that

$$
2^{-2n-4} \geq \sup_{I \ni t_n; I \subset [0,1]} \frac{1}{|I|} \int_I \int_0^\infty |\partial_t \widetilde{u}_n(t,r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t,
$$

and define sets E_n (for all large n) via,

$$
E_n := \{\tau \in [0,1] \; : \; \|\partial_t \widetilde{u}_n(t_n + \tau \sigma_{n,i_0})\|_{L^2}^2 \geq 2^{-n-2}\}.
$$

Thus,

$$
2^{-2n-4} \geq \frac{1}{\sigma_{n,i_0}} \int_{t_n}^{t_n + \sigma_{n,i_0}} \|\partial_t \widetilde{u}_n(t)\|_{L^2}^2 \, \mathrm{d}t = \int_0^1 \|\partial_t \widetilde{u}_n(t_n + \tau \sigma_{n,i_0})\|_{L^2}^2 \, \mathrm{d}\tau \geq |E_n| \, 2^{-n-2}
$$

which means that $|E_n| \leq 2^{-n-2}$ for all n large enough. Hence $|\cup_{n\geq 0} E_n| \leq \frac{1}{2}$, and thus any $\tau_0 \in [0,1] \setminus \cup_{n\geq 0} E_n$ satisfies (3.18).

Next, we will need to evolve the profiles for time $= \tau_0 \sigma_{n,i_0}$. To get in the setting of Lemma 2.13 we first need to truncate the sequence again, removing all profiles concentrating at a scales $\ll \sigma_{n,i_0}$. To this end, and following [19, Erratum], we denote by $\mathcal{K} = \mathcal{K}_{c,0} \cup \mathcal{K}_{\mathrm{L},0}$, where $\mathcal{K}_{c,0} \subset \mathcal{J}_{c,0}$ is the set of indices j such that,

$$
\exists C_j > 0 \quad \text{such that} \quad \lambda_{n,j} \leq C_j \sigma_{n,i_0},
$$

and letting $\epsilon_0 > 0$ be as in Lemma 2.8, $\mathcal{K}_{\mathrm{L},0} \subset \mathcal{I}_{\mathrm{L},0}$ is the set of indices i such that both

$$
E(\boldsymbol{v}_{\mathrm{NL}}^i) \geq \epsilon_0 \quad \text{and} \quad \exists C_i > 0 \quad \text{such that} \quad \max(\sigma_{n,i}, |t_{n,i}|) \leq C_i \sigma_{n,i_0}.
$$

Observe that $i_0, i_0 + 1 \notin \mathcal{K}_{\mathrm{L},0}$ and that by the pythagorean expansion of the nonlinear energy, $\mathcal{K}$ is a finite set.

Since $\sigma_{n,i_0} \ll |t_{n,i_0}|$ we can, arguing as in (3.12), find a scale σ_n such that $\sigma_{n,i_0} \ll \sigma_n \ll |t_{n,i_0}|$ and such that $E(\widetilde{u}_n; \sigma_n/4, 4\sigma_n) \to 0$ as $n \to \infty$. Using Lemma 2.1, and arguing as above, after passing to a subsequence we can find a sequence $y_n \in [\frac{3}{4}\sigma_n, \frac{5}{4}\sigma_n]$ and an integer ℓ_1 with $|\widetilde{u}_n(t_n, y_n) - \ell_1 \boldsymbol{\pi}| \to 0$. We then define a sequence $\check{\boldsymbol{u}}_n(t_n) \in \mathcal{E}_{\ell_1, m_1}$ by

$$
\check{\boldsymbol{u}}_n(t_n) := \chi_{\sigma_n} \ell_1 \boldsymbol{\pi} + (1 - \chi_{\sigma_n}) \widetilde{\boldsymbol{u}}_n(t_n).
$$

It follows that for any $J \geq \max(i; i \in \mathcal{K}_{\mathrm{L},0}) + 1$ we have

$$
\check{\boldsymbol{u}}_n(t_n) = m_1 \boldsymbol{\pi} + \sum_{j \in \mathcal{J}_{c,0} \setminus \mathcal{K}_{c,0}} \iota_j \big((Q(\tfrac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi} \big) + \sum_{i \leq J, \, i \in \mathcal{I}_{\mathrm{L},0} \setminus \mathcal{K}_{\mathrm{L},0}} \boldsymbol{v}_{\mathrm{L},n}^i(0) + \check{\boldsymbol{w}}_{n,0}^J + o_n(1)
$$

where we define $\tilde{w}_{n,0}^J(r) = (1 - \chi_{\sigma_n})\tilde{w}_{n,0}^J(r)$. We need to justify the $o_n(1)$ term above. First, it is clear the harmonic maps with indices $j \in \mathcal{K}_{c,0}$ satisfy $\|(1 - \chi_{\sigma_n/2})(\iota_j Q(\frac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi})\|_{\mathcal{E}} = o_n(1)$ since $j \in \mathcal{K}_{c,0}$ implies $\lambda_{n,j} \leq C_j \sigma_{n,i_0} \ll \sigma_n$. Next for those indices $i \in \mathcal{K}_{L,0}$ we claim that,

$$\|(1 - \chi_{\sigma_n/2})\boldsymbol{v}_L^i(-t_{n,i}/\sigma_{n,i})\|_{\mathcal{E}} \lesssim \|\boldsymbol{v}_L^i(-t_{n,i}/\sigma_{n,i})\|_{\mathcal{E}(r \geq \frac{1}{2}\sigma_n/\sigma_{n,i})} = o_n(1). \tag{3.19}$$

To prove the last inequality above note that since $i \in \mathcal{K}_{L,0}$ we have

$$\frac{1}{2}\frac{\sigma_n}{\sigma_{n,i}} = \frac{1}{2}\frac{\sigma_n}{\sigma_{n,i_0}}\frac{\sigma_{n,i_0}}{\sigma_{n,i}} \geq \frac{1}{2C_i}\frac{\sigma_n}{\sigma_{n,i_0}}\frac{|t_{n,i}|}{\sigma_{n,i}}$$

and now (3.19) follows from [15, Lemma 9] after noting again that $\sigma_n/\sigma_{n,i_0} \to \infty$ as $n \to \infty$, and using the equivalence between (2.3) and (2.4) outlined in Section 2.2.

Note that

$$\check{\boldsymbol{u}}_n(t_n, r) = \tilde{\boldsymbol{u}}_n(t_n, r) = \boldsymbol{u}_n(t_n, r) \quad \text{if} \quad 4\sigma_n \leq r \leq r_{2,n}$$

and thus, denoting by $\check{\boldsymbol{u}}_n(t)$ the wave map evolution of $\check{\boldsymbol{u}}_n$ we have by finite speed of propagation that for $s > 0$,

$$\check{\boldsymbol{u}}_n(t_n + s, r) = \tilde{\boldsymbol{u}}_n(t_n + s, r) = \boldsymbol{u}_n(t_n + s, r) \quad \text{if} \quad 4\sigma_n + s \leq r \leq r_{2,n} - s. \tag{3.20}$$

The point of these truncations is that we can now apply the nonlinear profile decomposition Lemma 2.13 to $\check{\boldsymbol{u}}_n(0)$ up to time $\tau_0 \sigma_{n,i_0}$, obtaining an error term $\boldsymbol{z}_n^J(t)$ satisfying for all $s \in [0, \tau_0 \sigma_{n,i_0}]$,

$$\check{\boldsymbol{u}}_n(t_n + s) = m_1\boldsymbol{\pi} + \sum_{j \in \mathcal{J}_{c,0} \setminus \mathcal{K}_{c,0}} \iota_j((Q(\frac{\cdot}{\lambda_{n,j}}), 0) - \boldsymbol{\pi}) + \sum_{i \leq J, \, i \in \mathcal{I}_{L,0} \setminus \mathcal{K}_{L,0}} \boldsymbol{v}_{NL,n}^i(s) + \check{\boldsymbol{w}}_n^J(s) + \boldsymbol{z}_n^J(s)$$

$$\lim_{J \to \infty} \limsup_{n \to \infty} \left(\sup_{t \in [0, \tau_0 \sigma_{n,i_0}]} \|\boldsymbol{z}_n^J(t)\|_{\mathcal{E}} + \|\boldsymbol{z}_n^J\|_{\mathcal{S}([0, \tau_0 \sigma_{n,i_0}])} \right) = 0.$$

Next observe that plugging in $s = \tau_0 \sigma_{n,i_0}$ above gives rise to *linear* profile decomposition for $\check{\boldsymbol{u}}_n(t_n + \tau_0 \sigma_{n,i_0})$ in the sense of Lemma 2.10, where the profiles are given by $(\boldsymbol{Q}, \lambda_{n,j})$ and $(\tilde{\boldsymbol{v}}_L^i, \tilde{\sigma}_{n,i}, \tilde{t}_{n,i}) = (\boldsymbol{v}_L^i, \sigma_{n,i}, t_{n,i} - \tau_0 \sigma_{n,i_0})$.

We apply Lemma 2.16 to the sequence,

$$(f_n, g_n) := \left(\check{\boldsymbol{u}}_n(t_n + \tau_0 \sigma_{n,i_0}, \sigma_{n,i_0} \cdot), \sigma_{n,i_0} \partial_t \check{\boldsymbol{u}}_n(t_n + \tau_0 \sigma_{n,i_0}, \sigma_{n,i_0} \cdot) \right)$$

$$- m_1\boldsymbol{\pi} - \sum_{j \in \mathcal{J}_{c,0} \setminus \mathcal{K}_{c,0}} (\iota_j Q(\frac{\sigma_{n,i_0}}{\lambda_{n,j}} \cdot), 0) - \boldsymbol{\pi})$$

with $\alpha_n = 4\frac{\sigma_n}{\sigma_{n,i_0}} + \tau_0$ and $s_n = \frac{t_{n,i_0}}{\sigma_{n,i_0}}$. By (3.20) and (3.18) we have $\|g_n\|_{L^2(r \geq \alpha_n)} \to 0$ as $n \to \infty$. Since $\sigma_n \ll |t_{n,i_0}|$ we also have $\frac{|s_n|}{\alpha_n} \to \infty$ as $n \to \infty$. Hence we may apply Lemma 2.16. On the one hand, by the way the profiles are obtained,

$$\vec{S}_L(s_n)(f_n, g_n) = \vec{S}_L(\frac{t_{n,i_0}}{\sigma_{n,i_0}})(f_n, g_n) = \vec{S}_L(\tau_0)\vec{S}_L(\frac{\tilde{t}_{n,i_0}}{\tilde{\sigma}_{n,i_0}})(f_n, g_n) \rightharpoonup (v_L^{i_0}(\tau_0), \partial_t v_L^{i_0}(\tau_0)) \in \mathcal{E}$$

but on the other hand, since $\tilde{t}_{n,i_0+1} = -t_{n,i_0} - \tau_0 \sigma_{n,i_0}$ and since $\sigma_{n,i_0} = \sigma_{n,i_0+1} = \tilde{\sigma}_{n,i_0+1}$ we have

$$\vec{S}_L(-s_n)(f_n, g_n) = \vec{S}_L(-\frac{t_{n,i_0}}{\sigma_{n,i_0}})(f_n, g_n)$$

$$= \vec{S}_L(\tau_0)\vec{S}_L(-\frac{\tilde{t}_{n,i_0+1}}{\tilde{\sigma}_{n,i_0+1}})(f_n, g_n) \rightharpoonup (v_L^{i_0+1}(\tau_0), \partial_t v_L^{i_0+1}(\tau_0)) = (v_L^{i_0}(-\tau_0), -\partial_t v_L^{i_0}(-\tau_0)) \in \mathcal{E}.$$

An application of Lemma 2.16 then gives $v_{\mathrm{L}}^{i_0}(\tau_0) = v_{\mathrm{L}}^{i_0}(-\tau_0)$ and $\partial_t v_{\mathrm{L}}^{i_0}(\tau_0) = \partial_t v_{\mathrm{L}}^{i_0}(-\tau_0)$, or in other words $v_{\mathrm{L}}^{i_0}(t) = v_{\mathrm{L}}^{i_0}(t + 2\tau_0)$, is periodic with period $2\tau_0$, which is impossible since $\boldsymbol{v}_{\mathrm{L}}^{i_0}(t)$ is a finite energy solution to (2.3), unless $\boldsymbol{v}_{\mathrm{L}}^{i_0} \equiv 0$, which contradicts our assumption. Thus, there are no nonzero profiles with indices in the set $\mathcal{I}_{\mathrm{L},0}$.

Step 8.(Vanishing properties of the error $\boldsymbol{w}_{n,0}^{J}$) We summarize where the argument stands after all of the previous steps. With $\widetilde{\boldsymbol{u}}_n$ defined in (3.16), we may relabel the indices in $\mathcal{J}_{c,0}$, so that $\vec{\lambda}_n = (\lambda_{n,1}, \ldots \lambda_{n,K_0})$ with $0 \leq K_0 \leq K_1$, and with $\lambda_{n,1} \ll \lambda_{n,2} \ll \cdots \ll \lambda_{n,K_0} \lesssim 1$ and signs $\vec{\iota} = (\iota_1, \ldots, \iota_{K_0})$, so that

$$\widetilde{\boldsymbol{u}}_n(t_n) = m_1 \boldsymbol{\pi} + \sum_{j=1}^{K_0} \iota_j((Q\big(\frac{\cdot}{\lambda_{n,j}}\big), 0) - \boldsymbol{\pi}) + \widetilde{\boldsymbol{w}}_{n,0},$$

where we have removed the index J in $\widetilde{\boldsymbol{w}}_{n,0}^{J}$ using the previous step since there are no nonzero outgoing/incoming profiles relevant to the region $r \leq r_{2,n}$. To conclude the proof, it will suffice to show the existence of a sequence $r_n \to \infty$, with $r_n \leq r_{2,n}$ so that after passing to a subsequence,

$$\|\widetilde{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(r \leq r_n)} \to 0 \quad \text{as} \quad n \to \infty.$$

Using the pythagorean expansion of the energy with (3.17), and the seventh bullet point in Lemma 2.10, we have

$$\|\dot{\widetilde{w}}_{n,0}\|_{L^2} + \|\widetilde{w}_{n,0}\|_{L^\infty} \to 0 \quad \text{as} \quad n \to \infty. \tag{3.21}$$

After passing to a subsequence of the $\boldsymbol{u}_n$, we claim there is a sequence $r_n \to \infty$ with the properties,

$$1 \ll r_n \leq r_{2,n}, \quad \|\widetilde{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(\frac{1}{8}r_n^{-1} \leq r \leq 8r_n)} \to 0 \quad \text{as} \quad n \to \infty. \tag{3.22}$$

The existence of such a sequence follows from the following property about $\widetilde{\boldsymbol{w}}_{n,0}$: for any sequence $\lambda_n \lesssim 1$ and any $A > 1$ we have,

$$\|\widetilde{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(\lambda_n A^{-1} \leq r \leq \lambda_n A)} \to 0 \quad \text{as} \quad n \to \infty. \tag{3.23}$$

The property (3.23) was proved in [10, Step 2., p.1973-1975, Proof of Theorem 3.5] and [42, Proof of (5.29) in Theorem 5.1] and we refer the reader to those works for details of the argument. The intuition is that at any scale $\lambda_n \lesssim 1$ at which $\widetilde{\boldsymbol{u}}_n$ carries energy we have already extracted a profile $Q_{\lambda_{n,j}}$ with $\lambda_{n,j} \simeq \lambda_n$. To prove (3.22) we consider the case $\lambda_n = 1$ in (3.23), and passing to a subsequence of the $\widetilde{\boldsymbol{u}}_n$, we obtain a sequence as in (3.22).

We truncate to the region $r \leq r_n$, following the same procedure used to define $\widetilde{\boldsymbol{u}}_n$ in (3.13), using now r_n in place of $r_{2,n}$. Indeed, set

$$\breve{\boldsymbol{u}}_n(t_n, r) := \chi_{2r_n}(r)\widetilde{\boldsymbol{u}}_n(t, r) + (1 - \chi_{2r_n}(r))m_1\boldsymbol{\pi}.$$

Defining $\breve{\boldsymbol{w}}_{n,0} := \chi_{2r_n}(r)\widetilde{\boldsymbol{w}}_{n,0} + (\chi_{2r_n}(r) - 1)\sum_{j=1}^{K_0} \iota_j((Q\big(\frac{\cdot}{\lambda_{n,j}}\big), 0) - \boldsymbol{\pi})$ and using that $\lambda_{n,K_0} \lesssim 1$ along with (3.21) and (3.22) we see that,

$$\breve{\boldsymbol{u}}_n(t_n) = m_1 \boldsymbol{\pi} + \sum_{j=1}^{K_0} \iota_j((Q\big(\frac{\cdot}{\lambda_{n,j}}\big), 0) - \boldsymbol{\pi}) + \breve{\boldsymbol{w}}_{n,0}, \quad \text{and}$$

$$\lim_{n \to \infty} \left(\|\breve{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(r_n^{-1} \leq r < \infty)} + \|\dot{\breve{w}}_{n,0}\|_{L^2} + \|\breve{w}_{n,0}\|_{L^\infty} \right) = 0. \tag{3.24}$$

Moreover, by (3.23) we see that for any sequence $\lambda_n \lesssim 1$ and any $A > 1$ that,

$$\lim_{n \to \infty} \|\breve{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(\lambda_n A^{-1} \leq r \leq \lambda_n A)} = 0. \tag{3.25}$$

Letting $\breve{\boldsymbol{u}}_n(t)$ denote the wave map evolution of $\breve{\boldsymbol{u}}_n(t_n)$ we see from (3.14) that

$$\breve{\boldsymbol{u}}_n(t, r) = \boldsymbol{u}_n(t, r) \quad \text{if} \quad r \leq r_n \tag{3.26}$$

and from (3.17) that

$$\lim_{n \to \infty} \sup_{I \ni t_n; I \subset [0,1]} \frac{1}{|I|} \int_I \int_0^\infty |\partial_t \breve{u}_n(t, r)|^2 \, r \, \mathrm{d}r \, \mathrm{d}t = 0, \tag{3.27}$$

$$\|\partial_t \breve{u}_n(t_n)\|_{L^2} \to 0 \quad \text{as} \quad n \to \infty.$$

It remains to prove the vanishing $\lim_{n \to \infty} \|\breve{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(r \leq r_n^{-1})} = 0$ and it is at this final step where we employ the Jia-Kenig virial functional. We see from (3.8) and (3.26) that $\breve{\boldsymbol{u}}_n$ satisfies

$$\limsup_{n \to \infty} \int_0^\infty \left(k^2 \frac{\sin^2(2\breve{u}_n(t_n))}{2r^2} - (\partial_r^2 \breve{u}_n(t_n) + \frac{1}{r} \partial_r \breve{u}_n(t_n)) \sin(2\breve{u}_n(t_n)) \right) \chi \, r \mathrm{d}r \leq 0. \tag{3.28}$$

In addition to the above, after possibly passing to a subsequence of the $\breve{\boldsymbol{u}}_n$, we claim that

$$\lim_{n \to \infty} \int_0^\infty \left(k^2 \frac{\sin^2(2\breve{u}_n(t_n))}{2r^2} - (\partial_r^2 \breve{u}_n(t_n) + \frac{1}{r} \partial_r \breve{u}_n(t_n)) \sin(2\breve{u}_n(t_n)) \right) (1 - \chi) \, r \mathrm{d}r = 0. \tag{3.29}$$

The cases $\lim_{n \to \infty} \lambda_{n,K_0} = 0$ and $\lim_{n \to \infty} \lambda_{n,K_0} > 0$ require slightly different arguments and we consider these separately. First, if $\lim_{n \to \infty} \lambda_{n,K_0} = 0$, then from (3.24) we have,

$$\lim_{n \to \infty} E(\breve{\boldsymbol{u}}_n(t_n); 1/2, \infty) = 0. \tag{3.30}$$

Using the above the first term in the integrand satisfies,

$$\int_0^\infty k^2 \frac{\sin^2(2\breve{u}_n(t_n))}{2r^2} (1 - \chi) \, r \mathrm{d}r \lesssim E(\breve{\boldsymbol{u}}_n(t_n); 1/2, \infty) \to 0 \quad \text{as} \quad n \to \infty.$$

Integration by parts in the second term gives,

$$-\int_0^\infty (\partial_r^2 \breve{u}_n + \frac{1}{r} \partial_r \breve{u}_n) \sin(2\breve{u}_n)(1 - \chi) \, r \mathrm{d}r = \int_0^\infty (\partial_r \breve{u}_n)^2 2 \cos(2\breve{u}_n)(1 - \chi) \, r \, \mathrm{d}r$$

$$- \int_0^\infty \partial_r \breve{u}_n \sin(2\breve{u}_n) \chi' \, r \, \mathrm{d}r.$$

The terms on the right above satisfy,

$$\left| \int_0^\infty (\partial_r \breve{u}_n)^2 2 \cos(2\breve{u}_n)(1 - \chi) \, r \, \mathrm{d}r \right| + \left| \int_0^\infty \partial_r \breve{u}_n \sin(2\breve{u}_n) \chi' \, r \, \mathrm{d}r \right| \lesssim E(\breve{\boldsymbol{u}}_n(t_n); 1/2, \infty)$$

and thus vanish as $n \to \infty$ by (3.30), proving (3.29) in the case $\lim_{n \to \infty} \lambda_{n,K_0} = 0$.

Next suppose $\lim_{n \to \infty} \lambda_{n,K_0} > 0$. Note that in this case,

$$\lim_{n \to \infty} E(\breve{\boldsymbol{u}}_n(t_n) - \iota_{K_0}(\boldsymbol{Q}_{\lambda_{n,K_0}} - \boldsymbol{\pi}); 1/2, \infty) = 0, \tag{3.31}$$

which follows from (3.24). To ease notation let $\boldsymbol{\psi}_n := \breve{\boldsymbol{u}}_n(t_n) - \iota_{K_0}(\boldsymbol{Q}_{\lambda_{n,K_0}} - \boldsymbol{\pi})$ and $\boldsymbol{\phi}_n := \iota_{K_0}(\boldsymbol{Q}_{\lambda_{n,K_0}} - \boldsymbol{\pi})$. Expanding $\sin(2\breve{u}_n(t_n)) = \sin 2\psi_n \cos 2\phi_n + \sin 2\phi_n(\cos 2\psi_n - 1) + \sin 2\phi_n$ and using the pointwise vanishing of $\partial_r^2 \phi_n + \frac{1}{r} \partial_r \phi_n - \frac{k^2}{2r^2} \sin 2\phi_n = 0$, we rewrite the term in parenthesis in (3.29) as,

$$k^2 \frac{\sin^2(2\breve{u}_n(t_n))}{2r^2} - (\partial_r^2 \breve{u}_n(t_n) + \frac{1}{r} \partial_r \breve{u}_n(t_n)) \sin(2\breve{u}_n(t_n))$$

$$= \sin(2\breve{u}_n(t_n)) \left(\frac{k^2}{2r^2} \sin 2\psi_n \cos 2\phi_n - \frac{k^2}{r^2} \sin^2 \psi_n \sin 2\phi_n - (\partial_r^2 \psi_n + \frac{1}{r} \partial_r \psi_n) \right).$$

The contributions of the first two terms above satisfy

$$\left| \int_0^\infty \sin(2\breve{u}_n(t_n)) \left(\frac{k^2}{2r^2} \sin 2\psi_n \cos 2\phi_n - \frac{k^2}{r^2} \sin^2 \psi_n \sin 2\phi_n \right) (1-\chi)\, r\mathrm{d}r \right| \lesssim \sqrt{E(\boldsymbol{\psi}_n; 1/2, \infty)}$$

and thus vanish as $n \to \infty$ by (3.31). Integrating by parts, the third term satisfies,

$$-\int_0^\infty \sin 2\breve{u}_n(t_n)(\partial_r^2 \psi_n + \frac{1}{r}\partial_r \psi_n)(1-\chi)\, r\mathrm{d}r = -\int_0^\infty \partial_r \psi_n \sin(2\breve{u}_n(t_n))\chi'\, r\mathrm{d}r$$

$$+ 2\int_0^\infty \partial_r \psi_n \partial_r \breve{u}_n(t_n) \cos(2\breve{u}_n(t_n))(1-\chi)\, r\mathrm{d}r.$$

By Cauchy Schwarz,

$$\left| \int_0^\infty \partial_r \psi_n \sin(2\breve{u}_n(t_n))\chi'\, r\mathrm{d}r \right| + \left| \int_0^\infty \partial_r \psi_n \partial_r \breve{u}_n(t_n) \cos(2\breve{u}_n(t_n))(1-\chi)\, r\mathrm{d}r \right|$$

$$\lesssim \sqrt{E(\boldsymbol{\psi}_n; 1/2, \infty)} \to 0 \quad \text{as} \quad n \to \infty,$$

by (3.31), proving (3.29) in the case $\lim_{n\to\infty} \lambda_{n,K_0} > 0$. Combining (3.28) and (3.29) gives

$$\limsup_{n\to\infty} \int_0^\infty \left(k^2 \frac{\sin^2(2\breve{u}_n(t_n))}{2r^2} - (\partial_r^2 \breve{u}_n(t_n) + \frac{1}{r}\partial_r \breve{u}_n(t_n)) \sin(2\breve{u}_n(t_n)) \right) r\mathrm{d}r \leq 0.$$

The above together with (3.24), (3.25), and (3.27) yield all of the hypothesis of Lemma 3.6 for the sequence $\breve{\boldsymbol{u}}_n(t_n)$. An application of Lemma 3.6 gives $\lim_{n\to\infty} \|\breve{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}} = 0$. It follows that that $\boldsymbol{\delta}_\infty(\breve{\boldsymbol{u}}_n(t_n)) \to 0$ as $n \to \infty$. By (3.26) we in fact have proved that $\boldsymbol{\delta}_{r_n}(\boldsymbol{u}_n(t_n)) \to 0$ as $n \to \infty$, (note that $\lambda_{n,K_0} \lesssim 1$ and $r_n \to \infty$ ensures that the final ratio $\lambda_{n,K_0}/\lambda_{n,K_0+1} \to 0$, where $\lambda_{n,K_0+1} := r_n$), completing the proof. $\qquad\square$

4. Decomposition of the solution and collision intervals

In the final two sections we prove Theorem 1 for equivariance classes $k \geq 2$. We reserve the case $k = 1$ for the appendix.

4.1. **Proximity to a multi-bubble and collisions.** For the remainder of the paper we fix a solution $\boldsymbol{u}(t) \in \mathcal{E}_{\ell,m}$ of (1.1), defined on the time interval $I_* = (0, T_0]$ in the blow-up case and on $I_* = [T_0, \infty)$ in the global case, for some $T_0 > 0$. We set $T_* := \infty$ in the global case and $T_* := 0$ in the blow-up case. Let $\boldsymbol{u}^*(t)$ be the radiation as defined in Theorem 1.10. More precisely, we let $m_\Delta := \lim_{t\to T_+} u(t, \frac{1}{2}t) \in \mathbb{Z}$ and shift the radiation so that $\boldsymbol{u}^*(t) \in \mathcal{E}_{0,m_\infty}$ for some $m_\infty \in \mathbb{Z}$, and for $r \gtrsim t$, $\boldsymbol{u}(t, r) \sim m_\Delta \boldsymbol{\pi} + \boldsymbol{u}^*(t, r)$. Note that $m_\infty = 0$ if $T_* = \infty$.

It is a crucial insight of [9,11,12] that $\boldsymbol{u}^*(t)$ is given for continuous time. Recall that Theorem 1.10 gives a function $\rho : I_* \to (0, \infty)$ such that

$$\lim_{t\to T_*} \left((\rho(t)/t)^k + \|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - m_\Delta \boldsymbol{\pi}\|_{\mathcal{E}(\rho(t),\infty)}^2 \right) = 0, \tag{4.1}$$

and that for any $\alpha \in (0, 1)$ we have

$$\lim_{t\to T_*} E(\boldsymbol{u}^*(t); 0, \alpha t) = 0. \tag{4.2}$$

By Theorem 1.15 there exists a time sequence $t_n \to T_*$ and an integer $N \geq 0$, which we now fix, such that $\boldsymbol{u}(t_n) - \boldsymbol{u}^*(t_n)$ approaches an N-bubble as $n \to \infty$. Roughly, our goal is to show that on the region $r \in (0, \rho(t))$, the solution $\boldsymbol{u}(t)$ approaches a continuously modulated N-bubble, noting that the radiation $\boldsymbol{u}^*(t)$ is negligible in this region. By convention, we will set $\lambda_{N+1}(t) := t$ to be the "scale" of the radiation and $\lambda_0(t) := 0$. Our argument requires the following localized version of the distance function to a multi-bubble.

Definition 4.1 (Proximity to a multi-bubble). For all $t \in I$, $\rho \in (0, \infty)$, and $K \in \{0, 1, \ldots, N\}$, we define the *localized multi-bubble proximity function* as

$$\mathbf{d}_K(t; \rho) := \inf_{\vec{\iota}, \vec{\lambda}} \left(\|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\iota}, \vec{\lambda})\|_{\mathcal{E}(\rho, \infty)}^2 + \sum_{j=K}^{N} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^k \right)^{\frac{1}{2}},$$

where $\vec{\iota} := (\iota_{K+1}, \ldots, \iota_N) \in \{-1, 1\}^{N-K}$, $\vec{\lambda} := (\lambda_{K+1}, \ldots, \lambda_N) \in (0, \infty)^{N-K}$, $\lambda_K := \rho$ and $\lambda_{N+1} := t$.

The *multi-bubble proximity function* is defined by $\mathbf{d}(t) := \mathbf{d}_0(t; 0)$.

Remark 4.2. We emphasize that if $\mathbf{d}_K(t; \rho)$ is small, this means that $\boldsymbol{u}(t) - \boldsymbol{u}^*(t)$ is close to $N - K$ bubbles in the exterior region $r \in (\rho, \infty)$.

We can now rephrase Theorem 1.15 in this notation: there exists a monotone sequence $t_n \to T_*$ such that

$$\lim_{n \to \infty} \mathbf{d}(t_n) = 0. \tag{4.3}$$

Even though this fact is certainly a starting point of our analysis, it will turn out that we cannot use it as a black box. Rather, we need to examine the proof and use more precise information provided by the analysis in [10, 42]; see Section 3.

We state and prove some simple consequences of the set-up above. We always assume $N \geq 1$, since the pure radiation case $N = 0$ (in fact, also the case $N = 1$) is already settled by Côte's and Jia's and Kenig's work [10, 42].

First, a direct consequence of (4.1) is that $\boldsymbol{u}(t) - \boldsymbol{u}^*(t)$ always approaches a 0-bubble in some exterior region. With $\rho_N(t) = \rho(t)$ given by the function in Theorem 1.10 the following lemma is immediate from the conventions of Definition 4.1

Lemma 4.3. *There exists a function* $\rho_N : I \to (0, \infty)$ *such that*

$$\lim_{t \to T_*} \mathbf{d}_N(t; \rho_N(t)) = 0. \tag{4.4}$$

Theorem 1 will be a quick consequence of showing that, in fact,

$$\lim_{t \to T_*} \mathbf{d}(t) = 0. \tag{4.5}$$

The approach which we adopt in order to prove (4.5) is to study colliding bubbles. A collision is defined as follows.

Definition 4.4 (Collision interval). Let $K \in \{0, 1, \ldots, N\}$. A compact time interval $[a, b] \subset I_*$ is a *collision interval* with parameters $0 < \epsilon < \eta$ and $N - K$ exterior bubbles if

- $\mathbf{d}(a) \leq \epsilon$ and $\mathbf{d}(b) \leq \epsilon$,
- there exists $c \in (a, b)$ such that $\mathbf{d}(c) \geq \eta$,
- there exists a function $\rho_K : [a, b] \to (0, \infty)$ such that $\mathbf{d}_K(t; \rho_K(t)) \leq \epsilon$ for all $t \in [a, b]$.

In this case, we write $[a, b] \in \mathcal{C}_K(\epsilon, \eta)$.

Definition 4.5 (Choice of K). We define K as the *smallest* nonnegative integer having the following property. There exist $\eta > 0$, a decreasing sequence $\epsilon_n \to 0$ and sequences $(a_n), (b_n)$ such that $[a_n, b_n] \in \mathcal{C}_K(\epsilon_n, \eta)$ for all $n \in \{1, 2, \ldots\}$.

Lemma 4.6 (Existence of $K \geq 1$). *If* (4.5) *is false, then K is well defined and $K \in \{1, \ldots, N\}$.*

Remark 4.7. The fact that $K \geq 1$ means that at least one bubble must lose its shape if (4.5) is false.

Proof of Lemma 4.6. Assume (4.5) does not hold, so that there exist $\eta > 0$ and a monotone sequence $s_n \to T_*$ such that

$$\mathbf{d}(s_n) \geq \eta, \qquad \text{for all } n.$$

We claim that there exist sequences $(\epsilon_n), (a_n), (b_n)$ such that $[a_n, b_n] \in \mathcal{C}_N(\epsilon_n, \eta)$. Indeed, (4.3) implies that there exist $\epsilon_n \to 0$, $a_n \leq s_n$ and $b_n \geq s_n$ such that $\mathbf{d}(a_n) \leq \epsilon_n$ and $\mathbf{d}(b_n) \leq \epsilon_n$. Note that $a_n \to T_*$ and $b_n \to T_*$. Let $\rho_N : [a_n, b_n] \to (0, \infty)$ be the function given by Lemma 4.3, restricted to the time interval $[a_n, b_n]$. Then (4.4) yields

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} \mathbf{d}_N(t; \rho_N(t)) = 0.$$

Upon adjusting the sequence ϵ_n, we obtain that all the requirements of Definition 4.4 are satisfied for $K = N$.

We now prove that $K \geq 1$. Suppose $K = 0$. By Definition 4.4 of a collision interval, there exist $\eta > 0$, and sequences $c_n \in [a_n, b_n]$ and $\rho_0(c_n) \geq 0$ such that $\mathbf{d}_0(c_n; \rho_0(c_n)) \leq \epsilon_n$ and at the same time $\mathbf{d}(c_n) \geq \eta$. We show that this is impossible.

Define $\boldsymbol{v}_n := \boldsymbol{u}(c_n) - \boldsymbol{u}^*(c_n)$. Since $\mathbf{d}_0(c_n; \rho_0(c_n)) \leq \epsilon_n$ we can find parameters, $\rho_0(c_n) \ll \lambda_{n,1} \ll \cdots \ll \lambda_{n,N}$ and signs $\vec{\iota}_n$ such that defining $\boldsymbol{g}_n = \boldsymbol{v}_n - \boldsymbol{Q}(m_\Delta, \vec{\iota}_n, \vec{\lambda}_n)$ we have

$$\mathbf{d}_0(c_n; \rho_0(c_n)) \simeq \|\boldsymbol{g}_n\|_{\mathcal{E}(\rho_0(c_n), \infty)}^2 + \sum_{j=0}^{N} \left(\frac{\lambda_{n,j}}{\lambda_{n,j+1}} \right)^k \lesssim \epsilon_n^2. \tag{4.6}$$

With $\rho(t)$ as in (4.1) we see that we must have $\lambda_{n,N} \ll \rho(c_n) \ll c_n$, and thus using (4.1) along with (4.6) and Lemma 2.22 we have

$$\begin{aligned}
E(\boldsymbol{u}(c_n); \rho_0(c_n), \infty) &= E(\boldsymbol{g}_n + \boldsymbol{u}^*(c_n) + \boldsymbol{Q}(m_\Delta, \vec{\iota}_n, \vec{\lambda}_n); \rho_0(c_n), \rho(c_n)) \\
&\quad + E(\boldsymbol{g}_n + \boldsymbol{u}^*(c_n) + \boldsymbol{Q}(m_\Delta, \vec{\iota}_n, \vec{\lambda}_n); \rho(c_n), \infty) \\
&= NE(\boldsymbol{Q}) + E(\boldsymbol{u}^*) + o_n(1)
\end{aligned}$$

Since by (1.7) we know that $E(\boldsymbol{u}(c_n)) = NE(\boldsymbol{Q}) + E(\boldsymbol{u}^*(c_n))$, we conclude from the previous line that,

$$E(\boldsymbol{u}(c_n); 0, \rho_0(c_n)) = o_n(1) \quad \text{as} \quad n \to \infty.$$

Using (4.2) and the fact that $\rho_0(c_n) \ll \rho(c_n)$ it follows that $E(\boldsymbol{v}_n; 0, \rho_0(c_n)) = o_n(1)$, and hence by (2.1) we conclude that

$$\|\boldsymbol{v}_n - \ell\boldsymbol{\pi}\|_{\mathcal{E}(0, \rho_0(c_n))} \lesssim E(\boldsymbol{v}_n; 0, \rho_0(c_n)) = o_n(1) \quad \text{as} \quad n \to \infty$$

Thus, combining the above with (4.6) we have $\mathbf{d}(c_n) = o_n(1)$ as $n \to \infty$, a contradiction. $\qquad \square$

In the remaining part of the paper, we argue by contradiction, fixing K to be the number provided by Lemma 4.6. We also let η, ϵ_n, a_n and b_n be some choice of objects satisfying the requirements of Definition 4.5. We fix choices of signs and scales for the $N - K$ "exterior" bubbles provided by Definition 4.1 in the following lemma.

Remark 4.8. For each collision interval there exists a time $c_n \in [a_n, b_n]$ with $\mathbf{d}(c_n) \geq \eta$ and we may assume without loss of generality that $\mathbf{d}(a_n) = \mathbf{d}(b_n) = \epsilon_n$ and $\mathbf{d}(t) \geq \epsilon_n$ for each $t \in [a_n, b_n]$. Indeed, given some initial choice of $[a_n, b_n] \in \mathcal{C}_K(\eta, \epsilon_n)$, we can find $a_n \leq \widetilde{a}_n < c_n$ and $c_n < \widetilde{b}_n \leq b_n$ so that $\mathbf{d}(a_n) = \mathbf{d}(b_n) = \epsilon_n$ and $\mathbf{d}(t) \geq \epsilon_n$ for each $t \in [\widetilde{a}_n, \widetilde{b}_n]$. Just set $a_n \leq \widetilde{a}_n := \inf\{t \leq c_n \mid \mathbf{d}(t) \geq \epsilon_n\}$ and similarly for $\widetilde{b}_n$.

Similarly, given some initial choice $\epsilon_n \to 0, \eta > 0$ and intervals $[a_n, b_n] \in \mathcal{C}_K(\eta, \epsilon_n)$ we are free to "enlarge" ϵ_n by choosing some other sequence $\epsilon_n \leq \widetilde{\epsilon}_n \to 0$, and new collision subintervals

$[\widetilde{a}_n, \widetilde{b}_n] \subset [a_n, b_n] \cap \mathcal{C}_K(\eta, \widetilde{\epsilon}_n)$ as in the previous paragraph. We will enlarge our initial choice of ϵ_n in this fashion several times over the course of the proof.

Lemma 4.9. *Let $K \geq 1$ be the number given by Lemma 4.6, and let η, ϵ_n, a_n and b_n be some choice of objects satisfying the requirements of Definition 4.5. Then there exist a sequence $\vec{\sigma}_n \in \{-1, 1\}^{N-K}$, a function $\vec{\mu} = (\mu_{K+1}, \ldots, \mu_N) \in C^1(\cup_{n \in \mathbb{N}} [a_n, b_n]; (0, \infty)^{N-K})$, a sequence $\nu_n \to 0$, and a sequence $m_n \in \mathbb{Z}$, so that defining the function,*

$$\nu : \cup_{n \in \mathbb{N}} [a_n, b_n] \to (0, \infty), \quad \nu(t) := \nu_n \mu_{K+1}(t), \quad for \quad t \in [a_n, b_n], \tag{4.7}$$

we have,

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} \Big(\mathbf{d}_K(t; \nu(t)) + E(\boldsymbol{u}(t), \nu(t), 2\nu(t)) \Big) = 0, \tag{4.8}$$

and defining $\boldsymbol{w}(t), \boldsymbol{h}(t)$ for $t \in \cup_n [a_n, b_n]$ by

$$\boldsymbol{w}(t) = (1 - \chi_{\nu(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) + \chi_{\nu(t)} m_n \boldsymbol{\pi} = m_\Delta \boldsymbol{\pi} + \sum_{j=K+1}^{N} \sigma_{n,j}(\boldsymbol{Q}_{\mu_j(t)} - \boldsymbol{\pi}) + \boldsymbol{h}(t), \tag{4.9}$$

we have, $\boldsymbol{w}(t) \in \mathcal{E}_{m_n, m_\Delta}$, $\boldsymbol{h}(t) \in \mathcal{E}$, and

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} \left(\|\boldsymbol{h}(t)\|_{\mathcal{E}}^2 + \left(\frac{\nu(t)}{\mu_{K+1}(t)} \right)^k + \sum_{j=K+1}^{N} \left(\frac{\mu_j(t)}{\mu_{j+1}(t)} \right)^k \right) = 0, \tag{4.10}$$

with the convention that $\mu_{N+1}(t) = t$. Finally, $\nu(t)$ satisfies the estimate,

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} |\nu'(t)| = 0. \tag{4.11}$$

Remark 4.10. One should think of $\nu(t)$ as the scale that separates the $N - K$ "exterior" bubbles, which are defined continuously on the union of the collision intervals $[a_n, b_n]$ from the K "interior" bubbles that are coherent at the endpoints of $[a_n, b_n]$, but come into collision somewhere inside the interval and lose their shape. In the case $K = N$, there are no exterior bubbles, $\mu_{K+1}(t) = t$ and $\nu_n \to 0$ is chosen using (4.1).

Proof. By Definition 4.1 for each n we can find scales $\rho_K(t) \ll \widetilde{\mu}_{K+1}(t) \ll \cdots \ll \widetilde{\mu}_N(t) \ll t$ and signs $\vec{\sigma}(t) \in \{-1, 1\}^{N-k}$ for $t \in [a_n, b_n]$, such that defining $\boldsymbol{h}_{\rho_K}(t)$ for $r \in (\rho_K(t), \infty)$ by

$$\boldsymbol{u}(t) - \boldsymbol{u}^*(t) = \boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\sigma}(t), \vec{\widetilde{\mu}}(t)) + \boldsymbol{h}_{\rho_K}(t)$$

we have,

$$\mathbf{d}(t; \rho_K(t)) \simeq \|\boldsymbol{h}_{\rho_K}(t)\|_{\mathcal{E}(\rho_K(t), \infty)}^2 + \sum_{j=K}^{N} \left(\frac{\widetilde{\mu}_j(t)}{\widetilde{\mu}_{j+1}(t)} \right)^k \lesssim \epsilon_n^2, \tag{4.12}$$

keeping the convention $\widetilde{\mu}_K(t) := \rho_K(t), \widetilde{\mu}_{N+1}(t) := t$. Using $\lim_{n \to \infty} \sup_{t \in [a,b]} \mathbf{d}_K(t; \rho_K(t)) = 0$ and the fact that

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} E(\boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\sigma}(t), \vec{\widetilde{\mu}}(t)); \alpha_n \widetilde{\mu}_{K+1}(t), \beta_n \widetilde{\mu}_{K+1}(t)) = 0,$$

for any two sequence $\alpha_n \ll \beta_n \ll 1$, we can choose a sequence $\nu_n \to 0$ with

$$\rho_K(t) \leq \nu_n \widetilde{\mu}_{K+1}(t), \quad and \quad \lim_{n \to \infty} \sup_{t \in [a_n, b_n]} E(\boldsymbol{u}(t) - \boldsymbol{u}^*(t); \frac{1}{4} \nu_n \widetilde{\mu}_{K+1}(t), 4\nu_n \widetilde{\mu}_{K+1}(t)) = 0.$$

Defining $\widetilde{\nu}(t) = \nu_n \widetilde{\mu}_{K+1}(t)$, it follows from Lemma 2.1 that we can find integers $m_n \in \mathbb{Z}$, which are independent of $t \in [a_n, b_n]$ due to continuity of the flow so that

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} \sup_{r \in (\frac{1}{4}\nu_n \widetilde{\mu}_{K+1}(t), 4\nu_n \widetilde{\mu}_{K+1}(t))} |u(t,r) - u^*(t,r) - m_n \pi| = 0,$$

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} \|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - m_n \pi\|_{\mathcal{E}((\frac{1}{4}\nu_n \widetilde{\mu}_{K+1}(t), 4\nu_n \widetilde{\mu}_{K+1}(t))} = 0$$

(4.13)

Thus, defining $\widetilde{\boldsymbol{w}}(t) \in \mathcal{E}_{m_n, m_\Delta}$ and, $\widetilde{\boldsymbol{h}}(t) \in \mathcal{E}$ for $t \in \cup_n [a_n, b_n]$, by

$$\widetilde{\boldsymbol{w}}(t) := (1 - \chi_{\widetilde{\nu}(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) + \chi_{\widetilde{\nu}(t)} m_n \boldsymbol{\pi}$$

$$= (1 - \chi_{\widetilde{\nu}(t)})m_\Delta \boldsymbol{\pi} + \chi_{\widetilde{\nu}(t)} \sum_{j=K+1}^{N} \sigma_j(t) \boldsymbol{\pi} + \chi_{\widetilde{\nu}(t)} m_n \boldsymbol{\pi} + \sum_{j=K+1}^{N} \sigma_j(t)(\boldsymbol{Q}_{\widetilde{\mu}_j(t)} - \boldsymbol{\pi}) + \widetilde{\boldsymbol{h}}(t)$$

(4.14)

$$= m_\Delta \boldsymbol{\pi} + \sum_{j=K+1}^{N} \sigma_j(t)(\boldsymbol{Q}_{\widetilde{\mu}_j(t)} - \boldsymbol{\pi}) + \widetilde{\boldsymbol{h}}(t)$$

we have using (4.12),

$$\sup_{t \in [a_n, b_n]} \left(\|\widetilde{\boldsymbol{h}}(t)\|_{\mathcal{E}}^2 + \sum_{j=K}^{N} \left(\frac{\widetilde{\mu}_j(t)}{\widetilde{\mu}_{j+1}(t)} \right)^k \right) \le \theta_n^2.$$

(4.15)

for some sequence $\theta_n \to 0$. We note that the last equality in (4.14) follows from the observation that we must have,

$$m_\Delta - \sum_{j=K+1}^{N} \sigma_j(t) = m_n$$

(4.16)

for any $t \in [a_n, b_n]$. We invoke Lemma 2.27 and continuity of the flow to conclude that for each n, the sign vector $\vec{\sigma}(t) = \vec{\sigma}_n$ is independent of $t \in [a_n, b_n]$, and the functions $\widetilde{\mu}_{K+1}(t), \ldots, \widetilde{\mu}_N(t)$ can be adjusted to be continuous functions of t. However, in the next sections we require differentiability of the function $\widetilde{\mu}_{K+1}(t)$, so we must modify it slightly.

Given a vector $\vec{\mu}(t) = (\mu_{K+1}(t), \ldots \mu_N(t))$, set,

$$\boldsymbol{w}(t, \vec{\mu}(t)) := (1 - \chi_{\nu_n \mu_{K+1}(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) + \chi_{\nu_n \mu_{K+1}(t)} m_n \boldsymbol{\pi}$$

Fixing t and suppressing it in the notation, and setting up for an argument as in the proof of Lemma 2.25, define

$$F(h, \vec{\mu}) := h - (w(\cdot, \vec{\widetilde{\mu}}) - \mathcal{Q}(m_\Delta, \vec{\sigma}_n, \vec{\widetilde{\mu}})) + w(\cdot, \vec{\mu}) - \mathcal{Q}(m_\Delta, \vec{\sigma}_n, \vec{\mu})$$

and note that $F(0, \vec{\widetilde{\mu}}) = 0$. Moreover,

$$\|F(h, \vec{\mu})\|_H \lesssim \|h\|_H + \sum_{j=K+1}^{N} \left| \frac{\mu_j}{\widetilde{\mu}_j} - 1 \right|$$

Define,

$$G(h, \vec{\mu}) := \left(\frac{1}{\mu_{K+1}} \left\langle \mathcal{Z}_{\underline{\mu}_{K+1}} \mid F(h, \vec{\mu}) \right\rangle, \ldots, \frac{1}{\mu_N} \left\langle \mathcal{Z}_{\underline{\mu}_N} \mid F(h, \vec{\mu}) \right\rangle \right)$$

and thus $G(0, \vec{\widetilde{\mu}}) = (0, \ldots, 0)$. Following the same scheme as the proof of Lemma 2.25 we obtain via Remark 2.26 a mapping $\varsigma : B_H(0; C_0 \theta_n) \to (0, \infty)^{N-K}$ such that for each $h \in B_H(0; C_0 \theta_n)$

we have

$$|\varsigma_j(h)/\widetilde{\mu}_j - 1| \lesssim \theta_n$$

and such that

$$G(h, \vec{\mu}) = 0 \iff \vec{\mu} = \varsigma(h)$$

Using (4.15) we define

$$h := F(\widetilde{h}, \varsigma(\widetilde{h})), \quad \vec{\mu} := \varsigma(\widetilde{h})$$

By construction we then have,

$$\boldsymbol{w}(t, \vec{\mu}(t)) = (1 - \chi_{\nu(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) + \chi_{\nu(t)} m_n \boldsymbol{\pi}$$
$$= \boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\sigma}_n, \vec{\mu}(t)) + \boldsymbol{h}(t)$$

for $\nu(t) := \nu_n \mu_{K+1}(t)$, and for each $j = K+1, \ldots, N$,

$$\sup_{t \in [a_n, b_n]} \left(\|\boldsymbol{h}(t)\|_{\mathcal{E}}^2 + \sum_{j=K}^{N} \left(\frac{\mu_j(t)}{\mu_{j+1}(t)} \right)^k \right) \lesssim \theta_n^2, \quad 0 = \left\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid h(t) \right\rangle \tag{4.17}$$

Note that (4.8) follows from the above and from (4.1). The point is that we can now use orthogonality conditions above to deduce the differentiability of $\mu(t)$. Indeed, noting the identity,

$$\partial_t h(t) = \partial_t w(t, \vec{\mu}(t)) - \partial_t \mathcal{Q}(m_\Delta, \vec{\sigma}_n, \vec{\mu}(t))$$

$$= \nu_n \mu'_{K+1}(t) \Lambda \chi_{\underline{\nu(t)}}(u(t) - u^*(t) - m_n \pi) + \dot{h}(t) + \sum_{j=K+1}^{N} \sigma_{n,j} \mu'_j(t) \Lambda Q_{\underline{\mu_j(t)}},$$

differentiation of the jth orthogonality condition for $h(t)$ gives for each $j = K+1, \ldots, N$

$$\sigma_{n,j} \mu'_j(t) \langle \mathcal{Z} \mid \Lambda Q \rangle + \sum_{i \neq j} \sigma_{n,i} \mu'_i(t) \left\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid \Lambda Q_{\underline{\mu_i(t)}} \right\rangle$$

$$+ \nu_n \mu'_{K+1}(t) \left\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid \Lambda \chi_{\underline{\nu(t)}}(u(t) - u^*(t) - m_n \pi) \right\rangle - \mu'_j(t) \left\langle [r \Lambda \mathcal{Z}]_{\underline{\mu_j(t)}} \mid r^{-1} h \right\rangle \tag{4.18}$$

$$= -\left\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid \dot{h}(t) \right\rangle,$$

which is a diagonally dominant first order differential system for $\vec{\mu}(t)$. To see this, write the above system as

$$\sum_{\ell=K+1}^{N} A_{j,\ell}(t) \mu'_\ell(t) = -\left\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid \dot{h}(t) \right\rangle$$

where,

$$A_{K+1,K+1} = \sigma_{n,K+1} \langle \mathcal{Z} \mid \Lambda Q \rangle + \nu_n \left\langle \mathcal{Z}_{\mu_{K+1}} \mid \Lambda \chi_{\underline{\nu(t)}}(u - u^* - m_n \pi) \right\rangle - \left\langle [r \Lambda \mathcal{Z}]_{\mu_{K+1}} \mid r^{-1} h \right\rangle$$

$$A_{j,j} = \sigma_{n,j} \langle \mathcal{Z} \mid \Lambda Q \rangle - \left\langle [r \Lambda \mathcal{Z}]_{\mu_j} \mid r^{-1} h \right\rangle \text{ if } j = K+2, \ldots, N,$$

$$A_{j,K+1} = \sigma_{n,K+1} \left\langle \mathcal{Z}_{\mu_j} \mid \Lambda Q_{\mu_i} \right\rangle + \nu_n \left\langle \mathcal{Z}_{\mu_{K+1}} \mid \Lambda \chi_{\underline{\nu(t)}}(u - u^* - m_n \pi) \right\rangle, \text{ if } j \neq K+1$$

$$A_{j,\ell} = \sigma_{n,\ell} \left\langle \mathcal{Z}_{\mu_j} \mid \Lambda Q_{\mu_\ell} \right\rangle j \neq \ell, \quad \ell \neq K+1$$

Using that $\langle \mathcal{Z} \mid \Lambda Q \rangle > 0$ along with (4.13) and (4.17) we see that the diagonal terms are all $O(1)$, whereas the off-diagonal terms are all small by (4.13), Corollary 2.21, and (4.17).

Next, fix any $t_0 \in \cup_n [a_n, b_n]$ so that (4.17) holds at the initial data $\vec{\mu}(t_0)$. The existence and uniqueness theorem gives a unique solution $\vec{\mu}_{\text{ode}} \in C^1(J)$ for $J \ni t_0$ a sufficiently small

neighborhood. As the scales were uniquely defined using the implicit function theorem at each fixed t and the solution of the ODE preserves the orthogonality conditions, we must have $\vec{\mu}(t) = \vec{\mu}_{\mathrm{ode}}(t)$ must agree. Hence $\vec{\mu}(t) \in C^1$. Finally, inverting (4.18) we obtain the estimates,

$$\left|\mu_j'(t)\right| \lesssim \|\dot{h}\|_{L^2} \lesssim \theta_n$$

Using the above with $j = K + 1$ yields (4.11). This completes the proof. $\qquad\square$

4.2. **Basic modulation.** On some subintervals of the collision interval $[a_n, b_n]$, mutual interactions between the bubbles dominate the evolution of the solution. We justify the *modulation inequalities* allowing to obtain explicit information on the solution on such time intervals. We stress that in our current approach the modulation concerns only the bubbles from 1 to K.

Lemma 4.11 (Basic modulation, $k \geq 2$). *There exist $C_0, \eta_0 > 0$ and a sequence $\zeta_n \to 0$ such that the following is true.*

Let $n \in \mathbb{N}$ and $J \subset [a_n, b_n]$ be an open time interval such that $\mathbf{d}(t) \leq \eta_0$ for all $t \in J$. Then, there exist $\vec{\iota} \in \{-1, 1\}^K$ (independent of $t \in J$), modulation parameters $\vec{\lambda} \in C^1(J; (0, \infty)^K)$, and $\mathbf{g}(t) \in \mathcal{E}$ satisfying, for all $t \in J$,

$$\chi_{\nu(t)}\mathbf{u}(t) + (1 - \chi_{\nu(t)})m_n\boldsymbol{\pi} = \mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}(t)) + \mathbf{g}(t), \tag{4.19}$$

$$0 = \left\langle \mathcal{Z}_{\underline{\lambda_j(t)}} \mid g(t) \right\rangle, \tag{4.20}$$

where $\nu(t)$ is as in (4.7) and m_n is as in Lemma 4.9. The estimates,

$$C_0^{-1}\mathbf{d}(t) - \zeta_n \leq \|\mathbf{g}(t)\|_{\mathcal{E}} + \sum_{j=1}^{K-1}\left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^{\frac{k}{2}} \leq C_0\mathbf{d}(t) + \zeta_n, \tag{4.21}$$

and

$$\|\mathbf{g}(t)\|_{\mathcal{E}} + \sum_{j \notin \mathcal{A}}\left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^{\frac{k}{2}} \leq C_0 \max_{j \in \mathcal{A}}\left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^{\frac{k}{2}} + \zeta_n, \tag{4.22}$$

hold, where

$$\mathcal{A} := \left\{j \in \{1, \ldots, K-1\} : \iota_j \neq \iota_{j+1}\right\}. \tag{4.23}$$

Moreover, for all $j \in \{1, \ldots, K\}$ and $t \in J$,

$$\left|\lambda_j'(t)\right| \leq C_0\|\dot{g}(t)\|_{L^2} + \zeta_n. \tag{4.24}$$

If $j \in \{1, \ldots, K\}$ we have

$$\left|\iota_j\lambda_j'(t) + \frac{1}{\langle \mathcal{Z} \mid \Lambda Q \rangle}\left\langle \mathcal{Z}_{\underline{\lambda_j(t)}} \mid \dot{g}(t) \right\rangle\right|$$
$$\leq C_0\|\mathbf{g}(t)\|_{\mathcal{E}}^2 + C_0\left(\left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^{k-1} + \left(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\right)^{k-1}\right)\|\dot{g}(t)\|_{L^2} + \zeta_n, \quad (4.25)$$

where, by convention, $\lambda_0(t) = 0, \lambda_{K+1}(t) = \infty$ for all $t \in J$.

We observe that Lemma 4.11 is sufficient to reduce to the case $K \geq 2$. More precisely, under the contradiction assumption that (4.5) fails, the set $\mathcal{A}$ as defined in (4.23) is non-empty.

Lemma 4.12. *If (4.5) is false, then both $N, K \geq 2$ and the set $\mathcal{A}$ defined in (4.23) is non-empty.*

Proof. Assume (4.5) is false and $\mathcal{A}$ is empty. For n large, we have $\mathbf{d}(a_n) = \epsilon_n \leq \eta_0$ as in Lemma 4.11. Define $e_n := \sup\{t \in [a_n, b_n] : \mathbf{d}(\tau) \leq \eta_0 \ \forall \tau \in [a_n, t]\}$. Since $\mathcal{A}$ is empty, we see from (4.21) and (4.22) that $\mathbf{d}(t) \lesssim \zeta_n \ll 1$ for all $t \in [a_n, e_n)$ and thus $e_n = b_n$, for large n. But this is a contradiction, as $[a_n, b_n] \in \mathcal{C}_K(\eta, \epsilon_n)$ means there must be a $c_n \in [a_n, b_n]$ with $\mathbf{d}(c_n) \geq \eta > 0$. Since $\mathcal{A}$ being empty is impossible, this implies that $N, K \geq 2$ in the event that (4.5) is false. $\qquad\square$

Proof of Lemma 4.11. **Step 1:**(The decomposition (4.19) and the estimates (4.21) and (4.22)) First, observe that by Lemma 4.9,

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} |E(\boldsymbol{u}(t); \nu(t), \infty) - E(\boldsymbol{u}^*) - (N - K)E(\boldsymbol{Q})| = 0. \tag{4.26}$$

Since $E(\boldsymbol{u}) = E(\boldsymbol{u}^*) + NE(\boldsymbol{Q})$ it follows from the above along with (4.8) that

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} |E(\boldsymbol{u}(t); 0, 2\nu(t)) - KE(\boldsymbol{Q})| = 0. \tag{4.27}$$

Using continuity of the flow, the fact that $\mathbf{d}(t) \leq \eta_0$ on J, Lemma 2.27, and by taking $\eta_0 > 0$ small enough, we obtain continuous functions $\vec{\widetilde{\lambda}}(t) = (\widetilde{\lambda}_1(t), \ldots, \widetilde{\lambda}_N(t))$ and signs $\vec{\iota}$ independent of $t \in J$, so that

$$\boldsymbol{u}(t) - \boldsymbol{u}^*(t) = \boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\iota}, \vec{\widetilde{\lambda}}(t)) + \widetilde{\boldsymbol{g}}(t),$$

and,

$$\mathbf{d}(t)^2 \leq \|\widetilde{\boldsymbol{g}}(t)\|_{\mathcal{E}}^2 + \sum_{j=1}^{N} \left(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\right)^k \leq 4\mathbf{d}(t)^2. \tag{4.28}$$

with as usual the convention that $\widetilde{\lambda}_{N+1}(t) = t$. It follows from (4.26) and (4.27) that,

$$\sup_{t \in J} \left[\left(\frac{\widetilde{\lambda}_K(t)}{\nu(t)}\right)^k + \left(\frac{\nu(t)}{\widetilde{\lambda}_{K+1}(t)}\right)^k\right] \lesssim \mathbf{d}(t)^2 + o_n(1) \ \text{ as } \ n \to \infty, \tag{4.29}$$

which means, roughly speaking, that there are K bubbles to the left of the curve $\nu(t)$ and $N - K$ bubbles to the right of the curve $\nu(t)$. Above, and in the rest of the proof, the notation $o_n(1)$ is used to denote a quantity tending to 0 with n that is independent of the subinterval $J \in [a_n, b_n]$. In all instances where $o_n(1)$ appears in the proof it arises from quantities that are estimated uniformly in $t \in [a_n, b_n]$.

For the purposes of this argument we denote by

$$\begin{aligned}
\boldsymbol{v}(t) &:= \boldsymbol{u}(t)\chi_{\nu(t)} + (1 - \chi_{\nu(t)})m_n\boldsymbol{\pi}, \\
\boldsymbol{w}(t) &:= (\boldsymbol{u}(t) - \boldsymbol{u}^*(t))(1 - \chi_{\nu(t)}) + \chi_{\nu(t)}m_n\boldsymbol{\pi}.
\end{aligned} \tag{4.30}$$

Noting that Lemma 4.9 together with (4.29) imply the identity,

$$(m_\Delta - \sum_{j=K+1}^{N} \iota_j)\boldsymbol{\pi} = m_n\boldsymbol{\pi},$$

we may express $\boldsymbol{v}(t)$ on $J \subset [a_n, b_n]$ as follows,

$$\boldsymbol{v}(t) = m_n\boldsymbol{\pi} + \sum_{j=1}^{K} \iota_j(\boldsymbol{Q}_{\widetilde{\lambda}_j(t)} - \boldsymbol{\pi})$$

$$- (1 - \chi_{\nu(t)}) \sum_{j=1}^{K} \iota_j(\boldsymbol{Q}_{\widetilde{\lambda}_j(t)} - \boldsymbol{\pi}) + \chi_{\nu(t)} \sum_{j=K+1}^{N} \iota_j \boldsymbol{Q}_{\widetilde{\lambda}_j(t)} + \chi_{\nu(t)}\boldsymbol{u}^*(t) + \chi_{\nu(t)}\widetilde{\boldsymbol{g}}(t).$$

Using (4.1) along with (4.28) and (4.29) we see that,

$$\|\boldsymbol{v}(t) - m_n \boldsymbol{\pi} - \sum_{j=1}^{K} \iota_j (\boldsymbol{Q}_{\widetilde{\lambda}_j(t)} - \boldsymbol{\pi})\|_{\mathcal{E}}^2 + \sum_{j=1}^{K} \Big(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\Big)^k \lesssim \mathbf{d}(t)^2 + o_n(1) \quad \text{as} \;\; n \to \infty.$$

This means that

$$\mathbf{d}_{m_n,K}(\boldsymbol{v}(t)) \lesssim \mathbf{d}(t) + o_n(1) \quad \text{as} \;\; n \to \infty$$

in the notation of Lemma 2.25. By taking $\eta_0 > 0$ small enough, and n large enough, we may apply Lemma 2.25, (as well as Lemma 2.27, which ensures the signs $\vec{\iota}$ stays fixed) at each $t \in J$, to obtain unique $\boldsymbol{g}(t) \in \mathcal{E}$, $\vec{\lambda}(t) \in (0, \infty)^K$ so that

$$\begin{aligned} \boldsymbol{v}(t) &= \boldsymbol{\mathcal{Q}}(m_n, \vec{\iota}, \vec{\lambda}(t)) + \boldsymbol{g}(t), \\ 0 &= \langle \mathcal{Z}_{\lambda_j(t)} \mid g(t) \rangle, \quad \forall j = 1, \ldots, K, \end{aligned} \tag{4.31}$$

where in this formula $\vec{\iota}, \vec{\lambda}$ are K-vectors, i.e., $\vec{\iota} = (\iota_1, \ldots, \iota_K)$, $\vec{\lambda}(t) = (\lambda_1(t), \ldots, \lambda_K(t))$. We note the estimate,

$$\mathbf{d}_{m_n,K}(\boldsymbol{v}_n(t))^2 \le \|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 + \sum_{j=1}^{K-1} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^k + \Big(\frac{\lambda_K(t)}{\nu(t)}\Big)^k \lesssim \mathbf{d}_{m_n,K}(\boldsymbol{v}(t))^2 + \mathbf{d}(t)^2 + o_n(1) \tag{4.32}$$

$$\lesssim \mathbf{d}(t)^2 + o_n(1),$$

as $n \to \infty$. Next, using (4.27) we see that

$$E(\boldsymbol{v}) \le KE(\boldsymbol{Q}) + o_n(1).$$

Therefore, the estimate (2.30) from Lemma 2.25 applied here yields,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 \lesssim \sup_{j \in \mathcal{A}} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^k + o_n(1)$$

where $\mathcal{A} = \{j \in \{1, \ldots, K-1\} : \iota_j \ne \iota_{j+1}\}$, proving (4.22).

Next, we prove the lower bound in (4.21). Note the identity,

$$\begin{aligned} \boldsymbol{u}(t) - \boldsymbol{u}^*(t) &= \boldsymbol{v}(t) - m_n \boldsymbol{\pi} + \boldsymbol{w}(t) - \chi_{\nu(t)} \boldsymbol{u}^*(t) \\ &= m_\Delta \boldsymbol{\pi} + \sum_{j=1}^{K} \iota_j (\boldsymbol{Q}_{\lambda_j(t)} - \boldsymbol{\pi}) + \sum_{j=K+1}^{N} \sigma_{n,j} (\boldsymbol{Q}_{\mu_j(t)} - \boldsymbol{\pi}) \\ &\quad + \boldsymbol{g}(t) + \boldsymbol{h}(t) - \chi_{\nu(t)} \boldsymbol{u}^*(t) \end{aligned} \tag{4.33}$$

which follows from (4.30), (4.9) and (4.16).

First we prove that $(\iota_{K+1}, \ldots, \iota_N) = (\sigma_{K+1}, \ldots, \sigma_N)$. From (4.9) and (4.10) we see that

$$\|\boldsymbol{w}(t) - m_\Delta \boldsymbol{\pi} - \sum_{j=K+1}^{N} \sigma_{n,j} (\boldsymbol{Q}_{\mu_j(t)} - \boldsymbol{\pi})\|_{\mathcal{E}}^2 + \Big(\frac{\nu(t)}{\mu_{K+1}(t)}\Big)^k + \sum_{j=K+1}^{N} \Big(\frac{\mu_j(t)}{\mu_{j+1}(t)}\Big)^k \lesssim \epsilon_n^2.$$

On the other hand, we see from (4.29) that,

$$\|\boldsymbol{w}(t) - m_\Delta \boldsymbol{\pi} - \sum_{j=K+1}^{N} \iota_j (\boldsymbol{Q}_{\widetilde{\lambda}_j(t)} - \boldsymbol{\pi})\|_{\mathcal{E}}^2 + \Big(\frac{\nu(t)}{\widetilde{\lambda}_{K+1}(t)}\Big)^k + \sum_{j=K+1}^{N} \Big(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\Big)^k \lesssim \mathbf{d}(t)^2 + o_n(1).$$

Hence, using Lemma 2.27 we see that for any $\theta_0 > 0$ we may take $\eta_0 > 0$ small enough so that $(\iota_{K+1}, \ldots, \iota_N) = (\sigma_{K+1}, \ldots, \sigma_N)$, and in addition we have

$$\left| \frac{\widetilde{\lambda}_j(t)}{\mu_{n,j}(t)} - 1 \right| \leq \theta_0 \quad \forall j = K+1, \ldots, N.$$

The above, together with (4.10) implies that

$$\sum_{j=K+1}^{N} \left(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)} \right)^k \lesssim \epsilon_n^2.$$

We may thus rewrite (4.33) as

$$\boldsymbol{u}(t) - \boldsymbol{u}^*(t) = m_\Delta \boldsymbol{\pi} + \sum_{j=1}^{K} \iota_j (\boldsymbol{Q}_{\lambda_j(t)} - \boldsymbol{\pi}) + \sum_{j=K+1}^{N} \iota_j (\boldsymbol{Q}_{\mu_j(t)} - \boldsymbol{\pi})$$
$$+ \boldsymbol{g}(t) + \boldsymbol{h}(t) - \chi_{\nu(t)} \boldsymbol{u}^*(t)$$

Noting that

$$\sup_{t \in [a_n, b_n]} \|\boldsymbol{u}^*(t) \chi_{\nu(t)}\|_{\mathcal{E}} = o_n(1) \quad \text{as} \quad n \to \infty,$$

the previous line together with (4.32) and (4.10) imply that,

$$\mathbf{d}(t)^2 \lesssim \mathbf{d}_{m_n,K}(\boldsymbol{v}(t))^2 + o_n(1) \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 + \sum_{j=1}^{K-1} \left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)} \right)^k + o_n(1) \quad \text{as} \quad n \to \infty,$$

which proves the lower bound in (4.21).

Step 2:(The dynamical estimates (4.24) and (4.25)) Momentarily assuming that $\vec{\lambda} \in C^1(J)$ (we will justify this assumption below) we record the computations,

$$\partial_t v(t) = \dot{g}(t) + (m_n \pi - u(t)) \frac{\nu'(t)}{\nu(t)} \Lambda \chi_{\nu(t)}, \quad \partial_t \mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}(t)) = -\sum_{j=1}^{K} \iota_j \lambda_j'(t) \Lambda Q_{\underline{\lambda_j(t)}}.$$

Differentiating in time the first line in (4.31) and using the above lead to the expression,

$$\partial_t g(t) = \dot{g}(t) + \sum_{j=1}^{K} \iota_j \lambda_j'(t) \Lambda Q_{\underline{\lambda_j(t)}} + (m_n \pi - u(t)) \frac{\nu'(t)}{\nu(t)} \Lambda \chi_{\nu(t)}. \tag{4.34}$$

We differentiate the orthogonality conditions (4.20) for each $j = 1, \ldots, K$,

$$0 = -\frac{\lambda_j'}{\lambda_j} \left\langle \underline{\Lambda \mathcal{Z}_{\underline{\lambda_j}}} \mid g \right\rangle + \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \partial_t g \right\rangle$$
$$= -\frac{\lambda_j'}{\lambda_j} \left\langle \underline{\Lambda \mathcal{Z}_{\underline{\lambda_j}}} \mid g \right\rangle + \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \dot{g} \right\rangle + \sum_{\ell=1}^{K} \iota_\ell \lambda_\ell' \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_\ell}} \right\rangle + \frac{\nu'}{\nu} \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid (m_n \pi - u) \Lambda \chi_\nu \right\rangle,$$

which we rearrange into the system,

$$\iota_j \lambda_j' \left(\langle \mathcal{Z} \mid \Lambda Q \rangle - \lambda_j^{-1} \langle \underline{\Lambda \mathcal{Z}_{\underline{\lambda_j}}} \mid g \rangle \right) + \sum_{i \neq j} \iota_i \lambda_i' \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}} \right\rangle$$
$$= - \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \dot{g} \right\rangle - \frac{\nu'}{\nu} \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid (m_n \pi - u) \Lambda \chi_\nu \right\rangle. \tag{4.35}$$

This is a diagonally dominant system, hence invertible, and we arrive at the estimate,

$$\left|\lambda_j'\right| \lesssim \|\dot{g}\|_{L^2} + o_n(1) \quad j = 1, \ldots, K, \tag{4.36}$$

after noting the estimates,

$$\left|\left\langle \mathcal{Z}_{\underline{\lambda}_j} \mid \dot{g} \right\rangle\right| \lesssim \|\dot{g}\|_{L^2}$$

$$\left|\frac{\nu'(t)}{\nu(t)} \left\langle \mathcal{Z}_{\underline{\lambda}_j} \mid (m_n\pi - u(t))\Lambda\chi_{\nu(t)} \right\rangle\right| \lesssim |\nu'| \frac{\lambda_j}{\nu} \|r^{-1}(m_n\pi - u(t))\Lambda\chi_{\nu(t)}\|_{L^2} = o_n(1),$$

where the last line follows from (4.11). Lastly, we note that the system (4.35) implies that $\vec{\lambda}(t)$ is a C^1 function on J. Indeed, arguing as in the end of the proof of Lemma 4.9, let $t_0 \in J$ be any time and let $\vec{\lambda}(t_0)$ be defined as in (4.31). Using the smallness (4.32) at time t_0, the system (4.35) admits a unique C^1 solution $\vec{\lambda}_{\text{ode}}(t)$ in a neighborhood of t_0. Due to the way the system (4.35) was derived, the orthogonality conditions in (4.31) hold with $\vec{\lambda}_{\text{ode}}(t)$. Since $\vec{\lambda}(t)$ was obtained uniquely via the implicit function theorem, we must have $\vec{\lambda}(t) = \vec{\lambda}_{\text{ode}}(t)$, which means that $\vec{\lambda}(t)$ is C^1.

Lastly, the estimates (4.25) are immediate from (4.35) using (4.36) along with the estimates,

$$\left\langle \mathcal{Z}_{\underline{\lambda}_j} \mid \Lambda Q_{\underline{\lambda}_i} \right\rangle \lesssim \begin{cases} \left(\frac{\lambda_j}{\lambda_i}\right)^{k+1} & \text{if } j < i \\ \left(\frac{\lambda_i}{\lambda_j}\right)^{k-1} & \text{if } j > i \end{cases}$$

$$\left|\lambda_j^{-1} \left\langle \underline{\Lambda}\mathcal{Z}_{\underline{\lambda}_j} \mid g \right\rangle\right| \lesssim \|g\|_H,$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.3. **Refined modulation.** Next, our goal is to gain precise dynamical control of the modulation parameters in the spirit of [35,41]. The idea is to construct a virial correction to the modulation parameters; see (4.47). We start by finding suitable truncation of the function $\frac{1}{2}r^2$, similar to [36, Lemma 4.6]. Since here we may have arbitrary number of bubbles, we need to localize this function both away from $r = 0$ and away from $r = \infty$. To make the exposition as uniform as possible, we restrict to equivariance classes $k \geq 2$ in this section, saving case $k = 1$, which introduces additional technical complications, for the appendix.

Lemma 4.13. *For any $c > 0$ and $R > 1$ there exists a function $q = q_{c,R} \in C^4((0,\infty))$ having the following properties:*
- *(P1) $q(r) = \frac{1}{2}r^2$ for all $r \in [R^{-1}, R]$,*
- *(P2) there exists $\widetilde{R} > 0$ (depending on c and R) such that $q(r) = \text{const}$ for $r \geq \widetilde{R}$ and $q(r) = \text{const}$ for $r \leq \widetilde{R}^{-1}$,*
- *(P3) $|q'(r)| \lesssim r$ and $|q''(r)| \lesssim 1$ for all $r > 0$, with constants independent of c and R,*
- *(P4) $q''(r) \geq -c$ and $\frac{1}{r}q'(r) \geq -c$ for all $r > 0$,*
- *(P5) $\left|\left(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}\right)^2 q(r)\right| \leq cr^{-2}$ for all $r > 0$,*
- *(P6) $\left|\left(\frac{q'(r)}{r}\right)'\right| \leq cr^{-1}$ for all $r > 0$.*

Proof. **Step 1:** We construct a function $q(r)$ satisfying the desired properties for all $r \geq 1$. In this step, without loss of generality we can assume $R = 1$. Let $c_1 > 0$ be small, to be chosen later and set

$$q_o(r) := \begin{cases} \frac{1}{2}r^2 & \text{if } r \leq 1 \\ \frac{1}{2}r^2 - c_1\psi_o(r) & \text{if } r \geq 1 \end{cases}$$

where $\psi_{\mathrm{o}}(r) = \frac{1}{2}r^2 \log r + \widetilde{\psi}_{\mathrm{o}}(r)$ and $\widetilde{\psi}_{\mathrm{o}}(r)$ is any smooth function satisfying,

$$\widetilde{\psi}_{\mathrm{o}}(1) = 0, \quad \widetilde{\psi}_{\mathrm{o}}'(1) = -\frac{1}{2}, \quad \widetilde{\psi}_{\mathrm{o}}''(1) = -\frac{3}{2}, \quad \widetilde{\psi}_{\mathrm{o}}'''(1) = -1, \quad \widetilde{\psi}_{\mathrm{o}}^{(4)}(1) = 1,$$

$$\left| \widetilde{\psi}_{\mathrm{o}}^{(j)}(r) \right| \lesssim r^{2-j} \quad \forall r \geq 1, \ j = 0, 1, \ldots 4$$

which ensures $0 = \psi_{\mathrm{o}}(1) = \psi_{\mathrm{o}}'(1) = \psi_{\mathrm{o}}''(1) = \psi_{\mathrm{o}}'''(1) = \psi_{\mathrm{o}}^{(4)}(1)$. To construct such a function it suffices to take a suitable linear combination of negative powers of r, for example. Set $R_0 := \exp(1/c_1)$. We check all the properties for $1 \leq r \leq R_0$. We have, $q_{\mathrm{o}}'(r) = r(1 - c_1 \log r) + O(c_1 r)$ and $q_{\mathrm{o}}''(r) = (1 - c_1 \log r) + O(c_1)$, so *(P4)* holds. Also,

$$\left| \left(\frac{\psi_{\mathrm{o}}'(r)}{r} \right)' \right| = \left| (\log r + \frac{1}{2})' \right| + O(r^{-1}) = O(r^{-1})$$

for an absolute constant, implying *(P6)*. *(P5)* follows from $\Delta^2(r^2 \log r) = 0$, with all the remaining terms estimated brutally. We now truncate at R_0, setting $e_j(r) := \frac{1}{j!}r^j \chi(r)$ for $j = 1, \ldots, 4$ and defining,

$$q(r) := \begin{cases} q_{\mathrm{o}}(r) \ \text{if} \ r \leq R_0 \\ q_{\mathrm{o}}(R_0) + \sum_{j=1}^4 q_{\mathrm{o}}^{(j)}(R_0) R_0^j e_j(-1 + r/R_0), \ \text{if} \ r \geq R_0 \end{cases}$$

Noting that $\left| q_{\mathrm{o}}^{(j)}(R_0) \right| \lesssim c_1 R_0^{2-j}$ for $j = 1, \ldots, 4$, we see that $q(r)$ inherits all the desired properties from $q_{\mathrm{o}}(r)$ and is constant after $3R_0$; see [36, Lemma 4.6] for additional details.

Step 2: We next find a function $q(r)$ with all the desired properties for $r \leq 1$. As above, we may assume here that $R = 1$. Let $c_1 > 0$ be small, to be chosen later, and set

$$q_{\mathrm{i}}(r) := \begin{cases} \frac{1}{2}r^2 \ \text{if} \ r \geq 1 \\ \frac{1}{2}r^2 + c_1 \psi_{\mathrm{i}}(r) \ \text{if} \ r \leq 1 \end{cases}$$

where $\psi_{\mathrm{i}}(r) = \frac{1}{2}r^2 \log r + \widetilde{\psi}_{\mathrm{i}}(r)$ and $\widetilde{\psi}_{\mathrm{i}}(r)$ is any smooth function satisfying,

$$\widetilde{\psi}_{\mathrm{i}}(1) = 0, \quad \widetilde{\psi}_{\mathrm{i}}'(1) = -\frac{1}{2}, \quad \widetilde{\psi}_{\mathrm{i}}''(1) = -\frac{3}{2}, \quad \widetilde{\psi}_{\mathrm{i}}'''(1) = -1, \quad \widetilde{\psi}_{\mathrm{i}}^{(4)}(1) = 1,$$

$$\left| \widetilde{\psi}_{\mathrm{i}}^{(j)}(r) \right| \lesssim r^{2-j} \quad \forall r \leq 1, \ j = 0, 1, \ldots 4$$

which ensures $0 = \psi_{\mathrm{i}}(1) = \psi_{\mathrm{i}}'(1) = \psi_{\mathrm{i}}''(1) = \psi_{\mathrm{i}}'''(1) = \psi_{\mathrm{i}}^{(4)}(1)$. To obtain such a function it suffices to take a suitable linear combination of positive powers of r, for example. Set $R_0^{-1} := \exp(-\frac{1}{c_1})$. One can check, as in Step 1, that all the properties hold for $R_0^{-1} \leq r \leq 1$, using that $1 + c_1 \log r \geq 0$ in this regime. Then truncate as in Step 1 to obtain the truncated function $q(r)$.

Step 3: The final function $q(r)$ is obtained by gluing together the two functions called q obtained in Steps 1, 2. $\qquad \square$

Definition 4.14 (Localized virial operator)**.** For each $\lambda > 0$ we set

$$A(\lambda)g(r) := q'\left(\frac{r}{\lambda}\right) \cdot \partial_r g(r), \tag{4.37}$$

$$\underline{A}(\lambda)g(r) := \left(\frac{1}{2\lambda}q''\left(\frac{r}{\lambda}\right) + \frac{1}{2r}q'\left(\frac{r}{\lambda}\right)\right)g(r) + q'\left(\frac{r}{\lambda}\right) \cdot \partial_r g(r). \tag{4.38}$$

These operators depend on c and R as in Lemma 4.13.

Note the similarity between A and $\frac{1}{\lambda}\Lambda$ and between $\underline{A}$ and $\frac{1}{\lambda}\underline{\Lambda}$. For technical reasons we introduce the space

$$X := \{g \in H \mid \frac{g}{r}, \partial_r g \in H\}.$$

Lemma 4.15 (Localized virial estimates). [36, Lemma 5.5] *For any $c_0 > 0$ there exist $c_1, R_1 > 0$, so that for all c, R as Lemma 4.13 with $c < c_1$, $R > R_1$ the operators $A(\lambda)$ and $\underline{A}(\lambda)$ defined in (4.37) and (4.38) have the following properties:*

- *the families $\{A(\lambda) : \lambda > 0\}$, $\{\underline{A}(\lambda) : \lambda > 0\}$, $\{\lambda\partial_\lambda A(\lambda) : \lambda > 0\}$ and $\{\lambda\partial_\lambda \underline{A}(\lambda) : \lambda > 0\}$ are bounded in $\mathscr{L}(H; L^2)$, with the bound depending only on the choice of the function $q(r)$,*

- *Let $g_1 = \mathcal{Q}(m, \vec{\iota}, \vec{\lambda})$ be an M-bubble configuration and let $g \in X$. Then, for all $\lambda > 0$,*

$$\left| \left\langle A(\lambda)g_1 \mid \frac{1}{r^2}\big(f(g_1 + g_2) - f(g_1) - f'(g_1)g_2\big) \right\rangle \right.$$
$$\left. + \left\langle A(\lambda)g_2 \mid \frac{1}{r^2}\big(f(g_1 + g_2) - f(g_1) - k^2 g_2\big) \right\rangle \right| \leq \frac{c_0}{\lambda} \|g_2\|_H^2, \quad (4.39)$$

- *For all $g \in X$ we have*

$$\langle \underline{A}(\lambda)g \mid \mathcal{L}_0 g \rangle \geq -\frac{c_0}{\lambda}\|g\|_H^2 + \frac{1}{\lambda}\int_{R^{-1}\lambda}^{R\lambda} \left((\partial_r g)^2 + \frac{k^2}{r^2}g^2 \right) r\mathrm{d}r, \quad (4.40)$$

- *For $\lambda, \mu > 0$ with either $\lambda/\mu \ll 1$ or $\mu/\lambda \ll 1$,*

$$\|\underline{\Lambda}\Lambda Q_{\underline{\lambda}} - \underline{A}(\lambda)\Lambda Q_\lambda\|_{L^2} \leq c_0, \quad (4.41)$$

$$\|\big(\frac{1}{\lambda}\Lambda - A(\lambda)\big)Q_\lambda\|_{L^\infty} \leq \frac{c_0}{\lambda}, \quad (4.42)$$

$$\|A(\lambda)Q_\mu\|_{L^\infty} + \|\underline{A}(\lambda)Q_\mu\|_{L^\infty} \lesssim \frac{1}{\lambda}\min\{(\lambda/\mu)^k, (\mu/\lambda)^k\} \quad (4.43)$$

$$\|A(\lambda)Q_\mu\|_{L^2} + \|\underline{A}(\lambda)Q_\mu\|_{L^2} \lesssim \min\{(\lambda/\mu)^k, (\mu/\lambda)^k\} \quad (4.44)$$

- *Lastly, the following localized coercivity estimate holds. Fix any smooth function $\mathcal{Z} \in L^2 \cap X$ such that $\langle \mathcal{Z} \mid \Lambda Q \rangle > 0$. For any $g \in H, \lambda > 0$ with $\langle g \mid \mathcal{Z}_{\underline{\lambda}} \rangle = 0$,*

$$\frac{1}{\lambda}\int_{R^{-1}\lambda}^{R\lambda} (\partial_r g)^2 + \frac{k^2}{r^2}g^2 \, r\mathrm{d}r + \frac{1}{\lambda}\int_0^\infty \big(\frac{1}{2}q''\big(\frac{r}{\lambda}\big) + \frac{\lambda}{2r}q'\big(\frac{r}{\lambda}\big)\big)\frac{k^2}{r^2}(f'(Q_\lambda) - 1)g^2 \, r \, \mathrm{d}r$$
$$\geq -\frac{c_0}{\lambda}\|g\|_H^2. \quad (4.45)$$

Proof. See [36, Lemmas 4.7 and 5.5] for the proof in the cases $k \geq 2$ and [65, Lemma 3.7 and Remark 3.8] for modifications to handle the case $k = 1$. $\qquad\square$

The modulation parameters $\vec{\lambda}(t)$ defined in Lemma 4.11 are imprecise proxies for the dynamics in the case $k = 2$ (and also $k = 1$; see the appendix) due to the fact that the orthogonality conditions were imposed relative to $\mathcal{Z} \neq \Lambda Q$. Indeed, we use (4.20) primarily to ensure coercivity, and thus the estimate (4.22), as well as the differentiability of $\vec{\lambda}(t)$. To access the dynamics of (1.1) we introduce a correction $\vec{\xi}(t)$ defined as follows. For each $t \in J \subset [a_n, b_n]$ as in Lemma 4.11 set,

$$\xi_j(t) := \begin{cases} \lambda_j(t) & \text{if } k \geq 3 \\ \lambda_j(t) - \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle \chi_{L\lambda_j(t)}\Lambda Q_{\underline{\lambda_j(t)}} \mid g(t) \rangle & \text{if } k = 2 \end{cases} \quad (4.46)$$

for each $j = 1, \ldots, K-1$, and where $L > 0$ is a large constant to be determined below. (Note that for $j = K$ we only require the brutal estimate (4.24)). We require yet another modification, since the dynamics of (1.1) truly enter after taking two derivatives of the modulation parameters and it is not clear how to derive useful estimates from the expression for $\xi_j''(t)$. So we introduce a refined

modulation parameter, which we view as a subtle correction to $\xi_j'(t)$. For each $t \in J \subset [a_n, b_n]$ as in Lemma 4.11 and for each $j \in \{1, \ldots, K\}$ define,

$$\beta_j(t) := -\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle \Lambda Q_{\underline{\lambda_j(t)}} \mid \dot{g}(t)\rangle - \frac{1}{\|\Lambda Q\|_{L^2}^2}\langle \underline{A}(\lambda_j(t))g(t) \mid \dot{g}(t)\rangle. \tag{4.47}$$

Note that $\beta_j(t)$ is similar to the function called $b(t)$ in [41].

Lemma 4.16 (Refined modulation). *Let $k \geq 2$ and $c_0 \in (0, 1)$. There exist $\eta_0 > 0, L > 0, c > 0, R > 1, C_0 > 0$ and a decreasing sequence $\delta_n \to 0$ so that the following is true. Let $J \subset [a_n, b_n]$ be an open time interval with*

$$\mathbf{d}(t) \leq \eta_0 \quad and \quad \max_{i \in \mathcal{A}}\big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{k/2} \geq \delta_n, \tag{4.48}$$

for all $t \in J$, where $\mathcal{A} := \{j \in \{1, \ldots, K-1\} \mid \iota_j \neq \iota_{j+1}\}$. Then, for all $t \in J$,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{i \notin \mathcal{A}}\big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{k/2} \leq C_0 \max_{i \in \mathcal{A}}\big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{k/2}, \tag{4.49}$$

and,

$$\frac{1}{C_0}\mathbf{d}(t) \leq \max_{i \in \mathcal{A}}\big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{k/2} \leq C_0\mathbf{d}(t). \tag{4.50}$$

Moreover, for all $j \in \{1, \ldots, K-1\}$ and $t \in J$,

$$|\xi_j(t)/\lambda_j(t) - 1| \leq c_0, \tag{4.51}$$

$$|\xi_j'(t) - \beta_j(t)| \leq c_0 \max_{i \in \mathcal{A}}\left(\frac{\lambda_i(t)}{\lambda_{i+1}(t)}\right)^{k/2}, \tag{4.52}$$

and,

$$\begin{aligned}
\beta_j'(t) \geq{}& \left(-\iota_j\iota_{j+1}\omega^2 - c_0\right)\frac{1}{\lambda_j(t)}\left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^k + \left(\iota_j\iota_{j-1}\omega^2 - c_0\right)\frac{1}{\lambda_j(t)}\left(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\right)^k \\
& - \frac{c_0}{\lambda_j(t)}\max_{i \in \mathcal{A}}\left(\frac{\lambda_i(t)}{\lambda_{i+1}(t)}\right)^k.
\end{aligned} \tag{4.53}$$

where, by convention, $\lambda_0(t) = 0, \lambda_{K+1}(t) = \infty$ for all $t \in J$, and $\omega^2 > 0$ is defined by

$$\omega^2 = \omega^2(k) := 8k^2\|\Lambda Q\|_{L^2}^{-2} = 4k^2\pi^{-1}\sin(\pi/k) > 0.$$

Remark 4.17. By (4.22), without loss of generality (upon enlarging ϵ_n), we can assume that

$$\eta_0 \geq \mathbf{d}(t) \geq \epsilon_n \quad \text{implies} \quad \max_{i \in \mathcal{A}}\big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{k/2} \geq \delta_n,$$

so that Lemma 4.16 can always be applied on the time intervals $J \subset [a_n, b_n]$ as long as $\mathbf{d}(t) \leq \eta_0$ on J.

Before beginning the proof of Lemma 4.16 we record the equation satisfied by $\boldsymbol{g}(t)$. Observe the identity,

$$\Delta\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}) = \frac{k^2}{r^2}\sum_{j=1}^{K}\iota_j f(Q_{\lambda_j}),$$

and hence,

$$(\partial_t^2 u)\chi_\nu = \chi_\nu \Delta u - \frac{k^2}{r^2} f(u)\chi_\nu$$

$$= \Delta\big(u\chi_\nu + (1-\chi_\nu)m_n\pi\big) - \frac{k^2}{r^2} f\big(u\chi_\nu + (1-\chi_\nu)m_n\pi\big)$$

$$- \frac{2}{r}\partial_r u \Lambda\chi_\nu + \frac{1}{r^2}(m_n\pi - u)(r^2\Delta\chi)_\nu + \frac{k^2}{r^2}\Big(f\big(u\chi_\nu + (1-\chi_\nu)m_n\pi\big)\big) - f(u)\chi_\nu\Big)$$

$$= \Delta g - \frac{k^2}{r^2}\Big(f(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}) + g) - \sum_{j=1}^{K} \iota_j f(Q_{\lambda_j})\Big)$$

$$- \frac{2}{r}\partial_r u \Lambda\chi_\nu + \frac{1}{r^2}(m_n\pi - u)(r^2\Delta\chi)_\nu + \frac{k^2}{r^2}\Big(f\big(u\chi_\nu + (1-\chi_\nu)m_n\pi\big)\big) - f(u)\chi_\nu\Big).$$

Recalling (4.34), we are led to the system of equations,

$$\partial_t g(t) = \dot{g}(t) + \sum_{j=1}^{K} \iota_j \lambda_j'(t)\Lambda Q_{\underline{\lambda_j(t)}} + \phi(u(t), \nu(t)) \tag{4.54}$$

$$\partial_t \dot{g}(t) = -\mathcal{L}_{\mathcal{Q}} g + f_{\mathbf{i}}(m_n, \iota, \vec{\lambda}) + f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) + \dot{\phi}(u(t), \nu(t)),$$

where,

$$\phi(u, \nu) := (m_n\pi - u)\frac{\nu'}{\nu}\Lambda\chi_\nu$$

$$\dot{\phi}(u, \nu) := -\frac{2}{r}\partial_r u \Lambda\chi_\nu + \frac{1}{r^2}(m_n\pi - u)(r^2\Delta\chi)_\nu \tag{4.55}$$

$$+ \frac{k^2}{r^2}\Big(f\big(u\chi_\nu + (1-\chi_\nu)m_n\pi\big)\big) - f(u)\chi_\nu\Big) - \frac{\nu'}{\nu}\Lambda\chi_\nu\partial_t u,$$

which we note are supported in $r \in (\nu, \infty)$, and

$$f_{\mathbf{i}}(m_n, \vec{\iota}, \vec{\lambda}) := -\frac{k^2}{r^2}\Big(f\big(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda})\big) - \sum_{j=1}^{K} \iota_j f(Q_{\lambda_j})\Big)$$

$$f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) := -\frac{k^2}{r^2}\Big(f\big(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}) + g\big) - f\big(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda})\big) - f'\big(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda})\big)g\Big).$$

The subscript $\mathbf{i}$ above stands for "interaction" and $\mathbf{q}$ stands for "quadratic." In particular, $f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)$ satisfies,

$$\|f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\|_{L^1} \lesssim \|g\|_H^2. \tag{4.56}$$

In one instance it will be convenient to rewrite the right-hand side of the equation for $\dot{g}$ as follows,

$$\partial_t \dot{g} = -\mathcal{L}_0 g + f_{\mathbf{i}}(m_n, \iota, \vec{\lambda}) + \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) + \dot{\phi}(u, \chi_\nu), \tag{4.57}$$

where $\widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)$ is defined by the formula,

$$\widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) := -\frac{k^2}{r^2}\Big(f\big(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}) + g\big) - f\big(\mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda})\big) - g\Big). \tag{4.58}$$

Proof of Lemma 4.16. First, we prove the estimates (4.49) and (4.50). Let ζ_n be the sequence given by Lemma 4.11 and let δ_n be any sequence such that $\zeta_n/\delta_n \to 0$ as $n \to \infty$. Using Lemma 4.11, estimate (4.49) follows from (4.22) and the estimate (4.50) follows from (4.21).

Note also that with this choice of δ_n and (4.49), the estimate (4.24) leads to,

$$\left|\lambda'_j(t)\right| \lesssim \max_{i \in \mathcal{A}} \left(\lambda_i(t)/\lambda_{i+1}(t)\right)^{k/2}. \tag{4.59}$$

Next, we treat (4.51), which is only relevant in the case $k = 2$. From (4.46) we see that,

$$|\xi_j/\lambda_j - 1| = |\|\Lambda Q\|_{L^2}^{-2}\lambda_j^{-1}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid g\rangle|$$

$$\lesssim \|g\|_{L^\infty}(1 + \int_1^L \Lambda Q(r)\, r\, \mathrm{d}r) \lesssim (1 + \log(L))\|g\|_H \le c_0$$

using (4.49) and (4.50) in the last line. Next we compute $\xi'_j(t)$.

For $k = 2$, from (4.46) we have

$$\xi'_j = \lambda'_j - \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \partial_t g\rangle$$
$$+ \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\frac{\lambda'_j}{\lambda_j}\langle\Lambda\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid g\rangle + \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\frac{\lambda'_j}{\lambda_j}\langle\chi_{L\lambda_j}\underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid g\rangle. \tag{4.60}$$

We examine each of the terms on the right above. The last two terms are negligible. Indeed, using $\|g\|_{L^\infty} \lesssim \|g\|_H$,

$$\left|\frac{\lambda'_j}{\lambda_j}\langle\Lambda\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid g\rangle\right| \lesssim |\lambda'_j|\,\|g\|_{L^\infty}\int_{2^{-1}L}^{2L}\Lambda Q(r)\, r\, \mathrm{d}r$$
$$\lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

and,

$$\left|\frac{\lambda'_j}{\lambda_j}\langle\chi_{L\lambda_j}\underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid g\rangle\right| \lesssim |\lambda'_j|\,\|g\|_{L^\infty}\int_0^{2L}\Lambda Q(r)\, r\, \mathrm{d}r \lesssim (1 + \log(L))\|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

which is small relative to $\|\boldsymbol{g}\|_{\mathcal{E}}$ because of (4.49). Using (4.54) in the second term in (4.60) gives

$$-\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \partial_t g\rangle = -\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle - \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \sum_{i=1}^K \iota_i\lambda'_i\Lambda Q_{\underline{\lambda_i}}\rangle$$
$$- \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \phi(u,\nu)\rangle.$$

The first term on the right satisfies,

$$-\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle = -\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle + \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle(1 - \chi_{L\lambda_j})\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle$$
$$= -\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle + o_L(1)\|\boldsymbol{g}\|_{\mathcal{E}}.$$

where the $o_L(1)$ term can be made as small as we like by taking $L > 0$ large. Using (4.59), the second term yields,

$$-\frac{\iota_j}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \sum_{i=1}^K \iota_i\lambda'_i\Lambda Q_{\underline{\lambda_i}}\rangle = -\lambda'_j$$

$$-\sum_{i \ne j}\frac{\iota_j\iota_i\lambda'_i}{\|\Lambda Q\|_{L^2}^2}\langle\chi_{L\lambda_j}\Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}}\rangle + \frac{\lambda'_j}{\|\Lambda Q\|_{L^2}^2}\langle(1 - \chi_{L\lambda_j})\Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_j}}\rangle$$

$$= -\lambda'_j + O((\lambda_{j-1}/\lambda_j) + (\lambda_j/\lambda_{j+1}) + o_L(1))\max_{i \in \mathcal{A}}\left(\lambda_i/\lambda_{i+1}\right)^{k/2}.$$

Finally, the third term vanishes due to the fact that for each $j < K$, $L\lambda_j \ll \lambda_K \ll \nu$, and hence

$$\left\langle \chi_{L\lambda_j} \Lambda Q_{\underline{\lambda_j}} \mid \phi(u,\nu) \right\rangle = 0.$$

Plugging all of this back into (4.60) we obtain,

$$\left| \xi_j'(t) + \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2} \left\langle \Lambda Q_{\underline{\lambda_j}} \mid \dot{g} \right\rangle \right| \leq c_0 \max_{i \in \mathcal{A}} \left(\lambda_i/\lambda_{i+1} \right)^{k/2}. \tag{4.61}$$

for $k = 2$, after fixing $L > 0$ sufficiently large. The same estimate for $k \geq 3$, i.e., when $\xi_j'(t) = \lambda_j'(t)$, is immediate from (4.25) since in this case we take $\mathcal{Z} = \Lambda Q$. Thus (4.61) holds for all $k \geq 2$. The estimate (4.52) is then immediate from (4.61), the definition of β_j, and the estimate,

$$\left| \frac{1}{\|\Lambda Q\|_{L^2}^2} \left\langle \underline{A}(\lambda_j)g \mid \dot{g} \right\rangle \right| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

which follows from the first bullet point in Lemma 4.15.

We prove (4.53). We compute,

$$\beta_j' = \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2} \frac{\lambda_j'}{\lambda_j} \left\langle \underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g} \right\rangle - \frac{\iota_j}{\|\Lambda Q\|_{L^2}^2} \left\langle \Lambda Q_{\underline{\lambda_j}} \mid \partial_t \dot{g} \right\rangle$$
$$- \frac{1}{\|\Lambda Q\|_{L^2}^2} \frac{\lambda_j'}{\lambda_j} \left\langle \lambda_j \partial_{\lambda_j} \underline{A}(\lambda_j)g \mid \dot{g} \right\rangle - \frac{1}{\|\Lambda Q\|_{L^2}^2} \left\langle \underline{A}(\lambda_j)\partial_t g \mid \dot{g} \right\rangle - \frac{1}{\|\Lambda Q\|_{L^2}^2} \left\langle \underline{A}(\lambda_j)g \mid \partial_t \dot{g} \right\rangle. \tag{4.62}$$

Using (4.54) we arrive at the expression,

$$-\left\langle \Lambda Q_{\underline{\lambda_j}} \mid \partial_t \dot{g} \right\rangle = \left\langle \Lambda Q_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{Q}} - \mathcal{L}_{\lambda_j})g \right\rangle - \left\langle \Lambda Q_{\underline{\lambda_j}} \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda}) \right\rangle$$
$$- \left\langle \Lambda Q_{\underline{\lambda_j}} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle - \left\langle \Lambda Q_{\underline{\lambda_j}} \mid \dot{\phi}(u,\nu) \right\rangle,$$

where in the first term on the right we used that $\mathcal{L}_{\lambda_j} \Lambda Q_{\underline{\lambda_j}} = 0$. Using (4.54) we obtain,

$$-\left\langle \underline{A}(\lambda_j)\partial_t g \mid \dot{g} \right\rangle = -\left\langle \underline{A}(\lambda_j)\dot{g} \mid \dot{g} \right\rangle - \sum_{i=1}^{K} \iota_i \lambda_i' \left\langle \underline{A}(\lambda_j)\Lambda Q_{\underline{\lambda_i}} \mid \dot{g} \right\rangle - \left\langle \underline{A}(\lambda_j)\phi(u,\nu) \mid \dot{g} \right\rangle$$
$$= -\iota_j \lambda_j' \left\langle \underline{A}(\lambda_j)\Lambda Q_{\underline{\lambda_j}} \mid \dot{g} \right\rangle - \sum_{i \neq j} \iota_i \lambda_i' \left\langle \underline{A}(\lambda_j)\Lambda Q_{\underline{\lambda_i}} \mid \dot{g} \right\rangle - \left\langle \underline{A}(\lambda_j)\phi(u,\nu) \mid \dot{g} \right\rangle$$

where we used that $\left\langle \underline{A}(\lambda_j)\dot{g} \mid \dot{g} \right\rangle = 0$. Finally, using (4.57) we have,

$$-\left\langle \underline{A}(\lambda_j)g \mid \partial_t \dot{g} \right\rangle = \left\langle \underline{A}(\lambda_j)g \mid \mathcal{L}_0 g \right\rangle - \left\langle \underline{A}(\lambda_j)g \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda}) \right\rangle$$
$$- \left\langle \underline{A}(\lambda_j)g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle - \left\langle \underline{A}(\lambda_j)g \mid \dot{\phi}(u,\nu) \right\rangle.$$

Plugging these back into (4.62) and rearranging we have,

$$
\begin{aligned}
\|\Lambda Q\|_{L^2}^2 \beta_j' = {}& -\frac{\iota_j}{\lambda_j}\left\langle \Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle + \left\langle \underline{A}(\lambda_j)g \mid \mathcal{L}_0 g\right\rangle \\
& + \left\langle (A(\lambda_j) - \underline{A}(\lambda_j))g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle \\
& + \left\langle \Lambda Q_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{Q}} - \mathcal{L}_{\lambda_j})g\right\rangle + \iota_j \frac{\lambda_j'}{\lambda_j}\left\langle (\frac{1}{\lambda_j}\underline{\Lambda} - \underline{A}(\lambda_j))\Lambda Q_{\lambda_j} \mid \dot{g}\right\rangle \\
& - \left\langle A(\lambda_j)\sum_{i=1}^{K}\iota_i Q_{\lambda_i} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle - \left\langle A(\lambda_j)g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle \\
& + \iota_j\left\langle (A(\lambda_j) - \frac{1}{\lambda_j}\Lambda)Q_{\lambda_j} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle - \frac{\lambda_j'}{\lambda_j}\left\langle \lambda_j \partial_{\lambda_j}\underline{A}(\lambda_j)g \mid \dot{g}\right\rangle \\
& + \sum_{i\neq j}\iota_i\left\langle A(\lambda_j)Q_{\lambda_i} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle \\
& - \sum_{i\neq j}\iota_i\lambda_i'\left\langle \underline{A}(\lambda_j)\Lambda Q_{\underline{\lambda_i}} \mid \dot{g}\right\rangle - \left\langle \underline{A}(\lambda_j)g \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle \\
& - \iota_j\left\langle \Lambda Q_{\underline{\lambda_j}} \mid \dot{\phi}(u, \nu)\right\rangle - \left\langle \underline{A}(\lambda_j)\phi(u, \nu) \mid \dot{g}\right\rangle - \left\langle \underline{A}(\lambda_j)g \mid \dot{\phi}(u, \nu)\right\rangle
\end{aligned}
\tag{4.63}
$$

We examine each of the terms on the right-hand side above. The leading order contribution comes from the first term, i.e., by Lemma 2.23

$$
-\frac{\iota_j}{\lambda_j\|\Lambda Q\|_{L^2}^2}\left\langle \Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle = -(\omega^2 + O(\eta_0^2))\frac{\iota_j\iota_{j+1}}{\lambda_j}\left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k + (\omega^2 + O(\eta_0^2))\frac{\iota_j\iota_{j-1}}{\lambda_j}\left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k
$$

The second and third terms together will have a sign, up to an acceptable error. First, using (4.40) we have,

$$
\left\langle \underline{A}(\lambda_j)g \mid \mathcal{L}_0 g\right\rangle \geq -\frac{c_0}{\lambda_j}\|g\|_H^2 + \frac{1}{\lambda_j}\int_{R^{-1}\lambda_j}^{R\lambda_j}\left((\partial_r g)^2 + \frac{k^2}{r^2}g^2\right)r\mathrm{d}r
$$

To treat the third term, we start by using the definition (4.58) to observe the identity,

$$
\begin{aligned}
\widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) = {}& -\frac{k^2}{r^2}(f'(Q_{\lambda_j}) - 1)g - \frac{k^2}{r^2}(f'(\mathcal{Q}(m_n, \vec{\iota}_j, \vec{\lambda}_j)) - f'(Q_{\lambda_j}))g \\
& + f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)
\end{aligned}
\tag{4.64}
$$

Next, by definition,

$$
(A(\lambda_j) - \underline{A}(\lambda_j))g = -\frac{1}{\lambda_j}\left(\frac{1}{2}q''(\frac{r}{\lambda_j}) + \frac{\lambda_j}{2r}q'(\frac{r}{\lambda_j})\right)g
$$

The contributions of the second two terms in (4.64) yield acceptable errors. Indeed,

$$
\begin{aligned}
\left|\left\langle (A(\lambda_j) - \underline{A}(\lambda_j))g \mid \frac{k^2}{r^2}(f'(\mathcal{Q}(m_n, \vec{\iota}_j, \vec{\lambda}_j)) - f'(Q_{\lambda_j}))g\right\rangle\right| & \\
\lesssim \frac{1}{\lambda_j}\int_{\widetilde{R}^{-1}\lambda_j}^{\widetilde{R}\lambda_j}g^2\left|f'(\mathcal{Q}(m_n, \vec{\iota}_j, \vec{\lambda}_j)) - f'(Q_{\lambda_j}))\right|\frac{\mathrm{d}r}{r} & \leq c_0\frac{\|g\|_H^2}{\lambda_j}
\end{aligned}
$$

with $\widetilde{R}$ as in Lemma 4.13, and by (4.56) and the definition of q from Lemma 4.13,

$$\left| \left\langle (A(\lambda_j) - \underline{A}(\lambda_j))g \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle \right| \lesssim \frac{1}{\lambda_j} \|g\|_{L^\infty} \|g\|_H^2 \le c_0 \frac{\|g\|_H^2}{\lambda_j}$$

Putting this together we obtain,

$$\left| \langle (A(\lambda_j) - \underline{A}(\lambda_j))g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \rangle - \frac{1}{\lambda_j} \int_0^\infty \left(\frac{1}{2} q''\left(\frac{r}{\lambda_j}\right) + \frac{\lambda_j}{2r} q'\left(\frac{r}{\lambda_j}\right) \right) \frac{k^2}{r^2} (f'(Q_{\lambda_j}) - 1) g^2 \, r \, dr \right|$$
$$\lesssim c_0 \frac{\|g\|_H^2}{\lambda_j}.$$

We show that the remaining terms contribute acceptable errors. For the fourth term a direct calculation gives,

$$\left| \left\langle \Lambda Q_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{Q}} - \mathcal{L}_{\lambda_j})g \right\rangle \right| \lesssim \frac{1}{\lambda_j} \|g\|_H \sum_{i \ne j} (\|r^{-1} \Lambda Q_{\lambda_j} \Lambda Q_{\lambda_i}^2\|_{L^2} + \|r^{-1} \Lambda Q_{\lambda_j}^2 \Lambda Q_{\lambda_i}\|_{L^2})$$
$$\lesssim \frac{1}{\lambda_j} \|g\|_H \left(\left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k + \left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k \right).$$

By (4.41) along with (4.24) we have,

$$\left| \iota_j \frac{\lambda_j'}{\lambda_j} \left\langle \left(\frac{1}{\lambda_j} \underline{\Lambda} - \underline{A}(\lambda_j) \right) \Lambda Q_{\lambda_j} \mid \dot{g} \right\rangle \right| \lesssim \frac{c_0}{\lambda_j} \|g\|_{\mathcal{E}}^2.$$

For the sixth term on the right-hand side of (4.63) we note that

$$A(\lambda_j) \sum_{i=1}^K \iota_i Q_{\lambda_i} = A(\lambda_j) \mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}),$$

and hence we may apply (4.39) with $g_1 = \mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda})$ and $g_2 = g$ to conclude that

$$\left| \left\langle A(\lambda_j) \sum_{i=1}^K \iota_i Q_{\lambda_i} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle + \left\langle A(\lambda_j)g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle \right| \le \frac{c_0}{\lambda_j} \|g\|_H^2,$$

which takes care of the sixth and seventh terms. By (4.42) and (4.56) we see that,

$$\left| \left\langle \left(A(\lambda_j) - \frac{1}{\lambda_j} \Lambda \right) Q_{\lambda_j} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle \right| \lesssim \frac{c_0}{\lambda_j} \|g\|_H^2.$$

Using the first bullet point in Lemma 4.15 and (4.24) we estimate the eighth term as follows,

$$\left| \frac{\lambda_j'}{\lambda_j} \left\langle \lambda_j \partial_{\lambda_j} \underline{A}(\lambda_j)g \mid \dot{g} \right\rangle \right| \lesssim \frac{1}{\lambda_j} \|g\|_{\mathcal{E}}^3 \le \frac{c_0}{\lambda_j} \|g\|_H^2.$$

Next, using (4.43) and (4.56) we have,

$$\left| \sum_{i \ne j} \iota_i \left\langle A(\lambda_j) Q_{\lambda_i} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g) \right\rangle \right| \lesssim \frac{c_0}{\lambda_j} \|g\|_H^2.$$

An application of (4.44) and (4.24) gives

$$\sum_{i \ne j} \left| \lambda_i' \left\langle \underline{A}(\lambda_j) \Lambda Q_{\underline{\lambda_i}} \mid \dot{g} \right\rangle \right| \le \frac{c_0}{\lambda_j} \|g\|_H^2.$$

Next, consider the twelfth term. Using the first bullet point in Lemma 4.15, and in particular the spatial localization of $\underline{A}(\lambda_j)$ we obtain

$$\left|\left\langle \underline{A}(\lambda_j)g \mid f_{\mathbf{i}}(m_n,\iota,\vec{\lambda})\right\rangle\right| \lesssim \|g\|_H\|f_{\mathbf{i}}(m_n,\iota,\vec{\lambda})\|_{L^2(\widetilde{R}^{-1}\lambda_j\leq r\leq\widetilde{R}\lambda_j)}.$$

Using the expansion (2.25) from Lemma 2.23 we have the pointwise estimate,

$$\left|f_{\mathbf{i}}(m_n,\iota,\vec{\lambda})\right| \lesssim r^{-2}\sum_{\ell\neq i}\Lambda Q_{\lambda_\ell}\Lambda Q_{\lambda_i}^2. \tag{4.65}$$

It follows that

$$\|f_{\mathbf{i}}(m_n,\iota,\vec{\lambda})\|_{L^2(\widetilde{R}^{-1}\lambda_j\leq r\leq\widetilde{R}\lambda_j)} \lesssim \frac{1}{\lambda_j}\left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k + \frac{1}{\lambda_j}\left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k.$$

We obtain

$$\left|\left\langle \underline{A}(\lambda_j)g \mid f_{\mathbf{i}}(m_n,\iota,\vec{\lambda})\right\rangle\right| \lesssim \frac{1}{\lambda_j}\|g\|_H\left(\left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k + \frac{1}{\lambda_j}\left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k\right).$$

Finally, we treat the last line of (4.63). First, using Lemma 4.9 and the definition of $\dot{\phi}$ in (4.55) we have

$$\left|\left\langle \Lambda Q_{\underline{\lambda_j}} \mid \dot{\phi}(u,\chi_\nu)\right\rangle\right| \lesssim \frac{1}{\lambda_j}\left(\frac{\lambda_j}{\nu}\right)^k E(\boldsymbol{u}(t);\nu(t),2\nu(t)) \lesssim \frac{\theta_n}{\lambda_j}. \tag{4.66}$$

for some sequence $\theta_n \to 0$ as $n \to \infty$. The last two terms in (4.63) vanish due to the support properties of $\underline{A}(\lambda_j), \phi(u,\nu), \dot{\phi}(u,\nu)$ and the fact that $\lambda_j \leq \lambda_K \ll \nu$.

Combining these estimates in (4.63) we obtain the inequality,

$$\beta' \geq \left(-\iota_j\iota_{j+1}\omega^2 - c_0\right)\frac{1}{\lambda_j}\left(\frac{\lambda_j}{\lambda_{j+1}}\right)^k + \left(\iota_j\iota_{j-1}\omega^2 - c_0\right)\frac{1}{\lambda_j}\left(\frac{\lambda_{j-1}}{\lambda_j}\right)^k$$
$$+ \frac{1}{\lambda_j}\int_{R^{-1}\lambda_j}^{R\lambda_j}\left((\partial_r g)^2 + \frac{k^2}{r^2}g^2\right)r\mathrm{d}r + \frac{1}{\lambda_j}\int_0^\infty \left(\frac{1}{2}q''\left(\frac{r}{\lambda_j}\right) + \frac{\lambda_j}{2r}q'\left(\frac{r}{\lambda_j}\right)\right)\frac{k^2}{r^2}(f'(Q_{\lambda_j})-1)g^2\,r\,\mathrm{d}r$$
$$- c_0\frac{\|\boldsymbol{g}\|_{\mathcal{E}}^2}{\lambda_j} - c_0\frac{\delta_n}{\lambda_j}.$$

where to obtain $c_0\delta_n$ in the last term we enlarged δ_n so as to ensure $\delta_n \gg \theta_n$ in the estimate (4.66). Finally, we use (4.45) on the second line above followed by (4.49) and (4.48) to conclude the proof. $\qquad\square$

Finally, we prove that, again by enlarging ϵ_n, we can control the error in the virial identity, see Lemma 2.4, by $\mathbf{d}$.

Lemma 4.18. *There exist $C_0, \eta_0 > 0$ depending only on k and N and a decreasing sequence $\epsilon_n \to 0$ such that*

$$|\Omega_{\rho(t)}(\boldsymbol{u}(t))| \leq C_0\mathbf{d}(t)$$

for all $t \in [a_n, b_n]$ such that $\epsilon_n \leq \mathbf{d}(t) \leq \eta_0$, $\rho(t) \leq \nu(t)$ and $|\rho'(t)| \leq 1$.

Proof. Since $\lim_{n\to\infty}\sup_{t\in[a_n,b_n]}\|\boldsymbol{u}(t)\|_{\mathcal{E}(\nu(t),2\nu(t))} = 0$, Lemma 4.11 yields

$$\|\boldsymbol{u}(t) - \boldsymbol{\mathcal{Q}}(m_n,\vec{\iota},\vec{\lambda}(t)) - \boldsymbol{g}(t)\|_{\mathcal{E}(0,2\nu(t))} \to 0, \qquad \text{as } n \to \infty.$$

Using Remark 4.17, (4.49) and (4.50) we have $\|\boldsymbol{g}(t)\|_{\mathcal{E}} \lesssim \mathbf{d}(t)$, hence, after choosing $\epsilon_n \to 0$ sufficiently large, it suffices to check that

$$|\Omega_{\rho(t)}(\boldsymbol{\mathcal{Q}}(m_n,\vec{\iota},\vec{\lambda}(t)))| \leq C_0\mathbf{d}(t),$$

which in turn will follow from

$$\int_0^\infty \left| (\partial_r \mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}(t)))^2 - k^2 \frac{\sin^2 \mathcal{Q}(m_n, \vec{\iota}, \vec{\lambda}(t))}{r^2} \right| r \mathrm{d}r \leq C_0 \mathbf{d}(t).$$

Recall that $\Lambda Q_\lambda = r \partial_r Q_\lambda = k \sin Q_\lambda$, so it suffices to estimate the cross terms. It is easy to check that

$$\left| \sin^2 \left(\sum a_i \right) - \sum \sin^2 a_i \right| \leq 4 \sum_{i \neq j} |\sin a_i| |\sin a_j|.$$

Invoking the bound

$$\int_0^\infty |\Lambda Q_\lambda(r) \Lambda Q_\mu(r)| \frac{\mathrm{d}r}{r} \lesssim (\lambda/\mu)^{k/2}$$

from [41, p. 1277], we obtain the claim. $\qquad\square$

5. Conclusion of the proof

5.1. The scale of the K-th bubble.
As mentioned in the Introduction, the K-th bubble is of particular importance. We introduce below a function μ which is well-defined on every $[a_n, b_n]$, and close to λ_K on time intervals where the solution approaches a multi-bubble configuration.

Definition 5.1 (The scale of the K-th bubble). For all $t \in I$, we set

$$\mu(t) := \sup \left\{ r : E(\boldsymbol{u}(t); r) = (N - K + 1/2) E(\boldsymbol{Q}) + E(\boldsymbol{u}^*) \right\}.$$

Note that $K > 0$ implies $0 < (N - K + 1/2) E(\boldsymbol{Q}) + E(\boldsymbol{u}^*) < E(\boldsymbol{u})$, hence $\mu(t)$ is a well-defined finite positive number for all $t \in I$. By Lemma 4.9,

$$\lim_{n \to \infty} \sup_{t \in [a_n, b_n]} \left| E(\boldsymbol{u}(t); \nu(t)) - (N - K) E(\boldsymbol{Q}) - E(\boldsymbol{u}^*) \right| = 0,$$

which implies $\mu(t) \leq \nu(t)$ for all n large enough and $t \in [a_n, b_n]$, thus $\mu(t) \ll \mu_{K+1}(t)$ as $n \to \infty$.

In the next lemma, we relate the localized distance function $\boldsymbol{\delta}_R$, defined by (3.1), computed at any scale $R \gg \mu(t)$, with the global distance function $\mathbf{d}(t)$.

Lemma 5.2. *The function μ defined above has the following properties:*
  (i) *its Lipschitz constant is ≤ 1,*
  (ii) *for any $\epsilon > 0$ there exist $0 < \delta \leq \eta_0$ and $n_0 \in \mathbb{N}$ such that $t \in [a_n, b_n]$ with $n \geq n_0$ and $\mathbf{d}(t) \leq \delta$ imply $|\mu(t)/\lambda_K(t) - 1| \leq \epsilon$, where $\lambda_K(t)$ is the modulation parameter defined in Lemma 4.11,*
  (iii) *if $(t_n)_n$ and $(r_n)_n$ are any sequences such that $t_n \in [a_n, b_n]$ for all n, $1 \ll r_n \ll \mu_{K+1}(t_n)/\mu(t_n)$ and $\lim_{n \to \infty} \boldsymbol{\delta}_{r_n \mu(t_n)}(t_n) = 0$, then $\lim_{n \to \infty} \mathbf{d}(t_n) = 0$.*

Proof. Let $s, t \in I$. We prove that $|\mu(s) - \mu(t)| \leq |s - t|$. Assume, without loss of generality, $\mu(t) \geq \mu(s)$. Of course, we can also assume $\mu(t) > |s - t|$. By (2.2),

$$E(\boldsymbol{u}(s); \mu(t) - |s - t|) \geq E(\boldsymbol{u}(t); \mu(t)) = (N - K + 1/2) E(\boldsymbol{Q}) + E(\boldsymbol{u}^*),$$

which implies $\mu(s) \geq \mu(t) - |s - t|$.

In order to prove (ii), it suffices to check that

$$E(\boldsymbol{u}(t); (1 + \epsilon)\lambda_K(t)) < (N - K + 1/2) E(\boldsymbol{Q}) + E(\boldsymbol{u}^*),$$
$$E(\boldsymbol{u}(t); (1 - \epsilon)\lambda_K(t)) > (N - K + 1/2) E(\boldsymbol{Q}) + E(\boldsymbol{u}^*).$$

By (4.26), this will follow from

$$E(\boldsymbol{u}(t); (1+\epsilon)\lambda_K(t), \nu(t)) < E(\boldsymbol{Q})/2,$$
$$E(\boldsymbol{u}(t); (1-\epsilon)\lambda_K(t), \nu(t)) > E(\boldsymbol{Q})/2.$$

We use (4.19). By (4.21), $\|\boldsymbol{g}\|_{\mathcal{E}} \ll 1$ when $\delta \ll 1$ and $n_0 \gg 1$. Thus, it suffices to see that

$$E(\boldsymbol{\mathcal{Q}}(m_n, \vec{\iota}, \vec{\lambda}); (1+\epsilon)\lambda_K(t), \nu(t)) < E(\boldsymbol{Q})/2,$$
$$E(\boldsymbol{\mathcal{Q}}(m_n, \vec{\iota}, \vec{\lambda}); (1-\epsilon)\lambda_K(t), \nu(t)) > E(\boldsymbol{Q})/2$$

whenever $\sum_{j=1}^{K} \lambda_j(t)/\lambda_{j+1}(t) \ll 1$, which is obtained directly from the definition of $\boldsymbol{\mathcal{Q}}$.

We now prove (iii). Let R_n be a sequence such that $r_n\mu(t_n) \ll R_n \ll \mu_{K+1}(t_n)$. Without loss of generality, we can assume $R_n \geq \nu(t_n)$, since it suffices to replace R_n by $\nu(t_n)$ for all n such that $R_n < \nu(t_n)$. Let $M_n, m_n, \vec{\iota}_n, \vec{\lambda}_n$ be parameters such that

$$\|u(t_n) - \mathcal{Q}(m_n, \vec{\iota}_n, \vec{\lambda}_n)\|^2_{H(r \leq r_n\mu(t_n))} + \|\dot{u}(t_n)\|^2_{L^2(r \leq r_n\mu(t_n))} + \sum_{j=1}^{M-1} \left(\frac{\lambda_{n,j}}{\lambda_{n,j+1}}\right)^k \to 0, \qquad (5.1)$$

which exist by the definition of the localized distance function (3.1). Since

$$\left(K - \frac{1}{2}\right)E(\boldsymbol{Q}) \leq \liminf_{n\to\infty} E(\boldsymbol{u}(t_n); 0, r_n\mu(t_n)) \leq \limsup_{n\to\infty} E(\boldsymbol{u}(t_n); 0, r_n\mu(t_n)) \leq K E(\boldsymbol{Q}),$$

we have $M_n = K$ for n large enough. We set $\lambda_{n,j} := \mu_j(t_n)$ and $\iota_{n,j} := \sigma_j$ for $j > K$. We claim that

$$\lim_{n\to\infty} \left(\|\boldsymbol{u}(t) - \boldsymbol{u}^*(t) - \boldsymbol{\mathcal{Q}}(m_\Delta, \vec{\iota}_n, \vec{\lambda}_n)\|^2_{\mathcal{E}} + \sum_{j=1}^{N} \left(\frac{\lambda_{n,j}}{\lambda_{n,j+1}}\right)^k\right) = 0.$$

By the definition of $\mathbf{d}$, the proof will be finished. First, we observe that $\lambda_{n,K} \ll r_n\mu(t_n)$, so $\lambda_{n,K}/\lambda_{n,K+1} \to 0$. In the region $r \leq r_n\mu(t_n)$, convergence follows from (5.1), since the energy of the exterior bubbles asymptotically vanishes there. In the region $r \geq R_n$, the energy of the interior bubbles vanishes, hence it suffices to apply Lemma 4.9 and recall that $R_n \geq \nu(t_n)$. In particular

$$\lim_{n\to\infty} E(\boldsymbol{u}(t_n); 0, r_n\mu(t_n)) = K E(\boldsymbol{Q}), \qquad \lim_{n\to\infty} E(\boldsymbol{u}(t_n); R_n) = (N-K)E(\boldsymbol{Q}) + E(\boldsymbol{u}^*),$$

which implies

$$\lim_{n\to\infty} E(\boldsymbol{u}(t_n); r_n\mu(t_n), R_n) = 0,$$

and (2.1) yields convergence of the error also in the region $r_n\mu(t_n) \leq r \leq R_n$. $\qquad \square$

Our next goal is to prove that the minimality of K (see Definition 4.5) implies a lower bound on the length of the collision intervals. First, we have the following fact.

Lemma 5.3. *If $m_n \in \mathbb{Z}$, $\iota_n \in \{-1, 1\}$, $0 < r_n \ll \mu_n \ll R_n$, $0 < t_n \ll \mu_n$ and $\boldsymbol{u}_n$ a sequence of solutions of (1.1) such that $\boldsymbol{u}_n(t)$ is defined for $t \in [0, t_n]$ and*

$$\lim_{n\to\infty} \|\boldsymbol{u}_n(0) - (m_n\boldsymbol{\pi} + \iota_n \boldsymbol{Q}_{\mu_n})\|_{\mathcal{E}(r_n, R_n)} = 0,$$

then

$$\lim_{n\to\infty} \sup_{t \in [0, t_n]} \|\boldsymbol{u}_n(t) - (m_n\boldsymbol{\pi} + \iota_n \boldsymbol{Q}_{\mu_n})\|_{\mathcal{E}(r_n+t, R_n-t)} = 0.$$

Proof. Without loss of generality, we can assume $m_n = 0$, $\iota_n = 1$ and $\mu_n = 1$. After these reductions, the conclusion directly follows from [9, Lemma 3.4]. $\qquad \square$

Lemma 5.4. *If $\eta_1 > 0$ is small enough, then for any $\eta \in (0, \eta_1]$ there exist $\epsilon \in (0, \eta)$ and $C_{\boldsymbol{u}} > 0$ having the following property. If $[c, d] \subset [a_n, b_n]$, $\mathbf{d}(c) \leq \epsilon$, $\mathbf{d}(d) \leq \epsilon$ and there exists $t_0 \in [c, d]$ such that $\mathbf{d}(t_0) \geq \eta$, then*

$$d - c \geq C_{\boldsymbol{u}}^{-1} \max(\mu(c), \mu(d)).$$

Proof. We argue by contradiction. If the statement is false, then there exist $\eta > 0$, a decreasing sequence (ϵ_n) tending to 0, an increasing sequence (C_n) tending to ∞ and intervals $[c_n, d_n] \subset [a_n, b_n]$ (up to passing to a subsequence in the sequence of the collision intervals $[a_n, b_n]$) such that $\mathbf{d}(c_n) \leq \epsilon_n$, $\mathbf{d}(d_n) \leq \epsilon_n$, there exists $t_n \in [c_n, d_n]$ such that $\mathbf{d}(t_n) \geq \eta$ and $d_n - c_n \leq C_n^{-1} \max(\mu(c_n), \mu(d_n))$. We will check that, up to adjusting the sequence ϵ_n, $[c_n, d_n] \in \mathcal{C}_{K-1}(\epsilon_n, \eta)$ for all n, contradicting Definition 4.5.

The first and second requirement in Definition 4.4 are clearly satisfied. It remains to construct a function $\rho_{K-1} : [c_n, d_n] \to [0, \infty)$ such that

$$\lim_{n \to \infty} \sup_{t \in [c_n, d_n]} \mathbf{d}_{K-1}(t; \rho_{K-1}(t)) = 0. \tag{5.2}$$

Assume $\mu(c_n) \geq \mu(d_n)$ (the proof in the opposite case is very similar). Let r_n be a sequence such that $\lambda_{K-1}(c_n) \ll r_n \ll \lambda_K(c_n)$ (recall that $\lambda_K(c_n)$ is at main order equal to $\mu(c_n)$ and that $\lambda_0(t) = 0$ by convention). Set $\rho_{K-1}(t) := r_n + (t - c_n)$ for $t \in [c_n, d_n]$. Recall that $\vec{\sigma}_n \in \{-1, 1\}^{N-K}$ and $\vec{\mu}(t) \in (0, \infty)^{N-K}$ are defined in Lemma 4.9. Let ι_n be the sign of the K-th bubble at time c_n, and set $\widetilde{\sigma} := (\iota_n, \vec{\sigma}_n) \in \{-1, 1\}^{N-(K-1)}$ and $\widetilde{\mu}(t) := (\mu(c_n), \vec{\mu}(t)) \in (0, \infty)^{N-(K-1)}$. Let R_n be a sequence such that $\nu_n(c_n) \ll R_n \ll \mu_{K+1}(c_n)$. Applying Lemma 5.3 with these sequences r_n, R_n and $\boldsymbol{u}_n(t) := \boldsymbol{u}(c_n + t)$, we obtain

$$\lim_{n \to \infty} \sup_{t \in [c_n, d_n]} \|\boldsymbol{u}(t) - \boldsymbol{\mathcal{Q}}(m_\Delta, \widetilde{\sigma}_n, \widetilde{\mu}(t))\|_{\mathcal{E}(\rho_{K-1}(t), \infty)} = 0,$$

implying (5.2). $\qquad\square$

Remark 5.5. We denote the constant $C_{\boldsymbol{u}}$ to stress that it depends on the solution $\boldsymbol{u}$ and is obtained in a non-constructive way as a consequence of the assumption that $\boldsymbol{u}$ does not satisfy the continuous time soliton resolution.

5.2. **Demolition of the multi-bubble.** Recall the following notion from Real Analysis. If $X \subset \mathbb{R}$, $U : X \to \mathbb{R} \cup \{+\infty\}$ and $t_0 \in X$, we say that t_0 is a *local minimal point from the right* if there exists $t_1 > t_0$ such that $U(t_0) \leq U(t)$ for all $t \in X \cap (t_0, t_1)$. Similarly, we say that t_0 is a *local minimal point from the left* if there exists $t_1 < t_0$ such that $U(t_0) \leq U(t)$ for all $t \in X \cap (t_1, t_0)$.

Definition 5.6 (Weighted interaction energy). On each collision interval $[a_n, b_n]$, we define the function $U : [a_n, b_n] \to \overline{\mathbb{R}}_+$ as follows:

- if $\mathbf{d}(t) \geq \eta_0$, then $U(t) := +\infty$.
- if $\mathbf{d}(t) < \eta_0$, then $U(t) := \max_{i \in \mathcal{A}} \left(2^{-i} \xi_i(t) / \lambda_{i+1}(t)\right)^k$, where λ_{i+1} and ξ_i are the modulation parameter and its refinement defined above, see Lemma 4.16.

Remark 5.7. Continuity of $\mathbf{d}$, ξ_i and λ_i implies that U is finite and continuous in a neighborhood of any point where it is finite. Moreover, applying (4.51) and (4.50), we obtain that there exists $C_1 > 0$ such that

$$\mathbf{d}(t) < \eta_0 \quad \text{implies} \quad C_1^{-1} \mathbf{d}(t)^2 \leq U(t) \leq C_1 \mathbf{d}(t)^2. \tag{5.3}$$

Lemma 5.8. *Let $k \geq 2$. If η_0 is small enough, then there exists $C_0 \geq 0$ depending only on k and N such that the following is true. If t_0 is a local minimal point from the right of U such that $U(t_0) < +\infty$ and $t_* \geq t_0$ is such that $U(t) < \infty$ for all $t \in [t_0, t_*]$, then*

$$\frac{3}{4}\lambda_K(t_0) \leq \lambda_K(t_*) \leq \frac{4}{3}\lambda_K(t_0), \tag{5.4}$$

$$\int_{t_0}^{t_*} \mathbf{d}(t)\mathrm{d}t \leq C_0\mathbf{d}(t_*)^{\frac{2}{k}}\lambda_K(t_0). \tag{5.5}$$

An analogous statement is true if t_* is a local minimal point from the left.

Remark 5.9. Since $\mathbf{d}(t_0) \leq \eta_0$ and $\mathbf{d}(t_*) \leq \eta_0$ are small, λ_K differs from μ by a small relative error, so in the formulation of the lemma we could just as well write μ instead of λ_K.

Proof of Lemma 5.8. **Step 1.** We can assume $t_* > t_0$. For $j \in \mathcal{A}$, denote $\widetilde{\xi}_j(t) := 2^{-j}\xi_j(t)/\lambda_{j+1}(t)$ and let

$$\mathcal{A}_0 := \{j \in \mathcal{A} : U(t_0) = \widetilde{\xi}_j(t_0)^k\} = \{j \in \mathcal{A} : \widetilde{\xi}_j(t_0) = \max_{i \in \mathcal{A}} \widetilde{\xi}_i(t_0)\},$$

$$\widetilde{\mathcal{A}}_0 := \{j \in \mathcal{A}_0 : \widetilde{\xi}_j'(t_0) \geq 0\}.$$

Since t_0 is a local minimal point from the right of U, $\widetilde{\mathcal{A}}_0 \neq \emptyset$. Let $j_0 := \min \widetilde{\mathcal{A}}_0 \in \mathcal{A}$.

We now define by induction a sequence of times $t_0 \leq t_1 \leq \ldots \leq t_{l_*} = t_*$ and a sequence of elements of $\mathcal{A}$, $j_0 > j_1 > \ldots > j_{l_*-1}$, in the following way. Assume $t_0 \leq t_1 \leq \ldots \leq t_{l-1}$ and $j_0 > j_1 > \ldots > j_{l-1}$ are already defined. We set

$$t_l := \sup\left\{t \in [t_{l-1}, t_*] : \widetilde{\xi}_j(\tau) \leq \widetilde{\xi}_{j_{l-1}}(\tau) \text{ for all } \tau \in [t_{l-1}, t) \text{ and } j \in \mathcal{A} \text{ such that } j < j_{l-1}\right\}.$$

If $t_l = t_*$, then we set $l_* := l$ and terminate the procedure. If not, let

$$\mathcal{A}_l := \{j \in \mathcal{A} : j < j_{l-1} \text{ and } \widetilde{\xi}_j(t_l) = \widetilde{\xi}_{j_{l-1}}(t_l)\}.$$

By the definition of t_l and continuity, $\mathcal{A}_l \neq \emptyset$. We set $j_l := \min \mathcal{A}_l$.
Step 2. We check that $t_l > t_{l-1}$ for $l = 1, \ldots, l_*$.

In order to prove that $t_1 > t_0$, we need to show that there exists $t > t_0$ such that $\widetilde{\xi}_j(\tau) \leq \widetilde{\xi}_{j_0}(\tau)$ for all $\tau \in [t_0, t)$ and $j \in \mathcal{A}$ such that $j < j_0$. Since $\mathcal{A}$ is a finite set, it suffices to check this separately for each $j \in \mathcal{A}$. If $j \notin \mathcal{A}_0$, the claim is clear, by continuity. If $j \in \mathcal{A}_0 \setminus \widetilde{\mathcal{A}}_0$, then $\widetilde{\xi}_j(t_0) = \widetilde{\xi}_{j_0}(t_0)$, $\widetilde{\xi}_j'(t_0) < 0$ and $\widetilde{\xi}_{j_0}'(t_0) \geq 0$, again implying the claim.

For $l \geq 1$, the definition of j_l implies that $\widetilde{\xi}_j(t_l) < \widetilde{\xi}_{j_l}(t_l)$ for all $j < j_l$. Writing $l - 1$ instead of l, we get $\widetilde{\xi}_j(t_{l-1}) < \widetilde{\xi}_{j_{l-1}}(t_{l-1})$ for all $j < j_{l-1}$, whenever $l \geq 2$. Thus, by continuity, $t_l > t_{l-1}$.
Step 3. By induction with respect to l, we show that there exists a constant C_0 depending only on k and N such that for all $l \in \{1, \ldots, l_*\}$ we have

$$\int_{t_{l-1}}^{t_l} \mathbf{d}(t)\mathrm{d}t \leq C_0\mathbf{d}(t_l)^{\frac{2}{k}}\lambda_{j_{l-1}+1}(t_{l-1}), \tag{5.6}$$

$$\widetilde{\xi}_{j_{l-1}}(t) \geq \frac{1}{2}\widetilde{\xi}_j(t), \qquad \text{for all } t \in (t_{l-1}, t_l) \text{ and } j > j_{l-1}.$$

Suppose (5.6) is proved for $l \in \{1, \dots, l_0\}$ and let $T_0 \in (t_{l_0}, t_{l_0+1}]$ be the largest number such that

$$\int_{t_{l_0}}^{T_0} \mathbf{d}(t)\mathrm{d}t \leq 2C_0 \mathbf{d}(T_0)^{\frac{2}{k}} \lambda_{j_{l_0}+1}(t_{l_0}), \tag{5.7}$$

$$\widetilde{\xi}_{j_{l_0}}(t) \geq \frac{1}{4}\widetilde{\xi}_j(t), \qquad \text{for all } t \in (t_{l_0}, T_0) \text{ and } j > j_{l_0}, \tag{5.8}$$

$$\xi_{j_{l_0}}(t) \geq \frac{3}{4}\xi_{j_{l_0}}(t_{l_0}), \qquad \text{for all } t \in (t_{l_0}, T_0). \tag{5.9}$$

It suffices to prove that

$$\int_{t_{l_0}}^{T_0} \mathbf{d}(t)\mathrm{d}t \leq C_0 \mathbf{d}(T_0)^{\frac{2}{k}} \lambda_{j_{l_0}+1}(t_{l_0}), \tag{5.9}$$

$$\widetilde{\xi}_{j_{l_0}}(t) \geq \frac{1}{2}\widetilde{\xi}_j(t), \qquad \text{for all } t \in (t_{l_0}, T_0) \text{ and } j > j_{l_0}, \tag{5.10}$$

$$\xi_{j_{l_0}}(t) \geq \frac{7}{8}\xi_{j_{l_0}}(t_{l_0}), \qquad \text{for all } t \in (t_{l_0}, T_0). \tag{5.11}$$

It will be convenient to assume $T_0 = t_* = t_{l_0+1}$, which is allowed. Also, in order to simplify the notation, we write l instead of l_0 in the induction step which follows.

The first observation is that if $j_l < j \leq j_{l-1}$, then $\lambda_j(t)$ is "almost constant" on the time interval (t_l, t_{l+1}). More precisely, we claim that

$$|\lambda_j(t)/\lambda_j(t_l) - 1| \leq c_0, \qquad \text{if } j > j_l \text{ and } t > t_l, \tag{5.12}$$

where c_0 can be made arbitrarily small by taking η_0 small enough. Indeed, $|\lambda_j'(t)| \lesssim \mathbf{d}(t)$, so (5.7) implies the claim (we stress again that C_0 will not depend on η_0).

The definition of t_{l+1} implies

$$\widetilde{\xi}_{j_l}(t) \geq \widetilde{\xi}_j(t), \qquad \text{for all } t \in (t_l, t_{l+1}) \text{ and } j \leq j_l, \tag{5.13}$$

so (5.8) yields

$$\max_{i \in \mathcal{A}} \widetilde{\xi}_i(t) \lesssim \widetilde{\xi}_{j_l}(t), \qquad \text{for all } t \in (t_l, t_{l+1}). \tag{5.14}$$

The bound (4.53) yields for all $t \in (t_l, t_{l+1})$

$$\lambda_{j_l}(t)\beta_{j_l}'(t) \geq \left(-\iota_{j_l}\iota_{j_l+1}\omega^2 - c_0\right)\left(2^{j_l}\widetilde{\xi}_{j_l}(t)\right)^k + \left(\iota_{j_l}\iota_{j_l-1}\omega^2 - c_0\right)\left(2^{j_l-1}\widetilde{\xi}_{j_l-1}(t)\right)^k$$
$$- c_0 \max_{i \in \mathcal{A}}\left(2^i\widetilde{\xi}_i(t)\right)^k,$$

with the convention $\xi_0(t) = 0$. By (5.13), $2^{j_l-1}\widetilde{\xi}_{j_l-1}(t) \leq \frac{1}{2}2^{j_l}\widetilde{\xi}_{j_l}(t)$. Taking c_0 small enough and applying (5.14), we obtain

$$\lambda_{j_l}(t)\beta_{j_l}'(t) \geq \frac{\omega^2}{4}\left(2^{j_l}\widetilde{\xi}_{j_l}(t)\right)^k \quad \Rightarrow \quad \beta_{j_l}'(t) \geq c_1 \frac{\xi_{j_l}(t)^{k-1}}{\lambda_{j_l+1}(t_l)^k}, \tag{5.15}$$

where $c_1 > 0$ depends only on k and N, and in the last step we used (5.12).

With $c_2 > 0$ to be determined, consider the auxiliary function

$$\phi(t) := \beta_{j_l}(t) + c_2\left(\xi_{j_l}(t)/\lambda_{j_l+1}(t_l)\right)^{\frac{k}{2}}.$$

The Chain Rule gives

$$\phi'(t) = \beta_{j_l}'(t) + c_2\frac{k}{2}\lambda_{j_l+1}(t_l)^{-\frac{k}{2}}\xi_{j_l}(t)^{\frac{k}{2}-1}\xi_{j_l}'(t).$$

We have $|\xi'_{j_l}(t)| \leq c_3(\xi_{j_l}(t)/\lambda_{j_l+1}(t_l))^{\frac{k}{2}}$, with c_3 depending only on k and N, hence (5.15) implies

$$\phi'(t) \geq \frac{c_3}{\lambda_{j_l+1}(t_l)}\left(\frac{\xi_{j_l}(t)}{\lambda_{j_l+1}(t_l)}\right)^{k-1} \geq \frac{c_4}{\lambda_{j_l+1}(t_l)}\phi(t)^{\frac{2k-2}{k}},$$

with c_2, c_3, c_4 depending only on k and N. The last inequality yields

$$\left(\lambda_{j_l+1}(t_l)\phi(t)^{2/k}\right)' \gtrsim \phi(t) \quad \Rightarrow \quad \int_{t_l}^{t_{l+1}} \phi(t)\mathrm{d}t \lesssim \lambda_{j_l+1}(t_l)\phi(t_{l+1})^{2/k} \lesssim \mathbf{d}(t_{l+1})^{2/k}\lambda_{j_l+1}(t_l). \tag{5.16}$$

If we consider $\widetilde{\phi}(t) := \beta_{j_l}(t) + \frac{c_2}{2}\left(\xi_{j_l}(t)/\lambda_{j_l+1}(t_l)\right)^{\frac{k}{2}}$ instead of ϕ, then the computation above shows that $\widetilde{\phi}$ is increasing. From (5.19), we have $\widetilde{\phi}(t_l) \geq 0$, so $\widetilde{\phi}(t) \geq 0$ for all $t \in (t_l, t_{l+1})$, implying $\mathbf{d}(t) \lesssim \phi(t)$. Thus, (5.16) yields (5.9) if C_0 is sufficiently large (but depending on k and N only).

We now prove (5.11). By the definition of t_l and the fact that $j_l < j_{l-1}$, we have $\widetilde{\xi}_{j_l}(\tau) \leq \widetilde{\xi}_{j_{l-1}}(\tau)$ for all $\tau \in [t_{l-1}, t_l]$. By the definition of j_l, $\widetilde{\xi}_{j_l}(t_l) = \widetilde{\xi}_{j_{l-1}}(t_l)$, in particular we have

$$\widetilde{\xi}'_{j_l}(t_l) \geq \widetilde{\xi}'_{j_{l-1}}(t_l). \tag{5.17}$$

Recalling that $\widetilde{\xi}_j(t) = 2^{-j}\xi_j(t)/\lambda_{j+1}(t)$, we find

$$2^{-j_l}\frac{\xi'_{j_l}(t_l)}{\lambda_{j_l+1}(t_l)} - 2^{-j_l}\frac{\xi_{j_l}(t_l)\lambda'_{j_l+1}(t_l)}{\lambda_{j_l+1}(t_l)^2} \geq 2^{-j_{l-1}}\frac{\xi'_{j_{l-1}}(t_l)}{\lambda_{j_{l-1}+1}(t_l)} - 2^{-j_{l-1}}\frac{\xi_{j_{l-1}}(t_l)\lambda'_{j_{l-1}+1}(t_l)}{\lambda_{j_{l-1}+1}(t_l)^2}. \tag{5.18}$$

Since $\lambda_{j_l+1}(t_l)/\lambda_{j_{l-1}+1}(t_l) + \xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l) + \xi_{j_{l-1}}(t_l)/\lambda_{j_{l-1}+1}(t_l)$ is small when η_0 is small and, see Lemma 4.16,

$$|\lambda'_{j_l+1}(t_l)| + |\xi'_{j_{l-1}}(t_l)| + |\lambda'_{j_{l-1}+1}(t_l)| \lesssim \max_{i \in \mathcal{A}}(\xi_i/\lambda_{i+1})^{k/2},$$

we obtain

$$\xi'_{j_l}(t_l) \geq -c_0 \max_{i \in \mathcal{A}}(\xi_i/\lambda_{i+1})^{k/2}, \tag{5.19}$$

where c_0 can be made arbitrarily small upon taking η_0 small.

By (5.19) and (4.52), we have $\beta_{j_l}(t_l) \geq -c_0(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))^{\frac{k}{2}}$, where c_0 can be made as small as needed, and

$$\beta'_{j_l}(t) \geq (3/4)^{k-1}c_1\frac{\xi_{j_l}(t_l)^{k-1}}{\lambda_{j_l+1}(t_l)^k}.$$

We deduce that $\xi'_{j_l}(t) \geq 0$ provided

$$t - t_l \geq \frac{2c_0}{c_1}(4/3)^{k-1}\xi_{j_l}(t_l)(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))^{-\frac{k}{2}}.$$

But, if the opposite inequality is satisfied, the bound $|\xi'_{j_l}(t)| \lesssim (\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))^{\frac{k}{2}}$ yields (5.11), if c_0 is small enough. In fact, the argument gives the bound with $\frac{7}{8}$ replaced by $1 - c_0$, where $c_0 > 0$ is as small as we want. Combining this with (5.12), we obtain in particular

$$\widetilde{\xi}_{j_l}(t_{l+1}) \geq (1 - c_0)\widetilde{\xi}_{j_l}(t_l), \tag{5.20}$$

with $c_0 > 0$ arbitrarily small.

Finally, we prove (5.10). By (5.11) and (5.12), it suffices to show that

$$\widetilde{\xi}_{j_l}(t_l) \geq \frac{3}{4}\widetilde{\xi}_j(t_l), \qquad \text{for all } j > j_l.$$

Let $l' \leq l$ be such that $j_{l'} < j \leq j_{l'-1}$. The definition of $j_{l'}$ yields $\widetilde{\xi}_{j_{l'}}(t_{l'}) \geq \widetilde{\xi}_j(t_{l'})$, so it suffices to check that

$$\widetilde{\xi}_{j_{\widetilde{l}}}(t_{\widetilde{l}}) \geq (3/4)^{\frac{1}{K}} \widetilde{\xi}_{j_{\widetilde{l}-1}}(t_{\widetilde{l}-1}), \qquad \text{for all } \widetilde{l},$$

with $c_0 > 0$ small, and use this inequality $l - l'$ times. The last inequality follows from (5.17) and (5.20).

Step 4. Taking the sum over l of (5.9), we get (5.5). The bound (5.4) follows from (5.12). $\square$

Starting from now, $\eta_0 > 0$ is fixed so that Lemma 5.8 holds and Lemma 5.4 can be applied with $\eta = \eta_0$. We also fix $\epsilon > 0$ to be the value given by Lemma 5.4 for $\eta = \eta_0$. Recall that $\mathbf{d}(a_n) = \mathbf{d}(b_n) = \epsilon_n$ and $\mathbf{d}(t) \geq \epsilon_n$ for all $t \in [a_n, b_n]$.

Lemma 5.10. *There exists $\theta_0 > 0$ such that for any sequence satisfying $\epsilon_n \ll \theta_n \leq \theta_0$ and for all n large enough there exist $N_n \in \mathbb{N}^*$ and a partition of the interval $[a_n, b_n]$*

$$a_n = e_{n,0}^L \leq e_{n,0}^R \leq c_{n,0}^R \leq d_{n,0}^R \leq f_{n,0}^R \leq f_{n,1}^L \leq d_{n,1}^L \leq c_{n,1}^L \leq e_{n,1}^L \leq \ldots \leq e_{n,N_n}^R = b_n,$$

having the following properties.

(1) For all $m \in \{0, 1, \ldots, N_n\}$ and $t \in [e_{n,m}^L, e_{n,m}^R]$, $\mathbf{d}(t) \leq \eta_0$, and

$$\int_{e_{n,m}^L}^{e_{n,m}^R} \mathbf{d}(t)\mathrm{d}t \leq C_2 \theta_n^{2/k} \min(\mu(e_{n,m}^L), \mu(e_{n,m}^R)), \tag{5.21}$$

where $C_2 \geq 0$ depends only on k and N.

(2) For all $m \in \{0, 1, \ldots, N_n - 1\}$ and $t \in [e_{n,m}^R, c_{n,m}^R] \cup [f_{n,m}^R, f_{n,m+1}^L] \cup [c_{n,m+1}^L, e_{n,m+1}^L]$, $\mathbf{d}(t) \geq \theta_n$.

(3) For all $m \in \{0, 1, \ldots, N_n - 1\}$ and $t \in [c_{n,m}^R, f_{n,m}^R] \cup [f_{n,m+1}^L, c_{n,m+1}^L]$, $\mathbf{d}(t) \geq \epsilon$.

(4) For all $m \in \{0, 1, \ldots, N_n - 1\}$, $\mathbf{d}(d_{n,m}^R) \geq \eta_0$ and $\mathbf{d}(d_{n,m+1}^L) \geq \eta_0$.

(5) For all $m \in \{0, 1, \ldots, N_n - 1\}$, $\mathbf{d}(c_{n,m}^R) = \mathbf{d}(c_{n,m+1}^L) = \epsilon$.

(6) For all $m \in \{0, 1, \ldots, N_n - 1\}$, either $\mathbf{d}(t) \geq \epsilon$ for all $t \in [c_{n,m}^R, c_{n,m+1}^L]$, or $\mathbf{d}(f_{n,m}^R) = \mathbf{d}(f_{n,m+1}^L) = \epsilon$.

(7) For all $m \in \{0, 1, \ldots, N_n - 1\}$,

$$\sup_{t \in [e_{n,m}^L, c_{n,m}^R]} \mu(t) \Big/ \inf_{t \in [e_{n,m}^L, c_{n,m}^R]} \mu(t) \leq 2,$$

$$\sup_{t \in [c_{n,m+1}^L, e_{n,m+1}^R]} \mu(t) \Big/ \inf_{t \in [c_{n,m+1}^L, e_{n,m+1}^R]} \mu(t) \leq 2.$$

Remark 5.11. The purpose of the Lemma 5.10 above is to partition any collision interval $[a_n, b_n]$ into subintervals, depending on the values of $\mathbf{d}(t)$. On the intervals $[e_{n,j}^L, e_{n,j}^R]$, the bound (5.21) will always be invoked. Outside of these intervals, the lower bounds on $\mathbf{d}(t)$, combined with a more or less direct application of Lemma 3.1, will be used. Of special importance are the intervals $[c_{n,j}^R, f_{n,j}^R]$ and $[f_{n,j}^L, c_{n,j}^L]$, since they allow to apply Lemma 5.4, leading to the crucial bound (5.32).

Lemma 5.4 could not be applied directly on the intervals $[e_{n,j}^R, e_{n,j+1}^L]$, because there is no uniform (independent of n) lower bound on $\mathbf{d}(t)$ on these intervals, unless $\theta_n \gtrsim 1$. But, in our application of (5.21) in the proof of Theorem 1, it will be necessary to have $\theta_n \to 0$, see (5.33). For this reason, the contribution of the intervals $[e_{n,j}^R, c_{n,j}^R]$, $[f_{n,j}^R, f_{n,j+1}^L]$ and $[c_{n,j+1}^L, e_{n,j+1}^L]$ will be estimated differently, see (5.29)–(5.31).

The intervals $[c_{n,j}^R, f_{n,j}^R]$ and $[f_{n,j}^L, c_{n,j}^L]$ correspond to what were called "compactness intervals" $[c_n, d_n]$ in the summary of the proof in Section 1.4, whereas $[e_{n,j}^L, e_{n,j}^R]$ correspond to the "modulation intervals". For simplicity, we have not mentioned the remaining intervals in Section 1.4.

They are the "intermediate intervals," on which we will use both Lemma 3.1 and estimates of the modulation parameters leading to the property (7) above.

A part of Lemma 5.10 will also play a role in the construction of an appropriate time-dependent cut-off radius for the virial identity, in Lemma 5.12 below.

Proof of Lemma 5.10. For all $t_0 \in [a_n, b_n]$ such that $U(t_0) < \infty$, let $J(t_0) \subset [a_n, b_n]$ be the union of all the open (relatively in $[a_n, b_n]$) intervals containing t_0 on which U is finite. Equivalently, we have one of the following three cases:

- $J(t_0) = (\widetilde{a}_n, \widetilde{b}_n)$, $t_0 \in (\widetilde{a}_n, \widetilde{b}_n)$, $\mathbf{d}(\widetilde{a}_n) = \mathbf{d}(\widetilde{b}_n) = \eta_0$ and $\mathbf{d}(t) < \eta_0$ for all $t \in (\widetilde{a}_n, \widetilde{b}_n)$,
- $J(t_0) = [a_n, \widetilde{b}_n)$, $t_0 \in [a_n, \widetilde{b}_n)$, $\mathbf{d}(\widetilde{b}_n) = \eta_0$ and $\mathbf{d}(t) < \eta_0$ for all $t \in [a_n, \widetilde{b}_n)$,
- $J(t_0) = (\widetilde{a}_n, b_n]$, $t_0 \in (\widetilde{a}_n, b_n]$, $\mathbf{d}(\widetilde{a}_n) = \eta_0$ and $\mathbf{d}(t) < \eta_0$ for all $t \in (\widetilde{a}_n, b_n]$.

Note that $\theta_n \gg \epsilon_n$ implies $\widetilde{a}_n > a_n$ and $\widetilde{b}_n < b_n$. Clearly, any two such intervals are either equal or disjoint.

Consider the set

$$A := \{t \in [a_n, b_n] : \mathbf{d}(t) \leq \theta_n\}.$$

Observe that $\{a_n, b_n\} \in A$. Since A is a compact set, there exist $N_n \in \mathbb{N}$ and a sequence

$$a_n \leq s_{n,0} < s_{n,1} < \ldots < s_{n,N_n} \leq b_n$$

such that

$$s_{n,m} \in A, \qquad A \subset \bigcup_{m=0}^{N_n} J(s_{n,m}). \tag{5.22}$$

Without loss of generality, we can assume $J(s_{n,m}) \cap J(s_{n,m'}) = \emptyset$ whenever $m \neq m'$ (it suffices to remove certain elements from the sequence). We choose $\theta_0 < \eta$, where η is given by Definition 4.5. Then Definition 4.4 implies that $[a_n, b_n] \neq A$, thus $N_n > 0$.

Observe, using (5.3), that $U(s_{n,m}) \leq C\mathbf{d}(s_{n,m})^2 = o_n(1)$, whereas $U(\widetilde{a}_n) \geq C_1^{-1}\mathbf{d}(\widetilde{a}_n)^2 \geq C_1^{-1}\eta_0^2$ and similarly $U(\widetilde{b}_n) \geq C_1^{-1}\eta_0^2$, which for n large enough implies that U, restricted to $J(s_{n,m})$, attains its global minimum. Let $t_{n,m} \in J(s_{n,m})$ be one of the points where this global minimum is attained, in particular we have $J(t_{n,m}) = J(s_{n,m})$ and one of the following three cases:

- $t_{n,m} \in (a_n, b_n)$ is a local minimal point of U,
- $t_{n,m} = a_n$ is a local minimal point from the right of U,
- $t_{n,m} = b_n$ is a local minimal point from the left of U.

Note also that, again by (5.3),

$$\mathbf{d}(t_{n,m}) \leq \sqrt{C_1 U(t_{n,m})} \leq \sqrt{C_1 U(s_{n,m})} \leq C_1 \mathbf{d}(s_{n,m}) \leq C_1 \theta_n, \tag{5.23}$$

where the last inequality follows since $s_{n,m} \in A$.

We set $e_{n,0}^L := a_n$ and $e_{n,N_n}^R := b_n$. Let $m \in \{0, 1, \ldots, N_n - 1\}$. Since $J(t_{n,m}) \cap J(t_{n,m+1}) = \emptyset$, there exists $t \in (t_{n,m}, t_{n,m+1})$ such that $U(t) = \infty$. Let $d_{n,m}^R$ be the smallest such t, and $d_{n,m+1}^L$ the largest one. Let $c_{n,m}^R$ be the smallest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (c_{n,m}^R, d_{n,m}^R)$. Similarly, let $c_{n,m+1}^L$ be the biggest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (d_{n,m+1}^L, c_{n,m+1}^L)$. Next, let $e_{n,m}^R$ be the smallest number such that $\mathbf{d}(t) \geq 2C_1\theta_n$ for all $t \in (e_{n,m}^R, c_{n,m}^R)$. If we take $\theta_n \leq \theta_0 < \frac{\epsilon}{2C_1}$, then we have $e_{n,m}^R < c_{n,m}^R$. It follows from (5.23) that $e_{n,m}^R > t_{n,m}$. Similarly, let $e_{n,m+1}^L$ be the biggest number such that $\mathbf{d}(t) \geq 2C_1\theta_n$ for all $t \in (c_{n,m+1}^L, e_{n,m+1}^L)$ (again, it follows that $e_{n,m+1}^L < t_{n,m+1}$). Finally, if $\mathbf{d}(t) \geq \epsilon$ for all $t \in (d_{n,m}^R, d_{n,m+1}^L)$, we set $f_{n,m}^R$ and $f_{n,m+1}^L$ arbitrarily, for example $f_{n,m}^R := d_{n,m}^R$ and $f_{n,m+1}^L := d_{n,m+1}^L$. If, on the contrary, there exists $t \in (d_{n,m}^R, d_{n,m+1}^L)$ such that $\mathbf{d}(t) < \epsilon$, we let $f_{n,m}^R$ be the biggest number such that

$\mathbf{d}(t) \geq \epsilon$ for all $t \in (d_{n,m}^R, f_{n,m}^R)$, and $f_{n,m+1}^L$ be the smallest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (f_{n,m+1}^L, d_{n,m+1}^L)$.

We check all the desired properties. For all $n \in \{0, 1, \ldots, N_n\}$, we have $e_{n,m}^L \leq t_{n,m} \leq e_{n,m}^R$. By the definition of $d_{n,m}^R$, for all $n \in \{0, 1, \ldots, N_n - 1\}$ we have $U(t) < \infty$ for all $t \in [t_{n,m}, d_{n,m}^R)$, in particular for all $t \in [t_{n,m}, e_{n,m}^R]$. Similarly, $U(t) < \infty$ for all $t \in [e_{n,m+1}^L, t_{n,m+1}]$. We also have $a_n \in J(t_{n,0})$, which implies $U(t) < \infty$ for all $t \in [a_n, t_{n,0}] = [e_{n,0}^L, t_{n,0}]$. Similarly, $U(t) < \infty$ for all $t \in [t_{n,N_n}, e_{n,N_n}^R]$, thus $U(t) < \infty$ for all $n \in \{0, 1, \ldots, N_n\}$ and $t \in [e_{n,m}^L, e_{n,m}^R]$. Moreover, if $t_{n,m} = e_{n,m}^L$ (which can only happen for $m = 0$), then $t_{n,m}$ is a local minimal point from the right of U, and if $t_{n,m} = e_{n,m}^R$ (which can only happen for $m = N_n$), then $t_{n,m}$ is a local minimal point from the left of U. Since $\mathbf{d}(e_{n,m}^L) \leq 2C_1\theta_n$ and $\mathbf{d}(e_{n,m}^R) \leq 2C_1\theta_n$, the property (1) follows from (5.5). The properties (3), (4), (5) and (6) follow directly from the construction. The property (2) is now equivalent to the following statement: if $\mathbf{d}(t_0) < \theta_n$, then there exists $m \in \{0, 1, \ldots, N_n\}$ such that $t_0 \in [e_{n,m}^L, e_{n,m}^R]$. But (5.22) implies that $t_0 \in J(s_{n,m}) = J(t_{n,m})$ for some m and, by construction, $\mathbf{d}(t) > \theta_n$ for all $t \in J(t_{n,m}) \setminus [e_{n,m}^L, e_{n,m}^R]$, so we obtain $t \in [e_{n,m}^L, e_{n,m}^R]$. Finally, using again Lemma 5.8, but on the time intervals $[t_{n,m}, c_{n,m}^R]$ and $[c_{n,m+1}^L, t_{n,m+1}]$, we deduce the property (7) from (5.4). $\qquad\square$

5.3. **End of the proof: virial inequality with a cut-off.** In this section, we conclude the proof, by integrating the virial identity on the time interval $[a_n, b_n]$. The radius where the cut-off is imposed has to be carefully chosen, which is the object of the next lemma.

Lemma 5.12. *There exist $\theta_0 > 0$ and a locally Lipschitz function $\rho : \cup_{n=1}^\infty [a_n, b_n] \to (0, \infty)$ having the following properties:*

(1) $\max(\rho(a_n)\|\partial_t u(a_n)\|_{L^2}, \rho(b_n)\|\partial_t u(b_n)\|_{L^2}) \ll \max(\mu(a_n), \mu(b_n))$ *as* $n \to \infty$,
(2) $\lim_{n\to\infty} \inf_{t\in[a_n,b_n]} (\rho(t)/\mu(t)) = \infty$ *and* $\lim_{n\to\infty} \sup_{t\in[a_n,b_n]} (\rho(t)/\mu_{K+1}(t)) = 0$,
(3) *if* $\mathbf{d}(t_0) \leq \frac{1}{2}\theta_0$, *then* $|\rho'(t)| \leq 1$ *for almost all t in a neighborhood of t_0*,
(4) $\lim_{n\to\infty} \sup_{t\in[a_n,b_n]} |\Omega_{\rho(t)}(\boldsymbol{u}(t))| = 0$.

Proof. We will define two functions $\rho^{(a)}, \rho^{(b)}$, and then set $\rho := \min(\rho^{(a)}, \rho^{(b)}, \nu)$. First, we let

$$\rho^{(a)}(a_n) := \min(R_n\mu(a_n), \nu(a_n)),$$

where $1 \ll R_n \ll \|\partial_t u(a_n)\|_{L^2}^{-1}$. Consider an auxiliary sequence

$$\delta_n := \sup_{t\in[a_n,b_n]} \|\boldsymbol{u}(t)\|_{\mathcal{E}(\min(\rho^{(a)}(a_n)+t-a_n, \nu(t)); 2\nu(t))}.$$

We have $\lim_{n\to\infty} \delta_n = 0$. Indeed, we see from the finite speed of propagation that

$$\limsup_{n\to\infty} E(\boldsymbol{u}(t); \rho^{(a)}(a_n) + t - a_n, \infty) \leq E(\boldsymbol{u}^*) + (N - K)E(\boldsymbol{Q}).$$

This and Lemma 4.9 yield

$$\lim_{n\to\infty} E(\boldsymbol{u}(t); \min(\rho^{(a)}(a_n) + t - a_n, \nu(t)); 2\nu(t)) = 0,$$

thus Lemma 2.1 implies $\delta_n \to 0$.

Let $\theta_0 > 0$ be given by Lemma 5.10 (assuming without loss of generality that $\theta_0 < \epsilon$), and divide $[a_n, b_n]$ into subintervals applying this lemma for the constant sequence $\theta_n = \theta_0$. We let $\rho^{(a)}$ be the piecewise affine function such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho^{(a)}(t) := 1 \text{ if } t \in [e_{n,m}^L, e_{n,m}^R], \qquad \frac{\mathrm{d}}{\mathrm{d}t}\rho^{(a)}(t) := \delta_n^{-\frac{1}{2}} \text{ otherwise.}$$

Note that if $t \in [a_n, b_n]$ and $t \notin [e_{n,m}^L, e_{n,m}^R]$ for all m, then $\mathbf{d}(t) \geq \theta_0$.

We check that $\lim_{n\to\infty} \inf_{t\in[a_n,b_n]} \left(\rho^{(a)}(t)/\mu(t)\right) = \infty$. First, suppose that $t \in [e_{n,m}^R, e_{n,m+1}^L]$ and $t - e_{n,m}^R \gtrsim \mu(e_{n,m}^R)$. Then $\mu(t) \leq \mu(e_{n,m}^R) + (t - e_{n,m}^R) \lesssim t - e_{n,m}^R$ and $\rho^{(a)}(t) \geq \delta_n^{-\frac{1}{2}}(t - e_{n,m}^R)$, so $\rho^{(a)}(t) \gg \mu(t)$.

By Lemma 5.4, $e_{n,m+1}^L - e_{n,m}^R \geq C_u \mu(e_{n,m}^R)$, so in particular we obtain $\rho^{(a)}(e_{n,m+1}^L) \gg \mu(e_{n,m+1}^L)$ for all $m \in \{0, 1, \ldots, N_n - 1\}$. Note that we also have $\rho^{(a)}(e_{n,0}^L) = \rho^{(a)}(a_n) \gg \mu(a_n) = \mu(e_{n,0}^L)$, by the choice of $\rho^{(a)}(a_n)$. Since, by the property (7) in Lemma 5.10, μ changes at most by a factor 2 on $[e_{n,m}^L, e_{n,m}^R]$ and $\rho^{(a)}$ is increasing, we have $\rho^{(a)}(e_{n,m}^R) \gg \mu(e_{n,m}^R)$.

Finally, if $t - e_{n,m}^R \leq \mu(e_{n,m}^R)$, then $\mu(t) \leq 2\mu(e_{n,m}^R)$, which again implies $\rho^{(a)}(t) \gg \mu(t)$.

The function $\rho^{(b)}$ is defined similarly, but integrating from b_n backwards. Properties (1), (2), (3) are clear. By the expression for $\Omega_{\rho(t)}(\boldsymbol{u}(t))$, see Lemma 2.4, we have

$$|\Omega_{\rho(t)}(\boldsymbol{u}(t))| \lesssim (1 + |\rho'(t)|)\|\boldsymbol{u}(t)\|_{\mathcal{E}(\rho(t), 2\rho(t))}^2 \lesssim \sqrt{\delta_n} \to 0,$$

which proves the property (4). $\qquad\square$

We need one more elementary result.

Lemma 5.13. *If $a, b \in \mathbb{R}$, $a < b$, $\mu : [a, b] \to (0, \infty)$ is a 1-Lipschitz function and $b - a \geq \frac{1}{4}\mu(a)$, then there exist $l \in \mathbb{N}$ and a sequence $a = a_0 < a_1 < \ldots < a_l < a_{l+1} = b$ such that*

$$\frac{1}{4}\mu(a_i) \leq a_{i+1} - a_i \leq \frac{3}{4}\mu(a_i), \qquad \text{for all } i \in \{0, 1, \ldots, l\}. \tag{5.24}$$

Proof. We set $a_0 := a$ and define inductively $a_{i+1} := a_i + \frac{1}{4}\mu(a_i)$, as long as $b - a_i > \frac{3}{4}\mu(a_i)$. If $b - a_i \leq \frac{3}{4}\mu(a_i)$, then we set $l := i$, $a_{l+1} := b$ and terminate the procedure. Since $a_{i+1} - a_i \geq \frac{1}{4}\min_{t\in[a,b]}\mu(t) > 0$ for all i, this is achieved in a finite number of steps. We have $a_{i+1} - a_i = \frac{1}{4}\mu(a_i)$ for all $i < l$, thus (5.24) holds for all $i < l$. By the definition of l, we have $b - a_l \leq \frac{3}{4}\mu(a_l)$, and we only need to prove that $b - a_l \geq \frac{1}{4}\mu(a_l)$.

If $l = 0$, then the assumptions of the lemma yield $b - a_0 = b - a \geq \mu(a) = \mu(a_0)$. If $l > 0$, then, by the definition of l, we have $b - a_{l-1} > \frac{3}{4}\mu(a_{l-1})$. Since μ is 1-Lipschitz, we also have $\mu(a_l) = \mu(a_{l-1} + \mu(a_{l-1})/4) \leq \mu(a_{l-1}) + \mu(a_{l-1})/4 = \frac{5}{4}\mu(a_{l-1})$, thus

$$b - a_l = b - a_{l-1} - \frac{1}{4}\mu(a_{l-1}) > \frac{3}{4}\mu(a_{l-1}) - \frac{1}{4}\mu(a_{l-1}) > \frac{5}{16}\mu(a_{l-1}) \geq \frac{1}{4}\mu(a_l).$$

$\qquad\square$

Remark 5.14. Note that (5.24) and the fact that μ is 1-Lipschitz imply $\inf_{t\in[a_i,a_{i+1}]}\mu(t) \geq \frac{1}{4}\mu(a_i)$ and $\sup_{t\in[a_i,a_{i+1}]}\mu(t) \leq \frac{7}{4}\mu(a_i)$, thus

$$\frac{1}{7}\sup_{t\in[a_i,a_{i+1}]}\mu(t) \leq a_{i+1} - a_i \leq 3\inf_{t\in[a_i,a_{i+1}]}\mu(t),$$

in other words the length of each subinterval is comparable with both the smallest and the largest value of μ on this subinterval.

Lemma 5.15. *Let ρ be the function given by Lemma 5.12 and set*

$$\mathfrak{v}(t) := \int_0^\infty \partial_t u(t) r \partial_r u(t) \chi_{\rho(t)} \, r\mathrm{d}r. \tag{5.25}$$

1. *There exists a sequence $\theta_n \to 0$ such that the following is true. If $[\widetilde{a}_n, \widetilde{b}_n] \subset [a_n, b_n]$ is such that*

$$\widetilde{b}_n - \widetilde{a}_n \geq \frac{1}{4}\mu(\widetilde{a}_n) \quad and \quad \mathbf{d}(t) \geq \theta_n \text{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

   then

$$\mathfrak{v}(\widetilde{b}_n) < \mathfrak{v}(\widetilde{a}_n). \tag{5.26}$$

2. *For any $c, \theta > 0$ there exists $\delta > 0$ such that if n is large enough, $[\widetilde{a}_n, \widetilde{b}_n] \subset [a_n, b_n]$,*

$$c\mu(\widetilde{a}_n) \leq \widetilde{b}_n - \widetilde{a}_n \quad and \quad \mathbf{d}(t) \geq \theta \text{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

   then

$$\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \leq -\delta \sup_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \mu(t). \tag{5.27}$$

Proof. By the virial identity, we obtain

$$\mathfrak{v}'(t) = -\int_0^\infty (\partial_t u(t))^2 \chi_{\rho(t)} \, r\mathrm{d}r + o_n(1). \tag{5.28}$$

We argue by contradiction. If the claim is false, then there exists $\theta > 0$ and an infinite sequence $[\widetilde{a}_n, \widetilde{b}_n] \subset [a_n, b_n]$ (as usual, we pass to a subsequence in n without changing the notation) such that

$$\widetilde{b}_n - \widetilde{a}_n \geq \frac{1}{4}\mu(\widetilde{a}_n) \quad and \quad \mathbf{d}(t) \geq \theta \text{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

and

$$\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \geq 0.$$

We claim that there exists a subinterval of $[\widetilde{a}_n, \widetilde{b}_n]$, which we will still denote $[\widetilde{a}_n, \widetilde{b}_n]$, such that

$$\frac{1}{4}\mu(\widetilde{a}_n) \leq \widetilde{b}_n - \widetilde{a}_n \leq \frac{3}{4}\mu(\widetilde{a}_n) \quad and \quad \mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \geq 0.$$

Indeed, by Lemma 5.13, there exist $l_n \in \mathbb{N}$ and $\widetilde{a}_n = \widetilde{a}_{n,0} < \widetilde{a}_{n,1} < \ldots < \widetilde{a}_{n,l_n+1} = \widetilde{b}_n$ such that

$$\frac{1}{4}\mu(\widetilde{a}_{n,i}) \leq \widetilde{a}_{n,i+1} - \widetilde{a}_{n,i} \leq \frac{3}{4}\mu(\widetilde{a}_{n,i}), \qquad \text{for all } i \in \{0, 1, \ldots, l_n\}.$$

If we had $\mathfrak{v}(\widetilde{a}_{n,i+1}) - \mathfrak{v}(\widetilde{a}_{n,i}) < 0$ for all $i \in \{0, 1, \ldots, l_n\}$, then summation with respect to i would yield $\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) < 0$. Hence, there exists $i \in \{0, 1, \ldots, l_n\}$ such that $\mathfrak{v}(\widetilde{a}_{n,i+1}) - \mathfrak{v}(\widetilde{a}_{n,i}) \geq 0$, and $[\widetilde{a}_{n,i}, \widetilde{a}_{n,i+1}]$ is the subinterval of $[\widetilde{a}_n, \widetilde{b}_n]$ whose existence we claimed.

Let $\widetilde{\rho}_n := \inf_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \rho(t)$. From (5.28), we have

$$\lim_{n \to \infty} \frac{1}{\widetilde{b}_n - \widetilde{a}_n} \int_{\widetilde{a}_n}^{\widetilde{b}_n} \int_0^{\frac{1}{2}\widetilde{\rho}_n} (\partial_t u(t))^2 \, r\mathrm{d}r = 0.$$

By Lemma 5.12, $\inf_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \mu_{K+1}(t) \gg \widetilde{\rho}_n \gg \inf_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \mu(t) \simeq \sup_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \mu(t)$, so Lemma 3.1 yields sequences $t_n \in [\widetilde{a}_n, \widetilde{b}_n]$ and $1 \ll r_n \ll \mu_{K+1}(t_n)/\mu(t_n)$ such that

$$\lim_{n \to \infty} \boldsymbol{\delta}_{r_n\mu(t_n)}(\boldsymbol{u}(t_n)) = 0,$$

which is impossible by Lemma 5.2 (iii). The first part of the lemma is proved.

In the second part, we can assume without loss of generality $\widetilde{b}_n - \widetilde{a}_n \leq \frac{3}{4}\mu(\widetilde{a}_n)$. Indeed, in the opposite case, we apply Lemma 5.13 and keep only one of the subintervals where μ attains its supremum, and on the remaining subintervals we use (5.26).

After this preliminary reduction, we argue again by contradiction. If the claim is false, then there exist $c, \theta > 0$, a sequence $\delta_n \to 0$ and a sequence $[\widetilde{a}_n, \widetilde{b}_n] \subset [a_n, b_n]$ (after extraction of a subsequence) such that

$$c\mu(\widetilde{a}_n) \leq \widetilde{b}_n - \widetilde{a}_n \leq \frac{3}{4}\mu(\widetilde{a}_n) \quad \text{and} \quad \mathbf{d}(t) \geq \theta \text{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

and

$$\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \geq -\delta_n \mu(\widetilde{a}_n)$$

(we use the fact that $\mu(\widetilde{a}_n)$ is comparable to $\sup_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \mu(t)$, see Remark 5.14).

Let $\widetilde{\rho}_n := \inf_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \rho(t)$. From (5.28), we have

$$\lim_{n \to \infty} \frac{1}{\widetilde{b}_n - \widetilde{a}_n} \int_{\widetilde{a}_n}^{\widetilde{b}_n} \int_0^{\frac{1}{2}\widetilde{\rho}_n} (\partial_t u(t))^2 \, r\mathrm{d}r = 0.$$

We now conclude as in the first part. $\qquad\square$

Proof of Theorem 1. Let θ_n be the sequence given by Lemma 5.15, part 1. We partition $[a_n, b_n]$ applying Lemma 5.10 for this sequence θ_n. Note that this partition is different than the one used in the proof of Lemma 5.12. We claim that for all $m \in \{0, 1, \ldots, N_n - 1\}$

$$\mathfrak{v}(c_{n,m}^R) - \mathfrak{v}(e_{n,m}^R) \leq o_n(1)\mu(c_{n,m}^R), \tag{5.29}$$

$$\mathfrak{v}(f_{n,m+1}^L) - \mathfrak{v}(f_{n,m}^R) \leq o_n(1)\mu(f_{n,m}^R), \tag{5.30}$$

$$\mathfrak{v}(e_{n,m+1}^L) - \mathfrak{v}(c_{n,m+1}^L) \leq o_n(1)\mu(c_{n,m+1}^L). \tag{5.31}$$

Here, $o_n(1)$ denotes a sequence of positive numbers converging to 0 when $n \to \infty$. In order to prove the first inequality, we observe that if $c_{n,m}^R - e_{n,m}^R \geq \frac{1}{4}\mu(e_{n,m}^R)$, then (5.26) applies and yields $\mathfrak{v}(c_{n,m}^R) - \mathfrak{v}(e_{n,m}^R) < 0$. We can thus assume $c_{n,m}^R - e_{n,m}^R \leq \frac{1}{4}\mu(e_{n,m}^R) \leq \frac{1}{2}\mu(c_{n,m}^R)$, where the last inequality follows from Lemma 5.10, property (7). But then (5.28) again implies the required bound. The proofs of the second and third bound are analogous.

We now analyse the compactness intervals $[c_{n,j}^R, f_{n,j}^R]$ and $[f_{n,j+1}^L, c_{n,j+1}^L]$. We claim that there exists $\delta > 0$ such that for all n large enough and $m \in \{0, 1, \ldots, N_n\}$

$$\mathfrak{v}(c_{n,m+1}^L) - \mathfrak{v}(c_{n,m}^R) \leq -\delta \max(\mu(c_{n,m}^R), \mu(c_{n,m+1}^L)). \tag{5.32}$$

We consider separately the two cases mentioned in Lemma 5.10, property (6). If $\mathbf{d}(t) \geq \epsilon$ for all $t \in [c_{n,m}^R, c_{n,m+1}^L]$, then Lemma 5.4 yields $c_{n,m+1}^L - c_{n,m}^R \geq C_{\boldsymbol{u}}^{-1}\mu(c_{n,m}^R)$, so we can apply (5.27), which proves (5.32). If $\mathbf{d}(f_{n,m}^R) = \epsilon$, then we apply the same argument on the time interval $[c_{n,m}^R, f_{n,m}^R]$ and obtain

$$\mathfrak{v}(f_{n,m}^R) - \mathfrak{v}(c_{n,m}^R) \leq -\delta \max(\mu(c_{n,m}^R), \mu(f_{n,m}^R)),$$

and similarly

$$\mathfrak{v}(c_{n,m+1}^L) - \mathfrak{v}(f_{n,m+1}^L) \leq -\delta \max(\mu(c_{n,m+1}^L), \mu(f_{n,m+1}^L)).$$

The bound (5.30) yields (5.32).

Finally, on the intervals $[e_{n,m}^L, e_{n,m}^R]$, for n large enough Lemma 5.12 yields $|\rho'(t)| \leq 1$ for almost all t, and Lemma 4.18 implies $|\mathfrak{v}'(t)| \lesssim \mathbf{d}(t)$. By Lemma 5.10, properties (1) and (7), we obtain

$$\begin{aligned} \mathfrak{v}(e_{n,m}^R) - \mathfrak{v}(e_{n,m}^L) &\leq o_n(1)\mu(c_{n,m}^R), \qquad \text{for all } m \in \{0, 1, \ldots, N_n - 1\}, \\ \mathfrak{v}(e_{n,m}^R) - \mathfrak{v}(e_{n,m}^L) &\leq o_n(1)\mu(c_{n,m}^L), \qquad \text{for all } m \in \{1, \ldots, N_n - 1, N_n\}. \end{aligned} \tag{5.33}$$

Taking the sum in m of (5.29), (5.31), (5.32) and (5.33), we deduce that there exists $\delta > 0$ and n arbitrarily large such that

$$\mathfrak{v}(b_n) - \mathfrak{v}(a_n) \leq -\delta \max(\mu(c_{n,0}^R), \mu(c_{n,N_n}^L)).$$

But $a_n = e_{n,0}^L$ and $b_n = e_{n,N_n}^R$, hence property (7) in Lemma 5.10 yields $\mu(a_n) \leq 2\mu(c_{n,0}^R)$ and $\mu(b_n) \leq 2\mu(c_{n,N_n}^L)$, thus

$$\mathfrak{v}(b_n) - \mathfrak{v}(a_n) \leq -\widetilde{\delta} \max(\mu(a_n), \mu(b_n)).$$

Lemma 5.12 (1) and (5.25) yield

$$|\mathfrak{v}(a_n)| \ll \mu(a_n), \qquad |\mathfrak{v}(b_n)| \ll \mu(b_n),$$

a contradiction which finishes the proof. $\qquad\square$

5.4. **Absence of elastic collisions.** This section is devoted to proving Proposition 1.7. Our proof closely follows Step 3 in our proof of [41, Theorem 1.6].

Proof of Proposition 1.7. Suppose that a solution of (1.1), $\boldsymbol{u}$, defined on its maximal time of existence $t \in (T_-, T_+)$, is a pure multi-bubble in both time directions in the sense of Definition 1.6, in other words

$$\lim_{t \to T_+} \mathbf{d}(t) = 0, \qquad \text{and} \qquad \lim_{t \to T_-} \mathbf{d}(t) = 0,$$

and the radiation $\boldsymbol{u}^* = \boldsymbol{u}_L^*$ or $\boldsymbol{u}^* = \boldsymbol{u}_0^*$ in both time directions satisfies $\boldsymbol{u}^* \equiv 0$. In this proof, all the N bubbles can be thought of as "interior" bubbles thus, whenever we invoke the results from the preceding sections, it should always be understood that $K = N$. Applying Lemma 2.25 with $\theta = 0$ and $M = N$, we obtain from (2.30) and (2.29) that

$$\mathbf{d}(t) \leq C \max_{j \in \mathcal{A}} \left(\frac{\lambda_j}{\lambda_{j+1}} \right)^{\frac{k}{2}}.$$

Let $\eta > 0$ be a small number to be chosen later and t_+ be such that $\mathbf{d}(t) \leq \eta$ for all $t \geq t_+$. If η is sufficiently small, then the modulation parameters are well-defined for $t \geq t_+$, so we can set

$$U(t) := \max_{i \in \mathcal{A}} \left(2^{-i} \xi_i(t) / \lambda_{i+1}(t) \right)^k, \qquad \text{for all } t \geq t_+,$$

cf. Definition 5.6. Since U is a positive continuous function and $\lim_{t \to T_+} U(t) = 0$, there exists an increasing sequence $t_n \to T_+$ such that t_n is a local minimal point from the left of U. Thus, Lemma 5.8 yields

$$\int_{t_+}^{t_n} \mathbf{d}(t) \mathrm{d}t \leq C_0 \mathbf{d}(t_+)^{\frac{2}{k}} \lambda_N(t_+),$$

and passing to the limit $n \to +\infty$ we get

$$\int_{t_+}^{T_+} \mathbf{d}(t) \mathrm{d}t \leq C_0 \mathbf{d}(t_+)^{\frac{2}{k}} \lambda_N(t_+). \tag{5.34}$$

By inspecting the proof of Lemma 4.11, one finds that in the present case it holds with $\zeta_n = 0$, in particular we have $|\lambda_N'(t)| \lesssim \mathbf{d}(t)$. This bound, together with (5.34), implies that $\lim_{t \to T_+} \lambda_N(t)$ is a finite positive number, thus $T_+ = +\infty$.

Analogously, $T_- = -\infty$ and $\lim_{t \to -\infty} \lambda_N(t) \in (0, +\infty)$ exists.

The remaining part of the argument is exactly the same as in [41], but we reproduce it here for the reader's convenience.

Let $\delta > 0$ be arbitrary. Inspecting the proof of Lemma 4.18, we see that in the present case it holds with $\epsilon_n = 0$, thus for any $R > 0$ we have $|\Omega_R(\boldsymbol{u}(t))| \le C_0 \boldsymbol{d}(t)$. From this bound and the estimates above, we obtain existence of $T_1, T_2 \in \mathbb{R}$ such that

$$\int_{-\infty}^{T_1} |\Omega_R(\boldsymbol{u}(t))| \mathrm{d}t \le \frac{1}{3}\delta,$$

$$\int_{T_2}^{+\infty} |\Omega_R(\boldsymbol{u}(t))| \mathrm{d}t \le \frac{1}{3}\delta$$

for any $R > 0$. On the other hand, because of the bound $|\Omega_R(\boldsymbol{u}(t))| \le C_0 E(\boldsymbol{u}(t); R, 2R)$ and since $[T_1, T_2]$ is a finite time interval, for all R sufficiently large we have

$$\int_{T_1}^{T_2} |\Omega_R(\boldsymbol{u}(t))| \mathrm{d}t \le \frac{1}{3}\delta,$$

in other words

$$\int_{\mathbb{R}} |\Omega_R(\boldsymbol{u}(t))| \mathrm{d}t \le \delta.$$

Integrating the virial identity from Lemma 2.4 with $\rho(t) = R$ over the real line, we obtain

$$\int_{-\infty}^{+\infty} \int_0^\infty (\partial_t u(t,r) \chi_R(r))^2 \, r \mathrm{d}r \mathrm{d}t \le \delta.$$

By letting $R \to +\infty$, we get

$$\int_{-\infty}^{+\infty} \int_0^\infty (\partial_t u(t,r))^2 \, r \mathrm{d}r \mathrm{d}t \le \delta,$$

which implies the $\boldsymbol{u}$ is stationary since δ is arbitrary. $\qquad\square$

APPENDIX A. MODIFICATIONS TO THE ARGUMENT IN THE CASE $k = 1$

In this section we outline the changes to the arguments in Section 4 and Section 5 needed to prove Theorem 1 for the equivariance class $k = 1$.

A.1. **Modulation and refined modulation.** The set-up in Sections 4.1 holds without modification for $k = 1$. To be precise the number $K \ge 1$ is defined as in Lemma 4.6, the collision intervals $[a_n, b_n] \in \mathcal{C}_K(\eta, \epsilon_n)$ are as in Definition 4.5, and the sequences of signs $\vec{\sigma}_n \in \{-1, 1\}^{N-K}$, scales $\vec{\mu}(t) \in (0, \infty)^{N-K}$, and integers $m_n \in \mathbb{Z}$ associated to the exterior bubbles, and the sequence $\nu_n \to 0$ and the function $\nu(t) = \nu_n \mu_{K+1}(t)$ are as in Lemma 4.9.

Lemma 4.11 also holds without modification. Let $J \subset [a_n, b_n]$ be any time interval on which $\boldsymbol{d}(t) \le \eta_0$, where η_0 is as in Lemma 4.11. Let $\vec{\iota} \in \{-1, 1\}^K, \vec{\lambda}(t) \in (0, \infty)^K$, and $\boldsymbol{g}(t) \in \mathcal{E}$ be as in the statement of Lemma 4.11. Let $L > 0$ be a parameter to be fixed below and define for each $j \in \{1, \dots, K-1\}$,

$$\xi_j(t) := \lambda_j(t) - \frac{\iota_j}{2\log(\frac{\lambda_{j+1}(t)}{\lambda_j(t)})} \Big\langle \chi_{L\sqrt{\lambda_j(t)\lambda_{j+1}(t)}} \Lambda Q_{\underline{\lambda_j(t)}} \mid g(t) + \sum_{i<j} \iota_i (Q_{\underline{\lambda_i(t)}} - \pi) \Big\rangle,$$

and,

$$\beta_j(t) := -\iota_j \Big\langle \chi_{L\sqrt{\xi_j(t)\lambda_{j+1}(t)}} \Lambda Q_{\underline{\lambda_j(t)}} \mid \dot{g}(t) \Big\rangle - \Big\langle \underline{A}(\lambda_j(t)) g(t) \mid \dot{g}(t) \Big\rangle . \tag{A.1}$$

Proposition A.1 (Refined modulation, $k = 1$). *Let $c_0 \in (0, 1)$ and $c_1 > 0$. There exists constants $L_0 = L_0(c_0, c_1) > 0$, $\eta_0 = \eta_0(c_0, c_1)$, as well as $c = c(c_0, c_1)$ and $R = R(c_0, c_1) > 1$ as in Lemma 4.13, a constant $C_0 > 0$, and a decreasing sequence $\epsilon_n \to 0$ so that the following is true.*

Suppose $L > L_0$ and $J \subset [a_n, b_n]$ is an open time interval with $\epsilon_n \leq \mathbf{d}(t) \leq \eta_0$ for all $t \in J$, where $\mathcal{A} := \{j \in \{1, \ldots, K-1\} \mid \iota_j \neq \iota_{j+1}\}$. Then, for all $t \in J$,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{i \notin \mathcal{A}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}} \leq \max_{i \in \mathcal{A}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}, \tag{A.2}$$

and,

$$\frac{1}{C_0}\mathbf{d}(t) \leq \max_{i \in \mathcal{A}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}} \leq C_0 \mathbf{d}(t), \tag{A.3}$$

$$\left| \frac{\xi_j(t)}{\lambda_{j+1}(t)} - \frac{\lambda_j(t)}{\lambda_{j+1}(t)} \right| \leq c_0 \mathbf{d}(t)^2. \tag{A.4}$$

Moreover, let $j \in \mathcal{A}$ be such that for all $t \in J$

$$c_1 \mathbf{d}(t) \leq \left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)} \right)^{\frac{1}{2}}. \tag{A.5}$$

Then for all $t \in J$,

$$|\xi'_j(t)| \left(\log \left(\frac{\lambda_{j+1}(t)}{\lambda_j(t)} \right) \right)^{\frac{1}{2}} \leq C_0 \max_{i \in \mathcal{A}} \sqrt{\frac{\lambda_i(t)}{\lambda_{i+1}(t)}}, \tag{A.6}$$

$$\left| \xi'_j(t) 2 \log(\frac{\lambda_{j+1}(t)}{\lambda_j(t)}) - \beta_j(t) \right| \leq \frac{C_0}{c_1} \max_{i \in \mathcal{A}} \sqrt{\frac{\lambda_i(t)}{\lambda_{i+1}(t)}} \tag{A.7}$$

and,

$$\beta'_j(t) \geq \left(-\iota_j \iota_{j+1} 8 - c_0 \right) \frac{1}{\lambda_{j+1}(t)} + \left(\iota_j \iota_{j-1} 8 - c_0 \right) \frac{\lambda_{j-1}(t)}{\lambda_j(t)^2}$$
$$- \frac{c_0}{\lambda_j(t)} \max_{i \in \mathcal{A}} \frac{\lambda_i(t)}{\lambda_{i+1}(t)}. \tag{A.8}$$

where, by convention, $\lambda_0(t) = 0, \lambda_{K+1}(t) = \infty$ for all $t \in J$.

Proof. The estimates (A.2) and (A.3) follow as in the proofs of the corresponding estimates in Lemma 4.16. We next prove (A.4). From the definition of $\xi_j(t)$,

$$\left| \frac{\xi_j}{\lambda_{j+1}} - \frac{\lambda_j}{\lambda_{j+1}} \right| \lesssim \left| \frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})} \lambda_{j+1}^{-1} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid g \rangle \right|$$
$$+ \left| \frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})} \lambda_{j+1}^{-1} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \sum_{i<j}(Q_{\lambda_i} - \pi) \rangle \right|$$

For the first term on the right we have,

$$\left| \frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})} \lambda_{j+1}^{-1} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid g \rangle \right| \lesssim_L \frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})} \|g\|_H (\lambda_j/\lambda_{j+1})^{\frac{1}{2}}$$

Next, for any $i < j$ we have,

$$\lambda_{j+1}^{-1} |\langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid (Q_{\lambda_i} - \pi) \rangle| \lesssim_L \frac{\lambda_j}{\lambda_{j+1}} \int_0^{L(\lambda_{j+1}/\lambda_j)^{\frac{1}{2}}} \Lambda Q(r) \left| Q_{\lambda_i/\lambda_j}(r) - \pi \right| r \, dr$$
$$\lesssim_L \frac{\lambda_j}{\lambda_{j+1}} \frac{\lambda_i}{\lambda_j} \log(\lambda_{j+1}/\lambda_j)$$

and hence,

$$\left| \frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})} \lambda_{j+1}^{-1} \langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \sum_{i<j}(Q_{\lambda_i} - \pi)\rangle \right| \lesssim_L \frac{\lambda_j}{\lambda_{j+1}} \sum_{i<j} \frac{\lambda_i}{\lambda_j}$$

and (A.4) follows.

Next using (A.2) and (4.24) for each j, we have

$$|\lambda_j'| \lesssim \max_{i\in\mathcal{A}}(\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}. \tag{A.9}$$

We show that in fact ξ_j' satisfies the improved estimate (A.6). We compute,

$$\begin{aligned}
\xi_j' = \lambda_j' &- \frac{\iota_j}{2(\log(\frac{\lambda_{j+1}}{\lambda_j}))^2}\left(\frac{\lambda_j'}{\lambda_j} - \frac{\lambda_{j+1}'}{\lambda_{j+1}}\right)\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i(Q_{\lambda_i}-\pi)\rangle \\
&+ \frac{\iota_j}{4\log(\frac{\lambda_{j+1}}{\lambda_j})}\left(\frac{\lambda_j'}{\lambda_j} + \frac{\lambda_{j+1}'}{\lambda_{j+1}}\right)\langle \Lambda\chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i(Q_{\lambda_i}-\pi)\rangle \\
&+ \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\frac{\lambda_j'}{\lambda_j}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i(Q_{\lambda_i}-\pi)\rangle \\
&- \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \partial_t g\rangle \\
&+ \sum_{i<j} \frac{\iota_i\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_i'\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}}\rangle
\end{aligned} \tag{A.10}$$

The second, third, and fourth terms on the right above contribute acceptable errors. Indeed,

$$\left| \frac{\iota_j}{2(\log(\frac{\lambda_{j+1}}{\lambda_j}))^2}\left(\frac{\lambda_j'}{\lambda_j} - \frac{\lambda_{j+1}'}{\lambda_{j+1}}\right)\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i(Q_{\lambda_i}-\pi)\rangle \right| \lesssim \frac{\max_{i\in\mathcal{A}}(\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}}{c_1(\log(\frac{\lambda_{j+1}}{\lambda_j}))^2}$$

$$\left| \frac{\iota_j}{4\log(\frac{\lambda_{j+1}}{\lambda_j})}\left(\frac{\lambda_j'}{\lambda_j} + \frac{\lambda_{j+1}'}{\lambda_{j+1}}\right)\langle \Lambda\chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i(Q_{\lambda_i}-\pi)\rangle \right| \lesssim \frac{\max_{i\in\mathcal{A}}(\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}}{c_1\log(\frac{\lambda_{j+1}}{\lambda_j})}$$

$$\left| \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\frac{\lambda_j'}{\lambda_j}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i(Q_{\lambda_i}-\pi)\rangle \right| \lesssim \frac{\max_{i\in\mathcal{A}}(\lambda_i(t)/\lambda_{i+1}(t))}{\log(\frac{\lambda_{j+1}}{\lambda_j})}$$

with the gain in the last line arising from the fact that $\underline{\Lambda}\Lambda Q \in L^1$; see (2.16). The leading order comes from the second to last term in (A.10). Using (4.54) gives

$$\begin{aligned}
-\frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \partial_t g\rangle = &-\frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle \\
&- \lambda_j'\frac{1}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_j}}\rangle \\
&- \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \sum_{i\neq j}\iota_i\lambda_i'\Lambda Q_{\underline{\lambda_i}}\rangle \\
&- \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \phi(u,\nu)\rangle.
\end{aligned}$$

We estimate the contribution of each of the terms on the right above to (A.10). The last term above vanishes due to the support properties of $\phi(u,\nu)$. Using (2.15), (A.9) on the second term

above, gives

$$\left| -\lambda_j' \frac{1}{2\log(\frac{\lambda_{j+1}}{\lambda_j})} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_j}} \rangle + \lambda_j' \right| \lesssim \frac{\max_{i \in \mathcal{A}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}}{\log(\frac{\lambda_{j+1}}{\lambda_j})}$$

which means this terms cancels the term λ' on the right-hand side of (A.10) up to an acceptable error. Next we write,

$$-\frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \sum_{i \neq j} \iota_i \lambda_i' \Lambda Q_{\underline{\lambda_i}} \rangle = -\sum_{i<j} \frac{\iota_i \iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})} \lambda_i' \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}} \rangle$$
$$-\sum_{i>j} \frac{\iota_i \iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})} \lambda_i' \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}} \rangle$$

The first term cancels the last term in (A.10). For the second term we estimate, if $i > j$,

$$|\langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \Lambda Q_{\underline{\lambda_i}} \rangle| \lesssim \lambda_j/\lambda_{j+1}$$

and thus, using (A.9) the second term in the previous equation contributes an acceptable error. Plugging all of these estimates back into (A.10) gives the estimate,

$$\left| \xi_j' + \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \dot{g} \rangle \right| \lesssim \frac{\max_{i \in \mathcal{A}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}}{c_1 \log(\frac{\lambda_{j+1}}{\lambda_j})} \tag{A.11}$$

Using (A.2) and $\|\chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}}\|_{L^2} \lesssim (\log(\frac{\lambda_{j+1}}{\lambda_j}))^{\frac{1}{2}}$, we deduce the estimate,

$$\left| \frac{\iota_j}{2\log(\frac{\lambda_{j+1}}{\lambda_j})} \langle \chi_{L\sqrt{\lambda_j \lambda_{j+1}}} \Lambda Q_{\underline{\lambda_j}} \mid \dot{g} \rangle \right| \lesssim \frac{\max_{i \in \mathcal{A}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}}}{(\log(\frac{\lambda_{j+1}}{\lambda_j}))^{\frac{1}{2}}}$$

which completes the proof of (A.6).

Next we compare β_j and $2\xi_j' \log(\lambda_{j+1}/\lambda_j)$. Using (A.1) we have,

$$\left| \langle \underline{A}(\lambda_j(t))g(t) \mid \dot{g}(t) \rangle \right| \lesssim \|g\|_{\mathcal{E}}^2 \lesssim \max_{i \in \mathcal{A}} (\lambda_i/\lambda_{i+1}),$$

We also note the estimate

$$\left| \langle (\chi_{L\sqrt{\lambda_j \lambda_{j+1}}} - \chi_{L\sqrt{\xi_j \lambda_{j+1}}}) \Lambda Q_{\underline{\lambda_j}} \mid \dot{g} \rangle \right| \ll \max_{i \in \mathcal{A}} (\lambda_i/\lambda_{i+1})^{\frac{1}{2}}.$$

which is a consequence of (A.4). Using (A.11) the estimate (A.7) follows.

Finally, the proof of the estimate (A.8) is nearly identical to the argument used to prove (4.53), differing only in a few places where the cut-off $\chi_{L\sqrt{\xi_j \lambda_{j+1}}}$ is involved. Arguing as in the proof

of (4.53) we arrive at the formula,

$$\beta_j' = -\frac{\iota_j}{\lambda_j}\left\langle \Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle + \left\langle \underline{A}(\lambda_j)g \mid \mathcal{L}_0 g\right\rangle + \left\langle (A(\lambda_j) - \underline{A}(\lambda_j))g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle$$

$$+ \left\langle \chi_{L\sqrt{\xi_j \lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{Q}} - \mathcal{L}_{\lambda_j})g\right\rangle + \iota_j\frac{\lambda_j'}{\lambda_j}\left\langle (\frac{1}{\lambda_j}\underline{\Lambda} - \underline{A}(\lambda_j))\Lambda Q_{\lambda_j} \mid \dot{g}\right\rangle$$

$$- \left\langle A(\lambda_j)\sum_{i=1}^{K}\iota_i Q_{\lambda_i} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle - \left\langle A(\lambda_j)g \mid \widetilde{f}_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle$$

$$+ \iota_j\left\langle (A(\lambda_j) - \frac{1}{\lambda_j}\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda)Q_{\lambda_j} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle - \frac{\lambda_j'}{\lambda_j}\left\langle \lambda_j\partial_{\lambda_j}\underline{A}(\lambda_j)g \mid \dot{g}\right\rangle$$

$$+ \sum_{i\neq j}\iota_i\left\langle A(\lambda_j)Q_{\lambda_i} \mid f_{\mathbf{q}}(m_n, \vec{\iota}, \vec{\lambda}, g)\right\rangle - \sum_{i\neq j}\iota_i\lambda_i'\left\langle \underline{A}(\lambda_j)\Lambda Q_{\underline{\lambda_i}} \mid \dot{g}\right\rangle - \left\langle \underline{A}(\lambda_j)g \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle$$

$$- \iota_j\left\langle \chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}} \mid \dot{\phi}(u,\nu)\right\rangle - \left\langle \underline{A}(\lambda_j)\phi(u,\nu) \mid \dot{g}\right\rangle - \left\langle \underline{A}(\lambda_j)g \mid \dot{\phi}(u,\nu)\right\rangle$$

$$+ \frac{\iota_j}{\lambda_j}\left\langle (1 - \chi_{L\sqrt{\xi_j\lambda_{j+1}}})\Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle - \iota_j\frac{\lambda_j'}{\lambda_j}\left\langle (1 - \chi_{L\sqrt{\xi_j\lambda_{j+1}}})\underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\right\rangle$$

$$+ \left\langle \mathcal{L}_{\lambda_j}(\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}}) \mid g\right\rangle + \frac{\iota_j}{2}(\frac{\xi_j'}{\xi_j} + \frac{\lambda_{j+1}'}{\lambda_{j+1}})\langle \Lambda\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle$$

All but the last four terms above are treated exactly as in the proof of (4.53). For the fourth-to-last term a direct computation using the estimate (4.65) gives,

$$\left|\frac{\iota_j}{\lambda_j}\left\langle (1 - \chi_{L\sqrt{\xi_j\lambda_{j+1}}})\Lambda Q_{\lambda_j} \mid f_{\mathbf{i}}(m_n, \iota, \vec{\lambda})\right\rangle\right| \ll \frac{1}{\lambda_j}\left(\frac{\lambda_j}{\lambda_{j+1}} + \frac{\lambda_{j-1}}{\lambda_j}\right).$$

For the third-to-last term, we use that $\underline{\Lambda}\Lambda Q \in L^2$, (A.9), and (A.2) to deduce that,

$$\left|\iota_j\frac{\lambda_j'}{\lambda_j}\left\langle (1 - \chi_{L\sqrt{\xi_j\lambda_{j+1}}})\underline{\Lambda}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\right\rangle\right| \ll \frac{1}{\lambda_j}\max_{i\in\mathcal{A}}(\lambda_i/\lambda_{i+1}),$$

The size of the constant $L > 0$ becomes relevant only in the second-to-last term. Indeed, since $\mathcal{L}\Lambda Q = 0$, we have,

$$\mathcal{L}_{\lambda_j}(\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}}) = \frac{1}{L^2\xi_j\lambda_{j+1}}(\Delta\chi)_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}} + 2\frac{1}{L\sqrt{\xi_j\lambda_{j+1}}}\chi_{L\sqrt{\xi_j\lambda_{j+1}}}'\frac{1}{r}\Lambda^2 Q_{\underline{\lambda_j}}.$$

And therefore, using (A.2) and (A.4) we obtain the estimate,

$$\left|\left\langle \mathcal{L}_{\lambda_j}(\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}}) \mid g\right\rangle\right| \lesssim \frac{1}{L}\frac{1}{\lambda_j}\max_{i\in\mathcal{A}}(\lambda_i/\lambda_{i+1})$$

for a uniform constant, independent of L. Taking $L > 1$ large enough relative to c_0 makes this an acceptable error. Finally, for the last term we use the improved estimate (A.6) for ξ_j' and (A.4) to obtain,

$$\left|\frac{\xi_j'}{\xi_j} + \frac{\lambda_{j+1}'}{\lambda_{j+1}}\right| \lesssim \frac{1}{\lambda_j}\left(|\xi_j'| + \frac{\lambda_j}{\lambda_{j+1}}|\lambda_{j+1}'|\right) \ll \frac{1}{\lambda_j}\max_{i\in\mathcal{A}}(\lambda_i/\lambda_{i+1})^{\frac{1}{2}},$$

and hence,

$$\left|\frac{\iota_j}{2}(\frac{\xi_j'}{\xi_j} + \frac{\lambda_{j+1}'}{\lambda_{j+1}})\langle \Lambda\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\rangle\right| \ll \frac{1}{\lambda_j}\max_{i\in\mathcal{A}}(\lambda_i/\lambda_{i+1}).$$

This completes the proof. □

We note that Lemma 4.18 and its proof remain valid for $k = 1$.

A.2. **Demolition of the multi-bubble.** We define the weighted interaction energy in the same way as for $k \geq 2$, see Definition 5.6.

Lemma A.2. *If c_0 in Proposition A.1 is taken sufficiently small, then there exists a constant $C_1 > 0$ such that*

$$\mathbf{d}(t) < \eta_0 \quad \text{implies} \quad C_1^{-1}\mathbf{d}(t)^2 \leq U(t) \leq C_1\mathbf{d}(t)^2.$$

Remark A.3. *In the case $k \geq 2$ the corresponding estimate (5.3) follows immediately from (4.51). For $k = 1$, since the bound (A.4) does not imply that $\xi_j(t) \simeq \lambda_j(t)$ for all $j \in \mathcal{A}$.*

Proof. Let t be such that $\mathbf{d}(t) < \eta_0$, and let $j_0 \in \mathcal{A}$ be such that

$$\lambda_{j_0}(t)/\lambda_{j_0+1}(t) = \max_{i \in \mathcal{A}} \lambda_i(t)/\lambda_{i+1}(t).$$

Then we deduce from (A.4) that

$$\xi_{j_0}(t)/\lambda_{j_0+1}(t) \gtrsim \lambda_{j_0}(t)/\lambda_{j_0+1}(t) \gtrsim \mathbf{d}(t)^2,$$

which yields the required lower bound on $U(t)$.

The upper bound follows directly from (A.4) and the fact that $\lambda_j(t)/\lambda_{j+1}(t) \lesssim \mathbf{d}(t)^2$ for all $j \in \mathcal{A}$. □

The analog of Lemma 5.8 for $k = 1$ is formulated as follows.

Lemma A.4. *If η_0 is small enough, then there exists $C_0 \geq 0$ depending only on k and N such that the following is true. If t_0 is a local minimal point from the right of U such that $U(t_0) < +\infty$ and $t_* \geq t_0$ is such that $U(t) < \infty$ for all $t \in [t_0, t_*]$, then*

$$\frac{3}{4}\lambda_K(t_0) \leq \lambda_K(t_*) \leq \frac{4}{3}\lambda_K(t_0),$$

$$\int_{t_0}^{t_*} \mathbf{d}(t)\mathrm{d}t \leq C_0\mathbf{d}(t_*)^2\sqrt{-\log \mathbf{d}(t_*)}\lambda_K(t_0).$$

An analogous statement is true if t_* is a local minimal point from the left.

Proof. Steps 1 and 2 are similar as for $k \geq 2$. Step 3 differs significantly, so let us indicate the necessary changes. First, the modulation estimates (A.6), (A.7) and (A.8) only hold under the assumption (A.5). However, note that in Step 3 this last assumption is satisfied on the time interval (t_{l_0}, T_0), on which the modulation estimates are used, see (5.14).

Instead of (5.19), we claim that

$$\xi'_{j_l}(t_l)\sqrt{-\log \widetilde{\xi}_{j_l}(t_l)} \simeq \xi'_{j_l}(t_l)\sqrt{-\log(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))} \geq -c_0 \max_{i \in \mathcal{A}} \sqrt{\xi_i(t_l)/\lambda_{i+1}(t_l)}, \quad \text{(A.12)}$$

where c_0 can be made arbitrarily small upon taking η_0 small. Indeed, recalling that $\widetilde{\xi}_j(t) = 2^{-j}\xi_j(t)/\lambda_{j+1}(t)$, (5.18) yields

$$\xi'_{j_l}(t_l) \gtrsim -\widetilde{\xi}_{j_l}(t_l)|\lambda'_{j_l+1}(t_l)| - \frac{\lambda_{j_l+1}(t_l)}{\lambda_{j_l-1+1}(t_l)}\left(|\xi'_{j_l-1}(t_l)| + \widetilde{\xi}_{j_l-1}(t_l)|\lambda'_{j_l-1+1}(t_l)|\right).$$

Since $\widetilde{\xi}_{j_l}(t_l)$ is small when η_0 is small, (A.9) yields

$$\widetilde{\xi}_{j_l}(t_l)\sqrt{-\log \widetilde{\xi}_{j_l}(t_l)}|\lambda'_{j_l+1}(t_l)| \leq c_0 \max_{i \in \mathcal{A}} \sqrt{\xi_i(t_l)/\lambda_{i+1}(t_l)}.$$

Since $\widetilde{\xi}_{j_l}(t_l) = \widetilde{\xi}_{j_{l-1}}(t_l)$, (A.6) yields

$$\sqrt{-\log \widetilde{\xi}_{j_l}(t_l)}|\xi'_{j_{l-1}}(t_l)| \lesssim \max_{i \in \mathcal{A}} \sqrt{\xi_i(t_l)/\lambda_{i+1}(t_l)}.$$

For the same reason, and using again (A.9),

$$\sqrt{-\log \widetilde{\xi}_{j_l}(t_l)}\widetilde{\xi}_{j_{l-1}}(t_l)|\lambda'_{j_{l-1}+1}(t_l)| \lesssim \max_{i \in \mathcal{A}} \sqrt{\xi_i(t_l)/\lambda_{i+1}(t_l)}.$$

Since $j_l < j_{l-1}$, $\lambda_{j_l+1}(t_l)/\lambda_{j_{l-1}+1}(t_l)$ is small when η_0 is small, so we get (A.12).

In (5.6), (5.7) and (5.9), we replace $\mathbf{d}(t)^{k/2}$ by $\mathbf{d}(t)^2\sqrt{-\mathbf{d}(t)}$. Next, we introduce the auxiliary function $\Phi(x) := \sqrt{-x \log x}$ for $0 < x < 1$. Note that

$$\sqrt{x} \sim \Phi(x)/\sqrt{-\log \Phi(x)},$$

$$\Phi'(x) = \frac{\sqrt{-\log x}}{2\sqrt{x}} + O((-x\log x)^{-1/2}) > 0.$$

With $c_2 > 0$ to be determined, consider the auxiliary function

$$\phi(t) := \beta_{j_l}(t) + c_2 \Phi\big(\xi_{j_l}(t)/\lambda_{j_l+1}(t_l)\big).$$

The Chain Rule gives

$$\phi'(t) = \beta'_{j_l}(t) + c_2 \frac{\xi'_{j_l}(t)}{\lambda_{j_l+1}(t_l)} \Phi'\Big(\frac{\xi_{j_l}(t)}{\lambda_{j_l+1}(t_l)}\Big).$$

By (A.6) and (5.12), we have $|\xi'_{j_l}(t)| \leq C_3(\xi_{j_l}(t)/\lambda_{j_l+1}(t_l))^{\frac{1}{2}}\log(-\xi_{j_l}(t)/\lambda_{j_l+1}(t_l))^{-1/2}$, with C_3 depending only on N, hence (A.8) implies

$$\phi'(t) \geq \frac{4}{\lambda_{j_l+1}(t_l)}, \tag{A.13}$$

provided c_2 is taken small enough (depending only on N). If we consider $\widetilde{\phi}(t) := \beta_{j_l}(t) + \frac{c_2}{2}\Phi\big(\xi_{j_l}(t)/\lambda_{j_l+1}(t_l)\big)$ instead of ϕ, then the computation above shows that $\widetilde{\phi}$ is increasing. From (A.12), we have $\widetilde{\phi}(t_l) \geq 0$, so $\widetilde{\phi}(t) \geq 0$ for all $t \in (t_l, t_{l+1})$, implying

$$\mathbf{d}(t) \lesssim \sqrt{\xi_{j_l}(t)/\lambda_{j_l+1}(t_l)} \lesssim \phi(t)/\sqrt{-\log \phi(t)}. \tag{A.14}$$

The bound (A.13) yields

$$\big(\lambda_{j_l+1}(t_l)\phi(t)^2/\sqrt{-\log \phi(t)}\big)' \gtrsim \phi(t)/\sqrt{-\log \phi(t)}.$$

We observe that $|\phi(t)| \lesssim \Phi(\mathbf{d}(t)^2)$, hence $\phi(t)^2/\sqrt{-\log \phi(t)} \lesssim \mathbf{d}(t)^2\sqrt{-\log \mathbf{d}(t)}$ and

$$\int_{t_l}^{t_{l+1}} \phi(t)/\sqrt{-\log \phi(t)}\mathrm{d}t \lesssim \lambda_{j_l+1}(t_l)\phi(t_{l+1})^2/\sqrt{-\log \phi(t_{l+1})} \lesssim \mathbf{d}(t_{l+1})^2\sqrt{-\log \mathbf{d}(t_{l+1})}\lambda_{j_l+1}(t_l).$$

Thus, (A.14) yields (5.9) (with $\mathbf{d}(t)^2\sqrt{-\log \mathbf{d}(t)}$ instead of $\mathbf{d}(t)^{2/k}$) if C_0 is sufficiently large (but depending on k and N only).

We now prove (5.11). By (A.12) and (A.7), we have $\beta_{j_l}(t_l) \geq -c_0\Phi(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))$, where c_0 can be made as small as needed, and

$$\beta'_{j_l}(t) \geq c_1/\lambda_{j_l+1}(t_l).$$

We deduce that $\xi'_{j_l}(t) \geq 0$ provided

$$t - t_l \geq \frac{2c_0}{c_1}\lambda_{j_l+1}(t_l)\Phi(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l)) = \frac{2c_0}{c_1}\sqrt{\xi_{j_l}(t_l)\lambda_{j_l+1}(t_l)}\sqrt{-\log(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))}.$$

But, if the opposite inequality is satisfied, the bound

$$|\xi'_{j_l}(t)| \lesssim \sqrt{\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l)}/\sqrt{-\log(\xi_{j_l}(t_l)/\lambda_{j_l+1}(t_l))}$$

yields (5.11), if c_0 is small enough.

The proof is then finished as for $k \geq 2$. $\qquad\square$

The remaining arguments of Section 5 apply without major changes. In (5.21), one should replace $\theta_n^{2/k}$ by $\theta_n^2\sqrt{-\log\theta_n}$.

A.3. Absence of pure multi-bubbles with $N \geq 2$ when $k = 1$.
This section is devoted to proving Proposition 1.8, which follows quickly from Lemma A.4.

Proof of Proposition 1.8. Suppose that $N \geq 2$ and $\boldsymbol{u}$ is a pure N-bubble in the future time direction and let $T_+ \in \mathbb{R} \cup \{\infty\}$ be its final time of existence. Let $T_0 < T_+$ be such that $\mathbf{d}(t) < \eta_0$ for all $t \in [T_0, T_+)$, so that $U(t) < \infty$, where U is the weighted interaction energy from Definition 5.6. Since we assume $N \geq 2$, we have $U(t) > 0$ for all t.

We have $\lim_{t \to T_+} U(t) = 0$, thus there exists a sequence $t_n \to T_+$ such that t_n is a local minimal point from the left of U. By Lemma A.4, for all n and $t \in [T_0, t_n)$ we have

$$\frac{3}{4}\lambda_N(t_n) \leq \lambda_N(t) \leq \frac{4}{3}\lambda_N(t_n),$$

thus for all $t \in [T_0, T_+)$

$$\frac{9}{16}\lambda_N(T_0) \leq \lambda_N(t) \leq \frac{16}{9}\lambda_N(T_0),$$

which implies that we cannot have $\lim_{t \to T_+}\lambda_N(t) = 0$, in particular $T_+ = \infty$.

Again by Lemma A.4, for all $t \in [T_0, t_n)$ we have

$$\int_t^{t_n}\mathbf{d}(\tau)\mathrm{d}\tau \leq C_0\mathbf{d}(t)^2\sqrt{-\log\mathbf{d}(t)}\lambda_N(t_n) \leq \frac{4}{3}C_0\mathbf{d}(t)^2\sqrt{-\log\mathbf{d}(t)}\lambda_N(T_0),$$

and letting $n \to \infty$ we obtain

$$\int_t^\infty\mathbf{d}(\tau)\mathrm{d}\tau \leq \frac{4}{3}C_0\mathbf{d}(t)^2\sqrt{-\log\mathbf{d}(t)}\lambda_N(T_0) \leq \mathbf{d}(t)^{\frac{3}{2}}\lambda_N(T_0),$$

where the last bound is justified by taking η_0 small enough. The Chain Rule yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\left(\int_t^\infty\mathbf{d}(\tau)\mathrm{d}\tau\right)^{\frac{1}{3}}\right) \leq -\frac{1}{3}\lambda(T_0)^{-\frac{2}{3}}$$

for all $t \geq T_0$, which is impossible. $\qquad\square$

REFERENCES

[1] H. Bahouri and P. Gérard. High frequency approximation of solutions to critical nonlinear wave equations. *Amer. J. Math.*, 121:131–175, 1999.

[2] E. B. Bogomol' nyi. The stability of classical solutions. *Soviet J. Nuclear Phys.*, 24(4):449–454, 1976.

[3] H. Brezis and J. M. Coron. Convergence of solutions of H-systems or how to blow bubbles. *Arch. Ration. Mech. Anal.*, 89:21–56, 1985.

[4] T. Cazenave, J. Shatah, and A. Tahvildar-Zadeh. Harmonic maps of the hyperbolic space and development of singularities in wave maps and Yang-Mills fields. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 68(3):315–349, 1998.

[5] E. Chiodaroli, J. Krieger, and J. Lührmann. Concentration compactness for critical radial wave maps. *Ann. PDE*, 4(1):Paper No. 8, 148, 2018.

[6] S.-N. Chow and J. K. Hale. *Methods of Bifurcation Theory*, volume 251 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1982.

[7] D. Christodoulou and A. S. Tahvildar-Zadeh. On the asymptotic behavior of spherically symmetric wave maps. *Duke Math. J.*, 71(1):31–69, 1993.

[8] D. Christodoulou and A. S. Tahvildar-Zadeh. On the regularity of spherically symmetric wave maps. *Comm. Pure Appl. Math.*, 46(7):1041–1091, 1993.

[9] R. Côte. Instability of nonconstant harmonic maps for the $(1+2)$-dimensional equivariant wave map system. *Int. Math. Res. Not.*, (57):3525–3549, 2005.

[10] R. Côte. On the soliton resolution for equivariant wave maps to the sphere. *Comm. Pure Appl. Math.*, 68(11):1946–2004, 2015.

[11] R. Côte, C. Kenig, A. Lawrie, and W. Schlag. Characterization of large energy solutions of the equivariant wave map problem: I. *Amer. J. Math.*, 137(1):139–207, 2015.

[12] R. Côte, C. Kenig, A. Lawrie, and W. Schlag. Characterization of large energy solutions of the equivariant wave map problem: II. *Amer. J. Math.*, 137(1):209–250, 2015.

[13] R. Cote, C. Kenig, A. Lawrie, and W. Schlag. Profiles for the radial focusing 4d energy-critical wave equation. *Comm. Math. Phys.*, 357(3):934–1008, 2018.

[14] R. Côte, C. Kenig, and F. Merle. Scattering below critical energy for the $4d$ Yang-Mills equation and for the $2d$ corotational wave map system. *Comm. Math. Phys.*, 284(1):203–225, 2008.

[15] R. Côte, C. Kenig, and W. Schlag. Energy partition for the linear radial wave equation. *Math. Ann.*, 358(3-4):573–607, 2014.

[16] T. Duyckaerts, H. Jia, C. Kenig, and F. Merle. Soliton resolution along a sequence of times for the focusing energy critical wave equation. *Geom. Funct. Anal.*, 27(4):798–862, 2017.

[17] T. Duyckaerts, H. Jia, C. Kenig, and F. Merle. Universality of blow up profile for small blow up solutions to the energy critical wave map equation. *Int. Math. Res. Not. IMRN*, (22):6961–7025, 2018.

[18] T. Duyckaerts, C. Kenig, Y. Martel, and F. Merle. Soliton resolution for critical co-rotational wave maps and radial cubic wave equation. *arXiv e-prints*, 2021.

[19] T. Duyckaerts, C. Kenig, and F. Merle. Erratum: Profiles of bounded radial solutions of the focusing, energy-critical wave equation. *arXiv e-prints*, 01 2015.

[20] T. Duyckaerts, C. Kenig, and F. Merle. Scattering profile for global solutions of the energy-critical wave equation. *J. Eur. Math. Soc. (JEMS)*, 21(7):2117–2162, 2019.

[21] T. Duyckaerts, C. E. Kenig, and F. Merle. Universality of blow-up profile for small radial type II blow-up solutions of the energy-critical wave equation. *J. Eur. Math. Soc.*, 13(3):533–599, 2011.

[22] T. Duyckaerts, C. E. Kenig, and F. Merle. Profiles of bounded radial solutions of the focusing, energy-critical wave equation. *Geom. Funct. Anal.*, 22(3):639–698, 2012.

[23] T. Duyckaerts, C. E. Kenig, and F. Merle. Universality of the blow-up profile for small type II blow-up solutions of the energy-critical wave equation: the nonradial case. *J. Eur. Math. Soc.*, 14(5):1389–1454, 2012.

[24] T. Duyckaerts, C. E. Kenig, and F. Merle. Classification of the radial solutions of the focusing, energy-critical wave equation. *Camb. J. Math.*, 1(1):75–144, 2013.

[25] T. Duyckaerts, C. E. Kenig, and F. Merle. Exterior energy bounds for the critical wave equation close to the ground state. *Comm. Math. Phys.*, 379:1113–1175, 2020.

[26] T. Duyckaerts, C. E. Kenig, and F. Merle. Decay estimates for nonradiative solutions of the energy-critical focusing wave equation. *J. Geom. Anal.*, 2021.

[27] T. Duyckaerts, C. E. Kenig, and F. Merle. Soliton resolution for the radial critical wave equation in all odd space dimensions. *Acta Math.*, to appear.

[28] T. Duyckaerts and F. Merle. Dynamics of threshold solutions for energy-critical wave equation. *Int. Math. Res. Pap. IMRP*, 2008.

[29] T. Duyckaerts and F. Merle. Dynamic of threshold solutions for energy-critical NLS. *GAFA*, 18(6):1787–1840, 2009.

[30] P. Eckhaus, W.; Schuur. The emergence of solitons of the korteweg-de vries equation from arbitrary initial conditions. *Math. Methods Appl. Sci.*, 5(1):97–116, 1983.

[31] E. Fermi, J. Pasta, and S. Ulam. *S. Los Alamos Report LA-1940*, 1955.

[32] P. Gérard. Description du défaut de compacité de l'injection de Sobolev. *ESAIM Control Optim. Calc. Var.*, 3:213–233, 1998.

[33] J. Ginibre and G. Velo. Generalized Strichartz inequalities for the wave equation. *J. Funct. Anal.*, 133:50–68, 1995.

[34] R. Grinis. Quantization of time-like energy for wave maps into spheres. *Comm. Math. Phys.*, 352(2):641–702, 2017.

[35] J. Jendrej. Construction of two-bubble solutions for the energy-critical NLS. *Anal. PDE*, 10(8):1923–1959, 2017.

[36] J. Jendrej. Construction of two-bubble solutions for energy-critical wave equations. *Amer. J. Math.*, 141(1):55–118, 2019.

[37] J. Jendrej, M. Kowalczyk, and A. Lawrie. Dynamics of strongly interacting kink-antikink pairs for scalar fields on a line. *to appear in Duke Math. J.*

[38] J. Jendrej and A. Lawrie. An asymptotic expansion of two-bubble wave maps in high equivariant classes. *to appear in Anal. PDE.*

[39] J. Jendrej and A. Lawrie. Continuous time soliton resolution for two-bubble equivariant wave maps. *to appear in Math. Res. Lett.*

[40] J. Jendrej and A. Lawrie. Uniqness of two-bubble wave maps in high equivariant classes. *to appear in Comm. Pure Appl. Math.*

[41] J. Jendrej and A. Lawrie. Two-bubble dynamics for threshold solutions to the wave maps equation. *Invent. Math.*, 213(3):1249–1325, 2018.

[42] H. Jia and C. Kenig. Asymptotic decomposition for semilinear wave and equivariant wave map equations. *Amer. J. Math.*, 139(6):1521–1603, 2017.

[43] C. Kenig and F. Merle. Global well-posedness, scattering and blow-up for the energy-critical focusing nonlinear wave equation. *Acta Math.*, 201(2):147–212, 2008.

[44] J. Krieger, K. Nakanishi, and W. Schlag. Global dynamics away from the ground state for the energy-critical nonlinear wave equation. *Amer. J. Math.*, 135(4):935–965, 2013.

[45] J. Krieger, K. Nakanishi, and W. Schlag. Center-stable manifold of the ground state in the energy space for the critical wave equation. *Math. Ann.*, 361(1-2):1–50, 2015.

[46] J. Krieger, W. Schlag, and D. Tataru. Renormalization and blow up for charge one equivariant wave critical wave maps. *Invent. Math.*, 171(3):543–615, 2008.

[47] H. Lindblad and C. D. Sogge. On existence and scattering with minimal regularity for semilinear wave equations. *J. Funct. Anal.*, 130(2):357–426, 1995.

[48] P.-L. Lions. The concentration-compactness principle in the Calculus of Variations. The locally compact case, part I. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 1:79–145, 1984.

[49] P.-L. Lions. The concentration-compactness principle in the Calculus of Variations. The limit case, part II. *Rev. Mat. Iberoam.*, 1:145–201, 1985.

[50] N. Manton and P. Sutcliffe. *Topological solitons*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, Cambridge, 2004.

[51] Y. Martel and F. Merle. Description of two soliton collision for the quartic gKdV equation. *Ann. of Math. (2)*, 174(2):757–857, 2011.

[52] Y. Martel and F. Merle. Inelastic interaction of nearly equal solitons for the quartic gKdV equation. *Invent. Math.*, 183(3):563–648, 2011.

[53] Y. Martel and F. Merle. Inelasticity of soliton collisions for the 5D energy critical wave equation. *Invent. Math.*, 214(3):1267–1363, 2018.

[54] F. Merle and L. Vega. Compactness at blow-up time for L^2 solutions of the critical nonlinear Schrödinger equation in 2D. *Int. Math. Res. Not. IMRN*, (8):399–425, 1998.

[55] C. Muscalu and W. Schlag. *Classical and Multilinear Harmonic Analysis, Volume I*. Cambridge University Press, 2013.

[56] J. Nahas. Scattering of wave maps from $\mathbb{R}^{2+1}$ to general targets. *Calc. Var. Partial Differential Equations*, 46(1-2):427–437, 2013.

[57] K. Nakanishi and W. Schlag. Global dynamics above the ground state energy for the focusing nonlinear Klein-Gordon equation. *J. Differential Equations*, 250(5):2299–2333, 2011.

[58] K. Nakanishi and W. Schlag. Global dynamics above the ground state for the nonlinear Klein-Gordon equation without a radial assumption. *Arch. Ration. Mech. Anal.*, 203(3):809–851, 2011.

[59] F. Planchon, J. Stalker, and A. S. Tahvildar-Zadeh. L^p estimates for the wave equation with the inverse-square potential. *Discrete Contin. Dyn. Syst.*, 9(2):427–442, 2003.

[60] J. Qing. On singularities of the heat flow for harmonic maps from surfaces into spheres. *Comm. Anal. Geom.*, 3(297–316), 1995.

[61] J. Qing and G. Tian. Bubbling of the heat flows for harmonic maps from surfaces. *Comm. Pure Appl. Math.*, 50(4):295–310, 1997.

[62] P. Raphaël and I. Rodnianski. Stable blow up dynamics for the critical co-rotational wave maps and equivariant Yang-Mills problems. *Publ. Math. Inst. Hautes Études Sci.*, pages 1–122, 2012.

[63] P. Raphaël and J. Szeftel. Existence and uniqueness of minimal mass blow up solutions to an inhomogeneous L^2-critical NLS. *J. Amer. Math. Soc.*, 24(2):471–546, 2011.

[64] I. Rodnianski and J. Sterbenz. On the formation of singularities in the critical $O(3)$ σ-model. *Ann. of Math.*, 172:187–242, 2010.

[65] C. Rodriguez. Threshold dynamics for corotational wave maps. *to appear in Anal. PDE.*

[66] C. Rodriguez. Profiles for the radial focusing energy-critical wave equation in odd dimensions. *Adv. Differential Equations*, 21(5/6):505–570, 2016.

[67] P. C. Schuur. *Asymptotic analysis of soliton problems, an inverse scattering approach.* Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1986.

[68] H. Segur and M. J. Ablowitz. Asymptotic solutions and conservation laws for the nonlinear Schrödinger equation I. *J. Math. Phys.*, 17(5):710–713, 1976.

[69] J. Shatah and M. Struwe. *Geometric wave equations.* Courant Lecture notes in Mathematics, New York University, Courant Institute of Mathematical Sciences, New York. American Mathematical Society, Providence RI, 1998.

[70] J. Shatah and A. Tahvildar-Zadeh. Regularity of harmonic maps from the Minkowski space into rotationally symmetric manifolds. *Comm. Pure Appl. Math.*, 45(8):947–971, 1992.

[71] J. Shatah and A. S. Tahvildar-Zadeh. On the Cauchy problem for equivariant wave maps. *Comm. Pure Appl. Math.*, 47(5):719–754, 1994.

[72] M. Struwe. On the evolution of harmonic mappings of Riemannian surfaces. *Comment. Math. Helv.*, 60(4):558–581, 1985.

[73] M. Struwe. Equivariant wave maps in two space dimensions. *Comm. Pure Appl. Math.*, 56(7):815–823, 2003.

[74] Michael Struwe. Radially symmetric wave maps from $(1 + 2)$-dimensional Minkowski space to the sphere. *Math. Z.*, 242(3):407–414, 2002.

[75] Michael Struwe. Radially symmetric wave maps from $(1+2)$-dimensional Minkowski space to general targets. *Calc. Var. Partial Differential Equations*, 16(4):431–437, 2003.

[76] T. Tao. Global regularity of wave maps II: Small energy in two dimensions. *Comm. Math. Phys.*, 224(2):443–544, 2001.

[77] T. Tao. A (concentration-)compact attractor for high-dimensional non-linear Schrödinger equations. *Dyn. Partial Differ. Equ.*, 4(1):1–53, 2007.

[78] P. Topping. Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow. *Ann. of Math.*, 159(2):465–534, 2004.

[79] P. M. Topping. Rigidity in the harmonic map heat flow. *J. Diff. Geom.*, 45(3):593–610, 1997.

[80] N. J. Zabusky and M. D. Kruskal. Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. *Physical Review Letters*, 15(6):240, 1965.

JACEK JENDREJ

CNRS and LAGA, Université Sorbonne Paris Nord
99 av Jean-Baptiste Clément, 93430 Villetaneuse, France
jendrej@math.univ-paris13.fr

ANDREW LAWRIE

Department of Mathematics, Massachusetts Institute of Technology
77 Massachusetts Ave, 2-267, Cambridge, MA 02139, U.S.A.
alawrie@mit.edu