Annals of PDE (2023) 9:18
https://doi.org/10.1007/s40818-023-00159-4

MANUSCRIPT

Soliton Resolution for the Energy-Critical Nonlinear Wave Equation in the Radial Case

Jacek Jendrej[1] · **Andrew Lawrie**[2]

Received: 28 April 2022 / Accepted: 6 September 2023

Abstract
We consider the focusing energy-critical nonlinear wave equation for radially symmetric initial data in space dimensions $D \geq 4$. This equation has a unique (up to sign and scale) nontrivial, finite energy stationary solution W, called the ground state. We prove that every finite energy solution with bounded energy norm resolves, continuously in time, into a finite superposition of asymptotically decoupled copies of the ground state and free radiation.

Keywords Soliton resolution · Multi-soliton · Wave maps · Energy-critical

Mathematics Subject Classification 35L71 (primary) · 35B40 · 37K40

Contents

1 Introduction
  1.1 Setting of the Problem
  1.2 Statement of the Results
  1.3 Summary of the Proof
  1.4 Notational Conventions
2 Preliminaries
  2.1 Virial Identities
  2.2 Local Cauchy Theory

J.Jendrej is supported by ANR-18-CE40-0028 project ESSED. A. Lawrie is supported by NSF grant DMS-1954455, a Sloan Research Fellowship, and the Solomon Buchsbaum Research Fund.

✉ Jacek Jendrej
jendrej@math.univ-paris13.fr

Andrew Lawrie
alawrie@mit.edu

[1] CNRS and LAGA, Université Sorbonne Paris Nord, 99 av Jean-Baptiste Clément, 93430 Villetaneuse, France

[2] Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Ave, 2-267, Cambridge, MA 02139, USA

Published online: 06 October 2023

2.3 Profile Decomposition
2.4 Multi-bubble Configurations
3 Localized Sequential Bubbling
3.1 Technical Lemmas
3.2 Proof of the Compactness Lemma
4 The Sequential Decomposition
4.1 Identification of the Radiation
4.2 Non-concentration of Self-similar Energy
4.3 The Sequential Decomposition
5 Decomposition of the Solution and Collision Intervals
5.1 Proximity to a Multi-bubble and Collisions
5.2 Basic Modulation
5.3 Refined Modulation
6 Conclusion of the Proof
6.1 The Scale of the K-th Bubble
6.2 Demolition of the Multi-bubble
6.3 End of the Proof: Virial Inequality with a Cut-Off
6.4 Absence of Elastic Collisions
Appendix A. Modifications to the Argument in the Case $D = 5$
A.1 Decomposition of the Solution
A.2 Conclusion of the Proof
Appendix B. Modifications to the Argument in the Case $D = 4$
B.1 Decomposition of the Solution
B.2 Conclusion of the Proof
References

1 Introduction

1.1 Setting of the Problem

We study the Cauchy problem for the focusing nonlinear wave equation in the energy-critical case and under the assumption of radial symmetry, i.e.,

$$
\begin{aligned}
\partial_t^2 u - \partial_r^2 u - \frac{D-1}{r}\partial_r u - |u|^{\frac{4}{D-2}} u = 0,\\
(u(T_0), \partial_t u(T_0)) = (u_0, \dot{u}_0),
\end{aligned}
\tag{1.1}
$$

where here $D \in \{4, 5, 6, \dots\}$ is the underlying spatial dimension, $u = u(t, r) \in \mathbb{R}$, where $r = |x| \in (0, \infty)$ is the radial coordinate in $\mathbb{R}^D$, and $T_0 \in \mathbb{R}$.

The conserved energy for (1.1) is given by

$$
\begin{aligned}
E(u(t), \partial_t u(t)) := \int_0^\infty \frac{1}{2}\Big[(\partial_t u(t))^2 + (\partial_r u(t))^2\Big] r^{D-1}\, \mathrm{d}r\\
- \int_0^\infty \frac{D-2}{2D} |u(t)|^{\frac{2D}{D-2}}\, r^{D-1}\, \mathrm{d}r.
\end{aligned}
$$

The Cauchy problem for (1.1) can be rephrased as a Hamiltonian system. To formulate it as such, we will write pairs of functions using boldface, $\boldsymbol{v} = (v, \dot{v})$, noting that the notation $\dot{v}$ does not, in general, refer to the time derivative of v but just to the second component of the vector $\boldsymbol{v}$. We see that (1.1) is equivalent to

$$\partial_t \boldsymbol{u}(t) = J \circ \mathrm{D}E(\boldsymbol{u}(t)), \qquad \boldsymbol{u}(T_0) = \boldsymbol{u}_0, \tag{1.2}$$

where

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

Above we have introduced the notation $f(u) := |u|^{\frac{4}{D-2}} u$.

Solutions to (1.1) are invariant under the scaling,

$$\boldsymbol{u}(t, r) \mapsto \boldsymbol{u}_\lambda(t, r) := (\lambda^{-\frac{D-2}{2}} u(t/\lambda, r/\lambda), \lambda^{-\frac{D}{2}} \partial_t u(t/\lambda, r/\lambda)), \quad \lambda > 0,$$

and (1.1) is called *energy-critical* because $E(\boldsymbol{u}) = E(\boldsymbol{u}_\lambda)$.

The linearization of (1.1) about the zero solution is the free scalar wave equation,

$$\partial_t^2 v - \partial_r^2 v - \frac{D-1}{r}\partial_r v = 0. \tag{1.3}$$

In this paper we study solutions with initial data in the energy space $\mathcal{E}$, which is defined via the norm,

$$\|\boldsymbol{u}_0\|_{\mathcal{E}}^2 := \int_0^\infty \Big[(\dot{u}_0(r))^2 + (\partial_r u_0(r))^2 + \frac{(u_0(r))^2}{r^2}\Big] r^{D-1}\, \mathrm{d}r.$$

Using Hardy's inequality, functions $\boldsymbol{u}(r)$ in $\mathcal{E}$ can be identified with radially symmetric functions $\boldsymbol{v}(x)$ in the space $\dot{H}^1 \times L^2(\mathbb{R}^D)$ in the usual way. We will sometimes use the notation,

$$\|u_0\|_H^2 := \int_0^\infty \Big[(\partial_r u_0(r))^2 + \frac{(u_0(r))^2}{r^2}\Big] r^{D-1}\, \mathrm{d}r,$$

and write $\mathcal{E} = H \times L^2$.

It is a classical result of Ginibre and Velo [32] that (1.1) is well-posed in the space $\mathcal{E}$. Solutions $\boldsymbol{u}(t)$ are defined in the Duhamel sense; see Section 2.2. To every $\boldsymbol{u}_0 \in \mathcal{E}$, viewed as initial data for (1.2) at time $T_0 = 0$, we can associate maximal forward and backward times of existence $T_+ \in (0, \infty]$ and $T_- \in [-\infty, 0)$, a maximal interval of existence $I_{\max} = (T_-, T_+)$ on which $\boldsymbol{u} \in C(I_{\max}; \mathcal{E})$ and $u \in S(J) \cap W(J)$ for every compact subinterval $J \subset I_{\max}$, where

$$\begin{aligned} S(J) &:= L^{\frac{2(D+1)}{D-2}}(J \times \mathbb{R}^D), \\ W(J) &:= L^{\frac{2(D+1)}{D-1}}\Big(J; \dot{B}^{\frac{1}{2}}_{\frac{2(D+1)}{D-1},2}(\mathbb{R}^D)\Big); \end{aligned}$$

see Section 2.2 for details. We will only consider solutions $\boldsymbol{u}(t) \in \mathcal{E}$ to (1.1) that satisfy,

$$\limsup_{t \to T_+} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty \quad \text{or} \quad \limsup_{t \to T_-} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

Solutions for which $\lim_{t \to T_+} \|\boldsymbol{u}(t)\|_{\mathcal{E}} = \infty$ are known to exist and are called type-I (or ODE-type) blow up solutions; see e.g., [5, 15, 51].

We define the Aubin-Talenti solution, $\boldsymbol{W}(x) := (W(x), 0)$ where $W : \mathbb{R}^D \to \mathbb{R}$, by

$$W(x) := \Big(1 + \frac{|x|^2}{D(D-2)}\Big)^{-\frac{D-2}{2}},$$

and note that $W(x)$ is the unique (up to sign, scaling, and translation), *non-negative* and nontrivial C^2 solution to

$$-\Delta W(x) = |W(x)|^{\frac{4}{D-2}} W(x), \quad x \in \mathbb{R}^D.$$

Abusing notation slightly and writing $W(x) = W(r)$ with $r = |x|$, we see that $\boldsymbol{W}(r)$ is a stationary solution to (1.1). In fact, it is the unique (up to sign and scaling) static radial nontrivial solution to (1.1) in $\mathcal{E}$. For each $\lambda > 0$, we write $\boldsymbol{W}_\lambda(r) := (\lambda^{-\frac{D-2}{2}} W(r/\lambda), 0)$.

1.2 Statement of the Results

Our main result is formulated as follows.

Theorem 1 (Soliton Resolution) *Let $D \geq 4$ and let $\boldsymbol{u}(t)$ be a finite energy solution to* (1.1) *with initial data $\boldsymbol{u}(0) = \boldsymbol{u}_0 \in \mathcal{E}$, defined on its maximal forward interval of existence* $[0, T_+)$. *Suppose that,*

$$\limsup_{t \to T_+} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

Then,

(Global solution) *if $T_+ = \infty$, there exist a time $T_0 > 0$, a solution $\boldsymbol{u}_{\mathrm{L}}^* \in C(\mathbb{R}; \mathcal{E})$ to the linear wave equation* (1.3)*, an integer $N \geq 0$, continuous functions $\lambda_1(t), \dots, \lambda_N(t) \in C^0([T_0, \infty))$, signs $\iota_1, \dots, \iota_N \in \{-1, 1\}$, and $\boldsymbol{g}(t) \in \mathcal{E}$ defined by*

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

such that

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

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

(Blow-up solution) *if* $T_+ < \infty$, *there exists a time* $T_0 < T_+$, *a function* $\boldsymbol{u}_0^* \in \mathcal{E}$, *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*

$$\boldsymbol{u}(t) = \sum_{j=1}^{N} \iota_j \boldsymbol{W}_{\lambda_j(t)} + \boldsymbol{u}_0^* + \boldsymbol{g}(t),$$

such that

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

where 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. Theorem 1 has been proved for (1.1) in a series of remarkable works by Duyckaerts, Kenig, and Merle in odd space dimensions $D \geq 3$; see [22] for dimension $D = 3$ and see [24–26] for odd space dimensions $D \geq 5$. The space dimension $D = 4$ was treated by Duyckaerts, Kenig, Martel, and Merle in [18], which also covers the 1-equivariant wave maps equation, and dimension $D = 6$ was treated by Collot, Duyckaerts, Kenig, and Merle [10]. All of these papers use, in some fashion, the method of energy channels. Roughly, energy channels refer to measurements of the portion of energy that a linear or nonlinear wave radiates outside fattened light cones $\{|x| > |t| + R\}$ for $R \geq 0$. The approach we take to prove Theorem 1 is independent of the method of energy channels. Our method of proof follows closely our recent preprint [41], which proved the analogue of Theorem 1 for the k-equivariant wave maps equation in all equivariance classes $k \in \mathbb{N}$.

Remark 1.2 The soliton resolution problem is inspired by the theory of completely integrable systems, e.g., [29, 72, 75], motivated by numerical simulations, [30, 88], and by the bubbling theory of harmonic maps in the elliptic and parabolic settings [66, 67, 83, 86, 87]; see also [16, 18, 26] for discussions on the history of the problem.

Remark 1.3 Equation 1.1, its counterpart in defocusing case, as well as the sub- and super-critical versions, have been classically studied; see for example the articles [3, 34, 43, 44, 59, 60, 64, 65, 69, 73, 74, 76, 77, 82, 84] and the books by Strauss [81], Sogge [80], and Statah, Struwe [78].

Remark 1.4 Kenig and Merle [45] gave the first general description of dynamics in a non-perturbative setting for the focusing energy critical NLW (non-radial), proving that solutions $\boldsymbol{u}$ with energy below the ground state energy scatter in both directions if $\|\nabla u_0\|_{L^2}^2 < \|\nabla W\|_{L^2}^2$ or blow up in finite time in both directions if $\|\nabla u_0\|_{L^2}^2 > \|\nabla W\|_{L^2}^2$. In [27] Duyckaerts and Merle classified the dynamics of solutions at the threshold energy $E = E(\boldsymbol{W})$. Characterizations of the dynamics for energies slightly above the ground state energy were given by Krieger, Nakanishi, and Schlag in [46, 47].

Remark 1.5 Theorem 1 is a qualitative description of the dynamics of all finite energy solutions to (1.1) with bounded critical norm. A natural, challenging question is to ask which types of configurations of solitons and radiation are realized in solutions. The first results in this direction were by Krieger and Schlag [48] who found a manifold of global-in-time solutions that decoupled into a static $\boldsymbol{W}$ and free radiation; see also the improvement by Beceanu [4]. The first construction of a solution developing a bubbling singularity (with one concentrating bubble) in finite time was done by Krieger, Schlag, and Tataru [50]; see also Hillairet and Raphael [35] for a different construction in dimension $D = 4$, and also [49].

In [37], 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 dimension $D = 6$. We expect that solutions that form 2-bubble in forward infinite time also exist in dimensions $D \geq 7$.

When multi-bubble solutions do occur in one time direction, it is natural to ask about the dynamics of those solutions in the opposite time direction. We answered this question in [40] in the setting of equivariant wave maps for the pure 2-bubble solution $\boldsymbol{u}_{(2)}$ constructed by the first author in [37]. 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*. Viewed 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, 39] that $\boldsymbol{u}_{(2)}(t)$ is the unique 2-bubble solution up to sign, translation, and scaling in equivariance classes $k \geq 4$. 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.

Before stating the result, we define pure multi-bubbles in forward or backward time.

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 $t \mapsto (u(-t), -\dot{u}(-t))$ is a pure multi-bubble in the forward time direction.

Proposition 1.7 (No elastic collisions of pure multi-bubbles) *Stationary solutions are the only pure multi-bubbles in both time directions.*

Remark 1.8 We note that Proposition 1.7 was also proved for odd dimensions $D \geq 3$ in [23–26] and in dimensions $D = 4, 6$ in [10, 18] by the method of energy channels. The case of $k = 1$ equivariant wave maps was treated in [18] using energy channels. Here the proof of Proposition 1.7 follows from the method introduced by the authors in [41] where we treated equivariant wave maps for all equivariance classes $k \geq 1$. See [55–57] for more regarding the inelastic soliton collision problem for non-integrable PDEs.

1.3 Summary of the Proof

The proof of Theorem 1 is built on two significant partial results; (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 (at least in the case of certain space dimensions). Because the existing literature does not cover all space dimensions, we sketch a unified proof of the sequential soliton resolution (see Theorem 1.14 below) as a consequence of what we call the Compactness Lemma (see Lemma 3.1, which is also used crucially in the proof of the main theorem), the identification of the radiation, and the fact that no energy can concentrate in the self-similar region of the light cone.

We discuss these results in more detail. To unify the blow-up and global-in-time settings we make the following conventions. Consider a finite energy solution $\boldsymbol{u}(t) \in \mathcal{E}$. 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. We assume that $\boldsymbol{u}(t)$ exhibits type II behavior in that,

$$\limsup_{t \to T_*} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty. \tag{1.4}$$

Step 1: Extraction of the radiation. Below we will use the notation $\mathcal{E}(r_1, r_2)$ to denote the localized energy norm

$$\|\boldsymbol{g}\|_{\mathcal{E}(r_1,r_2)}^2 := \int_{r_1}^{r_2} \left((\dot{g})^2 + (\partial_r g)^2 + \frac{g^2}{r^2} \right) r^{D-1} \mathrm{d}r. \tag{1.5}$$

By convention, $\mathcal{E}(r_0) := \mathcal{E}(r_0, \infty)$ for $r_0 > 0$. The local nonlinear energy is denoted by

$$E(\boldsymbol{u}_0; r_1, r_2) := \int_{r_1}^{r_2} \frac{1}{2}\Big[(\dot{u}_0)^2 + (\partial_r u_0)^2\Big] r^{D-1} \,\mathrm{d}r - \int_{r_1}^{r_2} \frac{D-2}{2D} |u_0|^{\frac{2D}{D-2}} \, r^{D-1} \,\mathrm{d}r.$$

We adopt similar conventions as for $\mathcal{E}$ regarding the omission of r_2, or both r_1 and r_2.

Theorem 1.9 (Properties of the radiation) [12, 21, 22, 42, 71] *Let $\boldsymbol{u} \in C(I_*; \mathcal{E})$ be a finite energy solution to* (1.1) *on an interval I_* as above and such that* (1.4) *holds. Then,* *there exists an open neighborhood J of T_* and finite energy solution $\boldsymbol{u}^*(t) \in C(J; \mathcal{E})$ to (1.1) called the radiation, and a function $\rho : I_* \to (0, \infty)$ that satisfy,*

$$\lim_{t \to T_*} \left((\rho(t)/t)^{\frac{D-2}{2}} + \|\boldsymbol{u}(t) - \boldsymbol{u}^*(t)\|^2_{\mathcal{E}(\rho(t))} \right) = 0,$$

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

$$\|\boldsymbol{u}^*(t)\|_{\mathcal{E}(0,\alpha t)} \to 0 \ \text{ as } \ t \to T_*.$$

Remark 1.10 In the global setting, i.e., $I_* = [T_0, \infty)$ the linear wave $\boldsymbol{u}^*_{\mathrm{L}} \in C(\mathbb{R}; \mathcal{E})$ that appears in Theorem 1 is the unique solution to the linear equation (1.3) satisfying,

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

see Proposition 4.2. In the finite time blow-up setting the final radiation $\boldsymbol{u}^*_0 \in \mathcal{E}$ that appears in Theorem 1 can be viewed as initial data for $\boldsymbol{u}^*(t)$, i.e., the radiation $\boldsymbol{u}^*(t)$ in Theorem 1.9 satisfies $\boldsymbol{u}(t, r) = \boldsymbol{u}^*(t, r)$ for $r > t$. We refer the reader to Section 4.3 for a sketch of the proof of Theorem 1.9 following the scheme of Duyckaerts, Kenig, and Merle [22] (see also the preliminary results in Sections 4.1 for the identification of the radiation and Section 4.2 for non-concentration of energy in the self-similar region of the cone, which follow the methods of [12, 42] by Côte, Kenig, the second author, and Schlag, and by Jia and Kenig).

Remark 1.11 The radiation field for (1.1) can be identified even outside radial symmetry; see the work of Duyckaerts, Kenig, and Merle [20]. The radiation field can be identified in several other contexts and by different means. For example, Tao accomplished this in [85] for certain high dimensional NLS.

Step 2: Sequential soliton resolution. A deep insight of Duyckaerts, Kenig, and Merle, proved in [22] for $D = 3$, 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 copies of the ground state, 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 [2]. However, after finding the profiles and their space-time concentration properties (in our case their scales) via the main result in [2], one must identify them as elliptic objects (solitons) by some means, and then prove that the error vanishes in the energy space, rather than the weaker form of compactness (vanishing of certain Strichartz norms) given by [2]. This was proved in the breakthrough paper [22] using linear energy channels (amongst other techniques). Rodriguez [71] extended this result to all odd dimensions $D \geq 3$. It was shown in [12, 13] that the scheme of proof from [22] could be extended to the subset of even space dimensions $D \equiv 0 \pmod 4$. Jia and Kenig then proved the sequential soliton resolution result for dimension $D = 6$ using a different scheme based on virial inequalities rather than energy channels. We follow the Jia-Kenig approach here to prove a general result, which we call the

Compactness Lemma 3.1, which we then combine with Theorem 1.9 to give a unified proof of the sequential resolution in all space dimensions $D \geq 4$; for the latter, see Section 4.3.

Before stating the sequential resolution result, we introduce some notation.

Definition 1.12 (Multi-bubble configuration) Given $M \in \{0, 1, \ldots\}$, $\vec{\iota} = (\iota_1, \ldots, \iota_M) \in \{-1, 1\}^M$ and an increasing sequence $\vec{\lambda} = (\lambda_1, \ldots, \lambda_M) \in (0, \infty)^M$, a *multi-bubble configuration* is defined by the formula

$$\mathcal{W}(\vec{\iota}, \vec{\lambda}; r) := \sum_{j=1}^{M} \iota_j \boldsymbol{W}_{\lambda_j}(r).$$

Remark 1.13 If $M = 0$, it should be understood that $\mathcal{W}(\vec{\iota}, \vec{\lambda}; r) = 0$ 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.

Theorem 1.14 (Sequential soliton resolution) [22, Theorems 1 and 4], [12, Theorems 1.1 and 1.3] [42, Theorem 1.1] [71, Theorems 1.1 and 1.3] *Let $\boldsymbol{u} \in C(I_*; \mathcal{E})$ be a finite energy solution to (1.1) on an interval I_* as above. Let the radiation $\boldsymbol{u}^*$ be as in Theorem 1.9. Then, there exists an integer $N \geq 0$, a sequence of times $t_n \to T_*$, a vector of signs $\vec{\iota} \in \{-1, 1\}^N$, and a sequence of scales $\vec{\lambda}_n \in (0, \infty)^N$ such that,*

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

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

Remark 1.15 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].

Step 3: Collision intervals and no-return analysis. The challenging nature of bridging the gap between Theorem 1.14, 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 1.14 gives no relationship between the lengths of the time intervals $[t_n, t_{n+1}]$ and the concentration scales $\vec{\lambda}_n$ of the bubbles 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 $\boldsymbol{u} \in C(I_*; \mathcal{E})$, a finite energy solution to (1.1) on the time interval I_* as defined above. Let $N \geq 0$ and the radiation $\boldsymbol{u}^*$ be as in Theorem 1.14. We define a *multi-bubble proximity function* at each $t \in I_*$ by

$$\mathbf{d}(t) := \inf_{\vec\iota,\vec\lambda}\Big(\|\boldsymbol{u}(t)-\boldsymbol{u}^*(t)-\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda)\|_{\mathcal{E}}^2+\sum_{j=1}^{N}\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}}\Big)^{\frac{1}{2}}, \tag{1.6}$$

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

With this notation, we see that Theorem 1.14 gives a monotone sequence of times $t_n\to T_*$ such that,

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

Theorem 1 is a direct 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 $\boldsymbol{u}-\boldsymbol{u}^*$ 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 [27, 28], Nakanishi and Schlag [62, 63], and Krieger, Nakanishi and Schlag [46, 47]. In those works a key role is played by exponential instability, where here we have in addition attractive nonlinear interactions between the solitons. This latter consideration, and indeed the overall scheme of the proof is based on our previous works [40, 41]. We remark that the argument in [40] marks the first time where modulation analysis of bubble interactions was used in the context of the soliton resolution problem.

The basic tool we use is the standard virial functional

$$\mathfrak{v}(t) := \Big\langle \partial_t u(t)\mid \chi_{\rho(t)}\Big(r\partial_r u(t)+\frac{D-2}{2}u(t)\Big)\Big\rangle,$$

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 [62, 63]). Here the inner product is,

$$\langle\phi\mid\psi\rangle := \int_0^\infty \phi(r)\psi(r)\,r^{D-1}\mathrm{d}r,\qquad \text{for }\phi,\psi:(0,\infty)\to\mathbb{R}. \tag{1.7}$$

Differentiating $\mathfrak{v}(t)$ in time we have,

$$\mathfrak{v}'(t) = -\int_0^\infty |\partial_t u(t,r)|^2\,\chi_{\rho(t)}(r)\,r^{D-1}\,\mathrm{d}r+\Omega_{\rho(t)}(\boldsymbol{u}(t)), \tag{1.8}$$

where $\Omega_{\rho(t)}(\boldsymbol{u}(t))$ is the error created by the cut-off. Importantly, this error has structure, see Lemmas 2.1 and 5.19, and satisfies the estimates,

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

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 $\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda)$.

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 nonlinear waves $\boldsymbol{u}_n(t) \in \mathcal{E}$ on time intervals $[0, \tau_n]$ with bounded energy, and a sequence $R_n \to \infty$ such that

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

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 give a caricature of the no-return analysis, pointing the reader to the technical arguments in Sections 5, 6 for the actual arguments. We would like to integrate (1.8) 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 (1.8) we obtain,

$$\begin{aligned} \int_{a_n}^{b_n} \int_0^{\rho(t)} |\partial_t u(t, r)|^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}t &\lesssim \rho(a_n)\mathbf{d}(a_n) + \rho(b_n)\mathbf{d}(b_n) \\ &\quad + \int_{a_n}^{b_n} \left|\Omega_{\rho(t)}(\boldsymbol{u}(t))\right| \, \mathrm{d}t. \end{aligned} \tag{1.9}$$

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,

$$\int_{c_n}^{d_n} \int_0^{\rho(t)} |\partial_t u(t, r)|^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}t \gtrsim d_n - c_n, \tag{1.10}$$

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 boundary terms $\rho(a_n)\mathbf{d}(a_n), \rho(b_n)\mathbf{d}(b_n) \ll d_n - c_n$. For 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 5.1. These allow us to distinguish between "interior" bubbles that come into collision, and "exterior" bubbles that 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 $K \in \{1, \dots, N\}$, we say that an interval $[a, b]$ is a collision interval with parameters $0 < \epsilon < \eta$ and $N - K$ exterior bubbles for some $1 \le K \le N$, if $\mathbf{d}(a), \mathbf{d}(b) \le \epsilon$, there exists $c \in [a, b]$ with $\mathbf{d}(c) \ge \eta$, 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 (1.6) (a localized version of $\mathbf{d}(t)$); see Defintion 5.4. We write in this case $[a, b] \in \mathcal{C}_K(\epsilon, \eta)$.

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 $[a_n, b_n] \in \mathcal{C}_K(\epsilon_n, \eta)$; see Section 5.1 for the proof that K is well-defined and ≥ 1, under the contradiction hypothesis.

We revisit (1.9) 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 $\vec{\lambda} = (\lambda_1, \dots, \lambda_K)$ and the exterior scales, which stay coherent, by $\vec{\mu} = (\vec{\mu}_{K+1}, \dots, \vec{\mu}_N)$. 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-norm considerations; see Definition 6.1. 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 6.3. Thus the first terms on the right-hand-side of (1.9) can be absorbed using (1.10) 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

$$
\begin{aligned}
&\boldsymbol{u}(t, r) = \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}(t); r) + \boldsymbol{g}(t, r), \quad \text{if } r \leq \rho(t), \\
&\left\langle \Lambda W_{\lambda_j(t)} \mid g(t) \right\rangle = 0, \quad \text{for} \quad j = 1, \dots, K,
\end{aligned}
$$

where $\Lambda := r\partial_r + \frac{D-2}{2}$ is the generator of the H-invariant scaling (note that for $D = 4, 5, 6$ the decomposition is slightly different due to the slow decay of ΛW). The orthogonality conditions and an expansion of the nonlinear energy of $\boldsymbol{u}(t)$ up to scale $\rho(t)$ lead to the coercivity estimate,

$$
\begin{aligned}
\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{j \notin \mathcal{S}} \left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)} \right)^{\frac{D-2}{4}} &\lesssim \max_{i \in \mathcal{S}} \left(\frac{\lambda_i(t)}{\lambda_{i+1}(t)} \right)^{\frac{D-2}{4}} \\
&+ \max_{1 \leq i \leq K} \left| a_i^{\pm}(t) \right| + o_n(1) \simeq \mathbf{d}(t) + o_n(1),
\end{aligned}
$$

where $\mathcal{S} = \{j \in 1, \dots, K-1 \,:\, \iota_j = \iota_{j+1}\}$ captures the non-alternating bubbles (which experience an attractive interaction force). The terms $a_j^{\pm}(t)$ on the right-hand side above are, roughly speaking, the projections of $\boldsymbol{g}(t)$ onto the unstable/stable directions related to the unique, simple negative eigenvalue associated to the linearization about W. 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 $\mathbf{d}(t) \geq \epsilon_n$ on the modulation intervals.

The growth of $\mathbf{d}(t)$ is then captured by the dynamics of adjacent bubbles with the same sign, or by the dynamics along the unstable/stable directions $a_j^{\pm}(t)$. In the case when the dynamics is driven by bubble interactions, precise information enters at the level of $\lambda_j''(t)$, since (1.1) is second order. 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 $\lambda_j' \simeq -\iota_j \|\Lambda W\|_{L^2}^{-2} \lambda_j^{-1} \langle \Lambda W_{\lambda_j} \mid \dot g\rangle$, defining

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

where $\underline{A}(\lambda)$ is a truncated (to scale λ) version of $\underline{\Lambda} = r\partial_r + \frac{D}{2}$, the generator of L^2 scaling. Roughly, we show in Sections 5.3 and 6.2, that $(\lambda_j(t), \beta_j(t), a_j^{\pm}(t))$ satisfy a system of differential inequalities that can be used to control the growth of $\mathbf{d}(t)$ until the solution exits a modulation interval. All the while, the Kth scale $\lambda_K(t)$ does not move much, and we obtain bounds of the form $\int_{a_n}^{c_n} \mathbf{d}(t) \mathrm{d}t \leq C_0\big(\mathbf{d}(a_n)^{\min\left(1, \frac{4}{D-2}\right)} \lambda_K(a_n) + \mathbf{d}(c_n)^{\min\left(1, \frac{4}{D-2}\right)} \lambda_K(c_n)\big)$, and an analogous bound on the interval $[d_n, b_n]$ (see the "ejection" Lemma 6.5). Thus the errors can be absorbed into the left-hand side of (1.9) and we obtain a contradiction. In dimensions $D \geq 5$, this proof follows closely the scheme from [41] together with an elegant "weighted sum" trick from [26, Section 6], which simplifies some of the ODE analysis from [41]; see Section 6.2. The analysis in dimension $D = 4$ is more complicated, due to the fact that the modulation inequalities for the jth scale are only valid on subintervals where the ratio $(\lambda_j(t)/\lambda_{j+1}(t))^{\frac{D-2}{4}}$ is comparable to $\mathbf{d}(t)$, and thus a weighted sum trick involving the dynamics of all the bubbles at once does not seem to apply. For this special case we introduce an induction scheme together with the notion of an "ignition condition", (see Definition B.4) which identifies the most relevant controllable index j on a given subinterval of the modulation interval; see Appendix B.

While other aspects of the proof adapt readily to dimension $D = 3$, this refined analysis of the modulation parameters introduces significant complications due to the slow decay of ΛW.

A similar, but simpler refined modulation analysis was performed in [40]. The use of refinements to modulation parameters to obtain dynamical control of interacting bubbles for an energy-critical equation was introduced by the first author in the context of a two-bubble construction for NLS in [36]. The notion of localized virial corrections to modulation parameters was first introduced by Raphaël and Szeftel in [68] in a different context.

1.4 Notational Conventions

The nonlinear energy is denoted E, $\mathcal{E}$ is the energy space.

Given a function $\phi(r)$ and $\lambda > 0$, we denote by $\phi_\lambda(r) = \lambda^{-\frac{D-2}{2}}\phi(r/\lambda)$, the H-invariant re-scaling, and by $\phi_{\underline{\lambda}}(r) = \lambda^{-\frac{D}{2}}\phi(r/\lambda)$ the L^2-invariant re-scaling. We denote by $\Lambda := r\partial_r + \frac{D-2}{2}$ and $\underline{\Lambda} := r\partial r + \frac{D}{2}$ the infinitesimal generators of these scalings. We denote $\langle \cdot \mid \cdot \rangle$ the radial $L^2(\mathbb{R}^D)$ inner product given by (1.7).

We denote by $f(u) := |u|^{\frac{4}{D-2}}u$ 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 dimension D 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 Virial Identities

We have the following virial identities.

Lemma 2.1 (Virial identities) *Let $\boldsymbol{u} \in C(I; \mathcal{E})$ 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 $t \in I$,*

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \partial_t u(t) \mid \chi_{\rho(t)}\, r\partial_r u(t)\right\rangle = &-\frac{D}{2}\int_0^\infty (\partial_t u(t,r))^2 \chi_{\rho(t)}(r)\, r^{D-1}\,\mathrm{d}r \\
&+\frac{D-2}{2}\int_0^\infty \left[(\partial_r u(t,r))^2 - |u(t,r)|^{\frac{2D}{D-2}}\right] \\
&\times \chi_{\rho(t)}(r)\, r^{D-1}\,\mathrm{d}r + \Omega_{1,\rho(t)}(\boldsymbol{u}(t)),
\end{aligned} \tag{2.1}
$$

and,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle \partial_t u(t) \mid \chi_{\rho(t)}\, u(t)\right\rangle &= \int_0^\infty (\partial_t u(t,r))^2 \chi_{\rho(t)}(r)\, r^{D-1}\,\mathrm{d}r \\ &\quad - \int_0^\infty \Big[(\partial_r u(t,r))^2 - |u(t,r)|^{\frac{2D}{D-2}}\Big]\chi_{\rho(t)}(r)\, r^{D-1}\,\mathrm{d}r \\ &\quad + \Omega_{2,\rho(t)}(\boldsymbol{u}(t))\end{aligned} \tag{2.2}$$

where

$$\begin{aligned}\Omega_{1,\rho(t)}(\boldsymbol{u}(t)) &:= -\frac{\rho'(t)}{\rho(t)}\int_0^\infty \partial_t u(t,r) r\partial_r u(t,r)(r\partial_r\chi)(r/\rho(t))\, r^{D-1}\,\mathrm{d}r \\ &\quad - \frac{1}{2}\int_0^\infty \Big((\partial_t u(t,r))^2 + (\partial_r u(t,r))^2 - \frac{D-2}{D}|u(t,r)|^{\frac{2D}{D-2}}\Big) \\ &\quad \times (r\partial_r\chi)(r/\rho(t)) r^{D-1}\,\mathrm{d}r, \\ \Omega_{2,\rho(t)}(\boldsymbol{u}(t)) &:= -\frac{\rho'(t)}{\rho(t)}\int_0^\infty \partial_t u(t,r) u(t,r)(r\partial_r\chi)(r/\rho(t))\, r^{D-1}\,\mathrm{d}r \\ &\quad - \int_0^\infty \partial_r u(t,r)\frac{u(t,r)}{r}(r\partial_r\chi)(r/\rho(t))\, r^{D-1}\,\mathrm{d}r.\end{aligned} \tag{2.3}$$

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

Remark 2.2 In practice we will make use of the following two linear combinations of the identities (2.1) and (2.2).

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big\langle \partial_t u(t) \mid \chi_{\rho(t)}\Big(r\partial_r u(t) + \frac{D-2}{2}u(t)\Big)\Big\rangle &= -\int_0^\infty (\partial_t u(t,r))^2 \chi_{\rho(t)}(r)\, r^{D-1}\,\mathrm{d}r \\ &\quad + \Omega_{1,\rho(t)}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,\rho(t)}(\boldsymbol{u}(t))\end{aligned} \tag{2.4}$$

and,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\Big\langle \partial_t u(t) \mid \chi_{\rho(t)}\Big(r\partial_r u(t) + \frac{D}{2}u(t)\Big)\Big\rangle &= -\int_0^\infty \Big[(\partial_r u(t,r))^2 - |u(t,r)|^{\frac{2D}{D-2}}\Big] \\ &\quad \chi_{\rho(t)}(r)\, r^{D-1}\,\mathrm{d}r \\ &\quad + \Omega_{1,\rho(t)}(\boldsymbol{u}(t)) + \frac{D}{2}\Omega_{2,\rho(t)}(\boldsymbol{u}(t)).\end{aligned}$$

Note that the multiplier $(r\partial_r + \frac{D-2}{2})u$ in the first identity (2.4) corresponds to the generator of $\dot{H}^1$-invariant dilations.

2.2 Local Cauchy Theory

In the remainder of this section, we follow the presentation in [71, Section 2 and Appendix A].

Given that we are restricting our attention to radially symmetric functions $v : \mathbb{R}^D \to \mathbb{R}$, we often abuse notation, writing $v(x) = v(r)$ with $r = |x|$, and denoting, for $p \geq 1$,

$$\|v\|_{L^p(\mathbb{R}^D)} = \left(\int_0^\infty |v(r)|^p \, r^{D-1} \, \mathrm{d}r \right)^{\frac{1}{p}}$$

which agrees with the usual definition of the L^p norm for functions on $\mathbb{R}^D$ up to the dimensional constant $c_D > 0$. For $0 < s < 1$ and $1 < p < \infty$, $\dot{B}^s_{p,2} = \dot{B}^s_{p,2}(\mathbb{R}^D)$ denotes the homogeneous Besov space with norm

$$\|u\|_{\dot{B}^s_{p,2}} := \left(\sum_{j \in \mathbb{Z}} 2^{2js} \|P_j u\|^2_{L^p} \right)^{\frac{1}{2}},$$

where P_j are the Littlewood-Paley projections. We recall that if $s < D/p$, then $\dot{B}^s_{p,2}$ is a Banach space, see [1, Theorem 2.25].

For any time interval $I \subset \mathbb{R}$, we introduce the Strichartz-type spaces and norms

$$\begin{aligned} S(I) &:= L^{\frac{2(D+1)}{D-2}}(I \times \mathbb{R}^D), \\ W(I) &:= L^{\frac{2(D+1)}{D-1}}\Big(I; \dot{B}^{\frac{1}{2}}_{\frac{2(D+1)}{D-1},2}(\mathbb{R}^D)\Big). \end{aligned}$$

We denote $\vec{S}_{\mathrm{L}}(t) = (S_{\mathrm{L}}(t), \dot{S}_{\mathrm{L}}(t))$ the free wave propagator, in other words for all $\boldsymbol{v}_0 = (v_0, \dot{v}_0)$ we have

$$\vec{S}_{\mathrm{L}}(t)\boldsymbol{v}_0 = \left(\cos(t|\nabla|)v_0 + \frac{\sin(t|\nabla|)}{|\nabla|}\dot{v}_0, -|\nabla|\sin(t|\nabla|)v_0 + \cos(t|\nabla|)\dot{v}_0 \right).$$

We say that $\boldsymbol{u}$ is a solution of (1.1) on a time interval $I \ni 0$ with initial data $\boldsymbol{u}_0 \in \mathcal{E}$ if

- $\boldsymbol{u} \in C^0(I; \mathcal{E})$,
- $\|u\|_{S(J)} + \|u\|_{W(J)} < \infty$ for all compact intervals $J \subset I$,
- u satisfies the Duhamel formula

$$u(t) = S_{\mathrm{L}}(t)\boldsymbol{u}_0 + \int_0^t S_{\mathrm{L}}(t-s)(0, f(u(s)))\mathrm{d}s.$$

Local well-posedness was obtained by Ginibre and Velo [32], who used a slightly different but equivalent notion of solution; see also [34, 65, 84]. We use the versions in [7, 45]. Key to the proof are Strichartz estimates for the wave equation; see, Lindblad, Sogge [52], and Ginibre, Velo [33].

Lemma 2.3 (Cauchy theory in $\mathcal{E}$) [45, Theorem 2.7] [7, Theorem 3.3] *There exists $\delta_0 > 0$ and functions $\epsilon, C : [0, \infty) \to (0, \infty)$ with $\epsilon(\delta) \to 0$ as $\delta \to 0$, with the following properties. Let $A > 0$ and $\boldsymbol{u}_0 = (u_0, u_1) \in \mathcal{E}$ with $\|\boldsymbol{u}_0\|_{\mathcal{E}} \le A$. Let $I \ni 0$ be an open interval such that*

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

Then there exists a unique solution $\boldsymbol{u}(t)$ to (1.1) in the space $C(I; \mathcal{E}) \cap S(I) \cap W(I)$ with initial data $\boldsymbol{u}(0) = \boldsymbol{u}_0$. The solution $\boldsymbol{u}(t)$ satisfies the bounds $\|u\|_{S(I)} \le C(A)\epsilon(\delta)$, and $\|\boldsymbol{u}\|_{L^\infty_t(I;\mathcal{E})} \le C(A)$. To each solution $\boldsymbol{u}(t)$ to (1.1) we can associate a maximal interval of existence $0 \in I_{\max}(\boldsymbol{u}) = (-T_-, T_+)$ such that for each compact subinterval $I' \subset I_{\max}$ we have $\|u\|_{S(I')} < \infty$ and, if, say $T_+ < \infty$, then $\lim_{T \to T_+} \|u\|_{S([0,T))} = \infty$.

The data to solution map is continuous in the following sense. Let $\boldsymbol{u}_0 \in \mathcal{E}$ and let $\boldsymbol{u}$ be the unique solution to (1.1) with initial data $\boldsymbol{u}_0$. Then for every $\epsilon > 0$ and $T_0 < T_+(\boldsymbol{u}_0)$ there exists $\delta > 0$ with the following property: for all $\boldsymbol{v}_0 \in \mathcal{E}$ with $\|\boldsymbol{u}_0 - \boldsymbol{v}_0\|_{\mathcal{E}} < \delta$ we have $T_0 < T_+(\boldsymbol{v}_0)$ and $\sup_{t\in[0,T_0]} \|\boldsymbol{u}(t) - \boldsymbol{v}(t)\|_{\mathcal{E}} < \epsilon$, where $\boldsymbol{v}(t)$ is the unique solution to (1.1) associated to $\boldsymbol{v}_0$.

The completeness of wave operators holds for small data: there exists ϵ_0 small enough so that if $\boldsymbol{u}_0 \in \mathcal{E}$ satisfies $\|\boldsymbol{u}_0\|_{\mathcal{E}} < \epsilon_0$, the solution $\boldsymbol{u}(t)$ given above is defined globally in time, satisfies the bound,

$$\sup_{t\in\mathbb{R}} \|\boldsymbol{u}(t)\|_{\mathcal{E}} + \|u\|_{S(\mathbb{R})} + \|u\|_{W(\mathbb{R})} \lesssim \|\boldsymbol{u}_0\|_{\mathcal{E}}, \tag{2.5}$$

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

$$\|\boldsymbol{u}(t) - \boldsymbol{u}_{\mathrm{L}}^{\pm}(t)\|_{\mathcal{E}} \to 0 \ \text{ as } \ t \to \pm\infty. \tag{2.6}$$

Conversely, the existence of wave operators holds, i.e., for any solution $\boldsymbol{v}_{\mathrm{L}} \in C(\mathbb{R}; \mathcal{E})$ to the free linear equation, there exists $T_0 > 0$ and a unique, global-in-forward time solution $\boldsymbol{u} \in C([T_0, \infty); \mathcal{E})$ to (1.1) such that (2.6) holds as $t \to \infty$. An analogous statement holds for negative times.

Remark 2.4 For dimensions $D \ge 6$ the continuous dependence on the initial data is not stated explicitly as part of [7, Theorem 3.3], but it does follow from their proof; see for example [7, Remark 4.3].

The following lemma is a consequence of the local Cauchy theory.

Lemma 2.5 *(Propagation of small $\mathcal{E}$ norm) There exist $\delta, C > 0$ with the following properties. Let $I \ni 0$ be a time interval and let $\boldsymbol{v} \in C(I; \mathcal{E})$ be the solution to (1.1) on I with initial data $\boldsymbol{v}(0) = \boldsymbol{v}_0$. If,*

$$\|\boldsymbol{v}_0\|_{\mathcal{E}} \le \delta,$$

then $0 \le E(\boldsymbol{v}_0) \le C\delta^2$ *and*

$$\sup_{t \in I} \|\boldsymbol{v}(t)\|_{\mathcal{E}}^2 \le C E(\boldsymbol{v}) \le C\delta^2.$$

Proof Abusing notation and identifying $\boldsymbol{v}_0 \in \mathcal{E}$ with $\boldsymbol{v}_0 \in \dot{H}^1 \times L^2(\mathbb{R}^D)$ we can express the nonlinear energy of $\boldsymbol{v}_0$ as

$$E(\boldsymbol{v}_0) = \frac{1}{2}\|\dot{v}_0\|_{L^2(\mathbb{R}^D)}^2 + \frac{1}{2}\|\nabla v_0\|_{L^2(\mathbb{R}^D)}^2 - \frac{D-2}{2D}\|v_0\|_{L^{\frac{2D}{D-2}}(\mathbb{R}^D)}^{\frac{2D}{D-2}}.$$

It is clear that $E(\boldsymbol{v}) = E(\boldsymbol{v}_0) \le C\|\boldsymbol{v}_0\|_{\mathcal{E}}^2 \le C\delta^2$. Since $\frac{2D}{D-2} > 2$, by the Sobolev inequality, $\|v\|_{L^{\frac{2D}{D-2}}(\mathbb{R}^D)} \le C_1\|\nabla v\|_{L^2(\mathbb{R}^D)}$ together with Hardy's inequality, $\| \, |x|^{-1} \, v\|_{L^2(\mathbb{R}^D)} \le C_2\|\nabla v\|_{L^2(\mathbb{R}^D)}^2$, which hold for all $v \in \dot{H}^1(\mathbb{R}^D)$, we see that by taking $\delta > 0$ small enough we can find $C > 0$ so that

$$E(\boldsymbol{v}) = E(\boldsymbol{v}_0) \ge C^{-1}\|\boldsymbol{v}_0\|_{\mathcal{E}}^2.$$

The remaining conclusion now follows from (2.5) restricted to the time interval I. □

Using the finite speed of propagation and the previous lemma, one obtains the following localized version.

Lemma 2.6 *(Propagation of small localized $\mathcal{E}$ norm) There exist $\delta, C > 0$ with the following properties. Let $I \ni 0$ be a time interval and let $\boldsymbol{u} \in C(I; \mathcal{E})$ be a solution to (1.1) on I with initial data $\boldsymbol{u}(0) = \boldsymbol{u}_0$. Let $0 < r_1 < r_2$. Suppose that*

$$\|\boldsymbol{u}_0\|_{\mathcal{E}(\frac{r_1}{4}, 4r_2)} \le \delta.$$

Then,

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(r_1+|t|, r_2-|t|)} \le C\delta,$$

for all $t \in I$ such that $r_1 + 2\,|t| < r_2$.

Proof Let $\varphi(r)$ be a smooth cut-off function such that $\varphi(r) = 1$ if $r \in [\frac{1}{2}r_1, 2r_2]$, $\varphi(r) = 0$ if $r \in (0, \frac{1}{4}r_1] \cup [4r_2, \infty)$ and such that $|\partial_r\varphi(r)| \le 4r_1^{-1}$ for $r \in [\frac{1}{4}r_1, \frac{1}{2}r_1]$ and $|\partial_r\varphi(r)| \le 4r_2^{-1}$ for $r \in [2r_2, 4r_2]$. Setting $\boldsymbol{v}_0 := \varphi\boldsymbol{u}_0$ it follows from the definition of the local $\mathcal{E}$-norm in (1.5) that $\|\boldsymbol{v}_0\|_{\mathcal{E}} \le C\|\boldsymbol{u}_0\|_{\mathcal{E}(\frac{r_1}{4}, 4r_2)}$ for some constant $C > 0$ independent of $\boldsymbol{u}_0, r_1, r_2$. Taking $\delta > 0$ sufficiently small we may apply Lemma 2.5 to the solution $\boldsymbol{v}(t)$ with initial data $\boldsymbol{v}(0) = \boldsymbol{v}_0$ The conclusion then follows by finite speed of propagation, which ensures that $\boldsymbol{u}(t, r) = \boldsymbol{v}(t, r)$ for all $(t, r) \in I \times (0, \infty)$ with $r \in (r_1 + |t|\,, r_2 - |t|)$ and $r_1 + 2\,|t| < r_2$. □

Lemma 2.7 (Short time evolution close to W) *Let $\iota \in \{-1, 1\}$. There exists $\delta_0 > 0$ and a function $\epsilon_0 : [0, \delta_0] \to [0, \infty)$ with $\epsilon_0(\delta) \to 0$ as $\delta \to 0$ with the following* *properties. Let* $\boldsymbol{v}_0 \in \mathcal{E}$ *and let* $\boldsymbol{v}(t)$ *denote the unique solution to* (1.1) *with* $\boldsymbol{v}_0(0) = \boldsymbol{v}_0$. *Let* $\mu_0 > 0$ *and suppose that*

$$\|\boldsymbol{v}_0 - \iota \boldsymbol{W}_{\mu_0}\|_{\mathcal{E}} = \delta \le \delta_0.$$

Then, $\mu_0 < T_+(\boldsymbol{v}_0)$ *and*

$$\sup_{t\in[0,\mu_0]} \|\boldsymbol{v}(t) - \iota \boldsymbol{W}_{\mu_0}\|_{\mathcal{E}} < \epsilon_0(\delta).$$

Proof By rescaling we may assume $\mu_0 = 1$. The result is then a particular case of the local Cauchy theory, in particular the continuity of the data to solution map at $\boldsymbol{W}$. □

We also require the following localized version.

Lemma 2.8 (Localized short time evolution close to W) *Let* $\iota \in \{-1, 1\}$. *There exists* $\delta_0 > 0$ *and a function* $\epsilon_0 : [0, \delta_0] \to [0, \infty)$ *with* $\epsilon_0(\delta) \to 0$ *as* $\delta \to 0$ *with the following properties. Let* $\boldsymbol{u}_0 \in \mathcal{E}$, *and let* $\boldsymbol{u}(t)$ *denote the unique solution to* (1.1) *with* $\boldsymbol{u}_0(0) = \boldsymbol{u}_0$. *Let* $\mu_0 > 0$, $0 < r_1 < r_2 < \infty$ *and suppose that*

$$\|\boldsymbol{u}_0 - \iota \boldsymbol{W}_{\mu_0}\|_{\mathcal{E}(\frac{1}{4}r_1, 4r_2)} = \delta \le \delta_0.$$

Then,

$$\|\boldsymbol{u}(t) - \iota \boldsymbol{W}_{\mu_0}\|_{\mathcal{E}(r_1+t, r_2-t)} < \epsilon_0(\delta)$$

for all $0 < t < \min(\mu_0, T_+(\boldsymbol{u}_0))$ *such that* $r_1 + 2t < r_2$.

Proof Let $\varphi(r)$ be as in the proof of Lemma 2.6 and define $\boldsymbol{v}_0 := \varphi \boldsymbol{u}_0 + (1-\varphi)\iota \boldsymbol{W}_{\mu_0}$. By taking δ_0 sufficiently small we see that $\boldsymbol{v}_0, \mu_0$ satisfy the hypothesis of Lemma 2.7. The conclusion then follows from the finite speed of propagation. □

We will make use of the following consequence of the previous four lemmas.

Lemma 2.9 *If* $\iota_n \in \{-1, 0, 1\}$, $0 < r_n \ll \mu_n \ll R_n$, $0 < t_n \ll \mu_n$ *and* $\boldsymbol{u}_n$ *is a sequence of solutions of* (1.1) *such that* $\boldsymbol{u}_n(t)$ *is defined for* $t \in [0, t_n]$ *and*

$$\lim_{n\to\infty} \|\boldsymbol{u}_n(0) - \iota_n \boldsymbol{W}_{\mu_n}\|_{\mathcal{E}(\frac{1}{4}r_n, 4R_n)} = 0,$$

then

$$\lim_{n\to\infty} \sup_{t\in[0,t_n]} \|\boldsymbol{u}_n(t) - \iota_n \boldsymbol{W}_{\mu_n}\|_{\mathcal{E}(r_n+t, R_n-t)} = 0.$$

Proof This is a direct consequence of Lemma 2.6 when $\iota_n = 0$ and Lemma 2.8 when $\iota_n \in \{-1, 1\}$. □

2.3 Profile Decomposition

The linear profile decomposition of Bahouri and Gérard [2] is an essential ingredient in the study of solutions to (1.1); see also [6, 31, 53, 54, 58].

Lemma 2.10 *(Linear profile decomposition) [2] Let $\boldsymbol{u}_n$ be a bounded sequence in $\mathcal{E}$, i.e., $\limsup_{n\to\infty} \|\boldsymbol{u}_n\|_{\mathcal{E}} < \infty$. Then, after passing to a subsequence, there exist sequences $\lambda_{n,j} \in (0,\infty)$, and $t_{n,j} \in \mathbb{R}$ and finite energy free waves $\boldsymbol{v}_{\mathrm{L}}^j \in \mathcal{E}$ such that for each $J \ge 1$,*

$$\boldsymbol{u}_n = \sum_{j=1}^{J} \Big(\lambda_{n,j}^{-\frac{D-2}{2}} v_{\mathrm{L}}^j \big(\frac{-t_{n,j}}{\lambda_{n,j}}, \frac{\cdot}{\lambda_{n,j}}\big), \lambda_{n,j}^{-\frac{D}{2}} \partial_t v_{\mathrm{L}}^j \big(\frac{-t_{n,j}}{\lambda_{n,j}}, \frac{\cdot}{\lambda_{n,j}}\big)\Big) + \boldsymbol{w}_{n,0}^J(\cdot)$$

where, denoting by $\boldsymbol{w}_{n,\mathrm{L}}^J(t)$ the free wave with initial data $\boldsymbol{w}_{n,0}^J$, the following hold:

- *for each j, either $t_{n,j} = 0$ for all n or $\lim_{n\to\infty} \frac{-t_{n,j}}{\lambda_{n,j}} = \pm\infty$. 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;*
- *the scales $\lambda_{n,j}$ and times $t_{n,j}$ satisfy,*

$$\frac{\lambda_{n,j}}{\lambda_{n,j'}} + \frac{\lambda_{n,j'}}{\lambda_{n,j}} + \frac{\left|t_{n,j} - t_{n,j'}\right|}{\lambda_{n,j}} \to \infty \ \ as \ \ n \to \infty;$$

  for each $j \ne j'$;
- *the error term $\boldsymbol{w}_{n,\mathrm{L}}^J$ satisfies,*

$$(\lambda_{n,j}^{\frac{D-2}{2}} w_{n,\mathrm{L}}^J(t_{n,j}, \lambda_{n,j}\cdot), \lambda_{n,j}^{\frac{D}{2}} \partial_t w_{n,\mathrm{L}}^J(t_{n,j}, \lambda_{n,j}\cdot)) \rightharpoonup 0 \ \ in \ \mathcal{E} \ \ as \ \ n \to \infty$$

  for each $J \ge 1$, each $1 \le j \le J$, and vanishes strongly in the sense that

$$\lim_{J\to\infty} \limsup_{n\to\infty} \Big(\|w_{n,\mathrm{L}}^J\|_{L_t^\infty L^{\frac{2D}{D-2}}(\mathbb{R}\times\mathbb{R}^D)} + \|w_{n,\mathrm{L}}^J\|_{S(\mathbb{R})}\Big) = 0;$$

- *the following Pythagorean decomposition of the free energy holds: for each $J \ge 1$,*

$$\begin{aligned} \|\boldsymbol{u}_n\|_{\dot H^1\times L^2}^2 = \sum_{j=1}^{J} \|(v_{\mathrm{L}}^j(-t_{n,j}/\lambda_{n,j}), \partial_t v_{\mathrm{L}}^j(-t_{n,j}/\lambda_{n,j}))\|_{\dot H^1\times L^2}^2 \\ + \|\boldsymbol{w}_{n,0}^J\|_{\dot H^1\times L^2}^2 + o_n(1) \end{aligned} \tag{2.7}$$

  as $n \to \infty$.

Remark 2.11 We call the triplets $(\boldsymbol{v}_{\mathrm{L}}^i, \lambda_{n,j}, t_{n,j})$ *profiles*. Following Bahouri and Gérard [2] we refer to the profiles $(\boldsymbol{v}_{\mathrm{L}}^j, \lambda_{n,j}, 0)$ as *centered*, to the profiles $(\boldsymbol{v}_{\mathrm{L}}^j, \lambda_{n,j}, t_{n,j})$ with $-t_{n,j}/\lambda_{n,j} \to \infty$ as $n \to \infty$ as *outgoing*, and those with $-t_{n,j}/\lambda_{n,j} \to -\infty$ as *incoming*.

Remark 2.12 In Section 3 we implicitly make use of nonlinear profile decompositions, in addition to Lemma 2.10, when we invoke arguments from [19]. We refer the reader to [19, Proposition 2.1] for the statement.

2.4 Multi-bubble Configurations

In this section we study properties of finite energy functions near a multi-bubble configuration, and we record several properties of the ground state W.

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

$$\mathcal{L}_{\mathcal{W}}\, g := \mathrm{D}^2\, E_{\mathbf{p}}(\mathcal{W}(\vec{\iota}, \vec{\lambda}))g = -\partial_r^2 g - \frac{D-1}{r}\partial_r g - f'(\mathcal{W}(\vec{\iota}, \vec{\lambda}))g,$$

where $f(z) := |u|^{\frac{4}{D-2}}\, u$ and $f'(z) = \frac{D+2}{D-2}\, |z|^{\frac{4}{D-2}}$, and $E_{\mathbf{p}}$ is the potential energy,

$$E_{\mathbf{p}}(u) = \int_0^\infty \frac{1}{2}(\partial_r u)^2 r^{D-1}\,\mathrm{d}r - \int_0^\infty \frac{D-2}{2D}\, |u|^{\frac{2D}{D-2}}\, r^{D-1}\,\mathrm{d}r.$$

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

$$\left\langle \mathrm{D}^2\, E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g} \mid \boldsymbol{g}\right\rangle = \int_0^\infty \Big(\dot{g}(r)^2 + (\partial_r g(r))^2 - f'(\mathcal{W}(\vec{\iota}, \vec{\lambda}))g(r)^2\Big) r\,\mathrm{d}r.$$

An important instance of the operator $\mathcal{L}_{\mathcal{W}}$ is given by linearizing (1.1) about a single copy of the ground state $\mathcal{W}(\vec{\iota}, \vec{\lambda}) = W_\lambda$. In this case we use the short-hand notation,

$$\mathcal{L}_\lambda := -\partial_r^2 - \frac{D-1}{r}\partial_r - f'(W_\lambda).$$

We write $\mathcal{L} := \mathcal{L}_1$.

We define the infinitesimal generators of $\dot{H}^1$-invariant dilations by Λ and in the L^2-invariant case we write $\underline{\Lambda}$, which are given by

$$\Lambda := r\partial_r + \frac{D-2}{2}, \quad \underline{\Lambda} := r\partial_r + \frac{D}{2}.$$

We have

$$\Lambda W(r) = \Big(\frac{D-2}{D} - \frac{r^2}{2D}\Big)\Big(1 + \frac{r^2}{D(D-2)}\Big)^{-\frac{D}{2}}.$$

Note that both W and ΛW satisfy,

$$|W(r)|\,,\ |\Lambda W(r)| \simeq 1 \ \text{ if } \ r \le 1, \ \text{ and } \ |W(r)|\,,\ |\Lambda W(r)| \simeq r^{-D+2} \ \text{ as } \ r \to \infty.$$

In fact,

$$\begin{aligned} W(r) &= 1 + O(r^2) \text{ if } r \ll 1 \\ &= (D(D-2))^{\frac{D-2}{2}} r^{-D+2} + O(r^{-D}) \text{ if } r \gg 1 \end{aligned}$$

and,

$$\begin{aligned} \Lambda W(r) &= \frac{D-2}{2} + O(r^2) \text{ if } r \ll 1 \\ &= -\frac{(D(D-2))^{\frac{D}{2}}}{2D} r^{-D+2} + O(r^{-D}) \text{ if } r \gg 1. \end{aligned}$$

When $D = 4$ we will use the extra decay in

$$\underline{\Lambda}\Lambda W(r) = (r\partial_r + 2)\Lambda W(r) = \frac{1}{4}\frac{3r^2 - 8}{(1 + \frac{1}{8}r^2)^3} \simeq r^{-4} \text{ as } r \to \infty \text{ if } D = 4. \tag{2.8}$$

We note that if $D \geq 5$ then $\langle \underline{\Lambda}\Lambda W \mid \Lambda W \rangle = 0$. If $D = 4$ then $\langle \underline{\Lambda}\Lambda W \mid \Lambda W \rangle = 32$.

We will use the following computations,

$$\begin{aligned} &\frac{D+2}{D-2}\int_0^\infty \Lambda W(r) W(r)^{\frac{4}{D-2}} r^{D-1}\,\mathrm{d}r = -\frac{D-2}{2D}(D(D-2))^{\frac{D}{2}} \\ &\frac{D+2}{D-2}(D(D-2))^{\frac{D-2}{2}}\int_0^\infty \Lambda W(r) W(r)^{\frac{4}{D-2}} r\,\mathrm{d}r = \frac{D-2}{2D}(D(D-2))^{\frac{D}{2}} \\ &\|\Lambda W\|_{L^2}^2 = \frac{2(D^2-4)(D(D-2))^{\frac{D}{2}}}{D^2(D-4)}\frac{\Gamma(1+\frac{D}{2})}{\Gamma(D)} \text{ if } D \geq 5. \end{aligned}$$

If $D = 4$,

$$\int_0^R (\Lambda W(r))^2 r^3\,\mathrm{d}r = 16\log R + O(1) \text{ as } R \to \infty. \tag{2.9}$$

If $\sigma \ll 1$ we have,

$$|\langle \Lambda W_{\underline{\sigma}} \mid \Lambda W \rangle| \lesssim \sigma^{\frac{D-4}{2}} \text{ if } D \geq 5.$$

For any $R > 0$,

$$\left|\langle \chi_{R\sqrt{\sigma}} \Lambda W_{\underline{\sigma}} \mid \Lambda W \rangle\right| \lesssim R^2 \sigma^{\frac{D-2}{2}} \text{ if } D \geq 4.$$

Next we discuss the spectral properties of $\mathcal{L}$. Importantly,

$$\mathcal{L}(\Lambda W) = \frac{\mathrm{d}}{\mathrm{d}\lambda}\restriction_{\lambda=1} (-\partial_r^2 W_\lambda - \frac{D-1}{r}\partial_r W_\lambda - f(W_\lambda)) = 0$$

and thus if $D \geq 5$, $\Lambda W \in L^2$ is a zero energy eigenfunction for $\mathcal{L}$ and a threshold resonance if $D = 4$. In fact, $\{f \in H \mid \mathcal{L} f = 0\} = \operatorname{span}\{\Lambda W\}$ (see [27, Proposition 5.5]). In addition to this fact, it was also shown in [27, Proposition 5.5] that $\mathcal{L}$ has a unique negative simple eigenvalue that we denote by $-\kappa^2 < 0$ (we take $\kappa > 0$). We denote by $\mathcal{Y}$ the associated positive eigenfunction such that $\|\mathcal{Y}\|_{L^2} = 1$. By elliptic regularity $\mathcal{Y}$ is smooth, and by Agmon estimates it decays exponentially. Using that $\mathcal{L}$ is symmetric we deduce that $\langle \mathcal{Y} \mid \Lambda W \rangle = 0$.

We follow the notations and set-up in [37, Section 3]. Define

$$\boldsymbol{\mathcal{Y}}^- := (\frac{1}{\kappa}\mathcal{Y}, -\mathcal{Y}), \quad \boldsymbol{\mathcal{Y}}^+ := (\frac{1}{\kappa}\mathcal{Y}, \mathcal{Y})$$

and,

$$\boldsymbol{\alpha}^- = \frac{\kappa}{2} J \boldsymbol{\mathcal{Y}}^+ = \frac{1}{2}(\kappa\mathcal{Y}, -\mathcal{Y}), \quad \boldsymbol{\alpha}^+ := -\frac{\kappa}{2} J \boldsymbol{\mathcal{Y}}^- = \frac{1}{2}(\kappa\mathcal{Y}, \mathcal{Y}).$$

Recalling that $J \circ \mathrm{D}^2 E(\boldsymbol{W}) = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathcal{L} & 0 \end{pmatrix}$ we see that

$$J \circ \mathrm{D}^2 E(\boldsymbol{W}) \boldsymbol{\mathcal{Y}}^- = -\kappa \boldsymbol{\mathcal{Y}}^-, \ \text{ and } \ J \circ \mathrm{D}^2 E(\boldsymbol{W}) \boldsymbol{\mathcal{Y}}^+ = \kappa \boldsymbol{\mathcal{Y}}^+$$

and for all $\boldsymbol{h} \in \mathcal{E}$,

$$\left\langle \boldsymbol{\alpha}^- \mid J \circ \mathrm{D}^2 E(\boldsymbol{W}) \boldsymbol{h} \right\rangle = -\kappa \left\langle \boldsymbol{\alpha}^- \mid \boldsymbol{h} \right\rangle, \quad \left\langle \boldsymbol{\alpha}^+ \mid J \circ \mathrm{D}^2 E(\boldsymbol{W}) \boldsymbol{h} \right\rangle = \kappa \left\langle \boldsymbol{\alpha}^+ \mid \boldsymbol{h} \right\rangle.$$

We view $\boldsymbol{\alpha}^\pm$ as linear forms on $\mathcal{E}$ and we note that $\left\langle \boldsymbol{\alpha}^- \mid \boldsymbol{\mathcal{Y}}^- \right\rangle = \left\langle \boldsymbol{\alpha}^+ \mid \boldsymbol{\mathcal{Y}}^+ \right\rangle = 1$ and $\left\langle \boldsymbol{\alpha}^- \mid \boldsymbol{\mathcal{Y}}^+ \right\rangle = \left\langle \boldsymbol{\alpha}^+ \mid \boldsymbol{\mathcal{Y}}^- \right\rangle = 0$. For $\lambda > 0$ the rescaled versions of these objects are defined as,

$$\boldsymbol{\mathcal{Y}}_\lambda^- := (\frac{1}{\kappa}\mathcal{Y}_\lambda, -\mathcal{Y}_{\underline{\lambda}}), \quad \boldsymbol{\mathcal{Y}}_\lambda^+ := (\frac{1}{\kappa}\mathcal{Y}_\lambda, \mathcal{Y}_{\underline{\lambda}})$$

and,

$$\boldsymbol{\alpha}_\lambda^- := \frac{\kappa}{2\lambda} J \boldsymbol{\mathcal{Y}}_\lambda^+ = \frac{1}{2}(\frac{\kappa}{\lambda}\mathcal{Y}_{\underline{\lambda}}, -\mathcal{Y}_{\underline{\lambda}}), \quad \boldsymbol{\alpha}_\lambda^+ := -\frac{\kappa}{2\lambda} J \boldsymbol{\mathcal{Y}}^- = \frac{1}{2}(\frac{\kappa}{\lambda}\mathcal{Y}_{\underline{\lambda}}, \mathcal{Y}_{\underline{\lambda}}). \tag{2.10}$$

These choices of scalings ensure that $\left\langle \boldsymbol{\alpha}_\lambda^- \mid \boldsymbol{\mathcal{Y}}_\lambda^- \right\rangle = \left\langle \boldsymbol{\alpha}_\lambda^+ \mid \boldsymbol{\mathcal{Y}}_\lambda^+ \right\rangle = 1$. We have,

$$J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda) \boldsymbol{\mathcal{Y}}_\lambda^- = -\frac{\kappa}{\lambda} \boldsymbol{\mathcal{Y}}_\lambda^-, \ \text{ and } \ J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda) \boldsymbol{\mathcal{Y}}_\lambda^+ = \frac{\kappa}{\lambda} \boldsymbol{\mathcal{Y}}_\lambda^+$$

and for all $\boldsymbol{h} \in \mathcal{E}$,

$$\left\langle \boldsymbol{\alpha}_\lambda^- \mid J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda) \boldsymbol{h} \right\rangle = -\frac{\kappa}{\lambda} \left\langle \boldsymbol{\alpha}_\lambda^- \mid \boldsymbol{h} \right\rangle, \quad \left\langle \boldsymbol{\alpha}_\lambda^+ \mid J \circ \mathrm{D}^2 E(\boldsymbol{W}_\lambda) \boldsymbol{h} \right\rangle = \frac{\kappa}{\lambda} \left\langle \boldsymbol{\alpha}_\lambda^+ \mid \boldsymbol{h} \right\rangle. \tag{2.11}$$

We define a smooth non-negative function $\mathcal{Z} \in C^\infty(0,\infty) \cap L^1((0,\infty), r^{D-1}\,\mathrm{d}r)$ as follows. First if $D \geq 7$ we simply define

$$\mathcal{Z}(r) := \Lambda W(r) \text{ if } D \geq 7$$

and note that

$$\langle \mathcal{Z} \mid \Lambda W \rangle > 0 \text{ and } \langle \mathcal{Z} \mid \mathcal{Y} \rangle = 0. \tag{2.12}$$

In fact the precise form of $\mathcal{Z}$ is not so important, rather only the properties in (2.12) and that it has sufficient decay and regularity. As $\Lambda W \notin \dot{H}^{-1}$ for $D \leq 6$ we cannot take $\mathcal{Z} = \Lambda W$. Rather if $4 \leq D \leq 6$ we fix any $\mathcal{Z} \in C_0^\infty(0,\infty)$ so that

$$\langle \mathcal{Z} \mid \Lambda W \rangle > 0 \text{ and } \langle \mathcal{Z} \mid \mathcal{Y} \rangle = 0.$$

We record the following localized coercivity lemma from [37].

Lemma 2.13 (Localized coercivity for $\mathcal{L}$) [37, Lemma 3.8] *Fix $D \geq 4$. There exist uniform constants $c < 1/2$, $C > 0$ with the following properties. Let $g \in H$. Then,*

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

If $R > 0$ is large enough then,

$$\begin{aligned}&(1-2c)\int_0^R (\partial_r g(r))^2\, r^{D-1}\mathrm{d}r + c\int_R^\infty (\partial_r g(r))^2\, r^{D-1}\mathrm{d}r\\ &\quad - \int_0^\infty f'(W(r))g(r)^2\, r^{D-1}\,\mathrm{d}r\\ &\quad \geq -C\langle \mathcal{Z} \mid g \rangle^2 - C\langle \mathcal{Y} \mid g \rangle^2 .\end{aligned}$$

If $\rho > 0$ is small enough, then

$$\begin{aligned}&(1-2c)\int_\rho^\infty (\partial_r g(r))^2\, r^{D-1}\mathrm{d}r + c\int_0^\rho (\partial_r g(r))^2\, r^{D-1}\mathrm{d}r\\ &\quad - \int_0^\infty f'(W(r))g(r)^2\, r^{D-1}\,\mathrm{d}r\\ &\quad \geq -C\langle \mathcal{Z} \mid g \rangle^2 - C\langle \mathcal{Y} \mid g \rangle^2 .\end{aligned}$$

As a consequence, (see for example [37, Proof of Lemma 3.9] for an analogous argument in the case of two bubbles) we have the following coercivity property of $\mathcal{L}_{\mathcal{W}}$.

Lemma 2.14 *Fix $D \geq 4$, $M \in \mathbb{N}$. There exist $\eta, c_0 > 0$ with the following properties. Consider the subset of M-bubble configurations $\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})$ for $\vec{\iota} \in \{-1,1\}^M$, $\vec{\lambda} \in (0,\infty)^M$ such that,*

$$\sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} \leq \eta^2. \tag{2.13}$$

Let $\boldsymbol{g} \in \mathcal{E}$ *be such that*

$$0 = \big\langle \mathcal{Z}_{\underline{\lambda_j}} \mid g \big\rangle \text{ for } j = 1, \dots M$$

for $\vec{\lambda}$ *as in* (2.13). *Then,*

$$\frac{1}{2}\big\langle \mathrm{D}^2\, E(\mathcal{W}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g} \mid \boldsymbol{g}\big\rangle + 2\sum_{j=1}^{M} \Big(\big\langle \alpha_{\lambda_j}^{-} \mid \boldsymbol{g}\big\rangle^2 + \big\langle \alpha_{\lambda_j}^{+} \mid \boldsymbol{g}\big\rangle^2\Big) \geq c_0 \|\boldsymbol{g}\|_{\mathcal{E}}^2. \tag{2.14}$$

Lemma 2.15 *Fix* $D \geq 4$, $M \in \mathbb{N}$. *For any* $\theta > 0$, *there exists* $\eta > 0$ *with the following property. Consider the subset of* M*-bubble configurations* $\boldsymbol{\mathcal{W}}(\iota, \vec{\lambda})$ *such that*

$$\sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} \leq \eta.$$

Then,

$$\Big| E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})) - M E(\boldsymbol{W}) + \frac{(D(D-2))^{\frac{D}{2}}}{D} \sum_{j=1}^{M-1} \iota_j \iota_{j+1} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} \Big|$$
$$\leq \theta \sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}}.$$

Moreover, there exists a uniform constant $C > 0$ *such that for any* $\boldsymbol{g} = (g, 0) \in \mathcal{E}$,

$$\Big|\Big\langle DE_{\mathbf{p}}(\mathcal{W}(\vec{\iota}, \vec{\lambda})) \mid g \Big\rangle\Big| \leq C \|(g, 0)\|_{\mathcal{E}} \sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}}.$$

Proof This is an explicit computation analogous to [41, Lemma 2.22]. □

The following modulation lemma plays an important role in our analysis. Before stating it, we define a proximity function to M-bubble configurations.

Definition 2.16 Fix M as in Definition 1.12 and let $\boldsymbol{v} \in \mathcal{E}$. Define,

$$\mathbf{d}(\boldsymbol{v}) := \inf_{\vec{\iota}, \vec{\lambda}} \Big(\|\boldsymbol{v} - \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} \Big)^{\frac{1}{2}},$$

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

Lemma 2.17 (Static modulation lemma) *Let $D \ge 4$ and $M \in \mathbb{N}$. There exists $\eta, C > 0$ with the following properties. Let $\theta > 0$, and let $\boldsymbol{v} \in \mathcal{E}$ be such that*

$$\mathbf{d}(\boldsymbol{v}) \le \eta, \quad \textit{and} \quad E(\boldsymbol{v}) \le ME(\boldsymbol{W}) + \theta^2.$$

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

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

along with the estimates,

$$\mathbf{d}(\boldsymbol{v})^2 \le \|\boldsymbol{g}\|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}} \Big)^{\frac{D-2}{2}} \le C\mathbf{d}(\boldsymbol{v})^2. \tag{2.15}$$

Defining the unstable/stable components of $\boldsymbol{g}$ by,

$$a_j^{\pm} := \langle \boldsymbol{\alpha}_{\lambda_j}^{\pm} \mid \boldsymbol{g} \rangle$$

we additionally have the estimates,

$$\|\boldsymbol{g}\|_{\mathcal{E}}^2 + \sum_{j \notin \mathcal{S}} \Big(\frac{\lambda_j}{\lambda_{j+1}} \Big)^{\frac{D-2}{2}} \le C \max_{j \in \mathcal{S}} \Big(\frac{\lambda_j}{\lambda_{j+1}} \Big)^{\frac{D-2}{2}} + C \max_{i \in \{1, \dots, M\}, \pm} |a_i^{\pm}|^2 + \theta^2, \tag{2.16}$$

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

Remark 2.18 Note that the scaling in the definition of $\boldsymbol{\alpha}_{\lambda_j}^{\pm}$ is chosen so that $|a_j^{\pm}| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}$, see (2.10).

Remark 2.19 We use the following, less standard, version of the implicit function theorem in the proof of Lemma 2.17.

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

$$\begin{aligned} \|L_0 - D_y G(x, y)\|_{\mathcal{L}(Y,Z)} &\le \frac{1}{3\|L_0^{-1}\|_{\mathcal{L}(Z,Y)}} \\ \|G(x, y_0)\|_Z &\le \frac{\delta_2}{3\|L_0^{-1}\|_{\mathcal{L}(Z,Y)}}, \end{aligned} \tag{2.17}$$

for all $\|x - x_0\|_X \le \delta_1$ and $\|y - y_0\|_Y \le \delta_2$. 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., [8, Section 2.2]. The essential point is that the bounds (2.17) give uniform control on the size of the open set where the Banach contraction mapping theorem is applied.

Proof of Lemma 2.17 The proof is a standard argument and is very similar to [40, Proof of Lemma 3.1] and [41, Proof of Lemma 2.24]. We refer to those papers for details and here only sketch the distinction in the estimate (2.16) where the stable/unstable directions enter.

To prove the estimate (2.16) we expand the nonlinear energy of $\boldsymbol{v}$,

$$
\begin{aligned}
ME(\boldsymbol{W}) + \theta^2 &\geq E(\boldsymbol{v}) = E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}) + \boldsymbol{g}) \\
&= E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})) + \Big\langle \mathrm{D}\, E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})) \mid \boldsymbol{g} \Big\rangle + \frac{1}{2} \Big\langle \mathrm{D}^2\, E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g} \mid \boldsymbol{g} \Big\rangle + O(\|\boldsymbol{g}\|_{\mathcal{E}}^3)
\end{aligned}
$$

and apply the conclusions of Lemma 2.14, in particular the estimate (2.14) and Lemma 2.15. □

Lemma 2.20 *Let $D \geq 4$. There exists $\eta > 0$ sufficiently small with the following property. Let $M, L \in \mathbb{N}$, $\vec{\iota} \in \{-1, 1\}^M$, $\vec{\sigma} \in \{-1, 1\}^L$, $\vec{\lambda} \in (0, \infty)^M$, $\vec{\mu} \in (0, \infty)^L$, and $\boldsymbol{w} = (w, 0)$ be such that $\|\boldsymbol{w}\|_{\mathcal{E}} < \infty$ and,*

$$
\begin{aligned}
\|\boldsymbol{w} - \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \Big(\frac{\lambda_j}{\lambda_{j+1}} \Big)^{\frac{D-2}{2}} &\leq \eta, \\
\|\boldsymbol{w} - \boldsymbol{\mathcal{W}}(\vec{\sigma}, \vec{\mu})\|_{\mathcal{E}}^2 + \sum_{j=1}^{L-1} \Big(\frac{\mu_j}{\mu_{j+1}} \Big)^{\frac{D-2}{2}} &\leq \eta.
\end{aligned}
$$

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

$$
\max_{j=1,\dots M} \Big| \frac{\lambda_j}{\mu_j} - 1 \Big| \leq \theta. \tag{2.18}
$$

Proof of Lemma 2.20 Let $g_\lambda := w - \mathcal{W}(\vec{\iota}, \vec{\lambda})$ and $g_\mu := w - \mathcal{W}(\vec{\sigma}, \vec{\mu})$. By expanding the nonlinear potential energy we have,

$$
E_{\mathbf{p}}(w) = E_{\mathbf{p}}(\mathcal{W}(\vec{\iota}, \vec{\lambda})) + \big\langle DE_{\mathbf{p}}(\mathcal{W}(\vec{\iota}, \vec{\lambda})) \mid g_\lambda \big\rangle + O(\|(g_\lambda, 0)\|_{\mathcal{E}}^2).
$$

Choosing $\eta > 0$ small enough so that Lemma 2.15 applies, we see that

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

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

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

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 $\boldsymbol{w}_n = (w_n, 0)$ with $\|\boldsymbol{w}_n\|_{\mathcal{E}} \le C$, and sequences $\vec{\iota}_n, \vec{\sigma}_n, \vec{\lambda}_n, \vec{\mu}_n$ so that,

$$\|\boldsymbol{w}_n - \boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n)\|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \Big(\frac{\lambda_{n,j}}{\lambda_{n,j+1}}\Big)^{\frac{D-2}{2}} = o_n(1) \text{ as } n \to \infty,$$
$$\|\boldsymbol{w}_n - \boldsymbol{\mathcal{W}}(\vec{\sigma}_n, \vec{\mu}_n)\|_{\mathcal{E}}^2 + \sum_{j=1}^{M-1} \Big(\frac{\mu_{n,j}}{\mu_{n,j+1}}\Big)^{\frac{D-2}{2}} = o_n(1) \text{ as } n \to \infty,$$

but with $\vec{\iota}_n \neq \vec{\sigma}_n$ for every n. Passing to a subsequence we may assume that there exists an index $j_0 \in \{1, \dots, M\}$ such that $\iota_{j,n} = \sigma_{j,n}$ for every $j > j_0$ and every n and $\iota_{j_0,n} \neq \sigma_{j_0,n}$ for every n. We first observe that,

$$\|\boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n) - \boldsymbol{\mathcal{W}}(\vec{\sigma}_n, \vec{\mu}_n)\|_{\mathcal{E}} \le \|\boldsymbol{w}_n - \boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n)\|_{\mathcal{E}} + \|\boldsymbol{w}_n - \boldsymbol{\mathcal{W}}(\vec{\sigma}_n, \vec{\mu}_n)\|_{\mathcal{E}} = o_n(1). \tag{2.19}$$

We first show that $j_0 < M$. Assume for contradiction that $j_0 = M$. Then, we may assume that $\iota_{n,M} = 1$, $\sigma_{n,M} = -1$ and $\lambda_{n,M} > \mu_{n,M}$ for all n. It follows that there exists a constant $c > 0$ for which

$$\Big|\mathcal{W}(\vec{\iota}_n, \vec{\lambda}_n) - \mathcal{W}(\vec{\sigma}_n, \vec{\mu}_n)\Big| \ge \frac{c}{\lambda_{n,M}^{\frac{D-2}{2}}} \quad \forall r \in [\lambda_{n,M}, 2\lambda_{n,M}],$$

for all n large enough. But then,

$$\|\boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n) - \boldsymbol{\mathcal{W}}(\vec{\sigma}_n, \vec{\mu}_n)\|_{\mathcal{E}}^2 \ge \int_{\lambda_{n,M}}^{2\lambda_{n,M}} \frac{c^2}{\lambda_{n,M}^{D-2}} r^{D-2} \frac{\mathrm{d}r}{r} \ge \frac{c^2}{D-2}(2^{D-2} - 1),$$

for all sufficiently large n, which contradicts (2.19). So $\iota_{1,n} = \sigma_{n,1}$ for all n. Thus $j_0 < M$. By a nearly identical argument we can show that we must have $|\lambda_{n,j}/\mu_{n,j} - 1| = o_n(1)$ for all $j > j_0$. Next, 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,

$$\Big|\mathcal{W}(\vec{\iota}_n, \vec{\lambda}_n) - \mathcal{W}(\vec{\sigma}_n, \vec{\mu}_n)\Big| \ge \frac{c}{\lambda_{n,j_0}^{\frac{D-2}{2}}} \quad \forall r \in [\lambda_{n,j_0}, 2\lambda_{n,j_0}],$$

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

Finally, we prove (2.18). Suppose (2.18) fails. Then there exists $\theta_0 > 0$ and sequences $\vec{\lambda}_n, \vec{\mu}_n$ such that

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

but

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

for all n. We arrive at a contradiction following the same logic as before. □

We require the following lemma, which gives the nonlinear interaction force between bubbles. Given an M-bubble configuration, $\mathcal{W}(\vec{\iota}, \vec{\lambda})$, we set

$$f_{\mathbf{i}}(\vec{\iota}, \vec{\lambda}) := f(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - \sum_{j=1}^{M} \iota_j f(W_{\lambda_j}). \tag{2.20}$$

Lemma 2.21 *Let $D \geq 4$, $M \in \mathbb{N}$. For any $\theta > 0$ there exists $\eta > 0$ with the following property. Let $\mathcal{W}(\vec{\iota}, \vec{\lambda})$ be an M-bubble configuration with*

$$\sum_{j=0}^{M} \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} \leq \eta,$$

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

$$\begin{aligned}
&\Big|\Big\langle \Lambda W_{\lambda_j} \mid f_{\mathbf{i}}(\vec{\iota}, \vec{\lambda})\Big\rangle - \iota_{j-1}\frac{D-2}{2D}(D(D-2))^{\frac{D}{2}}\Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}} \\
&+ \iota_{j+1}\frac{D-2}{2D}(D(D-2))^{\frac{D}{2}}\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}}\Big| \\
&\leq \theta\Big(\Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}} + \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}}\Big)
\end{aligned}$$

where here $f_{\mathbf{i}}(\vec{\iota}, \vec{\lambda})$ is defined in (2.20).

Proof This is an explicit computation analogous to the one in [41, Lemma 2.27]. □

3 Localized Sequential Bubbling

The goal of this section is to prove a localized sequential bubbling lemma for sequences of solutions to (1.1) with vanishing averaged kinetic energy on a (relatively) expanding region of space. The main result, and the arguments used to prove it are in the spirit of the main theorems of Duyckaerts, Kenig, and Merle in [22]. To prove the compactness lemma in all space dimensions via a unified approach, we use the virial inequalities of Jia and Kenig to obtain vanishing of the error instead of the channels-of-energy type arguments from [12, 22, 71], which in those works was limited to either odd space dimensions or the subset of even space dimensions that satisfy $D \equiv 0 \pmod 4$.

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

$$\boldsymbol{\delta}_R(\boldsymbol{u}) := \inf_{M,\vec{\iota},\vec{\lambda}} \Big(\|(u - \mathcal{W}(\vec{\iota}, \vec{\lambda}), \dot{u})\|^2_{\mathcal{E}(r \le R)} + \sum_{j=1}^{M} \Big(\frac{\lambda_j}{\lambda_{j+1}} \Big)^{\frac{D-2}{2}} \Big)^{\frac{1}{2}}, \tag{3.1}$$

where the infimum above is taken over all $M \in \{0, 1, 2, \dots\}$, 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 $\rho_n > 0$ be a sequence of positive numbers and let $\boldsymbol{u}_n \in C([0, \rho_n]; \mathcal{E})$ be a sequence of solutions to (1.1) on the time intervals $[0, \rho_n]$ such that*

$$\limsup_{n\to\infty} \sup_{t \in [0,\rho_n]} \|\boldsymbol{u}_n(t)\|_{\mathcal{E}} < \infty. \tag{3.2}$$

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

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

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 In fact, the proof provides a sequence $t_n \in [0, \rho_n]$, $r_n \le R_n$ with $r_n \to \infty$, a non-negative integer M independent of n, scales $\vec{\lambda}_n \in (0, \infty)^M$, and a vector of signs $\vec{\iota} \in \{-1, 1\}^M$ (also independent of n), such that

$$\lim_{n\to\infty} \Big(\|\boldsymbol{u}(t_n) - \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}_n)\|^2_{\mathcal{E}(r \le r_n \rho_n)} + \sum_{j=1}^{M} \Big(\frac{\lambda_{n,j}}{\lambda_{n,j+1}} \Big)^{\frac{D-2}{2}} \Big)^{\frac{1}{2}} = 0.$$

3.1 Technical Lemmas

The proof of Lemma 3.1 requires 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 $\widetilde{a}_{k,n} = \min\{a_{1,n}, \dots, a_{k,n}\}$. 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 $n_{b_n} \leq n < n_{b_n+1}$. 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

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

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

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

$$|\{\tau \in [0, 1] : Mf(\tau) > \alpha\}| \leq \frac{3}{\alpha} \int_0^1 f(t)\mathrm{d}t, \qquad \text{for all } \alpha > 0, \tag{3.3}$$

see [61, 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* $\limsup_{n\to\infty} \int_0^1 g_n(t)\mathrm{d}t \leq 0$. *Then there exists a sequence $t_n \in [0, 1]$ such that*

$$\lim_{n\to\infty} Mf_n(t_n) = 0, \qquad \lim_{n\to\infty} f_n(t_n) = 0, \qquad \limsup_{n\to\infty} g_n(t_n) \leq 0.$$

Proof Let α_n be a sequence such that $\int_0^1 f_n(t)\mathrm{d}t \ll \alpha_n \ll 1$. Let $A_n := \{t \in [0, 1] : Mf_n(t) + f_n(t) \leq \alpha_n\}$. By (3.3), $\lim_{n\to\infty} |A_n| = 1$. 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

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

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

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

$$\lim_{n\to\infty} \int_{\sigma_n}^{\tau_n} \Big[\int_0^\infty \Big(\partial_r^2 u_n + \frac{D-1}{r}\partial_r u_n + |u_n|^{\frac{4}{D-2}} u_n\Big)\Big(r\partial_r u_n + \frac{D}{2}u_n\Big)\chi\, r^{D-1}\,\mathrm{d}r \\ + \int_0^\infty \partial_t u_n\Big(r\partial_r\partial_t u_n + \frac{D}{2}\partial_t u_n\Big)\chi\, r^{D-1}\,\mathrm{d}r\Big]\mathrm{d}t = 0, \tag{3.4}$$

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

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

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

$$\lim_{n\to\infty} \int_0^{R_n} (\partial_t u_n(\sigma_n))^2\, r^{D-1}\mathrm{d}r = 0 \ \text{ and } \ \lim_{n\to\infty} \int_0^{R_n} (\partial_t u_n(\tau_n))^2\, r^{D-1}\mathrm{d}r = 0. \tag{3.5}$$

For $t \in [\sigma_n, \tau_n]$, we have the following Jia-Kenig virial identity; see [42, Lemma 2.2 and Lemma 2.6].

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big\langle \partial_t u_n \mid (r\partial_r u_n + \frac{D}{2}u_n)\chi\Big\rangle = \int_0^\infty \partial_t u_n\Big(r\partial_r\partial_t u_n + \frac{D}{2}\partial_t u_n\Big)\chi\, r^{D-1}\,\mathrm{d}r \\ + \int_0^\infty \Big(\partial_r^2 u_n + \frac{D-1}{r}\partial_r u_n + |u_n|^{\frac{4}{D-2}} u_n\Big)\Big(r\partial_r u_n + \frac{D}{2}u_n\Big)\chi\, r^{D-1}\,\mathrm{d}r. \tag{3.6}$$

By the Cauchy-Schwarz inequality, the assumption (3.2) and (3.5), we see that

$$\lim_{n\to\infty} \int_0^\infty \big(|\partial_t u_n(\sigma_n)||r\partial_r u_n(\sigma_n) + \frac{D}{2}u_n(\sigma_n)| \\ + |\partial_t u_n(\tau_n)||r\partial_r u_n(\tau_n) + \frac{D}{2}u(\tau_n)|\big)\chi\, r^{D-1}\mathrm{d}r = 0.$$

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

$$
\begin{aligned}
f_n(t) &:= \int_0^{R_n} |\partial_t u_n(t,r)|^2\, r^{D-1}\,\mathrm{d}r,\\
g_n(t) &:= -\int_0^\infty \Big(\partial_r^2 u_n(t) + \frac{D-1}{r}\partial_r u_n(t) + |u_n(t)|^{\frac{4}{D-2}}\, u_n(t)\Big)\\
&\qquad \Big(r\partial_r u_n(t) + \frac{D}{2}u_n(t)\Big)\chi\, r^{D-1}\,\mathrm{d}r\\
&\qquad - \int_0^\infty \partial_t u_n(t)\Big(r\partial_r\partial_t u_n(t) + \frac{D}{2}\partial_t u_n(t)\Big)\chi\, r^{D-1}\,\mathrm{d}r
\end{aligned}
$$

(integrating by parts we see that g_n 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,

$$
\begin{aligned}
&\lim_{n\to\infty} \sup_{I\ni t_n;\, I\subset[0,1]} \frac{1}{|I|}\int_I\int_0^{R_n} |\partial_t u_n(t,r)|^2\, r^{D-1}\,\mathrm{d}r\,\mathrm{d}t = 0,\\
\text{and}\quad &\lim_{n\to\infty}\int_0^{R_n} |\partial_t u_n(t_n,r)|^2\, r^{D-1}\,\mathrm{d}r = 0,
\end{aligned}
\tag{3.7}
$$

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

$$
\begin{aligned}
\limsup_{n\to\infty}\Big(-\int_0^\infty \Big[\Big(\partial_r^2 u_n(t_n) + \frac{D-1}{r}\partial_r u_n(t_n) + |u_n(t)|^{\frac{4}{D-2}}\, u_n(t_n)\Big)&\\
\Big(r\partial_r u_n(t_n) + \frac{D}{2}u_n(t_n)\Big)&\\
+ \partial_t u_n(t_n)\Big(r\partial_r\partial_t u_n(t_n) + \frac{D}{2}\partial_t u_n(t_n)\Big)\Big]\chi\, r^{D-1}\,\mathrm{d}r\Big) \le 0.&
\end{aligned}
\tag{3.8}
$$

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 $(\boldsymbol{v}_{\mathrm{L}}^j, t_{n,j}, \lambda_{n,j})$, and $\boldsymbol{w}_{n,0}^J$, so that, using the notation,

$$
\boldsymbol{v}_{\mathrm{L},n}^j(0) := \Big(\lambda_{n,j}^{-\frac{D-2}{2}}\, v_{\mathrm{L}}^j\Big(\frac{-t_{n,j}}{\lambda_{n,j}}, \frac{\cdot}{\lambda_{n,j}}\Big), \lambda_{n,j}^{-\frac{D}{2}}\,\partial_t v_{\mathrm{L}}^j\Big(\frac{-t_{n,j}}{\lambda_{n,j}}, \frac{\cdot}{\lambda_{n,j}}\Big)\Big),
$$

we have

$$
\boldsymbol{u}_n(t_n) = \sum_{j=1}^J \boldsymbol{v}_{\mathrm{L},n}^j(0) + \boldsymbol{w}_{n,0}^J \tag{3.9}
$$

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

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

$$\mathcal{J}_{c,0} := \{j \in \mathbb{N} \mid t_{n,j} = 0 \ \forall n, \ \text{ and } \ \lim_{n\to\infty} \lambda_{n,j} < \infty\},$$
$$\mathcal{J}_{c,\infty} := \{j \in \mathbb{N} \mid t_{n,j} = 0 \ \forall n, \ \text{ and } \ \lim_{n\to\infty} \lambda_{n,j} = \infty\}.$$

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

$$\lim_{n\to\infty} \|(\lambda_{n,j}^{-\frac{D-2}{2}} v_{\mathrm{L}}^{j}(0, \cdot/\lambda_{n,j}), \lambda_{n,j}^{-\frac{D}{2}} \dot{v}_{\mathrm{L}}^{j}(0, \cdot/\lambda_{n,j}))\|_{\mathcal{E}(0, r_{0,n})} = 0 \tag{3.10}$$

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

Step 5. (Incoming/outgoing profiles with $\lim_{n\to\infty} |t_{n,j}| = \infty$.) We next treat profiles $(\boldsymbol{v}_{\mathrm{L}}^{j}, t_{n,j}, \lambda_{n,j})$ that satisfy,

$$-\frac{t_{n,j}}{\lambda_{n,j}} \to \pm\infty.$$

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

$$\mathcal{J}_{\mathrm{L},0} := \{j \in \mathbb{N} \mid -\frac{t_{n,j}}{\lambda_{n,j}} \to \pm\infty \ \text{ and } \ -t_{n,j} \to t_{\infty,j} \in \mathbb{R}\},$$
$$\mathcal{J}_{\mathrm{L},\infty} := \{j \in \mathbb{N} \mid -\frac{t_{n,j}}{\lambda_{n,j}} \to \pm\infty \ \text{ and } \ |t_{n,j}| \to \infty\}.$$

We impose additional restrictions on the sequence r_n. We require that $r_n \leq \frac{1}{2}|t_{n,j}|$ for large enough n, depending on j, for each sequence $t_{n,j}$ in $\mathcal{J}_{\mathrm{L},\infty}$. 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 $r_{1,n} \leq \frac{1}{2}|t_{n,j}|$ for large enough n, depending on j, for each sequence $t_{n,j}$ in $\mathcal{J}_{\mathrm{L},\infty}$.

Since $\boldsymbol{v}_{\mathrm{L}}^{j}$ is a free wave we know that it asymptotically concentrates all of its energy near the light-cone. In fact,

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

which is proved in [21, Lemma 4.1] in odd space dimensions and is a direct consequence of [13, Theorem 5] in even space dimensions.

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

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

by the above.

Step 6. (Incoming/outgoing profiles with $\lim_{n\to\infty}\left|t_{n,j}\right| < \infty$.) Next, we consider profiles $(\boldsymbol{v}_{\mathrm{L}}^{j}, t_{n,j}, \lambda_{n,j})$ such that

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

that is, those in $\mathcal{J}_{\mathrm{L},0}$. We note that $\lambda_{n,i} \to 0$ as $n \to \infty$ for each $i \in \mathcal{J}_{\mathrm{L},0}$. We claim that any such profile must satisfy $\boldsymbol{v}_{\mathrm{L}}^{i} \equiv 0$. We use the argument given in [19, Erratum], modulo a few technicalities which reduce our situation to the one considered there.

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

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

for some $1 \ll A_n \ll r_{2,n}$. By Lemma 2.6 it suffices to have

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

and then replace A_n by its quarter, 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

$$\sum_{l=0}^{A_n-1}\|\boldsymbol{u}_n(0)\|^2_{\mathcal{E}(A_n^{-1}R_n^{(l)},A_n R_n^{(l)})} \leq \|\boldsymbol{u}_n(0)\|^2_{\mathcal{E}}.$$

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

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

proving (3.12)

We now pass to a new sequence $\widetilde{\boldsymbol{u}}_n$ with vanishing average kinetic energy on the whole space. Indeed, define

$$\widetilde{\boldsymbol{u}}_n(t_n) = \chi_{2r_{2,n}}\boldsymbol{u}_n(t_n).$$

Denoting by $\widetilde{\boldsymbol{u}}_n(t)$ the solutions to (1.1) with data $\widetilde{\boldsymbol{u}}_n(t_n)$ on the interval $t \in [0, 1]$ we can use the finite speed of propagation, the local Cauchy theory, and (3.12) to deduce that

$$\widetilde{\boldsymbol{u}}_n(t,r) = \boldsymbol{u}_n(t,r) \text{ if } r \le r_{2,n}, \quad \text{and } \limsup_{n\to\infty} \sup_{t\in[0,1]} \|\widetilde{\boldsymbol{u}}_n(t)\|_{\mathcal{E}(r_{2,n},\infty)} = 0. \tag{3.13}$$

From (3.9) we have,

$$\begin{aligned}\widetilde{\boldsymbol{u}}_n(t_n) = &\sum_{1\le j\le J;\, j\in\mathcal{J}_{c,0}\cup\mathcal{J}_{\mathrm{L},0}} \boldsymbol{v}^j_{\mathrm{L},n}(0) + \chi_{2r_{2,n}}\boldsymbol{w}^J_{n,0} \\ &+ \sum_{1\le j\le J;\, j\in\mathcal{J}_{c,0}\cup\mathcal{J}_{\mathrm{L},0}} (\chi_{2r_{2,n}}-1)\boldsymbol{v}^j_{\mathrm{L},n}(0) + \sum_{1\le j\le J;\, j\in\mathcal{J}_{c,\infty}\cup\mathcal{J}_{\mathrm{L},\infty}} \chi_{2r_{2,n}}\boldsymbol{v}^j_{\mathrm{L},n}(0).\end{aligned}$$

We claim that in fact $\widetilde{\boldsymbol{u}}_n(t_n)$ admits a profile decomposition in the sense of Lemma 2.10 of the form,

$$\widetilde{\boldsymbol{u}}_n(t_n) = \sum_{1\le j\le J;\, j\in\mathcal{J}_{c,0}\cup\mathcal{J}_{\mathrm{L},0}} \boldsymbol{v}^j_{\mathrm{L},n}(0) + \widetilde{\boldsymbol{w}}^J_{n,0}, \tag{3.14}$$

with the same profiles $(\boldsymbol{v}^j_{\mathrm{L}}, \lambda_{n,j}, t_{n,j})$ as in the decomposition for $\boldsymbol{u}_n(t_n)$ and where the error above, $\widetilde{\boldsymbol{w}}^J_{n,0}$, satisfies,

$$\widetilde{\boldsymbol{w}}^J_{n,0} = \chi_{2r_{2,n}}\boldsymbol{w}^J_{n,0} + o_n(1) \text{ as } n\to\infty.$$

The expansion (3.14) and the $o_n(1)$ above is justified as follows: let $\epsilon > 0$ and use (2.7) to find $J_0 > 0$ such that

$$\sum_{j>J_0} \|\boldsymbol{v}^j_{\mathrm{L},n}(0)\|^2_{\mathcal{E}} \le \epsilon.$$

Using (3.10), (3.11) we see that,

$$\sum_{j\le J_0,\, j\in\mathcal{J}_{c,\infty}\cup\mathcal{J}_{\mathrm{L},\infty}} \|\chi_{2r_{2,n}}\boldsymbol{v}^j_{\mathrm{L},n}(0)\|^2_{\mathcal{E}} \to 0 \text{ as } n\to\infty.$$

Using the same logic used to deduce (3.10), (3.11) we have (since $r_{2,n}\to\infty$),

$$\sum_{1\le j\le J_0;\, j\in\mathcal{J}_{c,0}\cup\mathcal{J}_{\mathrm{L},0}} \|(1-\chi_{2r_{2,n}})\boldsymbol{v}^j_{\mathrm{L},n}(0)\|^2_{\mathcal{E}} \to 0 \text{ as } n\to\infty,$$

from which the vanishing of the $o_n(1)$ term follows. It remains to deduce the vanishing properties of the error $\widetilde{\boldsymbol{w}}^J_{n,0}$, which follow directly from [21, Claim A.1 and Lemma 2.1] in the odd dimensional case and [13, Lemma 10 and 11] in the case of even dimensions.

Finally, we can use (3.7) and (3.13) to see that,

$$\begin{aligned}&\lim_{n\to\infty}\sup_{I\ni t_n; I\subset[0,1]}\frac{1}{|I|}\int_I\int_0^\infty|\partial_t\widetilde{u}_n(t,r)|^2\,r^{D-1}\,\mathrm{d}r\,\mathrm{d}t=0,\\&\|\partial_t\widetilde{u}_n(t_n)\|_{L^2}\to 0\ \text{ as }\ n\to\infty.\end{aligned}\tag{3.15}$$

Then following the exact argument in [19, Erratum], but applied to $\widetilde{\boldsymbol{u}}_n(t_n)$ we conclude that the set $\mathcal{J}_{\mathrm{L},0}$ is empty, i.e., all of the profiles $(\boldsymbol{v}_{\mathrm{L}}^j,\lambda_{n,j},t_{n,j})$ with $j\in\mathcal{J}_{\mathrm{L},0}$ satisfy, $\boldsymbol{v}_{\mathrm{L}}^j\equiv 0$.

Step 7. (Centered profiles at bounded scales.) To recap, we now have

$$\widetilde{\boldsymbol{u}}_n(t_n)=\sum_{1\le j\le J;\,j\in\mathcal{J}_{c,0}}\boldsymbol{v}_{\mathrm{L},n}^j(0)+\widetilde{\boldsymbol{w}}_{n,0}^J$$

where $\widetilde{\boldsymbol{u}}_n(t_n)$ satisfies (3.13) and all of the profiles $(\boldsymbol{v}_{\mathrm{L}}^j,\lambda_{n,j},0)$ have $t_{n,j}=0$ and $\lambda_{n,j}\lesssim 1$ for all n, j. Moreover, we have the vanishing in (3.15). We can now apply the exact same arguments of Duyckaerts, Kenig, and Merle [19, Proofs of Corollary 4.1 and Corollary 4.2] (see also the identical arguments by Rodriguez in [71] and Jia and Kenig [42]) to deduce that in fact either

$$\boldsymbol{v}_{\mathrm{L},n}^j(0)=\iota_j\boldsymbol{W}_{\lambda_{n,j}}\ \text{ or }\ \boldsymbol{v}_{\mathrm{L},n}^j(0)=\boldsymbol{0}$$

for $\iota_j\in\{-1,1\}$ for each $j\in\mathcal{J}_{c,0}$. By (2.7) there can only be finitely many of these profiles that are non-trivial, and thus after reordering the indices we can find $K_0\ge 0$ and $\lambda_{n,1}\ll\lambda_{n,2}\ll\cdots\ll\lambda_{n,K_0}\lesssim 1$ such that

$$\widetilde{\boldsymbol{u}}_n(t_n)=\sum_{1\le j\le K_0}\iota_j\boldsymbol{W}_{\lambda_{n,j}}+\widetilde{\boldsymbol{w}}_{n,0}\tag{3.16}$$

where $\widetilde{\boldsymbol{w}}_{n,0}:=\widetilde{\boldsymbol{w}}_{n,0}^{K_0}$. We note that the error $\widetilde{\boldsymbol{w}}_{n,0}$ satisfies,

$$\limsup_{n\to\infty}\left(\|\widetilde{w}_{n,0}\|_{L^{\frac{2D}{D-2}}}+\|\dot{\widetilde{w}}_{n,0}\|_{L^2}\right)=0\tag{3.17}$$

where the L^2 vanishing of $\dot{\widetilde{w}}_{n,0}$ follows from (3.15) and the decomposition (3.16).

Step 8. (Vanishing properties of the error.) We now select the final sequence by choosing $r_n\to\infty$ so that

$$r_n\le\frac{1}{2}r_{2,n},\quad\lim_{n\to\infty}\|\widetilde{w}_{n,0}\|_{H(\frac{1}{4}r_n^{-1},4r_n)}=0.\tag{3.18}$$

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,

$$\|\widetilde{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(\lambda_nA^{-1}\le r\le\lambda_nA)}\to 0\ \text{ as }\ n\to\infty.\tag{3.19}$$

The property (3.19) was proved in [11, 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 $W_{\lambda_{n,j}}$ with $\lambda_{n,j} \simeq \lambda_n$. To prove (3.18) we consider the case $\lambda_n = 1$ in (3.19), and passing to a subsequence of the $\widetilde{\boldsymbol{u}}_n$, we obtain a sequence as in (3.18).

We truncate to the region $r \le r_n$, following the same procedure used to define $\widetilde{\boldsymbol{u}}_n$ in Step 6, using now r_n in place of $r_{2,n}$. Indeed, set

$$\breve{\boldsymbol{u}}_n(t_n, r) := \chi_{2r_n}(r)\widetilde{\boldsymbol{u}}_n(t_n, r).$$

Setting $\breve{\boldsymbol{w}}_{n,0} := \chi_{2r_n}\widetilde{\boldsymbol{w}}_{n,0} + (\chi_{2r_n} - 1)\sum_{j=1}^{K_0} \iota_j \boldsymbol{W}_{\lambda_{n,j}}$ and using that $\lambda_{n,K_0} \lesssim 1$ along with (3.17) and (3.18) we see that,

$$\begin{aligned} &\breve{\boldsymbol{u}}_n(t_n) = \sum_{j=1}^{K_0} \iota_j \boldsymbol{W}_{\lambda_{n,j}} + \breve{\boldsymbol{w}}_{n,0}, \quad \text{and} \\ &\lim_{n\to\infty} \Big(\|\breve{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(r_n^{-1} \le r < \infty)} + \|\dot{\breve{w}}_{n,0}\|_{L^2} + \|\breve{w}_{n,0}\|_{L^{\frac{2D}{D-2}}} \Big) = 0. \end{aligned} \tag{3.20}$$

Letting $\breve{\boldsymbol{u}}_n(t)$ denote the nonlinear evolution of $\breve{\boldsymbol{u}}_n(t_n)$ we see from (3.13) that

$$\breve{\boldsymbol{u}}_n(t, r) = \boldsymbol{u}_n(t, r) \ \text{ if } \ r \le r_n \tag{3.21}$$

and from (3.15) that

$$\begin{aligned} &\lim_{n\to\infty} \sup_{I \ni t_n; I \subset [0,1]} \frac{1}{|I|} \int_I \int_0^\infty |\partial_t \breve{u}_n(t, r)|^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}t = 0, \\ &\|\partial_t \breve{u}_n(t_n)\|_{L^2} \to 0 \ \text{ as } \ n \to \infty. \end{aligned}$$

By (3.21) and (3.20) it remains to show the vanishing,

$$\lim_{n\to\infty} \|\breve{\boldsymbol{w}}_{n,0}\|_{\mathcal{E}(r \le r_n^{-1})} = 0.$$

It is at this stage where we use the Jia-Kenig virial functional. By (3.8) we have

$$\begin{aligned} \limsup_{n\to\infty} \Big(-\int_0^\infty \Big[&\Big(\partial_r^2 \breve{u}_n(t_n) + \frac{D-1}{r} \partial_r \breve{u}_n(t_n) + |\breve{u}_n(t)|^{\frac{4}{D-2}} u_n(t_n) \Big) \\ &\Big(r\partial_r \breve{u}_n(t_n) + \frac{D}{2} \breve{u}_n(t_n) \Big) \\ + \partial_t \breve{u}_n(t_n) &\Big(r\partial_r \partial_t \breve{u}_n(t_n) + \frac{D}{2} \partial_t \breve{u}_n(t_n) \Big) \Big] \chi \, r^{D-1} \, \mathrm{d}r \Big) \le 0. \end{aligned} \tag{3.22}$$

We distinguish between two cases, $\lim_{n\to\infty} \lambda_{n,K_0} = 0$ and $\lim_{n\to\infty} \lambda_{n,K_0} > 0$, which require slightly different arguments. First suppose $\lim_{n\to\infty} \lambda_{n,K_0} = 0$. Using (3.20) we see that in this case,

$$\lim_{n\to\infty} \|\breve{\boldsymbol{u}}_n\|_{\mathcal{E}(\frac{1}{2},\infty)} = 0.$$

Integration by parts in (3.22), we obtain,

$$\limsup_{n\to\infty}\bigg(\int_0^\infty \Big[(\partial_r \breve{u}_n(t_n))^2 - |\breve{u}_n(t_n)|^{\frac{2D}{D-2}}\Big]\chi\, r^{D-1}\,\mathrm{d}r \\ - \Omega_{1,1}(\breve{\boldsymbol{u}}_n(t_n)) - \frac{D}{2}\Omega_{2,1}(\breve{\boldsymbol{u}}_n(t_n))\bigg) \le 0$$

where $\Omega_{1,1}, \Omega_{2,1}$ are defined in (2.3). Using (3.20) along with Sobolev embedding and Hardy's inequality we obtain the vanishing of the errors terms $\Omega_{j,1}(\breve{\boldsymbol{u}}_n(t_n))$ above and we conclude that,

$$\limsup_{n\to\infty}\int_0^\infty \Big[(\partial_r \breve{u}_n(t_n))^2 - |\breve{u}_n(t_n)|^{\frac{2D}{D-2}}\Big]\chi\, r^{D-1}\,\mathrm{d}r \le 0. \tag{3.23}$$

Due to (3.20), the orthogonality of the profiles (i.e., $\lambda_{n,1} \ll \lambda_{n,2} \ll \dots \lambda_{n,K_0}$), the fact that $\lambda_{n,K_0} \to 0$, and the fact that the Jia-Kenig functional vanishes at W, i.e.,

$$\int_0^\infty \Big[(\partial_r W)^2 - |W|^{\frac{2D}{D-2}}\Big] r^{D-1}\,\mathrm{d}r = 0$$

we can conclude that

$$\limsup_{n\to\infty}\int_0^\infty \Big[(\partial_r \breve{w}_{0,n})^2 - \big|\breve{w}_{0,n}\big|^{\frac{2D}{D-2}}\Big]\chi\, r^{D-1}\,\mathrm{d}r \le 0.$$

But then we may use (3.20) to see that in fact

$$\lim_{n\to\infty}\int_0^\infty \big|\breve{w}_{0,n}\big|^{\frac{2D}{D-2}}\,\chi\, r^{D-1}\,\mathrm{d}r = 0.$$

Using this estimate in the previous line we conclude that

$$\lim_{n\to\infty}\int_0^\infty (\partial_r \breve{w}_{0,n})^2\chi\, r^{D-1}\,\mathrm{d}r = 0$$

and combining with (3.20) we have $\lim_{n\to\infty}\|\partial_r \breve{w}_{0,n}\|_{L^2} = 0$. By Hardy's inequality we deduce finally that $\lim_{n\to\infty}\|\breve{\boldsymbol{w}}_{0,n}\|_{\mathcal{E}} = 0$. Lastly, by (3.20) and the fact that $\boldsymbol{u}_n(t_n, r) = \breve{\boldsymbol{u}}_n(t_n, r)$ if $r \le r_n$, we see that $\lim_{n\to\infty}\boldsymbol{\delta}_{r_n}(\boldsymbol{u}_n(t_n)) = 0$, completing the proof in the case $\lim_{n\to\infty}\lambda_{n,K_0} = 0$.

Next, consider the case $\lim_{n\to\infty}\lambda_{n,K_0} = \lambda_\infty > 0$. We may assume that n is sufficiently large so that

$$\lambda_{n,K_0} \in \Big(\frac{1}{2}\lambda_\infty, 2\lambda_\infty\Big). \tag{3.24}$$

To ease notation, define $\boldsymbol{\psi}_n = \breve{\boldsymbol{u}}_n(t_n) - \iota_{K_0}\boldsymbol{W}_{\lambda_{n,K_0}}$ and $\phi_n := \iota_{K_0}\boldsymbol{W}_{\lambda_{n,K_0}}$. Defining the auxiliary sequence $\zeta_n := r_n^{-1} + \sqrt{\lambda_{n,K_0-1}}$, which satisfies $\zeta_n \to 0$ as $n \to \infty$, we see that

$$\lim_{n\to\infty} \|\boldsymbol{\psi}_n\|_{\mathcal{E}(\zeta_n;\infty)} = 0 \text{ and } \lim_{n\to\infty} \|\boldsymbol{\phi}_n\|_{\mathcal{E}(0,2\zeta_n)} = 0. \tag{3.25}$$

Note that integration by parts of the terms in the second line of (3.22) yields the term,

$$-\frac{1}{2}\int_0^\infty (\partial_t \breve{u}_n(t_n))^2 r\chi'(r)\, r^{D-1}\,\mathrm{d}r$$

which tends to zero as $n \to \infty$ by (3.20). For the first line in (3.22) we first express $\breve{u}_n(t_n) = \psi_n + \phi_n$. Noting the point-wise vanishing $\partial_r^2\phi_n + \frac{D-1}{r}\partial_r\phi_n + |\phi_n|^{\frac{4}{D-2}}\phi_n = 0$, and employing the notation $f(v) = |v|^{\frac{4}{D-2}}v$, we rewrite the first term in the integrand of the first line in (3.22) as,

$$\begin{aligned}\partial_r^2\breve{u}_n(t_n) + \frac{D-1}{r}\partial_r\breve{u}_n(t_n) + f(\breve{u}_n(t_n)) &= \partial_r^2\psi_n(t_n) + \frac{D-1}{r}\partial_r\psi_n + f(\psi_n)\\ &\quad + \Big(f(\psi_n+\phi_n) - f(\psi_n) - f(\phi_n)\Big).\end{aligned}$$

We then write,

$$\begin{aligned}&-\int_0^\infty \bigg[\Big(\partial_r^2\breve{u}_n(t_n) + \frac{D-1}{r}\partial_r\breve{u}_n(t_n) + |\breve{u}_n(t)|^{\frac{4}{D-2}}u_n(t_n)\Big)\\
&\qquad\qquad \Big(r\partial_r\breve{u}_n(t_n) + \frac{D}{2}\breve{u}_n(t_n)\Big)\chi\, r^{D-1}\,\mathrm{d}r\\
&\quad= -\int_0^\infty \bigg[\Big(\partial_r^2\psi_n(t_n) + \frac{D-1}{r}\partial_r\psi_n + f(\psi_n)\Big)\Big(r\partial_r\psi_n + \frac{D}{2}\psi_n\Big)\chi\, r^{D-1}\,\mathrm{d}r\\
&\quad -\int_0^\infty \bigg[\Big(\partial_r^2\psi_n(t_n) + \frac{D-1}{r}\partial_r\psi_n + f(\psi_n)\Big)\Big(r\partial_r\phi_n + \frac{D}{2}\phi_n\Big)\chi\, r^{D-1}\,\mathrm{d}r\\
&\quad -\int_0^\infty \bigg[\Big(f(\psi_n+\phi_n) - f(\psi_n) - f(\phi_n)\Big)\Big(r\partial_r\phi_n + \frac{D}{2}\phi_n\Big)\chi\, r^{D-1}\,\mathrm{d}r\\
&\quad -\int_0^\infty \bigg[\Big(f(\psi_n+\phi_n) - f(\psi_n) - f(\phi_n)\Big)\Big(r\partial_r\psi_n + \frac{D}{2}\psi_n\Big)\chi\, r^{D-1}\,\mathrm{d}r.\end{aligned}$$

The last three lines above tend to zero as $n \to \infty$ which can be seen by dividing the integrals into the regions $r \le \zeta_n$ and $r \in (\zeta_n, 2)$, using (3.25), (3.24), and additionally in the last two lines the pointwise estimate,

$$|f(\psi_n+\phi_n) - f(\psi_n) - f(\phi_n)| \lesssim |\psi_n|^{\frac{4}{D-2}}|\phi_n| + |\phi_n|^{\frac{4}{D-2}}|\psi_n|.$$

From (3.22) we then conclude that,

$$\limsup_{n\to\infty}\Big(-\int_0^\infty\Big(\partial_r^2\psi_n(t_n)+\frac{D-1}{r}\partial_r\psi_n+f(\psi_n)\Big)\Big(r\partial_r\psi_n+\frac{D}{2}\psi_n\Big)\chi\, r^{D-1}\,\mathrm{d}r\Big)\le 0.$$

Integration by parts and arguing as in the proof of (3.23) then yields,

$$\limsup_{n\to\infty}\int_0^\infty\Big[(\partial_r\psi_n)^2-|\psi_n|^{\frac{2D}{D-2}}\Big]\chi\, r^{D-1}\,\mathrm{d}r\le 0.$$

We have now reduced matters to the previous case as we have $\lim_{n\to\infty}\lambda_{n,K_0-1}=0$ and we conclude as before that $\lim_{n\to\infty}\|\breve{\boldsymbol{w}}_{0,n}\|_{\mathcal{E}}=0$ and hence, $\lim_{n\to\infty}\boldsymbol{\delta}_{r_n}(\boldsymbol{u}_n(t_n))=0$, completing the proof. □

4 The Sequential Decomposition

In this section we sketch the proof of Theorem 1.14, the sequential decomposition. We view this result as the consequence of three main ingredients: (1) the identification of the linear radiation $\boldsymbol{u}^*$, (2) a proof that no energy can concentrate in the self-similar region of the light cone, and (3) the compactness lemma proved in the previous section.

4.1 Identification of the Radiation

The results in this subsection were proved by Duyckaerts, Kenig, and Merle in [21, 22] in the case $D=3$. Following their approach, analogous results were obtained in [12] in dimension $D=4$, [42] in dimension $D=6$, and in [71] for all odd $D\ge 5$. The case of even dimensions $D>6$ follows from an identical argument as the one used in [71].

Proposition 4.1 *(Radiation in case of finite time blow-up) [21, Theorem 3.2] Let $\boldsymbol{u}\in C(I;\mathcal{E})$ be a solution to (1.1) defined on the time interval $I=(0,T]$ for some $T>0$ and blowing up in the type-II sense as $t\searrow 0$, that is, such that*

$$\sup_{t\in(0,T]}\|\boldsymbol{u}(t)\|_{\mathcal{E}}<\infty.$$

Then, there exists $\boldsymbol{u}_0^*\in\mathcal{E}$ such that

$$\boldsymbol{u}(t)\rightharpoonup\boldsymbol{u}_0^*\quad\text{weakly in } \mathcal{E}\ \text{ as }\ t\to 0,$$
$$\|\varphi(\boldsymbol{u}(t)-\boldsymbol{u}_0^*)\|_{\mathcal{E}}\to 0\ \text{ as }\ t\to 0,$$

where the latter limit holds for any $\varphi\in C_0^\infty(0,\infty)$. Moreover, there exists $T_0>0$ such that the solution $\boldsymbol{u}^*(t)$ of (1.1) with initial data $\boldsymbol{u}_0^*$ is defined on the interval $[0,T_0]$ and satisfies,

$$\boldsymbol{u}(t,r) = \boldsymbol{u}^*(t,r) \ \ \textit{if} \ \ r \geq t, \quad \forall t \in (0, T_0],$$
$$\lim_{t\to 0} E(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) = E(\boldsymbol{u}) - E(\boldsymbol{u}^*).$$

Proposition 4.2 (Radiation for a global-in-time solution) [22, Corollary 3.9], [12, Proposition 4.1] *Let $\boldsymbol{u} \in C(I; \mathcal{E})$ be a solution to (1.1) defined on the time interval $I = [T, \infty)$ for some $T \geq 0$ and such that*

$$\sup_{t\in[T,\infty)} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

Then, there exists a free wave $\boldsymbol{u}_{\mathrm{L}}^* \in C(\mathbb{R}; \mathcal{E})$ such that

$$\vec{S}_{\mathrm{L}}(-t)\boldsymbol{u}(t) \rightharpoonup \boldsymbol{u}_{\mathrm{L}}^*(0) \quad \textit{weakly in } \mathcal{E} \ \textit{as} \ t \to \infty,$$
$$\textit{and } \forall R \in \mathbb{R}, \quad \lim_{t\to\infty} \int_{t-R}^{\infty} \Big[(\partial_t(u(t) - u_{\mathrm{L}}^*(t)))^2 + (\partial_r(u(t) - u_{\mathrm{L}}^*(t)))^2 + \frac{(u(t) - u_{\mathrm{L}}^*(t))^2}{r^2} \Big] r^{D-1}\,\mathrm{d}r = 0.$$

Moreover, if we denote by $\boldsymbol{u}^*$ the unique solution to (1.1) given by Lemma 2.3 such that

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

then,

$$\lim_{t\to\infty} E(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) = E(\boldsymbol{u}) - E(\boldsymbol{u}^*).$$

Remark 4.3 We note that the proof of Proposition 4.1 given in [21] was given only in dimensions $D = 3, 4, 5$ (and for non-radially symmetric data), but it generalizes in a straight-forward way to higher space dimensions using the local Cauchy theory from Lemma 2.3. The proof of Proposition 4.2 is given in dimension $D = 3$ in [22] and was generalized to dimension $D = 4$ in [12] using technical tools related to profile decompositions in even space dimensions proved by Côte, Kenig, and Schlag in [13]. It was proved in all odd space dimensions in [71]. Again the proofs given in those references generalize to all even space dimensions, using [13] .

4.2 Non-concentration of Self-similar Energy

In this section we sketch the proof that finite energy solutions cannot concentrate linear energy in the self-similar region of the cone. As a consequence of this fact and virial identities, we deduce the vanishing of the averaged kinetic energy in the cone. The proof in this section closely follows the arguments given in [12] and [42], which in turn follow the scheme developed by Christodoulou and Tahvildar-Zadeh [9] and Shatah and Tahvildar-Zadeh [79] in the context of equivariant wave maps. We make one minor observation here, namely that the reductions performed in [12, 42] from the D-dimensional radially symmetric NLW (1.1) to a wave maps-type equation in two-space dimensions work equally well in all space dimensions $D \geq 3$, and thus the arguments from [42] (which generalized the Shatah, Tahvildar-Zadeh arguments to cover all finite energy solutions) apply directly[1].

Proposition 4.4 (No self-similar concentration for blow-up solutions) [42, Theorem 2.1] *Let $\boldsymbol{u} \in C(I; \mathcal{E})$ be a solution to* (1.1) *defined on the time interval $I = (0, T]$ for some $T > 0$ and blowing up in the type-II sense as $t \searrow 0$, that is, such that*

$$\sup_{t \in (0,T]} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

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

$$\lim_{t \searrow 0} \int_{\alpha t}^{t} \Big[(\partial_t u(t,r))^2 + (\partial_r u(t,r))^2 + \frac{(u(t,r))^2}{r^2}\Big] r^{D-1}\, \mathrm{d}r = 0.$$

Proposition 4.5 (No self-similar concentration for global solutions) [42, Theorem 2.4] *Let $\boldsymbol{u} \in C(I; \mathcal{E})$ be a solution to* (1.1) *defined on the time interval $I = [T, \infty)$ for some $T \geq 0$ and such that*

$$\sup_{t \in [T,\infty)} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

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

$$\lim_{R \to \infty} \limsup_{t \to \infty} \int_{\alpha t}^{t-R} \Big[(\partial_t u(t,r))^2 + (\partial_r u(t,r))^2 + \frac{(u(t,r))^2}{r^2}\Big] r^{D-1}\, \mathrm{d}r = 0.$$

Corollary 4.6 (Time-averaged vanishing of kinetic energy for blow-up solutions) *[42, Lemma 2.2] Let $\boldsymbol{u}$ be a solution to* (1.1) *satisfying the hypothesis of Proposition 4.4. Then,*

$$\lim_{\tau \searrow 0} \frac{1}{\tau} \int_0^{\tau} \int_0^{t} (\partial_t u(t,r))^2\, r^{D-1}\, \mathrm{d}r\, \mathrm{d}t = 0.$$

Corollary 4.7 (Time-averaged vanishing of kinetic energy for global solutions) *[42, Lemma 2.6] Let $\boldsymbol{u}$ be a solution to* (1.1) *satisfying the hypothesis of Proposition 4.5. Then,*

$$\lim_{\tau \to \infty} \frac{1}{\tau} \int_0^{\tau} \int_0^{\frac{t}{2}} (\partial_t u(t,r))^2\, r^{D-1}\, \mathrm{d}r\, \mathrm{d}t = 0.$$

[1] We note that the published version of [12] contained a gap in the proof of the corresponding results, as the arguments used to deduce Proposition 4.4 and Proposition 4.5 in that paper were performed only for smooth solutions. This gap was closed by an argument of the first author and was included in an appendix to [14]. An earlier argument by Jia and Kenig from [42] can also be used to close the gap in [12], and we refer to their approach here.

Remark 4.8 The proofs of Proposition 4.4 and Proposition 4.5 in [12, 42] are done for the cases $D = 4, 6$ and are based on the following reduction, which we generalize to cover all dimensions $D \geq 3$. Let

$$k := \frac{D-2}{2}.$$

Given $\boldsymbol{u}(t) \in \mathcal{E}$, set,

$$\boldsymbol{\psi}(t, r) := (r^k u(t, r), r^k \partial_t u(t, r)). \tag{4.1}$$

We see that $\boldsymbol{u}(t)$ solves (1.1) if and only if $\boldsymbol{\psi}(t, r)$ solves,

$$\partial_t^2 \psi - \partial_r^2 \psi - \frac{1}{r}\partial_r \psi + \frac{k^2 - |\psi|^{\frac{2}{k}}}{r^2}\psi = 0$$

which bears enough structural similarities to the equivariant wave maps equation that the main elements of the arguments from [9, 79] carry over. The key feature for our purposes, is that

$$F_k(\psi) := \frac{1}{2}|\psi|^2\left(k^2 - \frac{k}{k+1}|\psi|^{\frac{2}{k}}\right)$$

is positive when $|\psi(t, r)|$ is sufficiently small and hence so is the flux density,

$$\frac{1}{2}(\partial_t \psi(t, r) - \partial_r \psi(t, r))^2 + \frac{F_k(\psi(t, r))}{r^2} > 0.$$

Up to changing the values of some constants, the line-by-line arguments in [42, Proof of Theorem 2.1] and [42, Proof of Theorem 2.4] are valid in any dimension $D \geq 4$ with $\boldsymbol{\psi}$ defined as in (4.1).

Remark 4.9 The proof of Corollary 4.6 follows from the virial identity (2.4) with the cutoff at $\rho(t) = t/2$ together with Proposition 4.4. The exact argument in [42, Proof of Lemma 2.2, in particular Eq. (2.66)] applies in our setting as well. The proof of Corollary 4.7 is similar, using now Proposition 4.5, and follows from the exact argument in [42, Proof of Lemma 2.4, second displayed equation on page 1552].

4.3 The Sequential Decomposition

In this section we deduce Theorem 1.14, the sequential decomposition as a consequence of the Compactness Lemma 3.1 and the collection of results from earlier in this section.

In the remainder of the paper we unify the blow-up and global-in-time settings by making the following conventions. Consider a finite energy solution $\boldsymbol{u} \in C(I_*; \mathcal{E})$ on its maximal time of existence I_*. 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. We assume that $\boldsymbol{u}(t)$ exhibits type II behavior in that,

$$\lim_{t \to T_*} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

First we complete the proof of Theorem 1.9.

Proof of Theorem 1.9 We let $\boldsymbol{u}^*(t)$ be defined as in Proposition 4.1 in the case $T_* = 0$ and as in Proposition 4.2 in the case $T_* = \infty$. If $T_* = 0$ the conclusions of Theorem 1.9 are a direct consequence of Proposition 4.1 and Proposition 4.4. If $T_* = +\infty$ we first note that for any $\alpha \in (0, 1)$,

$$\|\boldsymbol{u}^*(t)\|_{\mathcal{E}(0,\alpha t)} \leq \|\boldsymbol{u}^*(t) - \boldsymbol{u}^*_{\mathrm{L}}(t)\|_{\mathcal{E}} + \|\boldsymbol{u}^*_{\mathrm{L}}(t)\|_{\mathcal{E}(0,\alpha t)} \to 0 \ \text{ as } \ t \to T_*$$

where the vanishing of the last term above is due to the asymptotic concentration of free waves near the light cone; see [21, Lemma 4.1] for odd D and [13, Theorem 5] for even D. Now apply Proposition 4.2 and Proposition 4.5. □

Proof of Theorem 1.14 By Corollary 4.6 if $T_* = 0$ or Corollary 4.7 if $T_* = \infty$ we have,

$$\lim_{\tau \to T_*} \frac{1}{\tau} \int_0^\tau \int_0^{\frac{t}{2}} (\partial_t u(t,r))^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}t = 0.$$

We claim there exists a sequence $\tau_n \to T_*$ such that,

$$\lim_{n \to \infty} \sup_{0 < \sigma < \tau_n} \frac{1}{\sigma} \int_{\tau_n}^{\tau_n + \sigma} \int_0^{\frac{t}{2}} (\partial_t u(t,r))^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}t = 0. \tag{4.2}$$

We show that the above is a consequence of the classical maximal function estimate (3.3). Indeed, define

$$\phi(t) = \int_0^{\frac{t}{2}} (\partial_t u(t,r))^2 \, r^{D-1} \, \mathrm{d}r, \quad \Psi(\tau) := \frac{1}{\tau} \int_0^\tau \phi(t) \, \mathrm{d}t.$$

Then (4.2) reduces to the following claim: If $\Psi(\tau) \to 0$ as $\tau \to T_*$, then there exists at least one sequence of times $\tau_n \to T_*$ such that $M\phi(\tau_n) \to 0$ as $n \to \infty$. Now considering intervals $J_n = (0, 1/n]$ if $T_* = 0$ or $J_n = [n/2, n]$ if $T_* = \infty$ apply the maximal function estimate (3.3) with $\alpha_n = 6\Psi(n^{-1})$ if $T_* = 0$ or $\alpha_n = 12\Psi(n)$ if $T_* = \infty$, noting that in both cases $\alpha_n \to 0$ as $n \to \infty$,

$$\Big|\{t \in J_n \ : \ M\phi(t) > \alpha_n\}\Big| \leq \frac{3}{\alpha_n} \int_{J_n} \phi(t) \, \mathrm{d}t \leq \frac{1}{2} |J_n|.$$

This means that $M\phi(t) \le \alpha_n \to 0$ for half of the points in J_n, from which we select the sequence $\tau_n \to T_*$.

Next, let $\rho(t)$ be as in Theorem 1.9. With τ_n as in (4.2) we set

$$\rho_n := \sup_{t \in [\tau_n, \tau_n + \rho(\tau_n)]} \rho(t) \ll \tau_n.$$

It follows from (4.2) that

$$\lim_{n\to\infty} \frac{1}{\rho_n} \int_{\tau_n}^{\tau_n+\rho_n} \int_0^{\frac{\tau_n}{2}} (\partial_t u(t,r))^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}t = 0.$$

Next defining $\boldsymbol{u}_n(s,r) := \boldsymbol{u}(\tau_n + s, r)$ and changing variables above we obtain a sequence of solutions $\boldsymbol{u}_n$ defined on intervals $[0, \rho_n]$ such that,

$$\lim_{n\to\infty} \frac{1}{\rho_n} \int_0^{\rho_n} \int_0^{\rho_n \frac{\tau_n}{2\rho_n}} (\partial_s u_n(s,r))^2 \, r^{D-1} \, \mathrm{d}r \, \mathrm{d}s = 0.$$

We can now apply the Compactness Lemma 3.1 since $R_n := \frac{\tau_n}{2\rho_n} \to \infty$ as $n \to \infty$. We obtain sequences $s_n \in [0, \rho_n]$ and $1 \ll r_n \ll \frac{\tau_n}{2\rho_n}$ for which $\boldsymbol{\delta}_{\rho_n r_n}(\boldsymbol{u}_n(s_n)) \to 0$ as $n \to \infty$. Passing back to the original variables we set $t_n = \tau_n + s_n$ and we have, $\boldsymbol{\delta}_{r_n\rho_n}(\boldsymbol{u}(t_n)) \to 0$ as $n \to \infty$. From (3.1) (and examining the proof of The Compactness Lemma 3.1, see Remark 3.2) we obtain an integer $K_0 \ge 0$, and scales $\lambda_{n,1} \ll \lambda_{n,2} \ll \cdots \ll \lambda_{n,K_0} \lesssim \rho_n \ll t_n$, and a vector of signs $\vec{\iota} \in \{-1, 1\}^{K_0}$ such that

$$\|\boldsymbol{u}(t_n) - \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}_n)\|_{\mathcal{E}(0, r_n\rho_n)} \to 0 \ \text{ as } \ n \to \infty.$$

Note that by construction $\rho(t_n) \ll r_n\rho(t_n) \ll r_n\rho_n \ll t_n$. Thus, from Theorem 1.9 we have,

$$\|\boldsymbol{u}(t_n) - \boldsymbol{u}^*(t_n)\|_{\mathcal{E}(\rho_n)} \to 0 \ \text{ and } \ \|\boldsymbol{u}^*(t_n)\|_{\mathcal{E}(0, r_n\rho_n)} \to 0 \ \text{ as } \ n \to \infty.$$

Combining the two last displayed equations completes the proof. □

5 Decomposition of the Solution and Collision Intervals

In the final two sections we prove Theorem 1 for dimensions $D \ge 6$. We reserve the cases $D = 4, 5$ for the appendix, as these low dimensions require a few technical modifications stemming from the slower decay of $W(r)$ as $r \to \infty$.

5.1 Proximity to a Multi-bubble and Collisions

For the remainder of the paper we fix a solution $\boldsymbol{u} \in C(I_*; \mathcal{E})$ 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 and we assume,

$$\lim_{t \to T_*} \|\boldsymbol{u}(t)\|_{\mathcal{E}} < \infty.$$

Let $\boldsymbol{u}^*(t)$ be the radiation as defined in Proposition 4.1 and Proposition 4.2. We will use crucially the fact that the radiation is given in continuous time. Note that combining the results of Proposition 4.1 and Proposition 4.4 in the blow-up case and Proposition 4.2 together with Proposition 4.5 give a function $\rho : I_* \to (0, \infty)$ such that

$$\lim_{t \to T_*} \left((\rho(t)/t)^{\frac{D-2}{2}} + \|\boldsymbol{u}(t) - \boldsymbol{u}^*(t)\|^2_{\mathcal{E}(\rho(t),\infty)} \right) = 0. \tag{5.1}$$

We also note that

$$\lim_{t \to T_*} \|\boldsymbol{u}^*(t)\|_{\mathcal{E}(0,\alpha t)} = 0, \tag{5.2}$$

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

By Theorem 1.14 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 5.1 (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

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

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 5.2 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 1.14 in this notation: there exists a monotone sequence $t_n \to T_*$ such that

$$\lim_{n \to \infty} \mathbf{d}(t_n) = 0. \tag{5.3}$$

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 [22] (this is done in 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 a consequence of the sequential decomposition (as observed by Duyckaerts, Kenig, and Merle in [22, Theorem 2, Theorem 5, Corollary 6]).

Next, a direct consequence of (5.1) 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 (5.1) the following lemma is immediate from the conventions of Definition 5.1

Lemma 5.3 *There exists a function* $\rho_N : I \to (0, \infty)$ *such that*

$$\lim_{t \to T_*} \mathbf{d}_N(t; \rho_N(t)) = 0. \tag{5.4}$$

Theorem 1 will be a quick consequence of showing that,

$$\lim_{t \to T_*} \mathbf{d}(t) = 0. \tag{5.5}$$

The approach which we adopt in order to prove (5.5) is to study colliding bubbles. A collision is defined as follows.

Definition 5.4 (Collision interval) Let $K \in \{0, 1, \dots, N\}$. 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 5.5 (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 $n \in \{1, 2, \dots\}$.

Lemma 5.6 (Existence of $K \geq 1$) *If* (5.5) *is false, then* K *is well defined and* $K \in \{1, \dots, N\}$.

Remark 5.7 The fact that $K \geq 1$ means that at least one bubble must lose its shape if (5.5) is false.

Proof of Lemma 5.6 Assume (5.5) 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 (ϵ_n), (a_n), (b_n) such that $[a_n, b_n] \in \mathcal{C}_N(\epsilon_n, \eta)$. Indeed, (5.3) implies that there exist $\epsilon_n \to 0$, $a_n \le s_n$ and $b_n \ge s_n$ such that $\mathbf{d}(a_n) \le \epsilon_n$ and $\mathbf{d}(b_n) \le \epsilon_n$. 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 5.3, restricted to the time interval $[a_n, b_n]$. Then (5.4) 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 5.4 are satisfied for $K = N$.

We now prove that $K \ge 1$. Suppose $K = 0$. The definition of a collision interval yields $\mathbf{d}_0(c_n; \rho_n) \le \epsilon_n$ for some sequence $\rho_n \ge 0$, and at the same time $\mathbf{d}(c_n) \ge \eta$ for some $\eta > 0$. Without loss of generality we may assume that c_n is a time at which $\eta \le \mathbf{d}(c_n) \le 2\eta$ for each n, and we may assume further that $\eta > 0$ is small relative to $\|\boldsymbol{W}\|_{\mathcal{E}}$. We show that this is impossible.

First, by Theorem 1.14 we know that

$$E(\boldsymbol{u}) = NE(\boldsymbol{W}) + E(\boldsymbol{u}^*). \tag{5.6}$$

On the other hand, since $\boldsymbol{d}_0(c_n, \rho_n) \le \epsilon_n$ we can find parameters, $\rho_n \ll \lambda_{n,1} \ll \cdots \ll \lambda_{n,N} \ll \rho(c_n) \ll c_n$ and signs $\vec{\iota}_n$ such that

$$\|\boldsymbol{u}(c_n) - \boldsymbol{u}^*(c_n) - \boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n)\|^2_{\mathcal{E}(\rho_n,\infty)} + \sum_{j=0}^{N} \Big(\frac{\lambda_{n,j}}{\lambda_{n,j+1}}\Big)^{\frac{D-2}{2}} \lesssim \epsilon_n^2. \tag{5.7}$$

Since $\rho_n \ll \lambda_{n,1} \ll \rho(c_n)$ we have by (5.1) that

$$\|\boldsymbol{u}^*(c_n) + \boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n)\|_{\mathcal{E}(\rho_n,2\rho_n)} = o_n(1)$$

and hence from the previous line,

$$\|\boldsymbol{u}(c_n)\|_{\mathcal{E}(\rho_n,2\rho_n)} = o_n(1). \tag{5.8}$$

Using the above along with (5.2), Lemma 2.15, and the asymptotic orthogonality of the various parameters we have,

$$E(\boldsymbol{u}(c_n); \rho_n, \infty) = NE(\boldsymbol{W}) + E(\boldsymbol{u}^*) + o_n(1) \ \text{ as } \ n \to \infty.$$

Using the above along with (5.6) we conclude that,

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

Now, let $\boldsymbol{v}_n := \boldsymbol{u}(c_n)\chi_{\rho_n}$. Using the above along with (5.8) we have shown that $E(\boldsymbol{v}_n) = o_n(1)$. We claim that we must have $\|\boldsymbol{v}_n\|_{\mathcal{E}} \simeq \eta$, which would give a contradiction with the critical Sobolev inequality, since $\eta > 0$ can be chosen small. To prove the claim, find parameters $\vec{\sigma}_n$, $\vec{\mu}_n$ such that

$$\eta \simeq \mathbf{d}(c_n) \simeq \Big(\|\boldsymbol{u}(c_n) - \boldsymbol{u}^*(c_n) - \boldsymbol{\mathcal{W}}(\vec{\sigma}_n, \vec{\mu}_n)\|_{\mathcal{E}}^2 + \sum_{j=0}^{N} \Big(\frac{\mu_{n,j}}{\mu_{n,j+1}} \Big)^{\frac{D-2}{2}} \Big)^{\frac{1}{2}}.$$

An application of Lemma 2.20 (taking $\eta > 0$ smaller if needed) together with the above and (5.7) yields that $\vec{\sigma}_n = \vec{\iota}_n$ and moreover that $\rho_n \ll \mu_{n,1} \ll \dots \mu_{n,N} \ll \rho(c_n) \ll c_n$. In fact, we have $|\lambda_{n,j}/\mu_{n,j} - 1| \lesssim \theta(\eta)$ for $\theta(\eta)$ as in Lemma 2.20 and thus by (5.7) we have,

$$\|\boldsymbol{u}(c_n) - \boldsymbol{u}^*(c_n) - \boldsymbol{\mathcal{W}}(\vec{\sigma}_n, \vec{\mu}_n)\|_{\mathcal{E}(\rho_n, \infty)}^2 + \sum_{j=0}^{N} \Big(\frac{\mu_{n,j}}{\mu_{n,j+1}} \Big)^{\frac{D-2}{2}} = o_n(1).$$

Since $\rho_n \ll \mu_{n,1}$ and $\mu_{n,N} \ll \rho(c_n)$ we also have,

$$\|\boldsymbol{u}^*(c_n) + \boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\mu}_n)\|_{\mathcal{E}(0,\rho_n)} = o_n(1).$$

From the previous three displayed equations we can conclude that $\|\boldsymbol{v}_n\|_{\mathcal{E}} \simeq \eta$, proving the claim, and establishing the contradiction. □

In the remaining part of the paper, we argue by contradiction, fixing K to be the number provided by Lemma 5.6. We also let η, ϵ_n, a_n and b_n be some choice of objects satisfying the requirements of Definition 5.5. We fix choices of signs and scales for the $N - K$ "exterior" bubbles provided by Definition 5.1 in the following lemma.

Remark 5.8 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 5.9 *Let $K \geq 1$ be the number given by Lemma 5.6, and let η, ϵ_n, a_n and b_n be some choice of objects satisfying the requirements of Definition 5.5. Then there exist a sequence $\vec{\sigma}_n \in \{-1, 1\}^{N-K}$, a function $\vec{\mu} = (\mu_{K+1}, \dots, \mu_N) \in C^1(\cup_{n\in\mathbb{N}}[a_n, b_n]; (0, \infty)^{N-K})$, a sequence $\nu_n \to 0$, and a sequence $m_n \in \mathbb{Z}$, so that defining the function,*

$$\nu : \cup_{n\in\mathbb{N}}[a_n, b_n] \to (0, \infty), \quad \nu(t) := \nu_n \mu_{K+1}(t), \tag{5.9}$$

we have,

$$\lim_{n\to\infty} \sup_{t\in[a_n,b_n]} \Big(\mathbf{d}_K(t; \nu(t)) + \|\boldsymbol{u}(t)\|_{\mathcal{E}(\nu(t)\leq r\leq 2\nu(t))} \Big) = 0, \tag{5.10}$$

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

$$\boldsymbol{w}(t) = (1 - \chi_{\nu(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) = \sum_{j=K+1}^{N} \sigma_{n,j} \boldsymbol{W}_{\mu_j(t)} + \boldsymbol{h}(t), \tag{5.11}$$

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

$$\lim_{n\to\infty} \sup_{t\in[a_n,b_n]} \Big(\|\boldsymbol{h}(t)\|_{\mathcal{E}}^2 + \Big(\frac{\nu(t)}{\mu_{K+1}(t)}\Big)^{\frac{D-2}{2}} + \sum_{j=K+1}^{N} \Big(\frac{\mu_j(t)}{\mu_{j+1}(t)}\Big)^{\frac{D-2}{2}} \Big) = 0, \tag{5.12}$$

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

$$\lim_{n\to\infty} \sup_{t\in[a_n,b_n]} \left|\nu'(t)\right| = 0. \tag{5.13}$$

Remark 5.10 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 (5.1).

Proof By Definition 5.1 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

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

we have,

$$\mathbf{d}(t; \rho_K(t)) \simeq \|\boldsymbol{h}_{\rho_K}(t)\|_{\mathcal{E}(\rho_K(t),\infty)}^2 + \sum_{j=K}^{N} \Big(\frac{\widetilde{\mu}_j(t)}{\widetilde{\mu}_{j+1}(t)}\Big)^{\frac{D-2}{2}} \lesssim \epsilon_n^2, \tag{5.14}$$

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

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

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

$$\rho_K(t) \le \nu_n\widetilde{\mu}_{K+1}(t), \quad \text{and} \quad \lim_{n\to\infty} \sup_{t\in[a_n,b_n]} \|\boldsymbol{u}(t) - \boldsymbol{u}^*(t)\|_{\mathcal{E}(\frac{1}{4}\nu_n\widetilde{\mu}_{K+1}(t)\le 4\nu_n\widetilde{\mu}_{K+1}(t))} = 0, \tag{5.15}$$

and define $\widetilde{\nu}(t) = \nu_n \widetilde{\mu}_{K+1}(t)$. Thus, defining $\widetilde{\boldsymbol{w}}(t), \widetilde{\boldsymbol{h}}(t) \in \mathcal{E}$ for $t \in \cup_n [a_n, b_n]$, by

$$\widetilde{\boldsymbol{w}}(t) := (1 - \chi_{\widetilde{\nu}(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t)) = \sum_{j=K+1}^{N} \sigma_j(t) \boldsymbol{W}_{\widetilde{\mu}_j(t)} + \widetilde{\boldsymbol{h}}(t)$$

we have using (5.14),

$$\sup_{t \in [a_n, b_n]} \Big(\|\widetilde{\boldsymbol{h}}(t)\|_{\mathcal{E}}^2 + \sum_{j=K}^{N} \Big(\frac{\widetilde{\mu}_j(t)}{\widetilde{\mu}_{j+1}(t)} \Big)^{\frac{D-2}{2}} \Big) \le \theta_n^2 \tag{5.16}$$

for some sequence $\theta_n \to 0$. We invoke Lemma 2.20 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 $\widetilde{\mu}_{K+1}(t), \dots, \widetilde{\mu}_N(t)$ 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 $\vec{\mu}(t) = (\mu_{K+1}(t), \dots \mu_N(t))$, set,

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

Fixing t and suppressing it in the notation, and setting up for an argument as in the proof of Lemma 2.17, define

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

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,

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

and thus $G(0, \vec{\widetilde{\mu}}) = (0, \dots, 0)$. Following the same scheme as the proof of Lemma 2.17 we obtain via Remark 2.19 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

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

and such that

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

Using (5.16) we define

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

By construction we then have,

$$w(t, \vec{\mu}(t)) = (1 - \chi_{\nu(t)})(u(t) - u^*(t)) = \mathcal{W}(\vec{\sigma}_n, \vec{\mu}(t)) + h(t)$$

for $\nu(t) := \nu_n \mu_{K+1}(t)$, and we define $\dot{h}$ by

$$\dot{h}(t) := \dot{w}(t, \vec{\mu}(t)).$$

Then for each $j = K+1, \dots, N$,

$$\sup_{t \in [a_n, b_n]} \Big(\|\boldsymbol{h}(t)\|_{\mathcal{E}}^2 + \sum_{j=K}^{N} \Big(\frac{\mu_j(t)}{\mu_{j+1}(t)} \Big)^{\frac{D-2}{2}} \Big) \lesssim \theta_n^2, \quad 0 = \Big\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid h(t) \Big\rangle. \tag{5.17}$$

Note that (5.10) follows from the above and from (5.1). The point is that we can now use orthogonality conditions above to deduce the differentiability of $\mu(t)$. Indeed, noting the identity,

$$\begin{aligned} \partial_t h(t) &= \partial_t w(t, \vec{\mu}(t)) - \partial_t \mathcal{W}(\vec{\sigma}_n, \vec{\mu}(t)) \\ &= \frac{\mu'_{K+1}(t)}{\mu_{K+1}(t)} (r\partial_r \chi)(\cdot/\nu(t))\big(u(t) - u^*(t)\big) + \dot{h}(t) + \sum_{j=K+1}^{N} \sigma_{n,j} \mu'_j(t) \Lambda W_{\underline{\mu_j(t)}}, \end{aligned}$$

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

$$\begin{aligned} &\sigma_{n,j}\mu'_j(t) \,\langle \mathcal{Z} \mid \Lambda W \rangle + \sum_{i \neq j, K+1 \le i \le N} \sigma_{n,i} \mu'_i(t) \Big\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid \Lambda W_{\underline{\mu_i(t)}} \Big\rangle \\ &\quad + \frac{\mu'_{K+1}(t)}{\mu_{K+1}(t)} \Big\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid (r\partial_r \chi)(\cdot/\nu(t))\big(u(t) - u^*(t)\big) \Big\rangle - \mu'_j(t) \Big\langle [r\Lambda \mathcal{Z}]_{\underline{\mu_j(t)}} \mid r^{-1} h \Big\rangle \\ &\quad = -\Big\langle \mathcal{Z}_{\underline{\mu_j(t)}} \mid \dot{h}(t) \Big\rangle, \end{aligned} \tag{5.18}$$

which, using (5.15) and (5.17), is a diagonally dominant first order differential system for $\vec{\mu}(t)$. Fix any $t_0 \in \cup_n [a_n, b_n]$ so that (5.17) holds at the initial data $\vec{\mu}(t_0)$. The existence and uniqueness theorem gives a unique solution $\vec{\mu}_{\rm ode} \in C^1(J)$ 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 $\vec{\mu}(t) = \vec{\mu}_{\mathrm{ode}}(t)$. Hence $\vec{\mu}(t) \in C^1$. Finally, inverting (5.18) 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 (5.13). This completes the proof. □

We require a few additional facts related to the scale $\nu(t)$. Observe that if $t_n \in [a_n, b_n]$ and $\nu(t_n) \le R_n \ll \mu_{K+1}(t_n)$, then

$$\lim_{n\to\infty} \|\boldsymbol{u}(t_n)\|_{\mathcal{E}(R_n, 2R_n)} = 0. \tag{5.19}$$

Also, if μ_n is a positive sequence such that $\lim_{n\to\infty} \boldsymbol{\delta}_{\mu_n}(t_n) = 0$, then

$$\lim_{n\to\infty} \|\boldsymbol{u}(t_n)\|_{\mathcal{E}(\frac{1}{2}\mu_n, \mu_n)} = 0. \tag{5.20}$$

Importantly, this choice of $\nu(t)$ gives us a way of relating the localized distance $\boldsymbol{\delta}_R$ from Section 3 with the global distance $\mathbf{d}$ on collision intervals.

Lemma 5.11 *There exists a constant $\eta_0 > 0$ having the following property. Let $t_n \in [a_n, b_n]$ and let μ_n be a positive sequence satisfying the conditions:*

(i) $\lim_{n\to\infty} \frac{\mu_n}{\mu_{K+1}(t_n)} = 0$,
(ii) $\mu_n \ge \nu(t_n)$ *or* $\|\boldsymbol{u}(t_n)\|_{\mathcal{E}(\mu_n, \nu(t_n))} \le \eta_0$,
(iii) $\lim_{n\to\infty} \boldsymbol{\delta}_{\mu_n}(t_n) = 0$.

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

Proof Let R_n be a sequence such that $\mu_n \ll R_n \ll \mu_{K+1}(t_n)$. Without loss of generality, we can assume $R_n \ge \nu(t_n)$, since it suffices to replace R_n by $\nu(t_n)$ for all n such that $R_n < \nu(t_n)$. Let $M_n, \vec{\iota}_n, \vec{\lambda}_n$ be parameters such that

$$\|u(t_n) - \mathcal{W}(\vec{\iota}_n, \vec{\lambda}_n)\|_{H(r\le\mu_n)}^2 + \|\dot{u}(t_n)\|_{L^2(r\le\mu_n)}^2 + \sum_{j=1}^{M_n-1} \left(\frac{\lambda_{n,j}}{\lambda_{n,j+1}}\right)^{\frac{D-2}{2}} \to 0, \tag{5.21}$$

which exist by the definition of the localized distance function (3.1). Set

$$\begin{aligned}
\boldsymbol{u}_n^{(i)} &:= \chi_{\frac{1}{2}\mu_n} \boldsymbol{u}(t_n),\\
\boldsymbol{u}_n^{(o)} &:= (1 - \chi_{R_n})\boldsymbol{u}(t_n),\\
\boldsymbol{u}_n^{(m)} &:= \boldsymbol{u}(t_n) - \boldsymbol{u}_n^{(i)} - \boldsymbol{u}_n^{(o)}.
\end{aligned}$$

Invoking (5.20), we have from (5.21) that

$$\lim_{n\to\infty} \|\boldsymbol{u}_n^{(i)} - \boldsymbol{\mathcal{W}}(\vec{\iota}_n, \vec{\lambda}_n)\|_{\mathcal{E}} = 0.$$

Assumption (ii), together with (5.20) and (5.19), yields

$$\|\boldsymbol{u}_n^{(m)}\|_{\mathcal{E}} \leq 2\eta_0, \qquad \text{for all } n \text{ large enough.}$$

We also have, again using (5.20) and (5.19),

$$\limsup_{n\to\infty} \left|E(\boldsymbol{u}(t_n)) - E(\boldsymbol{u}_n^{(i)}) - E(\boldsymbol{u}_n^{(m)}) - E(\boldsymbol{u}_n^{(o)})\right| = 0.$$

Since $\lim_{n\to\infty} E(\boldsymbol{u}_n^{(o)}) = (N-K)E(\boldsymbol{W}) + E(\boldsymbol{u}^*)$ and $0 \leq E(\boldsymbol{u}_n^{(m)}) \leq 2\eta_0$, the convergence above yields $M_n = K$ and $\lim_{n\to\infty} E(\boldsymbol{u}_n^{(m)}) = 0$. Using Sobolev embedding, we get $\lim_{n\to\infty} \|\boldsymbol{u}_n^{(m)}\|_{\mathcal{E}} = 0$, and the result follows. □

5.2 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 5.12 (Basic modulation, $D \geq 6$) *There exist $C_0, \eta_0 > 0$ and a sequence $\zeta_n \to 0$ such that the following is true.*

Let $J \subset [a_n, b_n]$ 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 $\boldsymbol{g}(t) \in \mathcal{E}$ satisfying, for all $t \in J$,

$$\chi(\cdot/\nu(t))\boldsymbol{u}(t) = \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}(t)) + \boldsymbol{g}(t), \tag{5.22}$$

$$0 = \left\langle \mathcal{Z}_{\underline{\lambda_j(t)}} \mid g(t)\right\rangle, \tag{5.23}$$

where $\nu(t)$ is as in (5.9). Define the stable/unstable components $a_j^-(t)$, $a_j^+(t)$ of $\boldsymbol{g}(t)$ by

$$a_j^{\pm}(t) := \left\langle \boldsymbol{\alpha}_{\lambda_j(t)}^{\pm} \mid \boldsymbol{g}(t)\right\rangle,$$

where $\boldsymbol{\alpha}_\lambda^{\pm}$ is as in (2.10).

The estimates,

$$C_0^{-1}\mathbf{d}(t) - \zeta_n \leq \|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{j=1}^{K-1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{4}} \leq C_0\mathbf{d}(t) + \zeta_n, \tag{5.24}$$

and

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{j \notin \mathcal{S}} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{4}} \le C_0 \max_{j \in \mathcal{S}} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{4}} + \max_{1 \le i \le K} |a_i^{\pm}(t)| + \zeta_n, \tag{5.25}$$

hold, where

$$\mathcal{S} := \big\{ j \in \{1, \dots, K-1\} : \iota_j = \iota_{j+1} \big\}.$$

Moreover, for all $j \in \{1, \dots, K\}$ and $t \in J$,

$$\Big|\lambda_j'(t)\Big| \le C_0 \|\dot{g}(t)\|_{L^2} + \zeta_n \tag{5.26}$$

and,

$$\begin{aligned} &\Big|\iota_j \lambda_j'(t) + \frac{1}{\langle \mathcal{Z} \mid W\rangle} \langle \mathcal{Z}_{\underline{\lambda_j(t)}} \mid \dot{g}(t)\rangle\Big| \\ &\le C_0 \|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 + C_0 \bigg(\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-4}{2}} + \Big(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\Big)^{\frac{D-4}{2}}\bigg) \|\dot{g}(t)\|_{L^2} + \zeta_n, \end{aligned} \tag{5.27}$$

where, by convention, $\lambda_0(t) = 0$, $\lambda_{K+1}(t) = \infty$ for all $t \in J$. Finally, we have

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t} a_j^{\pm}(t) \mp \frac{\kappa}{\lambda_j(t)} a_j^{\pm}(t)\Big| \le \frac{C_0}{\lambda_j(t)} \mathbf{d}(t)^2 + \frac{\zeta_n}{\lambda_j}. \tag{5.28}$$

Remark 5.13 We stress that the scaling in $\boldsymbol{\alpha}_{\lambda}^{\pm}$ is $\dot{H}^{-1} \times L^2$-invariant so as to ensure that

$$|a_j^{\pm}(t)| \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}.$$

We also remark that everything in Lemma 5.12, except for the estimates (5.27) and (5.28), holds without change in the lower dimensions $D = 4, 5$, and we refer to the appendix for suitable modifications of $\lambda_j(t)$, $a_j^{\pm}(t)$ in those cases.

Proof of Lemma 5.12 **Step 1.** (The decomposition (5.22) and the estimates (5.24) and (5.25).)

First, observe that by Lemma 5.9,

$$\sup_{t \in [a_n, b_n]} |E(\boldsymbol{u}(t) - \boldsymbol{u}^*(t); \nu(t), \infty) - (N - K) E(\boldsymbol{W})| = o_n(1) \text{ as } n \to \infty.$$

Since $E(\boldsymbol{u}) = E(\boldsymbol{u}^*) + NE(\boldsymbol{W})$ it follows from the above along with (5.10), (5.1), and (5.2) that

$$\sup_{t\in[a_n,b_n]} |E(\boldsymbol{u}(t); 0, 2\nu(t)) - KE(\boldsymbol{W})| = o_n(1) \ \text{ as } \ n \to \infty. \tag{5.29}$$

Using continuity of the flow, the fact that $\mathbf{d}(t) \leq \eta_0$ on J, Lemma 2.20, and by taking $\eta_0 > 0$ small enough, we obtain continuous functions $\vec{\widetilde{\lambda}}(t) = (\widetilde{\lambda}_1(t), \dots, \widetilde{\lambda}_N(t))$ and signs $\vec{\iota}$ independent of $t \in J$, so that

$$\boldsymbol{u}(t) - \boldsymbol{u}^*(t) = \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\widetilde{\lambda}}(t)) + \widetilde{\boldsymbol{g}}(t),$$

and,

$$\mathbf{d}(t)^2 \leq \|\widetilde{\boldsymbol{g}}(t)\|_{\mathcal{E}}^2 + \sum_{j=1}^{N} \Big(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\Big)^{\frac{D-2}{2}} \leq 4\mathbf{d}(t)^2, \tag{5.30}$$

with as usual the convention that $\widetilde{\lambda}_{N+1}(t) = t$. Recalling the properties of $\boldsymbol{w}(t) := (1 - \chi_{\nu(t)})(\boldsymbol{u}(t) - \boldsymbol{u}^*(t))$ from Lemma 5.9, in particular (5.9) and (5.12), and using Lemma 2.20 we see from the above that we must have,

$$\Big(\frac{\nu(t)}{\widetilde{\lambda}_{K+1}(t)}\Big)^{\frac{D-2}{2}} \lesssim \mathbf{d}(t)^2 + o_n(1) \ \text{ as } \ n \to \infty, \tag{5.31}$$

where here and in the remainder of this argument we use the notation $o_n(1)$ to mean a quantity that tends to zero as $n \to \infty$ that can be chosen independently of $t \in [a_n, b_n]$. Using similar logic along with (5.10) we see that we also have,

$$\Big(\frac{\widetilde{\lambda}_K(t)}{\nu(t)}\Big)^{\frac{D-2}{2}} \lesssim \mathbf{d}(t)^2 + o_n(1) \ \text{ as } \ n \to \infty. \tag{5.32}$$

Together, the previous two lines mean, 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 $\nu(t)$.

For the purposes of this argument we denote by

$$\boldsymbol{v}(t) := \boldsymbol{u}(t)\chi_{\nu(t)}, \quad \boldsymbol{w}(t) := (\boldsymbol{u}(t) - \boldsymbol{u}^*(t))(1 - \chi_{\nu(t)}). \tag{5.33}$$

We may express $\boldsymbol{v}(t)$ on $J \subset [a_n, b_n]$ as follows,

$$\begin{aligned}\boldsymbol{v}(t) &= \sum_{j=1}^{K} \iota_j \boldsymbol{W}_{\widetilde{\lambda}_j(t)} - (1 - \chi_{\nu(t)}) \sum_{j=1}^{K} \iota_j \boldsymbol{W}_{\widetilde{\lambda}_j(t)} \\ &\quad + \chi_{\nu(t)} \sum_{j=K+1}^{N} \iota_j \boldsymbol{W}_{\widetilde{\lambda}_j(t)} + \chi_{\nu(t)} \boldsymbol{u}^*(t) + \chi_{\nu(t)} \widetilde{\boldsymbol{g}}(t).\end{aligned}$$

Using (5.1) along with (5.30) and (5.32) we see that,

$$\|\boldsymbol{v}(t)-\sum_{j=1}^{K}\iota_j\boldsymbol{W}_{\widetilde{\lambda}_j(t)}\|_{\mathcal{E}}^2+\sum_{j=1}^{K}\Big(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\Big)^{\frac{D-2}{2}}\lesssim \mathbf{d}(t)^2+o_n(1)\ \text{ as }\ n\to\infty.$$

This means that

$$\mathbf{d}(\boldsymbol{v}(t))\lesssim \mathbf{d}(t)+o_n(1)\ \text{ as }\ n\to\infty,$$

where $\mathbf{d}(\boldsymbol{v})$ is as in the notation of Lemma 2.17. By taking $\eta_0>0$ small enough, and n large enough, we may apply Lemma 2.17, (as well as Lemma 2.20, which ensures the signs $\vec{\iota}$ stay fixed) at each $t\in J$, to obtain unique $\boldsymbol{g}(t)\in\mathcal{E}$, $\vec{\lambda}(t)\in(0,\infty)^K$ so that

$$\boldsymbol{v}(t)=\boldsymbol{\mathcal{W}}(\vec{\iota},\vec{\lambda}(t))+\boldsymbol{g}(t),\quad 0=\big\langle \underline{\mathcal{Z}_{\lambda_j(t)}}\mid \boldsymbol{g}(t)\big\rangle,\quad \forall j=1,\dots,K,\tag{5.34}$$

where in this formula $\vec{\iota},\vec{\lambda}$ are K-vectors, i.e., $\vec{\iota}=(\iota_1,\dots,\iota_K)$, $\vec{\lambda}(t)=(\lambda_1(t),\dots,\lambda_K(t))$. We note the estimate,

$$\begin{aligned}\mathbf{d}(\boldsymbol{v}(t))^2&\le\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2+\sum_{j=1}^{K-1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}}+\Big(\frac{\lambda_K(t)}{\nu(t)}\Big)^{\frac{D-2}{2}}\\&\le 4\mathbf{d}(\boldsymbol{v}(t))^2+\mathbf{d}(t)^2+o_n(1)\\&\lesssim\mathbf{d}(t)^2+o_n(1),\end{aligned}\tag{5.35}$$

as $n\to\infty$. Here the estimate for $\frac{\lambda_K(t)}{\nu(t)}$ is due to (5.32) and the way $\vec{\lambda}(t)$ is chosen in the proof of Lemma 2.17 (here we refer the reader to [41, Proof of Lemma 2.24]). Next, using (5.29) we see that

$$E(\boldsymbol{v})\le KE(\boldsymbol{W})+o_n(1).$$

Therefore, the estimate (2.16) from Lemma 2.17 applied here yields,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}}^2\lesssim\sup_{j\in\mathcal{S}}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}}+\max_{1\le i\le K}|a_j^{\pm}(t)|+o_n(1)$$

where $\mathcal{S}=\{j\in\{1,\dots,K-1\}\,:\,\iota_j=\iota_{j+1}\}$, proving (5.25).

Next, we prove the lower bound in (5.24). Note the identity,

$$\begin{aligned}\boldsymbol{u}(t)-\boldsymbol{u}^*(t)&=\boldsymbol{v}(t)+\boldsymbol{w}(t)-\chi_{\nu(t)}\boldsymbol{u}^*(t)\\&=\sum_{j=1}^{K}\iota_j\boldsymbol{W}_{\lambda_j(t)}+\sum_{j=K+1}^{N}\sigma_{n,j}\boldsymbol{W}_{\mu_j(t)}+\boldsymbol{g}(t)+\boldsymbol{h}(t)-\chi_{\nu(t)}\boldsymbol{u}^*(t),\end{aligned}\tag{5.36}$$

which follows from (5.33) and (5.11). First we prove that $(\iota_{K+1}, \dots, \iota_N) = (\sigma_{K+1}, \dots, \sigma_N)$. From (5.11) and (5.12) we see that

$$\|\boldsymbol{w}(t) - \sum_{j=K+1}^{N} \sigma_{n,j} \boldsymbol{W}_{\mu_j(t)}\|_{\mathcal{E}}^2 + \Big(\frac{\nu(t)}{\mu_{K+1}(t)}\Big)^{\frac{D-2}{2}} + \sum_{j=K+1}^{N} \Big(\frac{\mu_j(t)}{\mu_{j+1}(t)}\Big)^{\frac{D-2}{2}} = o_n(1) \ \text{ as } \ n \to \infty.$$

On the other hand, we see from (5.31) that,

$$\|\boldsymbol{w}(t) - \sum_{j=K+1}^{N} \iota_j \boldsymbol{W}_{\widetilde{\lambda}_j(t)}\|_{\mathcal{E}}^2 + \Big(\frac{\nu(t)}{\widetilde{\lambda}_{K+1}(t)}\Big)^{\frac{D-2}{2}} + \sum_{j=K+1}^{N} \Big(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\Big)^{\frac{D-2}{2}} \lesssim \mathbf{d}(t)^2 + o_n(1).$$

Hence, using Lemma 2.20 we see that for any $\theta_0 > 0$ we may take $\eta_0 > 0$ small enough so that $(\iota_{K+1}, \dots, \iota_N) = (\sigma_{K+1}, \dots, \sigma_N)$, and in addition we have

$$\Big|\frac{\widetilde{\lambda}_j(t)}{\mu_{n,j}(t)} - 1\Big| \leq \theta_0 \quad \forall j = K+1, \dots, N.$$

The above, together with (5.12) implies that

$$\sum_{j=K+1}^{N} \Big(\frac{\widetilde{\lambda}_j(t)}{\widetilde{\lambda}_{j+1}(t)}\Big)^{\frac{D-2}{2}} = o_n(1) \ \text{ as } \ n \to \infty.$$

We may thus rewrite (5.36) as

$$\boldsymbol{u}(t) - \boldsymbol{u}^*(t) = \sum_{j=1}^{K} \iota_j \boldsymbol{W}_{\lambda_j(t)} + \sum_{j=K+1}^{N} \iota_j \boldsymbol{W}_{\mu_j(t)} + \boldsymbol{g}(t) + \boldsymbol{h}(t) - \chi_{\nu(t)} \boldsymbol{u}^*(t).$$

Noting that

$$\sup_{t \in [a_n, b_n]} \|\boldsymbol{u}^*(t) \chi_{\nu(t)}\|_{\mathcal{E}} = o_n(1) \ \text{ as } \ n \to \infty,$$

the previous line together with (5.35) and (5.12) imply that,

$$\mathbf{d}(t)^2 \lesssim \mathbf{d}(\boldsymbol{v}(t))^2 + o_n(1) \lesssim \|\boldsymbol{g}(t)\|_{\mathcal{E}}^2 + \sum_{j=1}^{K-1} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} + o_n(1) \ \text{ as } \ n \to \infty,$$

which proves the lower bound in (5.24).

Step 2. (The dynamical estimates (5.26), (5.27), and (5.28).) Momentarily assuming that $\vec{\lambda} \in C^1(J)$ (we will justify this assumption below) we record the computations,

$$\partial_t v(t) = \dot{g}(t) - \frac{\nu'(t)}{\nu(t)}(r\partial_r\chi)(\cdot/\nu(t))u(t), \quad \partial_t \mathcal{W}(\vec{\iota}, \vec{\lambda}(t)) = -\sum_{j=1}^{K} \iota_j \lambda_j'(t)\Lambda W_{\underline{\lambda_j(t)}},$$

which lead to the expression,

$$\partial_t g(t) = \dot{g}(t) + \sum_{j=1}^{K} \iota_j \lambda_j'(t)\Lambda W_{\underline{\lambda_j(t)}} - u(t)\frac{\nu'(t)}{\nu(t)}(r\partial_r\chi)(\cdot/\nu(t)).$$

We differentiate the orthogonality conditions (5.23) for each $j = 1, \ldots, K$,

$$\begin{aligned} 0 &= -\frac{\lambda_j'}{\lambda_j}\left\langle \underline{\Lambda}\mathcal{Z}_{\underline{\lambda_j}} \mid g\right\rangle + \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \partial_t g\right\rangle \\ &= -\frac{\lambda_j'}{\lambda_j}\left\langle \underline{\Lambda}\mathcal{Z}_{\underline{\lambda_j}} \mid g\right\rangle + \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \dot{g}\right\rangle + \sum_{\ell=1}^{K} \iota_\ell \lambda_\ell' \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_\ell}}\right\rangle \\ &\quad - \frac{\nu'}{\nu}\left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid u(r\partial_r\chi)(\cdot/\nu)\right\rangle, \end{aligned}$$

which we rearrange into the system,

$$\begin{aligned} &\iota_j\lambda_j'\Big(\langle \mathcal{Z} \mid \Lambda W\rangle - \lambda_j^{-1}\big\langle \underline{\Lambda}\mathcal{Z}_{\underline{\lambda_j}} \mid g\big\rangle\Big) + \sum_{i\neq j} \iota_i\lambda_i' \left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\right\rangle \\ &= -\left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \dot{g}\right\rangle + \frac{\nu'}{\nu}\left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid u(r\partial_r\chi)(\cdot/\nu)\right\rangle. \end{aligned} \tag{5.37}$$

This is a diagonally dominant system, hence invertible, and we arrive at the estimate,

$$\left|\lambda_j'\right| \lesssim \|\dot{g}\|_{L^2} + o_n(1) \quad j = 1, \ldots, K, \tag{5.38}$$

after noting the estimates,

$$\begin{aligned} \left|\left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \dot{g}\right\rangle\right| &\lesssim \|\dot{g}\|_{L^2} \\ \left|\frac{\nu'(t)}{\nu(t)}\left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid u(t)(r\partial_r\chi)(\cdot/\nu(t))\right\rangle\right| &\lesssim \left|\nu'\right|\frac{\lambda_j}{\nu}\|r^{-1}u(t)(r\partial_r\chi)(\cdot/\nu(t))\|_{L^2} = o_n(1), \end{aligned}$$

where the last line follows from (5.13). Lastly, we note that the system (5.37) implies that $\vec{\lambda}(t)$ is a C^1 function on J. Indeed, arguing as in the end of the proof of Lemma 5.9, let $t_0 \in J$ be any time and let $\vec{\lambda}(t_0)$ be defined as in (5.34). Using the smallness (5.35) at time t_0, the system (5.37) admits a unique C^1 solution $\vec{\lambda}_{\mathrm{ode}}(t)$ in a neighborhood of

t_0. Due to the way the system (5.37) was derived, the orthogonality conditions in (5.34) hold with $\vec{\lambda}_{\mathrm{ode}}(t)$. Since $\vec{\lambda}(t)$ was obtained uniquely via the implicit function theorem, we must have $\vec{\lambda}(t) = \vec{\lambda}_{\mathrm{ode}}(t)$, which means that $\vec{\lambda}(t)$ is C^1.

The estimates (5.27) are immediate from (5.37) using (5.38) along with the estimates,

$$
|\left\langle \mathcal{Z}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}} \right\rangle| \lesssim \begin{cases} \left(\frac{\lambda_j}{\lambda_i}\right)^{\frac{D}{2}} & \text{if } j < i \\ \left(\frac{\lambda_i}{\lambda_j}\right)^{\frac{D-4}{2}} & \text{if } j > i \end{cases}
$$

$$
\left| \lambda_j^{-1} \left\langle \underline{\Lambda} \mathcal{Z}_{\underline{\lambda_j}} \mid g \right\rangle \right| \lesssim \|g\|_H,
$$

using here that $D - 4 \geq \frac{D-2}{2}$ as long as $D \geq 6$.

Lastly, we consider the estimates (5.28). We first write the equation for $\boldsymbol{g}(t)$.

$$
\partial_t \boldsymbol{g} = \partial_t \boldsymbol{v} - \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}).
$$

Noting that

$$
\begin{aligned}
\partial_t \boldsymbol{v} &= \chi(\cdot/\nu)\partial_t \boldsymbol{u} - \frac{\nu'}{\nu}(r\partial_r\chi)(\cdot/\nu)\boldsymbol{u} \\
&= \chi(\cdot/\nu) J \circ \mathrm{D}E(\boldsymbol{u}) - \frac{\nu'}{\nu}(r\partial_r\chi)(\cdot/\nu)\boldsymbol{u} \\
&= J \circ \mathrm{D}E(\chi(\cdot/\nu)\boldsymbol{u}) + \Big(\chi(\cdot/\nu) J \circ \mathrm{D}E(\boldsymbol{u}) - J \circ \mathrm{D}E(\chi(\cdot/\nu)\boldsymbol{u})\Big) \\
&\quad - \frac{\nu'}{\nu}(r\partial_r\chi)(\cdot/\nu)\boldsymbol{u},
\end{aligned}
$$

we arrive at,

$$
\begin{aligned}
\partial_t \boldsymbol{g} &= J \circ \mathrm{D}E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}) + \boldsymbol{g}) - \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}) \\
&\quad + \Big(\chi(\cdot/\nu) J \circ \mathrm{D}E(\boldsymbol{u}) - J \circ \mathrm{D}E(\chi(\cdot/\nu)\boldsymbol{u})\Big) - \frac{\nu'}{\nu}(r\partial_r\chi)(\cdot/\nu)\boldsymbol{u}.
\end{aligned} \tag{5.39}
$$

We compute,

$$
\frac{\mathrm{d}}{\mathrm{d}t} a_j^- = \langle \partial_t \boldsymbol{\alpha}_{\lambda_j}^- \mid \boldsymbol{g} \rangle + \langle \boldsymbol{\alpha}_{\lambda_j}^- \mid \partial_t \boldsymbol{g} \rangle.
$$

Expanding the first term on the right gives,

$$
\begin{aligned}
\langle \partial_t \boldsymbol{\alpha}_{\lambda_j}^- \mid \boldsymbol{g} \rangle &= \frac{\kappa}{2}\langle \partial_t(\lambda_j^{-1}\mathcal{Y}_{\underline{\lambda_j}}) \mid g \rangle + \frac{1}{2}\langle \partial_t(\mathcal{Y}_{\underline{\lambda_j}}) \mid \dot{g} \rangle \\
&= -\frac{\kappa}{2}\frac{\lambda_j'}{\lambda_j}\langle \lambda_j^{-1}\mathcal{Y}_{\underline{\lambda_j}} + \frac{1}{\lambda_j}(\underline{\Lambda}\mathcal{Y})_{\underline{\lambda_j}} \mid g \rangle - \frac{1}{2}\frac{\lambda_j'}{\lambda_j}\langle (\underline{\Lambda}\mathcal{Y})_{\underline{\lambda_j}}) \mid \dot{g} \rangle,
\end{aligned}
$$

and thus,

$$\left|\left\langle \partial_t \boldsymbol{\alpha}^-_{\lambda_j} \mid \boldsymbol{g}\right\rangle\right| \lesssim \frac{1}{\lambda_j}(\mathbf{d}(t)^2 + o_n(1)).$$

We use (5.39) to expand the second term,

$$\begin{aligned}
\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{g}\right\rangle = &\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J \circ \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g}\right\rangle \\
&+ \left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J \circ \Big(\mathrm{D}\, E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}) + \boldsymbol{g}) - \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g}\Big)\right\rangle \\
&- \left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\right\rangle \\
&+ \left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \Big(\chi(\cdot/\nu) J \circ \mathrm{D}\, E(\boldsymbol{u}) - J \circ \mathrm{D}\, E(\chi(\cdot/\nu)\boldsymbol{u})\Big)\right\rangle \\
&- \frac{\nu'}{\nu}\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid (r\partial_r \chi)(\cdot/\nu)\boldsymbol{u}\right\rangle.
\end{aligned} \tag{5.40}$$

By (2.11) the first term on the right gives the leading order,

$$\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J \circ \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g}\right\rangle = -\frac{\kappa}{\lambda_j} a^-_j .$$

Next, we expand,

$$\begin{aligned}
&\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J \circ \Big(\mathrm{D}\, E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}) + \boldsymbol{g}) - \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}))\boldsymbol{g}\Big)\right\rangle \\
&= -\frac{1}{2}\left\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f(\mathcal{W}(\vec{\iota}, \vec{\lambda}) + g) - f(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - f'(\mathcal{W}(\vec{\iota}, \vec{\lambda}))g\right\rangle \\
&- \frac{1}{2}\left\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - \sum_{i=1}^{K} \iota_i f(W_{\lambda_i})\right\rangle.
\end{aligned}$$

The first line satisfies,

$$\left|\left\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f(\mathcal{W}(\vec{\iota}, \vec{\lambda}) + g) - f(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - f'(\mathcal{W}(\vec{\iota}, \vec{\lambda}))g\right\rangle\right| \lesssim \frac{1}{\lambda_j}(\mathbf{d}(t)^2 + o_n(1)).$$

Noting that $f(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - \sum_{i=1}^{K} \iota_i f(W_{\lambda_i}) = f_{\mathbf{i}}(\vec{\iota}, \vec{\lambda})$, the same argument used to prove Lemma 2.21 gives,

$$\left|\left\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f_{\mathbf{i}}(\vec{\iota}, \vec{\lambda})\right\rangle\right| \lesssim \frac{1}{\lambda_j}\Big(\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + \Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}}\Big) \lesssim \frac{1}{\lambda_j}(\mathbf{d}(t)^2 + o_n(1)).$$

Consider now the third line in (5.40):

$$
\begin{aligned}
-\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\rangle &= \frac{\kappa}{2}\iota_j \frac{\lambda_j'}{\lambda_j}\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_j}}\rangle + \sum_{i\neq j} \iota_i \frac{\kappa}{2}\frac{\lambda_i'}{\lambda_j}\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\rangle \\
&= \sum_{i\neq j} \iota_i \frac{\kappa}{2}\frac{\lambda_i'}{\lambda_j}\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\rangle
\end{aligned}
$$

where in the last equality we used the vanishing of $\langle \mathcal{Y} \mid \Lambda W\rangle$. Noting the estimates

$$
\left|\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\rangle\right| \lesssim \begin{cases} \left(\frac{\lambda_i}{\lambda_j}\right)^{\frac{D-4}{2}} & \text{if } i < j \\ \left(\frac{\lambda_i}{\lambda_j}\right)^{\frac{D}{2}} & \text{if } i > j, \end{cases}
$$

and using here the fact that $D \geq 6$, we obtain,

$$
\left|\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\rangle\right| \lesssim \frac{1}{\lambda_j}(\mathbf{d}(t)^2 + o_n(1)).
$$

Finally, using (5.10) and (5.13) we see that the last two lines of (5.40) satisfy,

$$
\begin{aligned}
\left|\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \Big(\chi(\cdot/\nu) J \circ \mathrm{D} E(\boldsymbol{u}) - J \circ \mathrm{D} E(\chi(\cdot/\nu)\boldsymbol{u})\Big)\rangle\right| &\lesssim \frac{1}{\lambda_j} o_n(1), \\
\left|\frac{\nu'}{\nu}\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid (r\partial_r \chi)(\cdot/\nu)\boldsymbol{u}\rangle\right| &\lesssim \frac{1}{\lambda_j} o_n(1).
\end{aligned}
$$

This completes the proof. □

5.3 Refined Modulation

Next, our goal is to gain precise dynamical control of the modulation parameters in the spirit of [36, 40]. The idea is to construct a virial correction to the modulation parameters (see (5.50)). The idea of adding a correction term based on underlying symmetries (in our case scaling) to modulation parameters originates in Raphaël and Szeftel [70, Proposition 4.3]. We start by finding a suitable truncation of the function $\frac{1}{2}r^2$, similar to [37, Lemma 3.10]. Since here we may have an arbitrary number of bubbles, we need to localize this function both away from $r = 0$ and away from $r = \infty$.

Lemma 5.14 *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* $x \in \mathbb{R}^D$ *such that* $r \in [R^{-1}, R]$,

(P2) there exists $\widetilde{R} > 0$ *(depending on c and R) such that* $q(r) = \mathrm{const}$ *for* $r \geq \widetilde{R}$ *and* $q(r) = \mathrm{const}$ *for* $r \leq \widetilde{R}^{-1}$,

(P3) $|q'(r)| \lesssim r$ *and* $|q''(r)| \lesssim 1$ *for all* $r \in (0, \infty)$, *with constants independent of* c *and* R,
(P4) $q''(r) + \frac{D-1}{r} q'(r) \geq -c$ *for all* $r \in (0, \infty)$,
(P5) $|\Delta^2 q(r)| \leq cr^{-2}$,
(P6) $\left|\left(\frac{q'(r)}{r}\right)'\right| \leq cr^{-1}$ *for all* $r > 0$.

Proof See [41, Proof of Lemma 4.13]. The exact same function can be used here. □

Definition 5.15 (Localized virial operator) For each $\lambda > 0$ we set

$$A(\lambda)g(r) := q'\left(\frac{r}{\lambda}\right) \cdot \partial_r g(r) + \frac{D-2}{2D}\frac{1}{\lambda}\Delta q\left(\frac{r}{\lambda}\right)g(r), \tag{5.41}$$

$$\underline{A}(\lambda)g(r) := q'\left(\frac{r}{\lambda}\right) \cdot \partial_r g(r) + \frac{1}{2}\frac{1}{\lambda}\Delta q\left(\frac{r}{\lambda}\right)g(r) \tag{5.42}$$

where we use the notation $\Delta = \partial_r^2 + \frac{D-1}{r}\partial_r$ in the remainder of the paper. These operators depend on c and R as in Lemma 5.14.

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

$$X := \{\boldsymbol{g} \in \mathcal{E} \mid \partial_r g \in H\}.$$

Lemma 5.16 *(Localized virial estimates) [37, Lemma 3.12] For any* $c_0 > 0$ *there exist* $c_1, R_1 > 0$, *so that for all* c, R *as Lemma 5.14 with* $c < c_1$, $R > R_1$ *the operators* $A(\lambda)$ *and* $A_0(\lambda)$ *defined in* (5.41) *and* (5.42) *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* $\mathscr{L}(\dot{H}^1; L^2)$, *with the bound depending only on the choice of the function* $q(r)$.
- *Let* $\boldsymbol{g}_1 = \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})$ *be an* M*-bubble configuration and let* $\boldsymbol{g}_2 \in X$. *Then, for each* λ_j, $j \in \{1, \ldots, M\}$ *we have*

$$\begin{aligned} &\left\langle A(\lambda_j)g_1 \mid (f(g_1 + g_2) - f(g_1) - f'(g_1)g_2\right\rangle \\ &\quad = -\left\langle A(\lambda_j)g_2 \mid f(g_1 + g_2) - f(g_1)\right\rangle. \end{aligned} \tag{5.43}$$

- *For all* $\boldsymbol{g} \in X$ *we have*

$$\left\langle \underline{A}(\lambda)g \mid -(\partial_r^2 + \frac{D-1}{r}\partial_r)g\right\rangle \geq -\frac{c_0}{\lambda}\|\boldsymbol{g}\|_{\mathcal{E}}^2 + \frac{1}{\lambda}\int_{R^{-1}\lambda}^{R\lambda} (\partial_r g)^2\, r^{D-1}\, \mathrm{d}r \tag{5.44}$$

- *For* $\lambda, \mu > 0$ *with either* $\lambda/\mu \ll 1$ *or* $\mu/\lambda \ll 1$,

$$\|\underline{\Lambda}\Lambda W_{\underline{\lambda}} - \underline{A}(\lambda)\Lambda W_\lambda\|_{L^2} \le c_0, \tag{5.45}$$

$$\|\big(\frac{1}{\lambda}\Lambda - A(\lambda)\big)W_\lambda\|_{L^{\frac{2D}{D-2}}} \le \frac{c_0}{\lambda}, \tag{5.46}$$

$$\begin{aligned}\|A(\lambda)W_\mu\|_{L^{\frac{2D}{D-2}}} + \|\underline{A}(\lambda)W_\mu\|_{L^{\frac{2D}{D-2}}} &\lesssim \frac{1}{\lambda}\min\{(\lambda/\mu)^{\frac{D-2}{2}}, (\mu/\lambda)^{\frac{D-2}{2}}\}\\ \|A(\lambda)\Lambda W_\mu\|_{L^2} + \|\underline{A}(\lambda)\Lambda W_\mu\|_{L^2} &\lesssim \min\{(\lambda/\mu)^{\frac{D-2}{2}}, (\mu/\lambda)^{\frac{D-2}{2}}\}.\end{aligned} \tag{5.47}$$

- *Lastly, the following localized coercivity estimate holds. Fix any smooth function* $\mathcal{Z} \in L^2 \cap X$ *such that* $\langle \mathcal{Z} \mid \Lambda W\rangle > 0$ *and* $\langle \mathcal{Z} \mid \mathcal{Y}\rangle = 0$. *For any* $\boldsymbol{g} \in \mathcal{E}$, $\lambda > 0$ *with* $\langle g \mid \mathcal{Z}_{\underline{\lambda}}\rangle = 0$,

$$\begin{aligned}&\frac{1}{\lambda}\int_{R^{-1}\lambda}^{R\lambda}(\partial_r g)^2 r^{D-1}\,\mathrm{d}r - \frac{1}{\lambda}\int_0^\infty \frac{1}{D}\Delta q\big(\frac{r}{\lambda}\big)f'(W_\lambda)g^2\,r^{D-1}\,\mathrm{d}r\\ &\ge -\frac{c_0}{\lambda}\|\boldsymbol{g}\|_{\mathcal{E}}^2 - \frac{C_0}{\lambda}\big\langle\frac{1}{\lambda}\mathcal{Y}_{\underline{\lambda}} \mid g\big\rangle^2.\end{aligned} \tag{5.48}$$

Proof See [37, Lemma 3.12] for the proof for $D = 6$ and [41, Lemma 4.13] for the case of k-equivariant wave maps. That argument generalizes in a straightforward way to all $D \ge 4$. □

The modulation parameters $\vec{\lambda}(t)$ defined in Lemma 5.12 are imprecise proxies for the dynamics in the cases $3 \le D \le 6$ due to the fact that the orthogonality conditions were imposed relative to $\mathcal{Z} \ne \Lambda W$ (note that we will treat the cases $D = 3, 4, 5$ in the appendix). Indeed, we use 5.23 primarily to ensure coercivity, and thus the estimate (5.25), 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 5.12 set,

$$\xi_j(t) := \begin{cases}\lambda_j(t) \text{ if } D \ge 7\\ \lambda_j(t) - \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle\chi(\cdot/L\lambda_j(t))\Lambda W_{\underline{\lambda_j(t)}} \mid g(t)\big\rangle \text{ if } D = 6\end{cases} \tag{5.49}$$

for each $j = 1, \dots, K-1$, and where $L > 0$ is a large parameter to be determined below. (Note that for $j = K$ we only require the brutal estimate (5.26)). 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 5.12 and for each $j \in \{1, \dots, K\}$ define,

$$\beta_j(t) := -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle\Lambda W_{\underline{\lambda_j(t)}} \mid \dot{g}(t)\big\rangle - \frac{1}{\|\Lambda W\|_{L^2}^2}\big\langle\underline{A}(\lambda_j(t))g(t) \mid \dot{g}(t)\big\rangle. \tag{5.50}$$

The function $\beta_j(t)$ is identical to the function $\beta_j(t)$ in [41] and similar to the function called $b(t)$ in [40].

Lemma 5.17 (Refined modulation) *Let $D \geq 6$ and $c_0 \in (0,1)$. There exist $\eta_0 = \eta(c_0) > 0$, $L_0 = L_0(c_0)$, as well as $c = c(c_0) > 0$, $R = R(c_0) > 1$ as in Lemma 5.14, a constant $C_0 > 0$, 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*

$$\delta_n \leq \mathbf{d}(t) \leq \eta_0$$

for all $t \in J$. Let $\mathcal{S} := \{j \in \{1, \dots, K-1\} \mid \iota_j = \iota_{j+1}\}$. Then, for all $t \in J$,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{i \notin \mathcal{S}} \big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{\frac{D-2}{4}} \leq C_0 \max_{i \in \mathcal{S}} \big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{\frac{D-2}{4}} + \max_{1 \leq i \leq K} |a_i^{\pm}(t)|, \tag{5.51}$$

and,

$$\frac{1}{C_0}\mathbf{d}(t) \leq \max_{i \in \mathcal{S}} \big(\lambda_i(t)/\lambda_{i+1}(t)\big)^{\frac{D-2}{4}} + \max_{1 \leq i \leq K} |a_i^{\pm}(t)| \leq C_0 \mathbf{d}(t). \tag{5.52}$$

Moreover, for all $j \in \{1, \dots, K-1\}$, $t \in J$, and $L \geq L_0$,

$$|\xi_j(t)/\lambda_j(t) - 1| \leq c_0, \tag{5.53}$$
$$|\xi_j'(t) - \beta_j(t)| \leq c_0 \mathbf{d}(t) \tag{5.54}$$

and,

$$\begin{aligned} \beta_j'(t) \geq{}& \Big(\iota_j \iota_{j+1} \omega^2 - c_0\Big) \frac{1}{\lambda_j(t)} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} \\ &+ \Big(-\iota_j \iota_{j-1} \omega^2 - c_0\Big) \frac{1}{\lambda_j(t)} \Big(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\Big)^{\frac{D-2}{2}} \\ &- \frac{c_0}{\lambda_j(t)} \mathbf{d}(t)^2 - \frac{C_0}{\lambda_j(t)} \Big((a_j^+(t))^2 + (a_j^-(t))^2\Big), \end{aligned} \tag{5.55}$$

where, by convention, $\lambda_0(t) = 0$, $\lambda_{K+1}(t) = \infty$ for all $t \in J$, and $\omega^2 > 0$ is defined by

$$\omega^2 = \omega^2(D) := \frac{D-2}{2D}(D(D-2))^{\frac{D}{2}} \|\Lambda W\|_{L^2}^{-2} > 0. \tag{5.56}$$

Finally, for each $j \in \{1, \dots, K\}$,

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t} a_j^{\pm}(t) \mp \frac{\kappa}{\lambda_j(t)} a_j^{\pm}(t)\Big| \leq \frac{C_0}{\lambda_j(t)} \mathbf{d}(t)^2. \tag{5.57}$$

Remark 5.18 Without loss of generality (upon enlarging ϵ_n) we can assume that $\epsilon_n \geq \delta_n$ so that Lemma 5.17 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 (since $\mathbf{d}(t) \geq \epsilon_n$ for all $t \in [a_n, b_n]$; see Remark 5.8).

Before proving Lemma 5.17 we rewrite the equation satisfied by $\boldsymbol{g}(t)$ in (5.39) in components as follows,

$$\begin{aligned}\partial_t g(t) &= \dot{g}(t) + \sum_{j=1}^{K} \iota_j \lambda_j'(t) \Lambda W_{\underline{\lambda_j(t)}} + \phi(u(t), \nu(t)) \\ \partial_t \dot{g}(t) &= -\mathcal{L}_{\mathcal{W}} g + f_{\mathbf{i}}(\iota, \vec{\lambda}) + f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g) + \dot{\phi}(u(t), \nu(t)),\end{aligned} \tag{5.58}$$

where

$$\begin{aligned}\phi(u, \nu) &= -u \frac{\nu'}{\nu} (r \partial_r \chi)(\cdot/\nu) \\ \dot{\phi}(u, \nu) &= -\partial_t u \frac{\nu'}{\nu} (r \partial_r \chi)(\cdot/\nu) - (r^2 \Delta \chi)(\cdot/\nu) r^{-2} u \\ &\quad - 2\frac{1}{r} (r \partial_r \chi)(\cdot/\nu) \partial_r u + \chi(\cdot/\nu) f(u) - f(\chi(\cdot/\nu) u),\end{aligned} \tag{5.59}$$

which we note are supported in $r \in (\nu, \infty)$, and

$$\begin{aligned}f_{\mathbf{i}}(\vec{\iota}, \vec{\lambda}) &:= f\Big(\sum_{i=1}^{K} \iota_i W_{\lambda_i}\Big) - \sum_{i=1}^{K} \iota_i f(W_{\lambda_i}) \\ f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g) &= f\Big(\sum_{i=1}^{K} \iota_i W_{\lambda_i} + g\Big) - f\Big(\sum_{i=1}^{K} \iota_i W_{\lambda_i}\Big) - f'\Big(\sum_{i=1}^{K} \iota_i W_{\lambda_i}\Big) g.\end{aligned}$$

The subscript $\mathbf{i}$ above stands for "interaction" and $\mathbf{q}$ stands for "quadratic." For $\|\boldsymbol{g}\|_{\mathcal{E}} \leq 1$, the term $f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda})$ satisfies,

$$\| f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}) \|_{L^{\frac{2D}{D+2}}} \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2 + \|\boldsymbol{g}\|_{\mathcal{E}}^{\frac{D+2}{D-2}}. \tag{5.60}$$

The proof of (5.60) follows from the pointwise estimates,

$$|f(x_1 + x_2) - f(x_1) - f'(x_1) x_2| \leq |z_2|^2 \quad \text{if} \quad D = 6$$

and if $D \geq 7$,

$$|f(x_1 + x_2) - f(x_1) - f'(x_1) x_2| \leq \begin{cases} |x_2|^{\frac{D+2}{D-2}} & \forall x_1, x_2 \in \mathbb{R} \\ |x_1|^{-\frac{D-6}{D-2}} |x_2|^2 & \text{if} \ \ x_1 \neq 0. \end{cases} \tag{5.61}$$

See [36, Lemma 2.1] for the proof of the previous two estimates.

In one instance it will be convenient to write the equation for $\dot g$ as follows,

$$\partial_t \dot g = \Delta g + f_{\mathbf{i}}(\vec\iota, \vec\lambda) + \widetilde{f}_{\mathbf{q}}(\vec\iota, \vec\lambda, g) + \dot\phi(u, \nu) \tag{5.62}$$

with,

$$\widetilde{f}_{\mathbf{q}}(\vec\iota, \vec\lambda, \vec g) := f\Big(\sum_{i=1}^{K} \iota_i W_{\lambda_i} + g\Big) - f\Big(\sum_{i=1}^{K} \iota_i W_{\lambda_i}\Big). \tag{5.63}$$

Proof of Lemma 5.17 First, we prove the estimates (5.51) and (5.52). Let ζ_n be the sequence given by Lemma 5.12 and let δ_n be any sequence such that $\zeta_n/\delta_n \to 0$ as $n \to \infty$. Using Lemma 5.12, estimate (5.51) follows from (5.25) and the estimate (5.52) follows from (5.24).

Note also that with this choice of δ_n and (5.51), the estimate (5.26) leads to,

$$\left|\lambda_j'(t)\right| \lesssim \mathbf{d}(t). \tag{5.64}$$

Next, we treat (5.53), which is only relevant in the case $D = 6$. From (5.49) we see that,

$$\begin{aligned}|\xi_j/\lambda_j - 1| &\lesssim |\lambda_j^{-1}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid g\rangle| \\ &\lesssim \|g\|_{L^3}\lambda_j^{-1}\|\chi(L\cdot/\lambda_j)\Lambda W_{\underline{\lambda_j}}\|_{L^{\frac32}} \lesssim (\log L)^{\frac23}\|\boldsymbol{g}\|_{\mathcal{E}},\end{aligned}$$

which is small by taking η_0 sufficiently small (after L is fixed below).

Next we compute $\xi_j'(t)$. For $D = 6$, from (5.49) we have

$$\begin{aligned}\xi_j' &= \lambda_j' - \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \partial_t g\rangle \\ &\quad + \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\frac{\lambda_j'}{\lambda_j}\langle (r\partial_r\chi)(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid g\rangle + \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\frac{\lambda_j'}{\lambda_j}\langle \chi(\cdot/L\lambda_j)\underline{\Lambda}\Lambda W_{\underline{\lambda_j}} \mid g\rangle.\end{aligned} \tag{5.65}$$

We examine each of the terms on the right above. The last two terms are negligible. Indeed, using $\|g\|_{L^3(\mathbb{R}^6)} \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}$,

$$\left|\frac{\lambda_j'}{\lambda_j}\langle (r\partial_r\chi)(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid g\rangle\right| \lesssim \left|\lambda_j'\right| \|g\|_{L^3}\Big(\int_{2^{-1}L}^{2L} |\Lambda W(r)|^{\frac32} r^5\,\mathrm{d}r\Big)^{\frac23} \lesssim \mathbf{d}(t)^2,$$

and,

$$\left|\frac{\lambda_j'}{\lambda_j}\langle \chi(\cdot/L\lambda_j)\underline{\Lambda}\Lambda W_{\underline{\lambda_j}} \mid g\rangle\right| \lesssim \left|\lambda_j'\right| \|g\|_{L^3}\int_0^{2L} |\underline{\Lambda}\Lambda W(r)|^{\frac{3}{2}} r^5 \,\mathrm{d}r$$
$$\lesssim (1+\log(L))\mathbf{d}(t)^2.$$

Using (5.58) in the second term in (5.65) gives

$$\begin{aligned}
-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \partial_t g\rangle &= -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle \\
&\quad -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \sum_{i=1}^K \iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\rangle \\
&\quad -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \phi(u,\nu)\rangle.
\end{aligned}$$

The first term on the right satisfies,

$$\begin{aligned}
-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle &= -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle \\
&\quad +\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle (1-\chi(\cdot/L\lambda_j))\Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle \\
&= -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle + o_L(1)\|\boldsymbol{g}\|_{\mathcal{E}}
\end{aligned}$$

where the $o_L(1)$ term can be made as small as we like by taking $L > 0$ large. Using (5.64), the second term yields,

$$\begin{aligned}
&-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \sum_{i=1}^K \iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\rangle = -\lambda_j' \\
&\quad -\sum_{i\neq j}\frac{\iota_j\iota_i\lambda_i'}{\|\Lambda W\|_{L^2}^2}\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\rangle \\
&\quad +\frac{\lambda_j'}{\|\Lambda W\|_{L^2}^2}\langle (1-\chi(\cdot/L\lambda_j))\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_j}}\rangle \\
&\quad = -\lambda_j' + O((\lambda_{j-1}/\lambda_j) + (\lambda_j/\lambda_{j+1}) + o_L(1))\mathbf{d}(t).
\end{aligned}$$

Finally, the third term vanishes due to the fact that for each $j < K$, $L\lambda_j \ll \lambda_K \ll \nu$, and hence

$$\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \phi(u,\nu)\rangle = 0.$$

Plugging all of this back into (5.65) we obtain,

$$\left|\xi_j'(t) + \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle\right| \le c_0 \mathbf{d}(t) \tag{5.66}$$

for $D = 6$, after fixing $L > 0$ sufficiently large. The same estimate for $D \ge 7$, i.e., when $\xi_j'(t) = \lambda_j'(t)$, is immediate from (5.27) since in this case we take $\mathcal{Z} = \Lambda W$. Thus (5.66) holds for all $D \ge 6$. The estimate (5.54) is then immediate from (5.66), the definition of β_j, and the estimate,

$$\left|\frac{1}{\|\Lambda W\|_{L^2}^2}\left\langle \underline{A}(\lambda_j) g \mid \dot g\right\rangle\right| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2,$$

which follows from the first bullet point in Lemma 5.16.

We prove (5.55). We compute,

$$\begin{aligned}\beta_j' = &\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\frac{\lambda_j'}{\lambda_j}\langle \underline{\Lambda} \Lambda W_{\underline{\lambda_j}} \mid \dot g\rangle - \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\langle \Lambda W_{\underline{\lambda_j}} \mid \partial_t \dot g\rangle \\ &- \frac{1}{\|\Lambda W\|_{L^2}^2}\frac{\lambda_j'}{\lambda_j}\left\langle \lambda_j \partial_{\lambda_j} \underline{A}(\lambda_j) g \mid \dot g\right\rangle - \frac{1}{\|\Lambda W\|_{L^2}^2}\left\langle \underline{A}(\lambda_j)\partial_t g \mid \dot g\right\rangle \\ &- \frac{1}{\|\Lambda W\|_{L^2}^2}\left\langle \underline{A}(\lambda_j) g \mid \partial_t \dot g\right\rangle.\end{aligned} \tag{5.67}$$

Using (5.58) we arrive at the expression,

$$\begin{aligned}-\langle \Lambda W_{\underline{\lambda_j}} \mid \partial_t \dot g\rangle = &\left\langle \Lambda W_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{W}} - \mathcal{L}_{\lambda_j}) g\right\rangle - \left\langle \Lambda W_{\underline{\lambda_j}} \mid f_{\mathbf{i}}(\iota, \vec\lambda)\right\rangle \\ &- \left\langle \Lambda W_{\underline{\lambda_j}} \mid f_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\right\rangle - \left\langle \Lambda W_{\underline{\lambda_j}} \mid \dot\phi(u, \nu)\right\rangle,\end{aligned}$$

where in the first term on the right we used that $\mathcal{L}_{\lambda_j} \Lambda W_{\underline{\lambda_j}} = 0$. Using (5.58) we obtain,

$$\begin{aligned}-\left\langle \underline{A}(\lambda_j)\partial_t g \mid \dot g\right\rangle &= -\left\langle \underline{A}(\lambda_j)\dot g \mid \dot g\right\rangle - \sum_{i=1}^K \iota_i \lambda_i' \left\langle \underline{A}(\lambda_j) \Lambda W_{\underline{\lambda_i}} \mid \dot g\right\rangle - \left\langle \underline{A}(\lambda_j)\phi(u, \nu) \mid \dot g\right\rangle \\ &= -\iota_j \lambda_j' \left\langle \underline{A}(\lambda_j) \Lambda W_{\underline{\lambda_j}} \mid \dot g\right\rangle - \sum_{i \ne j} \iota_i \lambda_i' \left\langle \underline{A}(\lambda_j) \Lambda W_{\underline{\lambda_i}} \mid \dot g\right\rangle - \left\langle \underline{A}(\lambda_j)\phi(u, \nu) \mid \dot g\right\rangle\end{aligned}$$

where we used that $\left\langle \underline{A}(\lambda_j)\dot g \mid \dot g\right\rangle = 0$. Finally, using (5.62) we have,

$$\begin{aligned}-\left\langle \underline{A}(\lambda_j) g \mid \partial_t \dot g\right\rangle = &-\left\langle \underline{A}(\lambda_j) g \mid \Delta g\right\rangle - \left\langle \underline{A}(\lambda_j) g \mid f_{\mathbf{i}}(\iota, \vec\lambda)\right\rangle \\ &- \left\langle \underline{A}(\lambda_j) g \mid \widetilde{f_{\mathbf{q}}}(\vec\iota, \vec\lambda, g)\right\rangle - \left\langle \underline{A}(\lambda_j) g \mid \dot\phi(u, \nu)\right\rangle.\end{aligned}$$

Plugging these back into (5.67) and rearranging we have,

$$
\begin{aligned}
\|\Lambda W\|_{L^2}^2 \beta_j' = & -\frac{\iota_j}{\lambda_j}\Big\langle \Lambda W_{\lambda_j} \mid f_{\mathbf{i}}(\iota, \vec\lambda)\Big\rangle - \big\langle \underline{A}(\lambda_j) g \mid \Delta g\big\rangle \\
& + \Big\langle (A(\lambda_j) - \underline{A}(\lambda_j)) g \mid \widetilde{f}_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\Big\rangle \\
& + \Big\langle \Lambda W_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{W}} - \mathcal{L}_{\lambda_j}) g\Big\rangle + \iota_j \frac{\lambda_j'}{\lambda_j}\Big\langle \Big(\frac{1}{\lambda_j}\underline{\Lambda} - \underline{A}(\lambda_j)\Big)\Lambda W_{\lambda_j} \mid \dot g\Big\rangle \\
& - \Big\langle A(\lambda_j)\sum_{i=1}^K \iota_i W_{\lambda_i} \mid f_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\Big\rangle - \Big\langle A(\lambda_j) g \mid \widetilde{f}_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\Big\rangle \\
& + \iota_j \Big\langle \Big(A(\lambda_j) - \frac{1}{\lambda_j}\Lambda\Big) W_{\lambda_j} \mid f_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\Big\rangle - \frac{\lambda_j'}{\lambda_j}\big\langle \lambda_j \partial_{\lambda_j}\underline{A}(\lambda_j) g \mid \dot g\big\rangle \\
& + \sum_{i\neq j}\iota_i \Big\langle A(\lambda_j) W_{\lambda_i} \mid f_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\Big\rangle \\
& - \sum_{i\neq j}\iota_i\lambda_i' \Big\langle \underline{A}(\lambda_j)\Lambda W_{\underline{\lambda_i}} \mid \dot g\Big\rangle - \Big\langle \underline{A}(\lambda_j) g \mid f_{\mathbf{i}}(\iota, \vec\lambda)\Big\rangle \\
& - \iota_j\Big\langle \Lambda W_{\underline{\lambda_j}} \mid \dot\phi(u,\nu)\Big\rangle - \big\langle \underline{A}(\lambda_j)\phi(u,\nu) \mid \dot g\big\rangle - \big\langle \underline{A}(\lambda_j) g \mid \dot\phi(u,\nu)\big\rangle .
\end{aligned}
\tag{5.68}
$$

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 2.21

$$
\begin{aligned}
& -\frac{\iota_j}{\lambda_j\|\Lambda W\|_{L^2}^2}\Big\langle \Lambda W_{\lambda_j} \mid f_{\mathbf{i}}(\iota, \vec\lambda)\Big\rangle \\
& = -(\omega^2 + O(\eta_0^2))\frac{\iota_j\iota_{j+1}}{\lambda_j}\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + (\omega^2 + O(\eta_0^2))\frac{\iota_j\iota_{j-1}}{\lambda_j}\Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}} .
\end{aligned}
$$

The second and third terms together will have a sign, up to an acceptable error. First, using (5.44) we have,

$$
\big\langle \underline{A}(\lambda_j) g \mid -\Delta g\big\rangle \geq -\frac{c_0}{\lambda_j}\|g\|_H^2 + \frac{1}{\lambda_j}\int_{R^{-1}\lambda_j}^{R\lambda_j} (\partial_r g)^2 r^{D-1}\mathrm{d}r .
$$

To treat the third term, we start by using the definition (5.63) to observe the identity,

$$
\widetilde{f}_{\mathbf{q}}(\vec\iota, \vec\lambda, g) = f_{\mathbf{q}}(\vec\iota, \vec\lambda, g) + \Big(f'(\mathcal{W}(\vec\iota, \vec\lambda)) - f'(W_{\lambda_j})\Big) g + f'(W_{\lambda_j}) g .
$$

The first two terms above contribute acceptable errors. Indeed, using (5.60),

$$
\Big|\big\langle (A(\lambda_j) - A_{\frac12}(\lambda_j)) g \mid f_{\mathbf{q}}(\vec\iota, \vec\lambda, g)\big\rangle\Big| \lesssim \frac{1}{\lambda_j}\big(\|\boldsymbol{g}\|_{\mathcal{E}}^3 + \|\boldsymbol{g}\|_{\mathcal{E}}^{\frac{2D}{D-2}}\big),
$$

and since $\|f'(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - f'(W_{\lambda_j})\|_{L^{\frac{D}{2}}} \lesssim \eta_0$, we have,

$$\Big|\big\langle \lambda_j^{\frac{1}{2}}(A(\lambda_j) - A_{\frac{1}{2}}(\lambda_j))g \mid (f'(\mathcal{W}(\vec{\iota}, \vec{\lambda})) - f'(W_{\lambda_j}))g\big\rangle\Big| \le \frac{c_0}{\lambda_j}\|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

Putting this together with the fact that $(A(\lambda_j) - \underline{A}(\lambda_j))g = -\frac{1}{D\lambda_j}\Delta q(\cdot/\lambda_j)g$ we have,

$$\begin{aligned}&\big\langle (A(\lambda_j) - A_{\frac{1}{2}}(\lambda_j))g \mid \widetilde{f}_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)\big\rangle \\ &\quad \ge -\frac{1}{D\lambda_j}\int_0^\infty \Delta q(r/\lambda_j) f'(W_{\lambda_j})g^2\, r^{D-1}\,\mathrm{d}r - \frac{c_0}{\lambda_j}\|\boldsymbol{g}\|_{\mathcal{E}}^2.\end{aligned}$$

We show that the remaining terms contribute acceptable errors. For the fourth term a direct calculation gives,

$$\|\Lambda W_{\lambda_j}(f'(\mathcal{W}(\vec{\iota}, \vec{\lambda}) - f'(W_{\lambda_j}))\|_{L^{\frac{2D}{D+2}}} \lesssim \Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + \Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}}$$

and hence,

$$\Big|\Big\langle \Lambda W_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{W}} - \mathcal{L}_{\lambda_j})g\Big\rangle\Big| \lesssim \frac{1}{\lambda_j}\|g\|_H\Big(\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + \Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}}\Big).$$

By (5.45) along with (5.26) we have,

$$\Big|\iota_j\frac{\lambda_j'}{\lambda_j}\Big\langle\Big(\frac{1}{\lambda_j}\underline{\Lambda} - \underline{A}(\lambda_j)\Big)\Lambda W_{\lambda_j} \mid \dot{g}\Big\rangle\Big| \lesssim \frac{c_0}{\lambda_j}\mathbf{d}(t)\|\boldsymbol{g}\|_{\mathcal{E}}.$$

For the sixth term on the right-hand side of (5.68) we note that

$$A(\lambda_j)\sum_{i=1}^K \iota_i W_{\lambda_i} = A(\lambda_j)\mathcal{W}(\vec{\iota}, \vec{\lambda}),$$

and hence we may apply (5.43) with $g_1 = \mathcal{W}(\vec{\iota}, \vec{\lambda})$ and $g_2 = g$ to conclude that

$$\big\langle A(\lambda_j)\sum_{i=1}^K \iota_i W_{\lambda_i} \mid f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)\big\rangle + \Big\langle A(\lambda_j)g \mid \widetilde{f}_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)\Big\rangle = 0,$$

which takes care of the sixth and seventh terms. Next we consider the eighth term. We claim the estimate,

$$\Big|\Big\langle (A(\lambda_j) - \frac{1}{\lambda_j}\Lambda)W_{\lambda_j} \mid f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)\Big\rangle\Big| \le \frac{c_0}{\lambda_j}\mathbf{d}(t)^2. \tag{5.69}$$

When $D = 6$ this follows directly from (5.46) and (5.60). For dimensions $D \geq 7$ the brutal estimate (5.60) is not sufficient and we require a more careful analysis, based on the point-wise estimate (5.61). First, recalling the definition of $A(\lambda_j)$ we note that

$$
\begin{aligned}
\Big| \big\langle (A(\lambda_j) - \frac{1}{\lambda_j}\Lambda) W_{\lambda_j} \mid f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g) \big\rangle \Big| \lesssim \frac{1}{\lambda_j} \int_0^{\frac{1}{R}\lambda_j} |\Lambda W_{\lambda_j}| |f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)| \, r^{D-1} \mathrm{d}r \\
+ \frac{1}{\lambda_j} \int_{R\lambda_j}^{\infty} |\Lambda W_{\lambda_j}| |f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)| \, r^{D-1} \mathrm{d}r.
\end{aligned}
\tag{5.70}
$$

For the first integral on the right we introduce an auxiliary large parameter $L > 0$ and divide the integral into two regions $r \in (0, L\sqrt{\lambda_{j-1}\lambda_j})$ and $r \in [L\sqrt{\lambda_{j-1}\lambda_j}, R^{-1}\lambda_j]$. In the first region we use the first estimate in (5.61) for $x_1 = \mathcal{W}(\vec{\iota}, \vec{\lambda}; r)$ and $x_2 = g(\cdot, r)$ to obtain

$$
\begin{aligned}
&\frac{1}{\lambda_j} \int_0^{L\sqrt{\lambda_{j-1}\lambda_j}} |\Lambda W_{\lambda_j}| |f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)| \, r^{D-1} \mathrm{d}r \\
&\quad \lesssim \frac{1}{\lambda_j} \int_0^{L\sqrt{\lambda_{j-1}\lambda_j}} |\Lambda W_{\lambda_j}| \, |g|^{\frac{D+2}{D-2}} \, r^{D-1} \, \mathrm{d}r \\
&\quad \lesssim \frac{1}{\lambda_j} \Big(\int_0^{L\sqrt{\lambda_{j-1}\lambda_j}} |\Lambda W_{\lambda_j}|^{\frac{2D}{D-2}} \, r^{D-1} \, \mathrm{d}r \Big)^{\frac{D-2}{2}} \|\boldsymbol{g}\|_{\mathcal{E}}^{\frac{D+2}{D-2}} \\
&\quad \lesssim \frac{1}{\lambda_j} L^{\frac{D-2}{2}} \Big(\frac{\lambda_{j-1}}{\lambda_j} \Big)^{\frac{D-2}{4}} \|\boldsymbol{g}\|_{\mathcal{E}}^{\frac{D+2}{D-2}} \leq \frac{c_0}{\lambda_j} \mathbf{d}(t)^2
\end{aligned}
$$

by ensuring η_0 is sufficiently small relative to L. Next we observe that $L > 0$ can be taken sufficiently large so that the point-wise estimate,

$$
\big| \mathcal{W}(\vec{\iota}, \vec{\lambda}; r) \big| \gtrsim \lambda_j^{-\frac{D-2}{2}}
$$

holds uniformly in $r \in [L\sqrt{\lambda_{j-1}\lambda_j}, R^{-1}\lambda_j]$. Using the second inequality in (5.61) we then have,

$$
\begin{aligned}
&\frac{1}{\lambda_j} \int_{L\sqrt{\lambda_{j-1}\lambda_j}}^{\frac{1}{R}\lambda_j} |\Lambda W_{\lambda_j}| |f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g)| \, r^{D-1} \mathrm{d}r \\
&\quad \lesssim \frac{1}{\lambda_j} \int_{L\sqrt{\lambda_{j-1}\lambda_j}}^{\frac{1}{R}\lambda_j} |\Lambda W_{\lambda_j}| |\mathcal{W}(\vec{\iota}, \vec{\lambda}; r)|^{-\frac{D-6}{D-2}} \, |g|^2 \, r^{D-1} \, \mathrm{d}r \\
&\quad \lesssim \frac{1}{\lambda_j} \frac{1}{R^2} \int_0^{\frac{1}{R}\lambda_j} \frac{g^2}{r^2} \, r^{D-1} \, \mathrm{d}r \leq \frac{c_0}{\lambda_j} \mathbf{d}(t)^2,
\end{aligned}
$$

where the last line follows from taking R sufficiently large. The analysis of the second integral in (5.70) is analogous, this time dividing the region of integration

$r \in [R\lambda_j, \infty]$ into the regions $r \in [R\lambda_j, L^{-1}\sqrt{\lambda_j \lambda_{j+1}}]$ and $r \in (L^{-1}\sqrt{\lambda_j \lambda_{j+1}}, \infty)$, and using the point-wise estimate

$$\left|\mathcal{W}(\vec{\iota}, \vec{\lambda}; r)\right| \gtrsim \lambda_j^{\frac{D-2}{2}} r^{-(D-2)}$$

in the region $r \in [R\lambda_j, L^{-1}\sqrt{\lambda_j \lambda_{j+1}}]$, which holds as long as L is taken sufficiently large. This proves (5.69).

Using the first bullet point in Lemma 5.16 and (5.26) we estimate the ninth term as follows,

$$\left|\frac{\lambda_j'}{\lambda_j} \left\langle \lambda_j \partial_{\lambda_j} \underline{A}(\lambda_j) g \mid \dot{g} \right\rangle\right| \lesssim \frac{1}{\lambda_j} \mathbf{d}(t) \|\boldsymbol{g}\|_{\mathcal{E}}^2.$$

Next, using (5.47) and (5.60) we have,

$$\left|\sum_{i \neq j} \iota_i \left\langle A(\lambda_j) W_{\lambda_i} \mid f_{\mathbf{q}}(\vec{\iota}, \vec{\lambda}, g) \right\rangle\right| \leq \frac{c_0}{\lambda_j} \mathbf{d}(t)^2.$$

An application of (5.47) and (5.26) gives

$$\sum_{i \neq j} \left|\lambda_i' \left\langle \underline{A}(\lambda_j) \Lambda W_{\underline{\lambda_i}} \mid \dot{g} \right\rangle\right| \leq \frac{c_0}{\lambda_j} \|g\|_H^2.$$

Next, consider the twelfth term. Using the first bullet point in Lemma 5.16, and in particular the spatial localization of $\underline{A}(\lambda_j)$ we obtain

$$\left|\left\langle \underline{A}(\lambda_j) g \mid f_{\mathbf{i}}(\iota, \vec{\lambda}) \right\rangle\right| \lesssim \|g\|_H \|f_{\mathbf{i}}(\iota, \vec{\lambda})\|_{L^2(\widetilde{R}^{-1}\lambda_j \leq r \leq \widetilde{R}\lambda_j)}.$$

Using the estimate,

$$\|f_{\mathbf{i}}(\iota, \vec{\lambda})\|_{L^2(\widetilde{R}^{-1}\lambda_j \leq r \leq \widetilde{R}\lambda_j)} \lesssim \frac{1}{\lambda_j}\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + \frac{1}{\lambda_j}\Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}},$$

we obtain

$$\left|\left\langle \underline{A}(\lambda_j) g \mid f_{\mathbf{i}}(\iota, \vec{\lambda}) \right\rangle\right| \lesssim \frac{1}{\lambda_j} \|g\|_H \left(\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + \frac{1}{\lambda_j}\Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}}\right).$$

Finally, we treat the last line of (5.68). First, using Lemma 5.9 and the definition of $\dot{\phi}$ in (5.59) we have

$$\left|\left\langle \Lambda W_{\underline{\lambda_j}} \mid \dot{\phi}(u, \nu) \right\rangle\right| \lesssim \frac{1}{\lambda_j}\Big(\frac{\lambda_j}{\nu}\Big)^{\frac{D-2}{2}} \|\boldsymbol{u}(t)\|_{\mathcal{E}(\nu(t), 2\nu(t))} \lesssim \frac{\theta_n}{\lambda_j}\Big(\frac{\lambda_j}{\nu}\Big)^{\frac{D-2}{2}}$$

for some sequence $\theta_n \to 0$ as $n \to \infty$. The last two terms in (5.68) vanish due to the support properties of $\underline{A}(\lambda_j)$, $\phi(u, \nu)$, $\dot{\phi}(u, \nu)$ and the fact that $\lambda_j \le \lambda_K \ll \nu$.

Combining these estimates in (5.68) we obtain the inequality,

$$\begin{aligned}\beta' \geq &\Big(-\iota_j\iota_{j+1}\omega^2 - c_0\Big)\frac{1}{\lambda_j}\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{D-2}{2}} + \Big(\iota_j\iota_{j-1}\omega^2 - c_0\Big)\frac{1}{\lambda_j}\Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{D-2}{2}} \\ &+ \frac{1}{\lambda_j}\int_{R^{-1}\lambda_j}^{R\lambda_j}(\partial_r g)^2\, r\mathrm{d}r - \frac{1}{D\lambda_j}\int_0^\infty \Delta q(r/\lambda_j) f'(W_{\lambda_j})g^2\, r^{D-1}\,\mathrm{d}r - c_0\frac{\mathbf{d}(t)^2}{\lambda_j}.\end{aligned}$$

Finally, we use the localized coercivity estimate (5.48) on the second line above along with the estimates $\big\langle \frac{1}{\lambda_j}\mathcal{Y}_{\underline{\lambda_j}} \mid g\big\rangle^2 \lesssim (a_j^+)^2 + (a_j^-)^2$ to see that

$$\begin{aligned}&\frac{1}{\lambda_j}\int_{R^{-1}\lambda_j}^{R\lambda_j}(\partial_r g)^2\, r\mathrm{d}r - \frac{1}{D\lambda_j}\int_0^\infty \Delta q(r/\lambda_j) f'(W_{\lambda_j})g^2\, r^{D-1}\,\mathrm{d}r \\ &\geq -c_0\frac{\|g\|_H^2}{\lambda_j} - \frac{C_0}{\lambda_j}\Big((a_j^+)^2 + (a_j^-)^2\Big).\end{aligned}$$

This completes the proof of (5.55).

Finally, (5.57) follows from (5.28) and our choice of δ_n. □

Finally, we prove that, by again enlarging ϵ_n, we can control the error in the virial identity, see Lemma 2.1, by $\mathbf{d}$.

Lemma 5.19 *There exist $C_0, \eta_0 > 0$ depending only on D and N and a decreasing sequence $\epsilon_n \to 0$ such that*

$$|\Omega_{1,\rho(t)}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,\rho(t)}(\boldsymbol{u}(t))| \leq C_0\mathbf{d}(t)$$

for all $t \in [a_n, b_n]$ such that $\epsilon_n \le \mathbf{d}(t) \le \eta_0$, $\rho(t) \le \nu(t)$ and $|\rho'(t)| \le 1$.

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 5.12 yields

$$\|\boldsymbol{u}(t) - \boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}(t)) - \boldsymbol{g}(t)\|_{\mathcal{E}(0,2\nu(t))} \to 0, \qquad \text{as } n \to \infty.$$

Using Remark 5.18, (5.51) and (5.52) 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

$$|\Omega_{1,\rho(t)}(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}(t))) + \frac{D-2}{2}\Omega_{2,\rho(t)}(\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda}(t)))| \leq C_0\mathbf{d}(t),$$

which in turn will follow from

$$\Big| \int_0^\infty \Big(\frac{1}{2}(\partial_r \mathcal{W}(\vec\iota, \vec\lambda))^2 + \frac{D-2}{2D}|\mathcal{W}(\vec\iota, \vec\lambda)|^{\frac{2D}{D-2}} + \frac{D-2}{2}\partial_r \mathcal{W}(\vec\iota, \vec\lambda)\frac{\mathcal{W}(\vec\iota, \vec\lambda)}{r}\Big) \\ (r\partial_r \chi)(r/\rho) r^{D-1}\,\mathrm{d}r\Big| \\ \le C_0 \mathbf{d}(t).$$

Noting the identity,

$$\frac{1}{2}(\partial_r W(r))^2 + \frac{D-2}{2D}|W(r)|^{\frac{2D}{D-2}} + \frac{D-2}{2}\partial_r W(r)\frac{W(r)}{r} = 0$$

it suffices to estimates the cross terms in the integral above, and the desired bound follows from an explicit computation. □

6 Conclusion of the Proof

6.1 The Scale of the *K*-th Bubble

As mentioned in the Introduction, the K-th bubble is of particular importance. We introduce a function μ which is well-defined on every $[a_n, b_n]$, and of size comparable with λ_K on time intervals where the solution approaches a multi-bubble configuration.

Definition 6.1 (The scale of the K-th bubble) Fix $\kappa_1 > 0$ small enough. For all $t \in I$, we set

$$\mu(t) := \sup\big\{r \le \nu(t) : \|\boldsymbol{u}(t)\|_{\mathcal{E}(r,\nu(t))} = \kappa_1\big\}. \tag{6.1}$$

Note that, if κ_1 is sufficiently small, then $K > 0$ implies $\|\boldsymbol{u}(t)\|_{\mathcal{E}(0,\nu(t))} \ge 2\kappa_1$, hence $\mu(t)$ is a well-defined finite positive number for all $t \in I$. Since in the definition of $\mu(t)$ we can restrict to rational r, μ is a measurable function. Even if μ is not necessarily a continuous function, still $\mu(t)$ is well-defined for each individual value of t. We stress that $\mu(t) \le \nu(t)$ for all n large enough and $t \in [a_n, b_n]$, thus $\mu(t) \ll \mu_{K+1}(t)$ as $n \to \infty$.

We also introduce a specific "regularization" of μ. For a given collision interval $[a_n, b_n]$, we set

$$\mu_* : [a_n, b_n] :\to (0, \infty), \qquad \mu_*(t) := \inf_{s\in[a_n,b_n]} \big(4\mu(s) + |s-t|\big).$$

We choose not to include n in the notation. We stress that μ_* depends on n, which will be known from the context.

Lemma 6.2 *The function μ_* has the following properties:*

(i) its Lipschitz constant is ≤ 1,

(ii) *there exist* $\delta, \kappa_2 > 0$ *and* $n_0 \in \mathbb{N}$ *depending on* κ_1 *such that* $t \in [a_n, b_n]$ *with* $n \geq n_0$ *and* $\mathbf{d}(t) \leq \delta$ *imply* $\kappa_2\lambda_K(t) \leq \mu^*(t) \leq \kappa_2^{-1}\lambda_K(t)$*, where* $\lambda_K(t)$ *is the modulation parameter defined in Lemma 5.12,*

(iii) *if* $t_n \in [a_n, b_n]$*,* $1 \ll r_n \ll \mu_{K+1}(t_n)/\mu_*(t_n)$ *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 Statement (i) is clear.

Recall that $4\mu(t) \leq 4\nu(t)$, hence the definition of $\mu_*(t)$ yields

$$\mu_*(t) = \inf\big\{4\mu(s) + |s-t| : s \in [a_n, b_n] \cap [t - 4\nu(t), t + 4\nu(t)]\big\}.$$

Let $s \in [a_n, b_n] \cap [t - 4\nu(t), t + 4\nu(t)]$. By the definition of μ and (5.19), we have, for n large enough,

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(\mu(s), 16\nu(s))} \leq 2\kappa_1,$$

hence Lemma 2.6 yields

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(4\mu(s)+|t-s|, 4\nu(s)-|t-s|)} \leq C\kappa_1.$$

From (5.13) and $|s-t| \leq 4\nu(t)$, we deduce that $\nu(t) \leq 4\nu(s) - |t-s|$, thus

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(4\mu(s)+|t-s|, \nu(t))} \leq C\kappa_1.$$

Taking the supremum with respect to $s \in [a_n, b_n] \cap [t - 4\nu(t), t + 4\nu(t)]$, we obtain

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(\mu_*(t), \nu(t))} \leq C\kappa_1. \tag{6.2}$$

Statement (iii) now follows from Lemma 5.11 with $\mu_n = r_n\mu_*(t_n)$, provided that we choose $\kappa_1 \leq \eta_0/C$.

Let $\kappa_2 > 0$ be such that

$$\|W\|_{H(r\geq(4\kappa_2)^{-1})} \leq \frac{1}{2}\kappa_1, \qquad \|W\|_{H(r\geq\kappa_2)} \geq 2C\kappa_1,$$

where C is the constant in (6.2). It is clear that $\frac{1}{4}\mu_*(t) \leq \mu(t)$, hence (6.1) yields

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(\frac{1}{4}\mu_*(t), \nu(t))} \geq \kappa_1.$$

Thus, in order to prove that $\mu_*(t) \leq \kappa_2^{-1}\lambda_K(t)$, it suffices to check that

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}((4\kappa_2)^{-1}\lambda_K(t), \nu(t))} < \kappa_1.$$

We use (5.22). By (5.24), $\|\boldsymbol{g}\|_{\mathcal{E}} \ll 1$ when $\delta \ll 1$ and $n_0 \gg 1$. Thus, it suffices to see that

$$\|\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\|_{\mathcal{E}((4\kappa_2)^{-1}\lambda_K(t), \nu(t))} < \kappa_1,$$

whenever $\sum_{j=1}^{K} \lambda_j(t)/\lambda_{j+1}(t) \ll 1$, which is obtained directly from the definitions of $\boldsymbol{\mathcal{W}}$ and κ_2.

Similarly, using (6.2), we will have $\mu_*(t) \geq \kappa_2 \lambda_K(t)$ if we can prove that

$$\|\boldsymbol{u}(t)\|_{\mathcal{E}(\kappa_2\lambda_K(t),\nu(t))} > C\kappa_1.$$

But the last bound follows from

$$\|\boldsymbol{\mathcal{W}}(\vec{\iota}, \vec{\lambda})\|_{\mathcal{E}(\kappa_2\lambda_K(t),\nu(t))} > C\kappa_1$$

whenever $\sum_{j=1}^{K} \lambda_j(t)/\lambda_{j+1}(t) \ll 1$. □

Our next goal is to prove that the minimality of K (see Definition 5.5) implies a lower bound on the length of the collision intervals.

Lemma 6.3 *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 5.5.

The first and second requirement in Definition 5.4 are clearly satisfied. It remains to construct a function $\rho_{K-1} : [c_n, d_n] \to [0, \infty)$ such that

$$\lim_{n\to\infty} \sup_{t\in[c_n,d_n]} \mathbf{d}_{K-1}(t; \rho_{K-1}(t)) = 0. \tag{6.3}$$

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 $\kappa_2\mu_*(c_n) \leq \lambda_K(c_n) \leq \kappa_2^{-1}\mu_*(c_n)$ 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 5.9. 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 $\widetilde{\mu}(t) := (\lambda_K(c_n), \vec{\mu}(t)) \in (0, \infty)^{N-(K-1)}$. Let R_n be a sequence such that $\nu_n(c_n) \ll R_n \ll \mu_{K+1}(c_n)$. Applying Lemma 2.9 with these sequences r_n, R_n and $\boldsymbol{u}_n(t) := \boldsymbol{u}(c_n + t)$, we obtain

$$\lim_{n\to\infty} \sup_{t\in[c_n,d_n]} \|\boldsymbol{u}(t) - \boldsymbol{\mathcal{W}}(\widetilde{\sigma}_n, \widetilde{\mu}(t))\|_{\mathcal{E}(\rho_{K-1}(t),\infty)} = 0,$$

implying (6.3) □

Remark 6.4 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.

6.2 Demolition of the Multi-bubble

This paragraph is devoted to the analysis of the ODE system satisfied by the modulation parameters. We apply here the "weighted sum" trick from [26, Section 6].

Lemma 6.5 *Let $D \geq 6$. If η_0 is small enough, then there exists $C_0 \geq 0$ depending only on D and N such that the following is true. If $[t_1, t_2] \subset I_*$ is a finite time interval such that $\mathbf{d}(t) \leq \eta_0$ for all $t \in [t_1, t_2]$, then*

$$\sup_{t\in[t_1,t_2]} \lambda_K(t) \leq \frac{4}{3} \inf_{t\in[t_1,t_2]} \lambda_K(t), \tag{6.4}$$

$$\int_{t_1}^{t_2} \mathbf{d}(t)\mathrm{d}t \leq C_0\big(\mathbf{d}(t_1)^{\frac{4}{D-2}}\lambda_K(t_1) + \mathbf{d}(t_2)^{\frac{4}{D-2}}\lambda_K(t_2)\big). \tag{6.5}$$

Remark 6.6 Since $\mathbf{d}(t) \leq \eta_0$ is small, Lemma 6.2 (iii) yields $\lambda_K(t) \simeq \mu_*(t)$, so in the formulation of the lemma we could just as well write μ_* instead of λ_K.

Proof of Lemma 6.5 **Step 1.** First, we argue that (6.4) follows from (6.5). Without loss of generality, assume $\lambda_K(t_1) \geq \lambda_K(t_2)$. Since $|\lambda_K'(t)| \lesssim \mathbf{d}(t)$, see (5.26), and $\mathbf{d}(t)^{\frac{4}{D-2}} \lesssim \eta_0^{\frac{4}{D-2}}$ is small, (6.5) implies

$$\int_{t_1}^{t_2} |\lambda_K'(t)|\mathrm{d}t \leq \frac{1}{7}\lambda_K(t_1),$$

thus $\inf_{t\in[t_1,t_2]} \lambda_K(t) \geq \frac{6}{7}\lambda_K(t_1)$ and $\sup_{t\in[t_1,t_2]} \lambda_K(t) \leq \frac{8}{7}\lambda_K(t_1)$, so (6.4) follows. It remains to prove (6.5).

Step 2. Let $C_1 > 0$ be a large number chosen below and consider the auxiliary function

$$\phi(t) := \sum_{j\in\mathcal{S}} 2^{-j}\xi_j(t)\beta_j(t) - C_1\sum_{j=1}^{K}\lambda_j(t)a_j^-(t)^2 + C_1\sum_{j=1}^{K}\lambda_j(t)a_j^+(t)^2,$$

inspired by the function $A(t)$ from [26, Section 6]. We claim that for all $t \in [t_1, t_2]$

$$\phi'(t) \geq c_2\mathbf{d}(t)^2, \tag{6.6}$$

with $c_2 > 0$ depending only on D and N. The remaining part of Step 1 is devoted to proving this bound.

Using (5.54), (5.57) and recalling that $|\lambda_K'(t)| \lesssim \mathbf{d}(t)$, see (5.26), we obtain

$$\phi'(t) \geq \sum_{j\in\mathcal{S}} 2^{-j}\beta_j^2(t) + \sum_{j\in\mathcal{S}} 2^{-j}\lambda_j(t)\beta_j'(t) + C_1\nu\sum_{j=1}^{K}\big(a_j^-(t)^2 + a_j^+(t)^2\big) - c_0\mathbf{d}(t)^2. \tag{6.7}$$

We focus on the second term of the right hand side. Applying (5.55), we have

$$\begin{aligned}\sum_{j\in\mathcal{S}} \lambda_j(t)\beta_j'(t) \geq \omega^2 \sum_{j\in\mathcal{S}} 2^{-j}\bigg(\iota_j\iota_{j+1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} - \iota_j\iota_{j-1}\Big(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\Big)^{\frac{D-2}{2}}\bigg)\\ - C_2\sum_{k=1}^{K}\big(a_j^-(t)^2 + a_j^+(t)^2\big) - c_0\mathbf{d}(t)^2.\end{aligned}$$

Noting that $\iota_j\iota_{j+1} = 1$ if $j \in \mathcal{S}$, we rewrite the first sum on the right hand side as

$$\sum_{j\in\mathcal{S}} 2^{-j}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} - \sum_{j\in\mathcal{S}} 2^{-j}\iota_j\iota_{j-1}\Big(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\Big)^{\frac{D-2}{2}}.$$

Splitting the first sum into two equal terms, and shifting the index in the second sum, we obtain

$$\begin{aligned}&\sum_{j\in\mathcal{S}} 2^{-j-1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}}\\ &\quad + \bigg[\sum_{j\in\mathcal{S}} 2^{-j-1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} - \sum_{j+1\in\mathcal{S}} 2^{-j-1}\iota_j\iota_{j+1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}}\bigg].\end{aligned}$$

We will check that the number inside the square parenthesis is nonnegative. To see this, we rewrite it as

$$\sum_j 2^{-j-1}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}}\big[\mathbb{1}_{\mathcal{S}}(j) - \iota_j\iota_{j+1}\mathbb{1}_{\mathcal{S}}(j+1)\big].$$

If $j \notin \mathcal{S}$, then $\iota_j\iota_{j+1} = -1$, hence all the terms in the above sum are nonnegative. We have thus proved that

$$\sum_{j\in\mathcal{S}} \lambda_j(t)\beta_j'(t) \geq 2^{-N-1}\omega^2 \sum_{j\in\mathcal{S}}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} - C_2\sum_{k=1}^{K}\big(a_j^-(t)^2 + a_j^+(t)^2\big) - c_0\mathbf{d}(t)^2.$$

Taking $C_1 > C_2/\nu$ and using (6.7) together with (5.52), we get (6.6).

Step 3. Since ϕ is increasing, it has at most one zero, which we denote $t_3 \in [t_1, t_2]$. If $\phi(t) > 0$ for all $t \in [t_1, t_2]$, we set $t_3 := t_1$, and if $\phi(t) < 0$ for all $t \in [t_1, t_2]$, then we set $t_3 := t_2$. We will show that

$$\int_{t_3}^{t_2} \mathbf{d}(t)\mathrm{d}t \leq C_0 \mathbf{d}(t_2)^{\frac{4}{D-2}} \lambda_K(t_2). \tag{6.8}$$

By the symmetry of the problem, one can similarly bound the integral over $[t_1, t_3]$, and summing the two we get (6.5). Without loss of generality, we can assume $t_3 < t_2$, since otherwise (6.8) is trivial.

Observe that for all $t \in (t_3, t_2]$ we have

$$\begin{aligned} \frac{\phi(t)}{\lambda_K(t)} &\lesssim \sum_{j\in\mathcal{S}} \frac{\lambda_j(t)}{\lambda_K(t)}|\beta_j(t)| + \sum_{j=1}^{K} \frac{\lambda_j(t)}{\lambda_K(t)}\big(a_j^-(t)^2 + a_j^+(t)^2\big) \\ &\lesssim \sum_{j\in\mathcal{S}} \frac{\lambda_j(t)}{\lambda_{j+1}(t)}|\beta_j(t)| + \mathbf{d}(t)^2 \lesssim \mathbf{d}(t)^{\frac{D+2}{D-2}}. \end{aligned} \tag{6.9}$$

Combining this bound with (6.6), for all $t \in (t_3, t_2]$ we get

$$\phi'(t) \geq c_2\big(\phi(t)/\lambda_K(t)\big)^{\frac{D-2}{D+2}} \mathbf{d}(t),$$

thus

$$\lambda_K(t)^{\frac{D-2}{D+2}}\big(\phi(t)^{\frac{4}{D+2}}\big)' \gtrsim \mathbf{d}(t). \tag{6.10}$$

Using (6.9) and $|\lambda_K'(t)| \lesssim \mathbf{d}(t)$, we get

$$\begin{aligned} &\big(\lambda_K(t)^{\frac{D-2}{D+2}}\phi(t)^{\frac{4}{D+2}}\big)' - \lambda_K(t)^{\frac{D-2}{D+2}}\big(\phi(t)^{\frac{4}{D+2}}\big)' \\ &\quad= \frac{D-2}{D+2}\lambda_K'(t)\big(\phi(t)/\lambda_K(t)\big)^{\frac{4}{D+2}} \gtrsim -\mathbf{d}(t)^{\frac{D+2}{D-2}}. \end{aligned}$$

Since $\mathbf{d}(t)^{\frac{4}{D-2}} \lesssim \eta_0^{\frac{4}{D-2}}$ is small, (6.10) yields

$$\big(\lambda_K(t)^{\frac{D-2}{D+2}}\phi(t)^{\frac{4}{D+2}}\big)' \gtrsim \mathbf{d}(t)$$

which, integrated, gives

$$\int_{t_3}^{t_2} \mathbf{d}(t)\mathrm{d}t \lesssim \lambda_K(t_2)^{\frac{D-2}{D+2}}\phi(t_2)^{\frac{4}{D+2}} - \lambda_K(t_3)^{\frac{D-2}{D+2}}\phi(t_3)^{\frac{4}{D+2}} \leq \lambda_K(t_2)^{\frac{D-2}{D+2}}\phi(t_2)^{\frac{4}{D+2}}.$$

Invoking (6.9), we obtain (6.8). $\square$

Starting from now, $\eta_0 > 0$ is fixed so that Lemma 6.5 holds and Lemma 6.3 can be applied with $\eta = \eta_0$. We also fix $\epsilon > 0$ to be the value given by Lemma 6.3 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 6.7 *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 an $N_n \in \mathbb{N}^*$ and a partition of the interval $[a_n, b_n]$*

$$a_n = e^L_{n,0} \leq e^R_{n,0} \leq c^R_{n,0} \leq d^R_{n,0} \leq f^R_{n,0} \leq f^L_{n,1} \leq d^L_{n,1} \leq c^L_{n,1} \leq e^L_{n,1} \leq \ldots \leq e^R_{n,N_n} = b_n,$$

having the following properties.

(1) For all $m \in \{0, 1, \ldots, N_n\}$ and $t \in [e^L_{n,m}, e^R_{n,m}]$, $\mathbf{d}(t) \leq \eta_0$, and

$$\int_{e^L_{n,m}}^{e^R_{n,m}} \mathbf{d}(t)\mathrm{d}t \leq C_2 \theta_n^{4/(D-2)} \min(\mu_*(e^L_{n,m}), \mu_*(e^R_{n,m})), \tag{6.11}$$

where $C_2 \geq 0$ depends only on k and N.

(2) For all $m \in \{0, 1, \ldots, N_n - 1\}$ and $t \in [e^R_{n,m}, c^R_{n,m}] \cup [f^R_{n,m}, f^L_{n,m+1}] \cup [c^L_{n,m+1}, e^L_{n,m+1}]$, $\mathbf{d}(t) \geq \theta_n$.

(3) For all $m \in \{0, 1, \ldots, N_n - 1\}$ and $t \in [c^R_{n,m}, f^R_{n,m}] \cup [f^L_{n,m+1}, c^L_{n,m+1}]$, $\mathbf{d}(t) \geq \epsilon$.

(4) For all $m \in \{0, 1, \ldots, N_n - 1\}$, $\mathbf{d}(d^R_{n,m}) \geq \eta_0$ and $\mathbf{d}(d^L_{n,m+1}) \geq \eta_0$.

(5) For all $m \in \{0, 1, \ldots, N_n - 1\}$, $\mathbf{d}(c^R_{n,m}) = \mathbf{d}(c^L_{n,m+1}) = \epsilon$.

(6) For all $m \in \{0, 1, \ldots, N_n - 1\}$, either $\mathbf{d}(t) \geq \epsilon$ for all $t \in [c^R_{n,m}, c^L_{n,m+1}]$, or $\mathbf{d}(f^R_{n,m}) = \mathbf{d}(f^L_{n,m+1}) = \epsilon$.

(7) For all $m \in \{0, 1, \ldots, N_n - 1\}$,

$$\sup_{t \in [e^L_{n,m}, c^R_{n,m}]} \mu_*(t) \leq 2\kappa_2^{-2} \inf_{t \in [e^L_{n,m}, c^R_{n,m}]} \mu_*(t),$$

$$\sup_{t \in [c^L_{n,m+1}, e^R_{n,m+1}]} \mu_*(t) \leq 2\kappa_2^{-2} \inf_{t \in [c^L_{n,m+1}, e^R_{n,m+1}]} \mu_*(t).$$

Remark 6.8 The purpose of the Lemma 6.7 above is to partition any collision interval $[a_n, b_n]$ into subintervals, depending on the values of $\mathbf{d}(t)$. On the intervals $[e^L_{n,j}, e^R_{n,j}]$, the bound (6.11) 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 $[c^R_{n,j}, f^R_{n,j}]$ and $[f^L_{n,j}, c^L_{n,j}]$, since they allow to apply Lemma 6.3, leading to the crucial bound (6.21).

Lemma 6.3 could not be applied directly on the intervals $[e^R_{n,j}, e^L_{n,j+1}]$, 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 (6.11) in the proof of Theorem 1, it will be necessary to have $\theta_n \to 0$, see (6.22). For this reason, the contribution of the intervals $[e^R_{n,j}, c^R_{n,j}]$, $[f^R_{n,j}, f^L_{n,j+1}]$ and $[c^L_{n,j+1}, e^L_{n,j+1}]$ will be estimated differently, see (6.18)–(6.20).

The intervals $[c^R_{n,j}, f^R_{n,j}]$ and $[f^L_{n,j}, c^L_{n,j}]$ correspond to what were called "compactness intervals" $[c_n, d_n]$ in the summary of the proof in Section 1.3, whereas $[e^L_{n,j}, e^R_{n,j}]$

correspond to the "modulation intervals". For simplicity, we have not mentioned the remaining intervals in Section 1.3. 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 6.7 will also play a role in the construction of an appropriate time-dependent cut-off radius for the virial identity, in Lemma 6.9 below.

Proof For all $t_0 \in [a_n, b_n]$ such that $\mathbf{d}(t_0) < \eta_0$, 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 $\mathbf{d} < \eta_0$. 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

$$A := \{t \in [a_n, b_n] : \mathbf{d}(t) \leq \theta_n\}.$$

Since A is a compact set, there exists a finite sequence

$$a_n \leq s_{n,0} < s_{n,1} < \ldots < s_{n,N_n} \leq b_n$$

such that

$$s_{n,m} \in A, \qquad A \subset \bigcup_{m=0}^{N_n} J(s_{n,m}). \tag{6.12}$$

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).

Let $m \in \{0, 1, \ldots, N_n - 1\}$. Since $J(s_{n,m}) \cap J(s_{n,m+1}) = \emptyset$, there exists $t \in (s_{n,m}, s_{n,m+1})$ such that $\mathbf{d}(t) \geq \eta_0$. Let $d^R_{n,m}$ be the smallest such t, and $d^L_{n,m+1}$ the largest one. Let $c^R_{n,m}$ be the smallest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (c^R_{n,m}, d^R_{n,m})$. Similarly, let $c^L_{n,m+1}$ be the biggest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (d^L_{n,m+1}, c^L_{n,m+1})$. Next, let $e^R_{n,m}$ be the smallest number such that $\mathbf{d}(t) \geq 2\theta_n$ for all $t \in (e^R_{n,m}, c^R_{n,m})$. If we take $\theta_n < \frac{\epsilon}{2}$, then we have $e^R_{n,m} < c^R_{n,m}$. Since $\mathbf{d}(s_{n,m}) \leq \theta_n$, we have $e^R_{n,m} > s_{n,m}$. Similarly, let $e^L_{n,m+1}$ be the biggest number such that $\mathbf{d}(t) \geq 2\theta_n$ for all $t \in (c^L_{n,m+1}, e^L_{n,m+1})$ (again, it follows that $e^L_{n,m+1} < s_{n,m+1}$). Finally, if $\mathbf{d}(t) \geq \epsilon$ for all $t \in (d^R_{n,m}, d^L_{n,m+1})$, we set $f^R_{n,m}$ and $f^L_{n,m+1}$ arbitrarily, for example $f^R_{n,m} := d^R_{n,m}$ and $f^L_{n,m+1} := d^L_{n,m+1}$. If, on the contrary, there exists $t \in (d^R_{n,m}, d^L_{n,m+1})$ such that $\mathbf{d}(t) < \epsilon$, we let $f^R_{n,m}$ be the biggest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (d^R_{n,m}, f^R_{n,m})$, and $f^L_{n,m+1}$ be the smallest number such that $\mathbf{d}(t) \geq \epsilon$ for all $t \in (f^L_{n,m+1}, d^L_{n,m+1})$.

We check all the desired properties. For all $n \in \{0, 1, \dots, N_n\}$, we have $e^L_{n,m} \leq s_{n,m} \leq e^R_{n,m}$. Since $\mathbf{d}(e^L_{n,m}) \leq 2\theta_n$ and $\mathbf{d}(e^R_{n,m}) \leq 2\theta_n$, the property (1) follows from Lemma 6.5. 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 $m \in \{0, 1, \dots, N_n\}$ such that $t_0 \in [e^L_{n,m}, e^R_{n,m}]$. But (6.12) implies that $t_0 \in J(s_{n,m})$ for some m and, by construction, $\mathbf{d}(t) > \theta_n$ for all $t \in J(s_{n,m}) \setminus [e^L_{n,m}, e^R_{n,m}]$, so we obtain $t \in [e^L_{n,m}, e^R_{n,m}]$. Finally, note that $\mathbf{d}(t) \leq \eta_0$ for all $t \in [e^L_{n,m}, c^R_{n,m}] \cup [c^L_{n,m+1}, e^R_{n,m+1}]$, hence, using again Lemma 6.5, but on the time intervals $[e^L_{n,m}, c^R_{n,m}]$ and $[c^L_{n,m+1}, e^R_{n,m+1}]$, we deduce the property (7) from (6.4) and Lemma 6.2 (ii).

□

6.3 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 6.9 *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) $\lim_{n\to\infty} \inf_{t\in[a_n,b_n]} \big(\rho(t)/\mu_*(t)\big) = \infty$ *and* $\lim_{n\to\infty} \sup_{t\in[a_n,b_n]} \big(\rho(t)/\mu_{K+1}(t)\big) = 0$,
(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 claim that $\lim_{n\to\infty} \delta_n = 0$. Indeed, if $\rho^{(a)}(a_n) + t - a_n \geq \nu(t)$, then it suffices to recall (5.19). In the opposite case, (5.13) yields $t - a_n \leq \nu(t) \leq \nu(a_n) + o(t - a_n)$, hence $t - a_n \leq 2\nu(a_n)$. Since we have

$$\lim_{n\to\infty} \|\boldsymbol{u}(a_n)\|_{\mathcal{E}(\frac{1}{4}\rho^{(a)}(a_n);4\nu(a_n))} = 0,$$

it suffices to apply Lemma 2.6.

Let $\theta_0 > 0$ be given by Lemma 6.7, 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

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho^{(a)}(t) := 1 \text{ if } t \in [e_{n,m}^L, e_{n,m}^R], \qquad \frac{\mathrm{d}}{\mathrm{d}t}\rho^{(a)}(t) := \delta_n^{-\frac{1}{2}} \text{ otherwise.}$$

We check that $\lim_{n\to\infty}\inf_{t\in[a_n,b_n]}\big(\rho^{(a)}(t)/\mu_*(t)\big) = \infty$. 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 $\mu_*(t) \le \mu_*(e_{n,m}^R)+(t-e_{n,m}^R) \lesssim t-e_{n,m}^R$ and $\rho^{(a)}(t) \ge \delta_n^{-\frac{1}{2}}(t-e_{n,m}^R)$, so $\rho^{(a)}(t) \gg \mu_*(t)$.

By Lemma 6.3, $e_{n,m+1}^L - e_{n,m}^R \ge C_{\boldsymbol{u}}\mu_*(e_{n,m}^R)$, 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) of Lermma 6.7, μ_* changes at most by a factor $2\kappa_2^{-2}$ 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 $t - e_{n,m}^R \le \mu_*(e_{n,m}^R)$, then $\mu_*(t) \le 2\mu_*(e_{n,m}^R)$, 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 2.1, we have

$$|\Omega_{\rho(t)}(\boldsymbol{u}(t))| \lesssim (1 + |\rho'(t)|)\|\boldsymbol{u}(t)\|^2_{\mathcal{E}(\rho(t),2\rho(t))} \lesssim \sqrt{\delta_n} \to 0,$$

which proves the property (4).

□

We need one more elementary result.

Lemma 6.10 *If $\mu_* : [a, b] \to (0, \infty)$ is a 1-Lipschitz function and $b - a \ge \frac{1}{4}\mu_*(a)$, then there exists a sequence $a = a_0 < a_1 < \ldots < a_l < a_{l+1} = b$ such that*

$$\frac{1}{4}\mu_*(a_i) \le a_{i+1} - a_i \le \frac{3}{4}\mu_*(a_i), \qquad \textit{for all } i \in \{1, \ldots, l\}. \tag{6.13}$$

Proof We define inductively $a_{i+1} := a_i + \frac{1}{4}\mu_*(a_i)$, as long as $b - a_i > \frac{3}{4}\mu_*(a_i)$. We need to prove that $b - a_i > \frac{3}{4}\mu_*(a_i)$ implies $b - a_{i+1} > \frac{1}{4}\mu_*(a_{i+1})$.

Since μ_* is 1-Lipschitz, $\mu_*(a_{i+1}) = \mu_*(a_i + \mu_*(a_i)/4) \le \mu_*(a_i) + \mu_*(a_i)/4 = \frac{5}{4}\mu_*(a_i)$, thus

$$b - a_{i+1} = b - a_i - \frac{1}{4}\mu_*(a_i) > \frac{3}{4}\mu_*(a_i) - \frac{1}{4}\mu_*(a_i) > \frac{5}{16}\mu_*(a_i) \ge \frac{1}{4}\mu_*(a_{i+1}).$$

□

Remark 6.11 Note that (6.13) and the fact that μ_* is 1-Lipschitz imply $\inf_{t\in[a_i,a_{i+1}]}\mu_*(t) \ge \frac{1}{4}\mu_*(a_i)$ and $\sup_{t\in[a_i,a_{i+1}]}\mu_*(t) \le \frac{7}{4}\mu_*(a_i)$, thus

$$\frac{1}{7}\sup_{t\in[a_i,a_{i+1}]}\mu_*(t) \le a_{i+1} - a_i \le 3\inf_{t\in[a_i,a_{i+1}]}\mu_*(t),$$

in other words the length of each subinterval is comparable with both the smallest and the largest value of μ_* on this subinterval.

Lemma 6.12 *Let ρ be the function given by Lemma 6.9 and set*

$$\mathfrak{v}(t) := \int_0^\infty \partial_t u(t) r \partial_r u(t) \chi_{\rho(t)}\, r\mathrm{d}r. \tag{6.14}$$

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*

$$\widetilde{b}_n - \widetilde{a}_n \geq \frac{1}{4}\mu_*(\widetilde{a}_n) \quad \textit{and} \quad \mathbf{d}(t) \geq \theta_n \textit{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

then

$$\mathfrak{v}(\widetilde{b}_n) < \mathfrak{v}(\widetilde{a}_n). \tag{6.15}$$

2. *For any $c, \theta > 0$ there exists $\delta > 0$ such that if n is large enough, $[\widetilde{a}_n, \widetilde{b}_n] \subset [a_n, b_n]$,*

$$c\mu_*(\widetilde{a}_n) \leq \widetilde{b}_n - \widetilde{a}_n \quad \textit{and} \quad \mathbf{d}(t) \geq \theta \textit{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

then

$$\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \leq -\delta \sup_{t \in [\widetilde{a}_n, \widetilde{b}_n]} \mu_*(t). \tag{6.16}$$

Proof By the virial identity, we obtain

$$\mathfrak{v}'(t) = -\int_0^\infty (\partial_t u(t))^2 \chi_{\rho(t)}\, r^{D-1}\mathrm{d}r + o_n(1). \tag{6.17}$$

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

$$\widetilde{b}_n - \widetilde{a}_n \geq \frac{1}{4}\mu_*(\widetilde{a}_n) \quad \text{and} \quad \mathbf{d}(t) \geq \theta \text{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

and

$$\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \geq 0.$$

By Lemma 6.10, there exists a subinterval of $[\widetilde{a}_n, \widetilde{b}_n]$, which we still denote $[\widetilde{a}_n, \widetilde{b}_n]$, such that

$$\frac{1}{4}\mu_*(\widetilde{a}_n) \leq \widetilde{b}_n - \widetilde{a}_n \leq \frac{3}{4}\mu_*(\widetilde{a}_n) \quad \text{and} \quad \mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \geq 0.$$

Let $\widetilde{\rho}_n := \inf_{t\in[\widetilde{a}_n,\widetilde{b}_n]} \rho(t)$. From (6.17), we have

$$\lim_{n\to\infty} \frac{1}{\widetilde{b}_n - \widetilde{a}_n} \int_{\widetilde{a}_n}^{\widetilde{b}_n} \int_0^{\frac{1}{2}\widetilde{\rho}_n} (\partial_t u(t))^2\, r^{D-1} \mathrm{d}r = 0.$$

By Lemma 6.9, $\inf_{t\in[\widetilde{a}_n,\widetilde{b}_n]} \mu_{K+1}(t) \gg \widetilde{\rho}_n \gg \inf_{t\in[\widetilde{a}_n,\widetilde{b}_n]} \mu_*(t) \simeq \sup_{t\in[\widetilde{a}_n,\widetilde{b}_n]} \mu_*(t)$, 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 6.2 (iii). The first part of the lemma is proved.

In the second part, we can assume without loss of generality $\widetilde{b}_n - \widetilde{a}_n \le \frac{3}{4}\mu_*(\widetilde{a}_n)$. Indeed, in the opposite case, we apply Lemma 6.10 and keep only one of the subintervals where μ_* attains its supremum, and on the remaining subintervals we use (6.15).

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

$$c\mu_*(\widetilde{a}_n) \le \widetilde{b}_n - \widetilde{a}_n \le \frac{3}{4}\mu_*(\widetilde{a}_n) \quad \text{and} \quad \mathbf{d}(t) \ge \theta \text{ for all } t \in [\widetilde{a}_n, \widetilde{b}_n],$$

and

$$\mathfrak{v}(\widetilde{b}_n) - \mathfrak{v}(\widetilde{a}_n) \ge -\delta_n\mu_*(\widetilde{a}_n)$$

(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 6.11).

Let $\widetilde{\rho}_n := \inf_{t\in[\widetilde{a}_n,\widetilde{b}_n]} \rho(t)$. From (6.17), we have

$$\lim_{n\to\infty} \frac{1}{\widetilde{b}_n - \widetilde{a}_n} \int_{\widetilde{a}_n}^{\widetilde{b}_n} \int_0^{\frac{1}{2}\widetilde{\rho}_n} (\partial_t u(t))^2\, r^{D-1} \mathrm{d}r = 0.$$

We now conclude as in the first part. □

Proof of Theorem 1 Let θ_n be the sequence given by Lemma 6.12, part 1. We partition $[a_n, b_n]$ applying Lemma 6.7 for this sequence θ_n. Note that this partition is different than the one used in the proof of Lemma 6.9. We claim that for all $m \in \{0, 1, \ldots, N_n - 1\}$

$$\mathfrak{v}(c_{n,m}^R) - \mathfrak{v}(e_{n,m}^R) \le o_n(1)\mu_*(c_{n,m}^R), \tag{6.18}$$

$$\mathfrak{v}(f_{n,m+1}^L) - \mathfrak{v}(f_{n,m}^R) \le o_n(1)\mu_*(f_{n,m}^R), \tag{6.19}$$

$$\mathfrak{v}(e_{n,m+1}^L) - \mathfrak{v}(c_{n,m+1}^L) \le o_n(1)\mu_*(c_{n,m+1}^L). \tag{6.20}$$

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 $c^R_{n,m} - e^R_{n,m} \geq \frac{1}{4}\mu_*(e^R_{n,m})$, then (6.15) applies and yields $\mathfrak{v}(c^R_{n,m}) - \mathfrak{v}(e^R_{n,m}) < 0$. We can thus assume $c^R_{n,m} - e^R_{n,m} \leq \frac{1}{4}\mu_*(e^R_{n,m}) \leq \frac{\kappa_2^{-2}}{2}\mu_*(c^R_{n,m})$, where the last inequality follows from Lemma 6.7, property (7). But then (6.17) again implies the required bound. The proofs of the second and third bound are analogous.

We now analyse the compactness intervals $[c^R_{n,j}, f^R_{n,j}]$ and $[f^L_{n,j+1}, c^L_{n,j+1}]$. We claim that there exists $\delta > 0$ such that for all n large enough and $m \in \{0, 1, \ldots, N_n\}$

$$\mathfrak{v}(c^L_{n,m+1}) - \mathfrak{v}(c^R_{n,m}) \leq -\delta \max(\mu_*(c^R_{n,m}), \mu_*(c^L_{n,m+1})). \tag{6.21}$$

We consider separately the two cases mentioned in Lemma 6.7, property (6). If $\mathbf{d}(t) \geq \epsilon$ for all $t \in [c^R_{n,m}, c^L_{n,m+1}]$, then Lemma 6.3 yields $c^L_{n,m+1} - c^R_{n,m} \geq C_{\boldsymbol{u}}^{-1}\mu_*(c^R_{n,m})$, so we can apply (6.16), which proves (6.21). If $\mathbf{d}(f^R_{n,m}) = \epsilon$, then we apply the same argument on the time interval $[c^R_{n,m}, f^R_{n,m}]$ and obtain

$$\mathfrak{v}(f^R_{n,m}) - \mathfrak{v}(c^R_{n,m}) \leq -\delta \max(\mu_*(c^R_{n,m}), \mu_*(f^R_{n,m})),$$

and similarly

$$\mathfrak{v}(c^L_{n,m+1}) - \mathfrak{v}(f^L_{n,m+1}) \leq -\delta \max(\mu_*(c^L_{n,m+1}), \mu_*(f^L_{n,m+1})).$$

The bound (6.19) yields (6.21).

Finally, on the intervals $[e^L_{n,m}, e^R_{n,m}]$, for n large enough Lemma 6.9 yields $|\rho'(t)| \leq 1$ for almost all t, and Lemma 5.19 implies $|\mathfrak{v}'(t)| \lesssim \mathbf{d}(t)$. By Lemma 6.7, properties (1) and (7), we obtain

$$\begin{aligned}
\mathfrak{v}(e^R_{n,m}) - \mathfrak{v}(e^L_{n,m}) &\leq o_n(1)\mu_*(c^R_{n,m}), && \text{for all } m \in \{0, 1, \ldots, N_n - 1\}, \\
\mathfrak{v}(e^R_{n,m}) - \mathfrak{v}(e^L_{n,m}) &\leq o_n(1)\mu_*(c^L_{n,m}), && \text{for all } m \in \{1, \ldots, N_n - 1, N_n\}.
\end{aligned} \tag{6.22}$$

Taking the sum in m of (6.18), (6.20), (6.21) and (6.22), we deduce that there exists $\delta > 0$ and n arbitrarily large such that

$$\mathfrak{v}(b_n) - \mathfrak{v}(a_n) \leq -\delta \max(\mu_*(c^R_{n,0}), \mu_*(c^L_{n,N_n})).$$

But $\mu_*(a_n) \simeq \mu_*(c^R_{n,0})$ and $\mu_*(b_n) \simeq \mu_*(c^L_{n,N_n})$, hence

$$\mathfrak{v}(b_n) - \mathfrak{v}(a_n) \leq -\widetilde{\delta} \max(\mu_*(a_n), \mu_*(b_n)).$$

Lemma 6.9 (1) and (6.14) yield

$$|\mathfrak{v}(a_n)| \ll \mu_*(a_n), \qquad |\mathfrak{v}(b_n)| \ll \mu_*(b_n),$$

a contradiction which finishes the proof. □

6.4 Absence of Elastic Collisions

This section is devoted to proving Proposition 1.7 Our proof closely follows Step 3 in our proof of [40, Theorem 1.6].

Proof of Proposition 1.7 Suppose that a solution $\boldsymbol{u}$ of (1.1), 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 2.17 with $\theta = 0$ and $M = N$, we obtain from (2.16) and (2.15) that

$$\mathbf{d}(t)^2 \le C\Big(\sum_{j \in \mathcal{S}} \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} + \sum_{k=1}^{K} \big(a_k^-(t)^2 + a_k^+(t)^2\big)\Big).$$

Inspecting the proof of Lemma 5.12, it follows that the last inequality and the fact that $\boldsymbol{u}^* = 0$ imply that Lemma 5.12 holds with $\zeta_n = 0$. Similarly, Lemma 5.17 holds with $\delta_n = 0$.

Let $\eta > 0$ be a small number to be chosen later and t_+ be such that $\mathbf{d}(t) \le \eta$ for all $t \ge t_+$. Lemma 6.5 yields

$$\int_{t_+}^{t} \mathbf{d}(t)\mathrm{d}t \le C_0\big(\mathbf{d}(t_+)^{\frac{4}{D-2}}\lambda_N(t_+), \mathbf{d}(t)^{\frac{4}{D-2}}\lambda_N(t)\big)$$

and passing to the limit $t \to T_+$ we get

$$\int_{t_+}^{T_+} \mathbf{d}(t)\mathrm{d}t \le C_0 \mathbf{d}(t_+)^{\frac{4}{D-2}}\lambda_N(t_+). \tag{6.23}$$

From the bound $|\lambda_N'(t)| \lesssim \mathbf{d}(t)$, see (5.26) with $\zeta_n = 0$, together with (6.23), 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 [40], but we reproduce it here for the reader's convenience.

Let $\delta > 0$ be arbitrary. Inspecting the proof of Lemma 5.19, we see that in the present case it holds with $\epsilon_n = 0$, thus for any $R > 0$ we have $\big|\Omega_{1,R}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,R}(\boldsymbol{u}(t))\big| \le C_0 \mathbf{d}(t)$. From this bound and the estimates above, we obtain existence of $T_1, T_2 \in \mathbb{R}$ such that

$$\int_{-\infty}^{T_1} \Big|\Omega_{1,R}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,R}(\boldsymbol{u}(t))\Big|\mathrm{d}t \le \frac{1}{3}\delta,$$
$$\int_{T_2}^{+\infty} \Big|\Omega_{1,R}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,R}(\boldsymbol{u}(t))\Big|\mathrm{d}t \le \frac{1}{3}\delta$$

for any $R > 0$. On the other hand, because of the bound $|\Omega_{j,R}(\boldsymbol{u}(t))| \le C_0\|\boldsymbol{u}(t)\|_{\mathcal{E}(R,2R)}$ and since $[T_1, T_2]$ is a finite time interval, for all R sufficiently large we have

$$\int_{T_1}^{T_2} \Big|\Omega_{1,R}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,R}(\boldsymbol{u}(t))\Big|\mathrm{d}t \le \frac{1}{3}\delta,$$

in other words

$$\int_{\mathbb{R}} \Big|\Omega_{1,R}(\boldsymbol{u}(t)) + \frac{D-2}{2}\Omega_{2,R}(\boldsymbol{u}(t))\Big|\mathrm{d}t \le \delta.$$

Integrating the virial identity (2.4) with $\rho(t) = R$ over the real line, we obtain

$$\int_{-\infty}^{+\infty}\int_0^{\infty} (\partial_t u(t,r)\chi_R(r))^2\, r^{D-1}\mathrm{d}r\mathrm{d}t \le \delta.$$

By letting $R \to +\infty$, we get

$$\int_{-\infty}^{+\infty}\int_0^{\infty} (\partial_t u(t,r))^2\, r^{D-1}\mathrm{d}r\mathrm{d}t \le \delta,$$

which implies the $\boldsymbol{u}$ is stationary since δ is arbitrary. □

Appendix A. Modifications to the Argument in the Case $D = 5$

In this section we outline the technical changes to the arguments in Section 5 needed to prove Theorem 1 dimensions $D = 5$.

A.1 Decomposition of the Solution

The set-up in Sections 5.1 holds without modification for $D = 5$. The number $K \ge 1$ is defined as in Lemma 5.6, the collision intervals $[a_n, b_n] \in \mathcal{C}_K(\eta, \epsilon_n)$ are as in Definition 5.5, and the sequences of signs $\vec{\sigma}_n \in \{-1, 1\}^{N-K}$, scales $\vec{\mu}(t) \in (0, \infty)^{N-K}$, and the sequence $\nu_n \to 0$ and the function $\nu(t) = \nu_n\mu_{K+1}(t)$ are as in Lemma 5.9.

Lemma 5.12 also holds with a minor modification to the stable/unstable components. Let $J \subset [a_n, b_n]$ be any time interval on which $\mathbf{d}(t) \le \eta_0$, where η_0 is as in Lemma 5.12. Let $\vec{\iota} \in \{-1, 1\}^K$, $\vec{\lambda}(t) \in (0, \infty)^K$, $\boldsymbol{g}(t) \in \mathcal{E}$, and $a_j^{\pm}(t)$ be as in the statement of Lemma 5.12. Define for each $1 \le j \le K$, the modified stable/unstable components,

$$\widetilde{a}_j^{\pm}(t) := \big\langle \boldsymbol{\alpha}^{\pm}_{\lambda_j(t)} \mid \boldsymbol{g}(t) + \sum_{i<j} \iota_i \boldsymbol{W}_{\lambda_i(t)} \big\rangle.$$

The estimate (5.28) will hold for $\widetilde{a}_j^{\pm}(t)$ rather than for $a_j^{\pm}(t)$, see (A.7) and (A.8) below. We make a similar modification (i.e., removing the interior bubbles from $\boldsymbol{g}(t)$) to the refined modulation parameter $\xi_j(t)$. For each $j \in \{1, \dots, K-1\}$, we set

$$\xi_j(t) = \lambda_j(t) - \frac{\iota_j}{\|\Lambda W\|_{L^2}^2} \big\langle \chi(\cdot/L\lambda_j(t)) \Lambda W_{\underline{\lambda_j(t)}} \mid g(t) + \sum_{i<j} \iota_i W_{\lambda_i(t)} \big\rangle$$

where $L > 0$ is a large parameter to be determined below. The refined modulation parameter $\beta_j(t)$ requires no modifications and is defined as in (5.50) for all $j \in \{1, \dots, K-1\}$.

With these definitions, the following analogue of Lemma 5.17 holds.

Proposition A.1 (Refined modulation, $D = 5$) *Let $c_0 \in (0, 1)$. There exists constants $L_0 = L_0(c_0) > 0$, $\eta_0 = \eta_0(c_0)$, as well as $c = c(c_0)$ and $R = R(c_0) > 1$ as in Lemma 5.14, 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 $\epsilon_n \le \mathbf{d}(t) \le \eta_0$ for all $t \in J$, where $\mathcal{S} := \{j \in \{1, \dots, K-1\} \mid \iota_j = \iota_{j+1}\}$. Then, for all $t \in J$,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{i \notin \mathcal{S}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{3}{4}} \le \max_{i \in \mathcal{S}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{3}{4}} + \max_{1 \le i \le K} |a_i^{\pm}(t)| \tag{A.1}$$

and,

$$\frac{1}{C_0}\mathbf{d}(t) \le \max_{i \in \mathcal{S}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{3}{4}} + \max_{1 \le i \le K} |a_i^{\pm}(t)| \le C_0 \mathbf{d}(t), \tag{A.2}$$

and,

$$\left| \frac{\xi_j(t)}{\lambda_j(t)} - 1 \right| \le c_0. \tag{A.3}$$

Moreover, for all $t \in J$,

$$\left| \xi_j'(t) \right| \le C_0 \mathbf{d}(t), \tag{A.4}$$

$$\left| \xi_j'(t) - \beta_j(t) \right| \le c_0 \mathbf{d}(t), \tag{A.5}$$

and,

$$\begin{aligned}\beta_j'(t) \geq \Big(\iota_j\iota_{j+1}\omega^2 - c_0\Big)\frac{1}{\lambda_j(t)}\left(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\right)^{\frac{3}{2}} \\ + \Big(-\iota_j\iota_{j-1}\omega^2 - c_0\Big)\frac{1}{\lambda_j(t)}\left(\frac{\lambda_{j-1}(t)}{\lambda_j(t)}\right)^{\frac{3}{2}} \\ - \frac{c_0}{\lambda_j(t)}\mathbf{d}(t)^2 - \frac{C_0}{\lambda_j(t)}\Big((a_j^+(t))^2 + (a_j^-(t))^2\Big),\end{aligned} \tag{A.6}$$

where, by convention, $\lambda_0(t) = 0, \lambda_{K+1}(t) = \infty$ *for all* $t \in J$*, and* $\omega^2 > 0$ *is as in* (5.56) *Finally, for each* $j \in \{1, \ldots, K\}$*,*

$$\Big|\widetilde{a}_j^{\pm}(t) - a_j^{\pm}(t)\Big| \leq C_0\mathbf{d}(t)^2 \tag{A.7}$$

and

$$\Big|\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{a}_j^{\pm}(t) \mp \frac{\kappa}{\lambda_j(t)}\widetilde{a}_j^{\pm}(t)\Big| \leq \frac{C_0}{\lambda_j(t)}\mathbf{d}(t)^2. \tag{A.8}$$

Proof The estimates (A.1) and (A.2) follow as in the proofs of the corresponding estimates in Lemma 5.17. Next, we have,

$$\begin{aligned}\Big|\frac{\xi_j}{\lambda_j} - 1\Big| &\lesssim \frac{1}{\lambda_j}\|\chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}}\|_{L^{\frac{10}{7}}}\|g\|_{L^{\frac{10}{3}}} + \sum_{i<j}\frac{1}{\lambda_j}\Big|\big\langle\chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid W_{\lambda_i}\big\rangle\Big| \\ &\lesssim L^{\frac{1}{2}}\|\boldsymbol{g}\|_{\mathcal{E}} + L^2\sum_{i<j}\left(\frac{\lambda_i}{\lambda_j}\right)^{\frac{3}{2}},\end{aligned}$$

which proves (A.3) as long as η_0 is sufficiently small compared to L.

Next, we compute $\xi_j'(t)$. We have,

$$\begin{aligned}\xi_j' = \lambda_j' &- \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle\chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \partial_t g\big\rangle \\ &+ \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\Big\langle\chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \sum_{i<j}\iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\Big\rangle \\ &+ \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\frac{\lambda_j'}{\lambda_j}\Big\langle(r\partial_r\chi)(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i W_{\lambda_i(t)}\Big\rangle \\ &+ \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\frac{\lambda_j'}{\lambda_j}\Big\langle\chi(\cdot/L\lambda_j)\underline{\Lambda}\Lambda W_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i W_{\lambda_i(t)}\Big\rangle.\end{aligned} \tag{A.9}$$

The last two terms above are acceptable errors. Indeed,

$$\begin{aligned}&\Big|\frac{\lambda_j'}{\lambda_j}\big\langle (r\partial_r\chi)(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid g+\sum_{i<j}\iota_i W_{\lambda_i(t)}\big\rangle\Big|\\&\lesssim \frac{|\lambda_j'|}{\lambda_j}\Big(\|(r\partial_r\chi)(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}}\|_{L^{\frac{10}{7}}}\|g\|_{L^{\frac{10}{3}}}+\sum_{i<j}\Big|\big\langle (r\partial_r\chi)(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid W_{\lambda_i}\big\rangle\Big|\Big)\\&\lesssim \Big(L^{\frac12}\|\boldsymbol{g}\|_{\mathcal{E}}+L^2\sum_{i<j}\Big(\frac{\lambda_i}{\lambda_j}\Big)^{\frac32}\Big)\mathbf{d}(t)\end{aligned}$$

and similarly,

$$\Big|\frac{\lambda_j'}{\lambda_j}\big\langle \chi(\cdot/L\lambda_j)\underline{\Lambda}\Lambda W_{\underline{\lambda_j}} \mid g+\sum_{i<j}\iota_i W_{\lambda_i(t)}\big\rangle\Big| \lesssim \Big(L^{\frac12}\|\boldsymbol{g}\|_{\mathcal{E}}+L^2\sum_{i<j}\Big(\frac{\lambda_i}{\lambda_j}\Big)^{\frac32}\Big)\mathbf{d}(t).$$

Using (5.58) in the second term in (A.9) gives

$$\begin{aligned}-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \partial_t g\big\rangle &= -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle\\&\quad-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \sum_{i=1}^K\iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\big\rangle\\&\quad-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \phi(u,\nu)\big\rangle.\end{aligned}$$

The first term on the right satisfies,

$$\begin{aligned}&-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle\\&= -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle+\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle (1-\chi(\cdot/L\lambda_j))\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle\\&= -\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle+o_L(1)\|\boldsymbol{g}\|_{\mathcal{E}}\end{aligned}$$

where the $o_L(1)$ term can be made as small as we like by taking $L>0$ large. Using (5.64), the second term yields,

$$
\begin{aligned}
&-\frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \sum_{i=1}^{K}\iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\big\rangle \\
&\qquad = -\lambda_j' - \sum_{i\neq j}\frac{\iota_j\iota_i\lambda_i'}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\big\rangle \\
&\qquad\quad + \frac{\lambda_j'}{\|\Lambda W\|_{L^2}^2}\big\langle (1-\chi(\cdot/L\lambda_j))\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_j}}\big\rangle \\
&\qquad = -\lambda_j' - \sum_{i<j}\frac{\iota_j\iota_i\lambda_i'}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\big\rangle \\
&\qquad\quad + O((\lambda_j/\lambda_{j+1})^{\frac{5}{2}} + o_L(1))\mathbf{d}(t).
\end{aligned}
$$

Finally, the third term vanishes due to the fact that for each $j < K$, $L\lambda_j \ll \lambda_K \ll \nu$, and hence

$$
\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \phi(u,\nu)\big\rangle = 0.
$$

Plugging this all back into (A.9) gives,

$$
\Big|\xi_j' + \frac{\iota_j}{\|\Lambda W\|_{L^2}^2}\big\langle \chi(\cdot/L\lambda_j)\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle\Big| \le c_0\mathbf{d}(t)
$$

after first choosing L sufficiently large, and then η_0 sufficiently small. The estimates (A.4) is immediate, and (A.5) now follows as in the proof of (5.54) in Lemma 5.17.

The estimate (A.6) is proved exactly as in the proof of (5.55) in Lemma 5.17.

Next, we have,

$$
\big|a_j^\pm - \widetilde{a}_j^\pm\big| \lesssim \Big|\big\langle \boldsymbol{\alpha}_{\lambda_j}^\pm \mid \sum_{i<j}\iota_i \boldsymbol{W}_{\lambda_i}\big\rangle\Big| \lesssim \sum_{i<j}\frac{\lambda_i}{\lambda_j}\Big|\big\langle \mathcal{Y}_{\underline{\lambda_j}} \mid W_{\underline{\lambda_i}}\big\rangle\Big| \lesssim \sum_{i<j}\Big(\frac{\lambda_i}{\lambda_j}\Big)^{\frac{3}{2}},
$$

which proves (A.7).

Lastly, we prove (A.8), which is analogous to the proof of (5.28), but now using (A.1) and (A.2), and noting an extra cancellation of the contribution of the interior bubbles. We compute,

$$
\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{a}_j^- = \big\langle \partial_t\boldsymbol{\alpha}_{\lambda_j}^- \mid \boldsymbol{g} + \sum_{i<j}\iota_i\boldsymbol{W}_{\lambda_i}\big\rangle + \big\langle \boldsymbol{\alpha}_{\lambda_j}^- \mid \partial_t\boldsymbol{g}\big\rangle - \sum_{i<j}\iota_i\frac{\kappa}{2}\frac{\lambda_i'}{\lambda_j}\big\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\big\rangle. \quad \text{(A.10)}
$$

Expanding the first term on the right gives,

$$\begin{aligned}\langle \partial_t \boldsymbol{\alpha}^-_{\lambda_j} \mid \boldsymbol{g}\rangle &= \frac{\kappa}{2}\langle \partial_t(\lambda_j^{-1}\mathcal{Y}_{\underline{\lambda_j}}) \mid g + \sum_{i<j}\iota_i W_{\lambda_i}\rangle + \frac{1}{2}\langle \partial_t(\mathcal{Y}_{\underline{\lambda_j}}) \mid \dot g\rangle \\ &= -\frac{\kappa}{2}\frac{\lambda_j'}{\lambda_j}\langle \lambda_j^{-1}\mathcal{Y}_{\underline{\lambda_j}} + \frac{1}{\lambda_j}(\underline{\Lambda}\mathcal{Y})_{\underline{\lambda_j}} \mid g + \sum_{i<j}\iota_i W_{\lambda_i}\rangle - \frac{1}{2}\frac{\lambda_j'}{\lambda_j}\langle (\underline{\Lambda}\mathcal{Y})_{\underline{\lambda_j}}) \mid \dot g\rangle\end{aligned}$$

and thus,

$$\left|\langle \partial_t \boldsymbol{\alpha}^-_{\lambda_j} \mid \boldsymbol{g}\rangle\right| \lesssim \frac{1}{\lambda_j}\mathbf{d}(t)^2.$$

We use (5.39) to expand the second term,

$$\begin{aligned}\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{g}\rangle = \langle \boldsymbol{\alpha}^-_{\lambda_j} &\mid J\circ \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda))\boldsymbol{g}\rangle \\ &+\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J\circ\Big(\mathrm{D}\,E(\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda)+\boldsymbol{g}) - \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda))\boldsymbol{g}\Big)\rangle \\ &-\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda)\rangle \\ &+\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \Big(\chi(\cdot/\nu)J\circ \mathrm{D}\,E(\boldsymbol{u}) - J\circ \mathrm{D}\,E(\chi(\cdot/\nu)\boldsymbol{u})\Big)\rangle \\ &-\frac{\nu'}{\nu}\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid (r\partial_r\chi)(\cdot/\nu)\boldsymbol{u}\rangle.\end{aligned} \tag{A.11}$$

By (2.11) the first term on the right gives the leading order,

$$\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J\circ \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda))\boldsymbol{g}\rangle = -\frac{\kappa}{\lambda_j}a_j^-.$$

Next, we expand,

$$\begin{aligned}&\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid J\circ\Big(\mathrm{D}\,E(\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda)+\boldsymbol{g}) - \mathrm{D}^2 E(\boldsymbol{\mathcal{W}}(\vec\iota,\vec\lambda))\boldsymbol{g}\Big)\rangle \\ &= -\frac{1}{2}\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f(\mathcal{W}(\vec\iota,\vec\lambda)+g) - f(\mathcal{W}(\vec\iota,\vec\lambda)) - f'(\mathcal{W}(\vec\iota,\vec\lambda))g\rangle \\ &-\frac{1}{2}\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f(\mathcal{W}(\vec\iota,\vec\lambda)) - \sum_{i=1}^K \iota_i f(W_{\lambda_i})\rangle.\end{aligned}$$

The first line satisfies,

$$\left|\langle \mathcal{Y}_{\underline{\lambda}_j} \mid f(\mathcal{W}(\vec\iota,\vec\lambda)+g) - f(\mathcal{W}(\vec\iota,\vec\lambda)) - f'(\mathcal{W}(\vec\iota,\vec\lambda))g\rangle\right| \lesssim \frac{1}{\lambda_j}(\mathbf{d}(t)^2 + o_n(1)).$$

Noting that $f(\mathcal{W}(\vec{\iota},\vec{\lambda})) - \sum_{i=1}^{K} \iota_i f(W_{\lambda_i}) = f_{\mathbf{i}}(\vec{\iota},\vec{\lambda})$, the same argument used to prove Lemma 2.21 gives,

$$\left|\left\langle \mathcal{Y}_{\underline{\lambda_j}} \mid f_{\mathbf{i}}(\vec{\iota},\vec{\lambda})\right\rangle\right| \lesssim \frac{1}{\lambda_j}\Big(\Big(\frac{\lambda_j}{\lambda_{j+1}}\Big)^{\frac{3}{2}} + \Big(\frac{\lambda_{j-1}}{\lambda_j}\Big)^{\frac{3}{2}}\Big) \lesssim \frac{1}{\lambda_j}\mathbf{d}(t)^2.$$

Consider now the third line in (A.11).

$$\begin{aligned}
-\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota},\vec{\lambda})\right\rangle &= \frac{\kappa}{2}\iota_j \frac{\lambda_j'}{\lambda_j}\left\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_j}}\right\rangle + \sum_{i\neq j} \iota_i \frac{\kappa}{2}\frac{\lambda_i'}{\lambda_j}\left\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\right\rangle \\
&= \sum_{i\neq j} \iota_i \frac{\kappa}{2}\frac{\lambda_i'}{\lambda_j}\left\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\right\rangle
\end{aligned}$$

where in the last equality we used the vanishing $\left\langle \mathcal{Y} \mid \Lambda W\right\rangle$. Noting the estimates,

$$\left|\left\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\right\rangle\right| \lesssim \Big(\frac{\lambda_i}{\lambda_j}\Big)^{\frac{5}{2}} \text{ if } i > j$$

we obtain,

$$\left| -\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \partial_t \boldsymbol{\mathcal{W}}(\vec{\iota},\vec{\lambda})\right\rangle - \sum_{i\neq j} \iota_i \frac{\kappa}{2}\frac{\lambda_i'}{\lambda_j}\left\langle \mathcal{Y}_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\right\rangle\right| \lesssim \frac{1}{\lambda_j}\mathbf{d}(t)^2.$$

Using (5.10) and (5.13) we see that the last two lines of (A.11) satisfy,

$$\begin{aligned}
\left|\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid \Big(\chi(\cdot/\nu) J \circ \mathrm{D} E(\boldsymbol{u}) - J \circ \mathrm{D} E(\chi(\cdot/\nu)\boldsymbol{u})\Big)\right\rangle\right| &\lesssim \frac{1}{\lambda_j} o_n(1), \\
\left|\frac{\nu'}{\nu}\left\langle \boldsymbol{\alpha}^-_{\lambda_j} \mid (r\partial_r \chi)(\cdot/\nu)\boldsymbol{u}\right\rangle\right| &\lesssim \frac{1}{\lambda_j} o_n(1).
\end{aligned}$$

Plugging this all back into (A.10) and using (A.7) we obtain,

$$\left|\widetilde{a}_j^- + \frac{\kappa}{\lambda_j}\widetilde{a}_j^-\right| \lesssim \frac{1}{\lambda_j}(\mathbf{d}(t)^2 + o_n(1)).$$

This completes the proof after ensuring ϵ_n is large enough so that the $o_n(1)$ term above can be absorbed into $\mathbf{d}(t)$. □

A.2 Conclusion of the Proof

Using the modulation estimates above, we can prove the following analog of Lemma 6.5.

Lemma A.2 *Let $D = 5$. If η_0 is small enough, then there exists $C_0 \geq 0$ depending only on N such that the following is true. If $[t_1, t_2] \subset I_*$ is a finite time interval such that $\mathbf{d}(t) \leq \eta_0$ for all $t \in [t_1, t_2]$, then*

$$\sup_{t\in[t_1,t_2]} \lambda_K(t) \leq \frac{4}{3} \inf_{t\in[t_1,t_2]} \lambda_K(t),$$
$$\int_{t_1}^{t_2} \mathbf{d}(t)\mathrm{d}t \leq C_0\big(\mathbf{d}(t_1)\lambda_K(t_1) + \mathbf{d}(t_2)\lambda_K(t_2)\big).$$

Sketch of a proof **Step 1.** is exactly the same as for Lemma 6.5.
Step 2. Let $C_1 > 0$ be a large number chosen below and consider the auxiliary function

$$\phi(t) := \sum_{j\in\mathcal{S}} 2^{-j}\xi_j(t)\beta_j(t) - C_1\sum_{j=1}^{K}\lambda_j(t)\widetilde{a}_j^-(t)^2 + C_1\sum_{j=1}^{K}\lambda_j(t)\widetilde{a}_j^+(t)^2.$$

We claim that for all $t \in [t_1, t_2]$

$$\phi'(t) \geq c_2\mathbf{d}(t)^2, \tag{A.12}$$

with $c_2 > 0$ depending only on N. The remaining part of Step 1 is devoted to proving this bound.

Using (A.5), (A.8) and recalling that $|\lambda_K'(t)| \lesssim \mathbf{d}(t)$, we obtain

$$\phi'(t) \geq \sum_{j\in\mathcal{S}} 2^{-j}\beta_j^2(t) + \sum_{j\in\mathcal{S}} 2^{-j}\lambda_j(t)\beta_j'(t) + C_1\nu\sum_{j=1}^{K}\big(\widetilde{a}_j^-(t)^2 + \widetilde{a}_j^+(t)^2\big) - c_0\mathbf{d}(t)^2, \tag{A.13}$$

where $c_0 > 0$ can be made arbitrarily small. We focus on the second term of the right hand side. Like in Step 2. of the proof of Lemma 6.5, only using (A.6) instead of (5.55), we obtain

$$\sum_{j\in\mathcal{S}}\lambda_j(t)\beta_j'(t) \geq 2^{-N-1}\omega^2\sum_{j\in\mathcal{S}}\Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{D-2}{2}} - C_2\sum_{k=1}^{K}\big(a_j^-(t)^2 + a_j^+(t)^2\big) - c_0\mathbf{d}(t)^2.$$

The bound (A.7) implies that (A.13) holds with with $\widetilde{a}_j^\pm$ replaced by $a_j^\pm$. Taking $C_1 > C_2/\nu$ and using (A.2), we get (A.12).
Step 3. As in Step 3. of the proof of Lemma 6.5, it suffices to check that if $t_3 \in [t_1, t_2]$ and $\phi(t) > 0$ for all $t \in (t_3, t_2]$, then

$$\int_{t_3}^{t_2}\mathbf{d}(t)\mathrm{d}t \leq C_0\mathbf{d}(t_2)\lambda_K(t_2). \tag{A.14}$$

Observe that for all $t \in (t_3, t_2]$ we have

$$\begin{aligned}\frac{\phi(t)}{\lambda_K(t)} &\lesssim \sum_{j\in\mathcal{S}} \frac{\lambda_j(t)}{\lambda_K(t)}|\beta_j(t)| + \sum_{j=1}^{K} \frac{\lambda_j(t)}{\lambda_K(t)}\big(\widetilde{a}_j^-(t)^2 + \widetilde{a}_j^+(t)^2\big) \\ &\lesssim \sum_{j\in\mathcal{S}} \frac{\lambda_j(t)}{\lambda_{j+1}(t)}|\beta_j(t)| + \mathbf{d}(t)^2 \lesssim \mathbf{d}(t)^{\frac{4}{3}+1} + \mathbf{d}(t)^2 \lesssim \mathbf{d}(t)^2.\end{aligned} \tag{A.15}$$

Combining this bound with (A.12), for all $t \in (t_3, t_2]$ we get

$$\phi'(t) \geq c_2\sqrt{\phi(t)/\lambda_K(t)}\mathbf{d}(t),$$

thus

$$\sqrt{\lambda_K(t)}\big(\sqrt{\phi(t)}\big)' \gtrsim \mathbf{d}(t). \tag{A.16}$$

Using (A.15) and $|\lambda_K'(t)| \lesssim \mathbf{d}(t)$, we get

$$\big(\sqrt{\lambda_K(t)\phi(t)}\big)' - \sqrt{\lambda_K(t)}\big(\sqrt{\phi(t)}\big)' = \frac{1}{2}\lambda_K'(t)\sqrt{\phi(t)/\lambda_K(t)} \gtrsim -\mathbf{d}(t)^2.$$

Since $\mathbf{d}(t)$ is small, (A.16) yields

$$\Big(\sqrt{\lambda_K(t)\phi(t)^{\frac{4}{D+2}}}\Big)' \gtrsim \mathbf{d}(t)$$

which, integrated, gives

$$\int_{t^3}^{t_2} \mathbf{d}(t)\mathrm{d}t \lesssim \sqrt{\lambda_K(t_2)\phi(t_2)} - \sqrt{\lambda_K(t_3)\phi(t_3)} \leq \sqrt{\lambda_K(t_2)\phi(t_2)}.$$

Invoking (A.15), we obtain (A.14). □

Using Lemma A.2 in lieu of Lemma 6.5, the remaining arguments in Section 6 hold for $D = 5$ without changes.

Appendix B. Modifications to the Argument in the Case $D = 4$

In this section we outline the changes to the arguments in Section 5 and Section 6 needed to prove Theorem 1 dimension $D = 4$.

B.1 Decomposition of the Solution

The set-up in Sections 5.1 holds without modification for $D = 4$. The number $K \geq 1$ is defined as in Lemma 5.6, the collision intervals $[a_n, b_n] \in \mathcal{C}_K(\eta, \epsilon_n)$ are as in Definition 5.5, and the sequences of signs $\vec{\sigma}_n \in \{-1, 1\}^{N-K}$, scales $\vec{\mu}(t) \in (0, \infty)^{N-K}$, and the sequence $\nu_n \to 0$ and the function $\nu(t) = \nu_n \mu_{K+1}(t)$ are as in Lemma 5.9.

Lemma 5.12 also holds with a minor modification to the stable/unstable components. Let $J \subset [a_n, b_n]$ be any time interval on which $\mathbf{d}(t) \le \eta_0$, where η_0 is as in Lemma 5.12. Let $\vec{\iota} \in \{-1, 1\}^K$, $\vec{\lambda}(t) \in (0, \infty)^K$, $\boldsymbol{g}(t) \in \mathcal{E}$, and $a_j^{\pm}(t)$ be as in the statement of Lemma 5.12. Define for each $1 \le j \le K$, the modified stable/unstable components,

$$\widetilde{a}_j^{\pm}(t) := \big\langle \boldsymbol{\alpha}^{\pm}_{\lambda_j(t)} \mid \boldsymbol{g}(t) + \sum_{i<j} \iota_i \boldsymbol{W}_{\lambda_i(t)} \big\rangle.$$

Let $L > 0$ be a parameter to be fixed below and for each $j \in \{1, \dots, K-1\}$ set,

$$\xi_j(t) := \lambda_j(t) - \frac{\iota_j}{8\log(\frac{\lambda_{j+1}(t)}{\lambda_j(t)})} \big\langle \chi_{L\sqrt{\lambda_j(t)\lambda_{j+1}(t)}} \Lambda W_{\underline{\lambda_j(t)}} \mid g(t) + \sum_{i<j} \iota_i W_{\lambda_i(t)} \big\rangle,$$

and,

$$\beta_j(t) := -\iota_j \big\langle \chi_{L\sqrt{\xi_j(t)\lambda_{j+1}(t)}} \Lambda W_{\underline{\lambda_j(t)}} \mid \dot{g}(t) \big\rangle - \big\langle \underline{A}(\lambda_j(t)) g(t) \mid \dot{g}(t) \big\rangle . \tag{B.1}$$

Proposition B.1 (Refined modulation, $D = 4$) *Let $c_0 \in (0, 1)$ and $c_1 > 0$. There exists $L_0 = L_0(c_0, c_1) > 0$, $\eta_0 = \eta_0(c_0, c_1)$, as well as $c = c(c_0, c_1)$ and $R = R(c_0, c_1) > 1$ as in Lemma 5.14, 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 $\epsilon_n \le \mathbf{d}(t) \le \eta_0$ for all $t \in J$, where $\mathcal{S} := \{j \in \{1, \dots, K-1\} \mid \iota_j = \iota_{j+1}\}$. Then, for all $t \in J$,

$$\|\boldsymbol{g}(t)\|_{\mathcal{E}} + \sum_{i \notin \mathcal{S}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}} \lesssim \max_{i \in \mathcal{S}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}} + \max_{1 \le i \le K} |a_i^{\pm}(t)|, \tag{B.2}$$

and,

$$\frac{1}{C_0}\mathbf{d}(t) \le \max_{i \in \mathcal{S}} (\lambda_i(t)/\lambda_{i+1}(t))^{\frac{1}{2}} + \max_{1 \le i \le K} |a_i^{\pm}(t)| \le C_0 \mathbf{d}(t), \tag{B.3}$$

$$\Big| \frac{\xi_j(t)}{\lambda_{j+1}(t)} - \frac{\lambda_j(t)}{\lambda_{j+1}(t)} \Big| \le C_0 \frac{\mathbf{d}(t)^2}{\log(\lambda_{j+1}/\lambda_j)}, \tag{B.4}$$

as well as,

$$\Big| a_j^{\pm}(t) - \widetilde{a}_j^{\pm}(t) \Big| \le C_0 \mathbf{d}(t)^2, \tag{B.5}$$

and,

$$\Big| \frac{\mathrm{d}}{\mathrm{d}t} \widetilde{a}_j^{\pm}(t) \mp \frac{\kappa}{\lambda_j(t)} \widetilde{a}_j^{\pm}(t) \Big| \le \frac{C_0}{\lambda_j(t)} \mathbf{d}(t)^2. \tag{B.6}$$

Moreover, let $j \in \mathcal{S}$ be such that for all $t \in J$

$$c_1 \mathbf{d}(t) \le \Big(\frac{\lambda_j(t)}{\lambda_{j+1}(t)}\Big)^{\frac{1}{2}}. \tag{B.7}$$

Then for all $t \in J$,

$$\Big|\xi_j'(t)\Big| \Big(\log\Big(\frac{\lambda_{j+1}(t)}{\lambda_j(t)}\Big)\Big)^{\frac{1}{2}} \le C_0 \mathbf{d}(t), \tag{B.8}$$

$$\Big|\xi_j'(t) 8 \log(\frac{\lambda_{j+1}(t)}{\lambda_j(t)}) - \beta_j(t)\Big| \le C_0 \mathbf{d}(t), \tag{B.9}$$

and,

$$\begin{aligned}\beta_j'(t) \ge \Big(\iota_j\iota_{j+1} 16 - c_0\Big)\frac{1}{\lambda_{j+1}(t)} + \Big(-\iota_j\iota_{j-1}16 - c_0\Big)\frac{\lambda_{j-1}(t)}{\lambda_j(t)^2} \\ - \frac{c_0}{\lambda_j(t)}\mathbf{d}(t)^2 - \frac{C_0}{\lambda_j(t)}\Big((a_j^+(t))^2 + (a_j^-(t))^2\Big)\end{aligned} \tag{B.10}$$

where, by convention, $\lambda_0(t) = 0$, $\lambda_{K+1}(t) = \infty$ for all $t \in J$.

Remark B.2 Proposition B.1 and its proof are nearly identical to [41, Proposition A.1] and its proof, which treat the case $k = 1$ for the energy-critical equivariant wave map equation.

Proof The estimates (B.2) and (B.3) follow as in the proofs of the corresponding estimates in Lemma 5.17. We next prove (B.4). From the definition of $\xi_j(t)$,

$$\begin{aligned}\Big|\frac{\xi_j}{\lambda_{j+1}} - \frac{\lambda_j}{\lambda_{j+1}}\Big| \lesssim \Big|\frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_{j+1}^{-1}\big\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}}\Lambda W_{\underline{\lambda_j}} \mid g\big\rangle\Big| \\ + \Big|\frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_{j+1}^{-1}\big\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}}\Lambda W_{\underline{\lambda_j}} \mid \sum_{i<j} W_{\lambda_i}\big\rangle\Big|.\end{aligned}$$

For the first term on the right we have,

$$\Big|\frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_{j+1}^{-1}\big\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}}\Lambda W_{\underline{\lambda_j}} \mid g\big\rangle\Big| \lesssim_L \frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})}\|g\|_H (\lambda_j/\lambda_{j+1})^{\frac{1}{2}}.$$

Next, for any $i < j$ we have,

$$\begin{aligned}\lambda_{j+1}^{-1}|\big\langle \chi_{L\sqrt{\lambda_j\lambda_{j+1}}}\Lambda W_{\underline{\lambda_j}} \mid W_{\lambda_i}\big\rangle| &\lesssim_L \frac{\lambda_j}{\lambda_{j+1}}\int_0^{L(\lambda_{j+1}/\lambda_j)^{\frac{1}{2}}} |\Lambda W(r)||W_{\lambda_i/\lambda_j}(r)|\, r^3\,\mathrm{d}r \\ &\lesssim_L \frac{\lambda_j}{\lambda_{j+1}}\frac{\lambda_i}{\lambda_j}\log(\lambda_{j+1}/\lambda_j)\end{aligned}$$

and hence,

$$\Big|\frac{1}{\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_{j+1}^{-1}\big\langle\chi_{L\sqrt{\lambda_j\lambda_{j+1}}}\Lambda W_{\underline{\lambda_j}}\mid\sum_{i<j}W_{\lambda_i}\big\rangle\Big|\lesssim_L\frac{\lambda_j}{\lambda_{j+1}}\sum_{i<j}\frac{\lambda_i}{\lambda_j}$$

and (B.4) follows.

Next using (B.2) and (5.26) for each j, we have

$$\big|\lambda_j'\big|\lesssim\mathbf{d}(t). \tag{B.11}$$

We show that in fact ξ_j' satisfies the improved estimate (B.8). We compute,

$$\begin{aligned}\xi_j'=\lambda_j'&-\frac{\iota_j}{8(\log(\frac{\lambda_{j+1}}{\lambda_j}))^2}\Big(\frac{\lambda_j'}{\lambda_j}-\frac{\lambda_{j+1}'}{\lambda_{j+1}}\Big)\big\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid g+\sum_{i<j}\iota_iW_{\lambda_i}\big\rangle\\&+\frac{\iota_j}{16\log(\frac{\lambda_{j+1}}{\lambda_j})}\Big(\frac{\lambda_j'}{\lambda_j}+\frac{\lambda_{j+1}'}{\lambda_{j+1}}\Big)\big\langle(r\partial_r\chi)(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid g+\sum_{i<j}\iota_iW_{\lambda_i}\big\rangle\\&+\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\frac{\lambda_j'}{\lambda_j}\big\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\underline{\Lambda}\Lambda W_{\underline{\lambda_j}}\mid g+\sum_{i<j}\iota_iW_{\lambda_i}\big\rangle\\&-\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid\partial_tg\big\rangle\\&+\sum_{i<j}\frac{\iota_i\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_i'\big\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid\Lambda W_{\underline{\lambda_i}}\big\rangle.\end{aligned} \tag{B.12}$$

The second, third, and fourth terms on the right above contribute acceptable errors. Indeed,

$$\begin{aligned}&\Big|\frac{\iota_j}{8(\log(\frac{\lambda_{j+1}}{\lambda_j}))^2}\Big(\frac{\lambda_j'}{\lambda_j}-\frac{\lambda_{j+1}'}{\lambda_{j+1}}\Big)\big\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid g+\sum_{i<j}\iota_iW_{\lambda_i}\big\rangle\Big|\\&\quad\lesssim\frac{\mathbf{d}(t)}{c_1(\log(\frac{\lambda_{j+1}}{\lambda_j}))^2}\\&\Big|\frac{\iota_j}{16\log(\frac{\lambda_{j+1}}{\lambda_j})}\Big(\frac{\lambda_j'}{\lambda_j}+\frac{\lambda_{j+1}'}{\lambda_{j+1}}\Big)\big\langle(r\partial_r\chi)(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid g+\sum_{i<j}\iota_iW_{\lambda_i}\big\rangle\Big|\\&\quad\lesssim\frac{\mathbf{d}(t)}{c_1\log(\frac{\lambda_{j+1}}{\lambda_j})}\Big|\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\frac{\lambda_j'}{\lambda_j}\big\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\underline{\Lambda}\Lambda W_{\underline{\lambda_j}}\mid g+\sum_{i<j}\iota_iW_{\lambda_i}\big\rangle\Big|\\&\quad\lesssim\frac{\mathbf{d}(t)^2}{\log(\frac{\lambda_{j+1}}{\lambda_j})}\end{aligned}$$

with the gain in the last line arising from the fact that $\underline{\Lambda}\Lambda W \in L^{\frac{4}{3}}$; see (2.8). The leading order comes from the second to last term in (B.12). Using (5.58) gives

$$\begin{aligned}
&-\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \partial_t g\big\rangle \\
&\quad= -\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle \\
&\qquad - \lambda_j'\frac{1}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_j}}\big\rangle \\
&\qquad - \frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \sum_{i\neq j}\iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\big\rangle \\
&\qquad - \frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \phi(u,\nu)\big\rangle.
\end{aligned}$$

We estimate the contribution of each of the terms on the right above to (B.12). The last term above vanishes due to the support properties of $\phi(u,\nu)$. Using (2.9), (B.11) on the second term above, gives

$$\Big| - \lambda_j'\frac{1}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_j}}\big\rangle + \lambda_j'| \lesssim \frac{\mathbf{d}(t)}{\log(\frac{\lambda_{j+1}}{\lambda_j})},$$

which means this terms cancels the term λ' on the right-hand side of (B.12) up to an acceptable error. Next we write,

$$\begin{aligned}
&-\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \sum_{i\neq j}\iota_i\lambda_i'\Lambda W_{\underline{\lambda_i}}\big\rangle \\
&\quad= -\sum_{i<j}\frac{\iota_i\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_i'\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\big\rangle \\
&\qquad -\sum_{i>j}\frac{\iota_i\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\lambda_i'\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\big\rangle.
\end{aligned}$$

The first term cancels the last term in (B.12). For the second term we estimate, if $i > j$,

$$|\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \Lambda W_{\underline{\lambda_i}}\big\rangle| \lesssim \lambda_j/\lambda_{j+1}$$

and thus, using (B.11) the second term in the previous equation contributes an acceptable error. Plugging all of these estimates back into (B.12) gives the estimate,

$$\Big|\xi_j' + \frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\big\langle \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle\Big| \lesssim \frac{\mathbf{d}(t)}{c_1\log(\frac{\lambda_{j+1}}{\lambda_j})}. \tag{B.13}$$

Using (B.2) and $\|\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\|_{L^2} \lesssim_L (\log(\frac{\lambda_{j+1}}{\lambda_j}))^{\frac{1}{2}}$, we deduce the estimate,

$$\Big|\frac{\iota_j}{8\log(\frac{\lambda_{j+1}}{\lambda_j})}\langle\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \dot{g}\rangle\Big| \lesssim_L \frac{\mathbf{d}(t)}{(\log(\frac{\lambda_{j+1}}{\lambda_j}))^{\frac{1}{2}}},$$

which completes the proof of (B.8).

Next we compare β_j and $2\xi_j'\log(\lambda_{j+1}/\lambda_j)$. Using (B.1) we have,

$$\Big|\big\langle\underline{A}(\lambda_j(t))g(t) \mid \dot{g}(t)\big\rangle\Big| \lesssim \|\boldsymbol{g}\|_{\mathcal{E}}^2 \lesssim \mathbf{d}(t)^2.$$

We also note the estimate

$$\Big|\big\langle(\chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}}) - \chi(\cdot/L\sqrt{\lambda_j\lambda_{j+1}})\Lambda Q_{\underline{\lambda_j}} \mid \dot{g}\big\rangle\Big| \leq \frac{1}{c_1^2}\mathbf{d}(t)^2,$$

which is a consequence of (B.4). Using (B.13) the estimate (B.9) follows.

Finally, the proof of the estimate (B.10) is nearly identical to the argument used to prove (5.55), 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 (5.55) we arrive at the formula,

$$\begin{aligned}
\beta_j' = &-\frac{\iota_j}{\lambda_j}\Big\langle\Lambda W_{\lambda_j} \mid f_{\mathbf{i}}(\iota,\vec{\lambda})\Big\rangle + \big\langle\underline{A}(\lambda_j)g \mid -\Delta g\big\rangle \\
&+\Big\langle(A(\lambda_j) - \underline{A}(\lambda_j))g \mid \widetilde{f}_{\mathbf{q}}(\vec{\iota},\vec{\lambda},g)\Big\rangle \\
&+\Big\langle\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid (\mathcal{L}_{\mathcal{W}} - \mathcal{L}_{\lambda_j})g\Big\rangle \\
&+\iota_j\frac{\lambda_j'}{\lambda_j}\Big\langle\Big(\frac{1}{\lambda_j}\underline{\Lambda} - \underline{A}(\lambda_j)\Big)\Lambda W_{\lambda_j} \mid \dot{g}\Big\rangle \\
&-\Big\langle A(\lambda_j)\sum_{i=1}^{K}\iota_i W_{\lambda_i} \mid f_{\mathbf{q}}(\vec{\iota},\vec{\lambda},g)\Big\rangle - \Big\langle A(\lambda_j)g \mid \widetilde{f}_{\mathbf{q}}(\vec{\iota},\vec{\lambda},g)\Big\rangle \\
&+\iota_j\Big\langle(A(\lambda_j) - \frac{1}{\lambda_j}\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda)W_{\lambda_j} \mid f_{\mathbf{q}}(\vec{\iota},\vec{\lambda},g)\Big\rangle \\
&-\frac{\lambda_j'}{\lambda_j}\big\langle\lambda_j\partial_{\lambda_j}\underline{A}(\lambda_j)g \mid \dot{g}\big\rangle \\
&+\sum_{i\neq j}\iota_i\Big\langle A(\lambda_j)W_{\lambda_i} \mid f_{\mathbf{q}}(\vec{\iota},\vec{\lambda},g)\Big\rangle - \sum_{i\neq j}\iota_i\lambda_i'\Big\langle\underline{A}(\lambda_j)\Lambda W_{\underline{\lambda_i}} \mid \dot{g}\Big\rangle \\
&-\Big\langle\underline{A}(\lambda_j)g \mid f_{\mathbf{i}}(\iota,\vec{\lambda})\Big\rangle
\end{aligned}$$

$$
\begin{aligned}
&-\iota_j \Big\langle \chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \dot\phi(u,\nu)\Big\rangle - \big\langle \underline{A}(\lambda_j)\phi(u,\nu) \mid \dot g\big\rangle \\
&-\big\langle \underline{A}(\lambda_j) g \mid \dot\phi(u,\nu)\big\rangle \\
&+\frac{\iota_j}{\lambda_j}\Big\langle (1-\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}}))\Lambda W_{\lambda_j} \mid f_{\mathbf{i}}(\iota,\vec\lambda)\Big\rangle \\
&-\iota_j \frac{\lambda_j'}{\lambda_j}\Big\langle (1-\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}}))\underline{\Lambda}\Lambda W_{\underline{\lambda_j}} \mid \dot g\Big\rangle \\
&+\Big\langle \mathcal{L}_{\lambda_j}(\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}) \mid g\Big\rangle \\
&+\frac{\iota_j}{2}\Big(\frac{\xi_j'}{\xi_j}+\frac{\lambda_{j+1}'}{\lambda_{j+1}}\Big)\big\langle \Lambda\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} \mid \dot g\big\rangle.
\end{aligned}
$$

All but the last four terms above are treated exactly as in the proof of (5.55). For the fourth-to-last term a direct computation gives,

$$
\left|\frac{\iota_j}{\lambda_j}\Big\langle (1-\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}}))\Lambda W_{\lambda_j} \mid f_{\mathbf{i}}(\iota,\vec\lambda)\Big\rangle\right| \ll \frac{1}{\lambda_j}\Big(\frac{\lambda_j}{\lambda_{j+1}}+\frac{\lambda_{j-1}}{\lambda_j}\Big).
$$

For the third-to-last term, we use that $\underline{\Lambda}\Lambda W \in L^2$ (see (2.8)), (B.11), and (B.2) to deduce that,

$$
\left|\iota_j\frac{\lambda_j'}{\lambda_j}\Big\langle (1-\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}}))\underline{\Lambda}\Lambda W_{\underline{\lambda_j}} \mid \dot g\Big\rangle\right| \ll \frac{1}{\lambda_j}\mathbf{d}(t)^2.
$$

The size of the constant $L > 0$ becomes relevant only in the second-to-last term. Indeed, since $\mathcal{L}\Lambda W = 0$, we have,

$$
\begin{aligned}
&\mathcal{L}_{\lambda_j}(\chi_{L\sqrt{\xi_j\lambda_{j+1}}}\Lambda W_{\underline{\lambda_j}}) \\
&\quad= \frac{1}{L^2\xi_j\lambda_{j+1}}(\Delta\chi)(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}} + \frac{2}{L\sqrt{\xi_j\lambda_{j+1}}}\chi'(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\frac{(r\partial_r\Lambda W)_{\underline{\lambda_j}}}{r}.
\end{aligned}
$$

And therefore, using (B.2) and (B.4) we obtain the estimate,

$$
\left|\Big\langle \mathcal{L}_{\lambda_j}(\chi(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}) \mid g\Big\rangle\right| \lesssim \frac{1}{L}\frac{1}{\lambda_j}\mathbf{d}(t)^2
$$

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 (B.8) for ξ_j' and (B.4) to obtain,

$$
\left|\frac{\xi_j'}{\xi_j}+\frac{\lambda_{j+1}'}{\lambda_{j+1}}\right| \lesssim \frac{1}{\lambda_j}\Big(\left|\xi_j'\right| + \frac{\lambda_j}{\lambda_{j+1}}\left|\lambda_{j+1}'\right|\Big) \ll \frac{1}{\lambda_j}\mathbf{d}(t),
$$

and hence,

$$\Big|\frac{\iota_j}{2}\Big(\frac{\xi_j'}{\xi_j}+\frac{\lambda_{j+1}'}{\lambda_{j+1}}\Big)\big\langle(r\partial_r\chi)(\cdot/L\sqrt{\xi_j\lambda_{j+1}})\Lambda W_{\underline{\lambda_j}}\mid \dot g\big\rangle\Big| \ll \frac{1}{\lambda_j}\mathbf{d}(t).$$

This completes the proof of (B.10). Lastly, we note that the estimates (B.5) and (B.6) follow from exactly the same arguments used to prove (A.7) and (A.8) in Proposition A.1. ◻

We note that Lemma 5.19 and its proof remain valid for $D = 4$.

B.2 Conclusion of the Proof

We have the following analog of Lemma 6.5.

Lemma B.3 *Let $D = 4$. If η_0 is small enough, then there exists $C_0 \ge 0$ depending only on N such that the following is true. If $[t_1, t_2] \subset I_*$ is a finite time interval such that $\mathbf{d}(t) \le \eta_0$ for all $t \in [t_1, t_2]$, then*

$$\sup_{t\in[t_1,t_2]} \lambda_K(t) \le \frac{4}{3}\inf_{t\in[t_1,t_2]} \lambda_K(t), \tag{B.14}$$

$$\int_{t_1}^{t_2} \mathbf{d}(t)\mathrm{d}t \le C_0\big(\mathbf{d}(t_1)\lambda_K(t_1) + \mathbf{d}(t_2)\lambda_K(t_2)\big). \tag{B.15}$$

Due to the fact that some of the estimates in Proposition B.1 hold only under the additional assumption (B.7), we were not able to adapt to the current setting the proof for $D \ge 5$ given above. We will provide a different proof, closer to [41, Section 5].

We introduce below the notion of *ignition condition*. We stress that the definition which follows is meaningful for *any* continuous functions, not necessarily the ones given by the modulation.

Definition B.4 Let I be a time interval, $K \in \mathbb{N}$, $\iota_j \in \{-1, 1\}$ and $\lambda_j, a_j^-, a_j^+ \in C(I)$ for all $1 \le j \le K$. Set

$$\mathcal{S} := \{i : 1 \le i \le K-1 \text{ and } \iota_i = \iota_{i+1}\},$$

$$\mathbf{d}_{\mathrm{par}}(t) := \sqrt{\sum_{i\in\mathcal{S}}\frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{1\le i\le K}\big(a_i^+(t)^2 + a_i^-(t)^2\big)}.$$

We say that $\iota_1, \ldots, \iota_K, \lambda_1, \ldots, \lambda_K, a_1^-, \ldots, a_K^-, a_1^+, \ldots, a_K^+$ satisfy the *ignition condition* with parameters $c_1, c_2, C_2 > 0$ if for any $I = [t_1, t_2]$, $t_0 \in I$ and $j \in \{1, \ldots, K\}$ satisfying at least one of the two pairs of conditions:

- $$\sum_{i\in\mathcal{S}, i<j}\frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=1}^{j-1} a_i^{\pm}(t)^2 \le c_2\mathbf{d}_{\mathrm{par}}(t)^2, \qquad \text{for all } t \in [t_1, t_2], \tag{B.16}$$

and

$$a_j^+(t_0)^2 + a_j^-(t_0)^2 \ge c_1 \mathbf{d}_{\mathrm{par}}(t_0)^2, \tag{B.17}$$

•

$$\sum_{i\in\mathcal{S}, i<j-1} \frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=1}^{j-1} a_i^{\pm}(t)^2 \le c_2 \mathbf{d}_{\mathrm{par}}(t)^2, \qquad \text{for all } t \in [t_1, t_2], \tag{B.18}$$

and

$$\iota_{j-1}\iota_j \frac{\lambda_{j-1}(t_0)}{\lambda_j(t_0)} \ge c_1 \mathbf{d}_{\mathrm{par}}(t_0)^2, \tag{B.19}$$

there is at least one of the bounds:

$$\int_{t_1}^{t_0} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \le C_2 \mathbf{d}_{\mathrm{par}}(t_1)\lambda_K(t_1) \tag{B.20}$$

or

$$\int_{t_0}^{t_2} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \le C_2 \mathbf{d}_{\mathrm{par}}(t_2)\lambda_K(t_2). \tag{B.21}$$

Remark B.5 If the ignition condition is satisfied with given parameters $c_1, c_2, C_2 > 0$, then it is also satisfied with any parameters $(\widetilde{c}_1, \widetilde{c}_2, \widetilde{C}_2)$ such that $\widetilde{c}_1 \ge c_1$, $\widetilde{c}_2 \le c_2$ and $\widetilde{C}_2 \ge C_2$.

Remark B.6 By convention, $\lambda_0(t) = 0$ for all $t \in [t_1, t_2]$, thus (B.19) never holds for $j = 1$. Clearly, (B.19) cannot hold either if $j - 1 \notin \mathcal{S}$.

Lemma B.7 *If $(\iota_1, \ldots, \iota_K, \lambda_1, \ldots, \lambda_K, a_1^-, \ldots, a_K^-, a_1^+, \ldots, a_K^+)$ satisfy the ignition condition with parameters c_1, c_2, C_2 and I is a time interval such that*

$$\frac{\iota_1\iota_2 + 1}{2}\frac{\lambda_1(t)}{\lambda_2(t)} + a_1^-(t)^2 + a_1^+(t)^2 \le \frac{1}{2}\min(1, c_2)\mathbf{d}_{\mathrm{par}}(t)^2 \qquad \textit{for all } t \in I, \tag{B.22}$$

then $(\iota_2, \ldots, \iota_K, \lambda_2, \ldots, \lambda_K, a_2^-, \ldots, a_K^-, a_2^+, \ldots, a_K^+)$ satisfy the ignition condition with parameters $(2c_1, \frac{1}{2}\min(1, c_2), 2C_2)$ on the interval I.

Proof Set

$$\mathbf{d}^\sharp(t) := \sqrt{\frac{\iota_1\iota_2+1}{2}\frac{\lambda_1(t)}{\lambda_2(t)} + a_1^-(t)^2 + a_1^+(t)^2},$$

$$\mathbf{d}^\flat(t) := \sqrt{\sum_{i\in\mathcal{S}\setminus\{1\}}\frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=2}^K\big(a_i^+(t)^2 + a_i^-(t)^2\big)} = \sqrt{\mathbf{d}_{\mathrm{par}}(t)^2 - \mathbf{d}^\sharp(t)^2}.$$

By assumption (B.22), we have $\mathbf{d}^\sharp(t)^2 \le \frac{1}{2}\mathbf{d}_{\mathrm{par}}(t)^2$, which implies $\mathbf{d}_{\mathrm{par}}(t) \le \sqrt{2}\mathbf{d}^\flat(t)$.

We verify the ignition condition for $(\iota_2, \ldots, \iota_K, \lambda_2, \ldots, \lambda_K, a_2^-, \ldots, a_K^-, a_2^+, \ldots, a_K^+)$. If $j \ge 2$ and

$$\sum_{i\in\mathcal{S}, 1<i<j}\frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=2}^{j-1} a_i^\pm(t)^2 \le \frac{1}{2}\min(1, c_2)\mathbf{d}^\flat(t)^2, \qquad \text{for all } t \in [t_1, t_2],$$

then adding $\mathbf{d}^\sharp(t)^2$ to both sides and using (B.22), we get (B.16). Also, $a_j^+(t_0)^2 + a_j^-(t_0)^2 \ge 2c_1\mathbf{d}^\flat(t_0)$ implies (B.17). Since we assume $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=1}^K$ satisfy the ignition condition, we obtain at least one of the bounds (B.20), (B.21). Since $\mathbf{d}^\flat(t) \le \mathbf{d}_{\mathrm{par}}(t) \le \sqrt{2}\mathbf{d}^\flat(t)$, we obtain the same bound with with $\mathbf{d}^\flat$ instead of $\mathbf{d}_{\mathrm{par}}$ and $2C_2$ instead of C_2. The case where bounds (B.18) and (B.19) hold is similar. □

Lemma B.8 *For all $K \in \mathbb{N}$ and functions $c_2, C_2 : (0, c_1^*] \to (0, \infty)$, increasing and decreasing respectively, there exist $c_{1,\min}, C_1 > 0$ such that if for all $c_1 \ge c_{1,\min}$, $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=1}^K$ satisfy the ignition condition with parameters $(c_1, c_2(c_1), C_2(c_1))$ on a time interval $I = [t_1, t_2]$, then*

$$\int_I \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \le C_1(\lambda_K(t_1)\mathbf{d}_{\mathrm{par}}(t_1) + \lambda_K(t_2)\mathbf{d}_{\mathrm{par}}(t_2)).$$

Proof Induction with respect to K.

Step 1. For $K = 1$, we let $c_{1,\min} = \frac{1}{2}$. We will only use the fact that the ignition condition is satisfied with parameters $c_1 = \frac{1}{2}, c_2, C_2$ for some $c_2, C_2 > 0$. The conditions (B.16) and (B.17) hold for $j = 1$, thus for all t_0 we have (B.20) or (B.21). Let

$$A_- := \{t_0 \in I : \text{(B.20) holds}\}, \qquad A_+ := \{t_0 \in I : \text{(B.21) holds}\},$$

so that A_- and A_+ are closed sets, and $I = A_- \cup A_+$. We define

$$t_1^\sharp := \max A_-, \qquad t_2^\sharp := \min A_+.$$

We adopt the convention that $t_1^\sharp = t_1$ if $A_- = \emptyset$, and similarly $t_2^\sharp = t_2$ if $A_+ = \emptyset$. With these conventions, we find that $t_1^\sharp \geq t_2^\sharp$. By the ignition condition, we have

$$\begin{aligned}\int_{t_1}^{t_2} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t &\leq \int_{t_1}^{t_1^\sharp} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t + \int_{t_2^\sharp}^{t_2} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \\ &\leq C_2(c_{1,\min})\big(\mathbf{d}_{\mathrm{par}}(t_1)\lambda_K(t_1) + \mathbf{d}_{\mathrm{par}}(t_2)\lambda_K(t_2)\big),\end{aligned}$$

which settles the base case $K = 1$.

Step 2. We continue with the induction step. Set $c_2^\flat(c_1) := \frac{1}{2}\min(1, c_2(c_1/2))$ and $C_2^\flat(c_1) := 2C_2(c_1/2)$ for all $c_1 \in (0, c_1^*]$. Let $c_{1,\min}^\flat > 0$ be the number given by the induction hypothesis (for $K - 1$ instead of K) for these functions $c_2^\flat$ and $C_2^\flat$. We set $c_3 := \frac{1}{2}c_2^\flat(c_{1,\min}^\flat)$ and $c_{1,\min} := \min\big(\frac{1}{2}c_2^\flat(c_{1,\min}^\flat), c_2(c_3)\big)$.

Assume $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=1}^K$ satisfy the ignition condition for all $c_1 \geq c_{1,\min}$, and let

$$A := \{t \in [t_1, t_2] : a_1^+(t)^2 + a_1^-(t)^2 \geq c_{1,\min}\mathbf{d}_{\mathrm{par}}(t)^2\}.$$

By the ignition condition, (B.20) or (B.21) holds for all $t_0 \in A$, with $C_2 = C_2(c_{1,\min})$. Let

$$A_- := \{t_0 \in A : \text{(B.20) holds}\}, \qquad A_+ := \{t_0 \in A : \text{(B.21) holds}\},$$

so that A_- and A_+ are closed sets, and $A = A_- \cup A_+$. We define

$$t_1^\sharp := \max\ A_-, \qquad t_2^\sharp := \min\ A_+.$$

We adopt the convention that $t_1^\sharp = t_1$ if $A_- = \emptyset$, and similarly $t_2^\sharp = t_2$ if $A_+ = \emptyset$. It is not excluded either that $t_1^\sharp \geq t_2^\sharp$.

By the ignition condition, we have

$$\int_{t_1}^{t_1^\sharp} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t + \int_{t_2^\sharp}^{t_2} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \leq C_2(c_{1,\min})\big(\mathbf{d}_{\mathrm{par}}(t_1)\lambda_K(t_1) + \mathbf{d}_{\mathrm{par}}(t_2)\lambda_K(t_2)\big),$$

so it remains to consider the interval $(t_1^\sharp, t_2^\sharp)$. Notice that $(t_1^\sharp, t_2^\sharp) \subset I \setminus A$.

Step 3. We treat separately the cases $\iota_1\iota_2 = -1$ and $\iota_1\iota_2 = 1$. In the former case, we set $t_1^\flat = t_1^\sharp$, $t_2^\flat = t_2^\sharp$, and go to the next step.

Assume $\iota_1\iota_2 = 1$ and let

$$B := \Big\{t \in [t_1^\sharp, t_2^\sharp] : \frac{\lambda_1(t)}{\lambda_2(t)} \geq c_3\mathbf{d}_{\mathrm{par}}(t)^2\Big\}.$$

We have $c_{1,\min} \leq c_2(c_3)$, thus $a_1^+(t)^2 + a_1^-(t)^2 \leq c_2(c_3)$ for all $t \in [t_1^\sharp, t_2^\sharp]$, so that the ignition condition implies that (B.20) or (B.21) holds for all $t_0 \in A$, with $c_2 = c_2(c_3)$ and $C_2 = C_2(c_3)$. Let

$$B_- := \{t_0 \in B : (\text{B.20}) \text{ holds}\}, \qquad B_+ := \{t_0 \in B : (\text{B.21}) \text{ holds}\},$$

so that B_- and B_+ are closed sets, and $B = B_- \cup B_+$. We define

$$t_1^\flat := \max B_-, \qquad t_2^\flat := \min B_+.$$

We adopt the convention that $t_1^\flat = t_1^\sharp$ if $B_- = \emptyset$, and similarly $t_2^\flat = t_2^\sharp$ if $B_+ = \emptyset$. It is not excluded either that $t_1^\flat \geq t_2^\flat$.

By the ignition condition, we have

$$\int_{t_1^\sharp}^{t_1^\flat} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t + \int_{t_2^\flat}^{t_2^\sharp} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \leq C_2(c_3)\big(\mathbf{d}_{\mathrm{par}}(t_1^\sharp)\lambda_K(t_1^\sharp) + \mathbf{d}_{\mathrm{par}}(t_2^\sharp)\lambda_K(t_2^\sharp)\big),$$

so it remains to consider the interval $(t_1^\flat, t_2^\flat)$. Notice that $(t_1^\flat, t_2^\flat) \subset I \setminus (A \cup B)$.

Step 4. We check that for all $c_1^\flat \geq c_{1,\min}^\flat$, $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=2}^K$ satisfy the ignition condition with parameters $(c_1^\flat, c_2^\flat(c_1^\flat), C_2^\flat(c_1^\flat))$ on the interval $[t_1^\flat, t_2^\flat]$. We have

$$a_1^+(t)^2 + a_1^-(t)^2 + \frac{\iota_1\iota_2 + 1}{2}\frac{\lambda_1(t)}{\lambda_2(t)} \leq (c_{1,\min} + c_3)\mathbf{d}_{\mathrm{par}}(t)^2 \leq c_2^\flat(c_{1,\min}^\flat)\mathbf{d}_{\mathrm{par}}(t)^2. \tag{B.23}$$

By the definition of $c_2^\flat$, we also have

$$c_2^\flat(c_{1,\min}^\flat) \leq \frac{1}{2}\min(1, c_2(c_{1,\min}^\flat/2)) \leq \frac{1}{2}\min(1, c_2(c_1^\flat/2)),$$

where in the last step we used the fact that c_2 is an increasing function. Thus, by Lemma B.7, $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=2}^K$ satisfy the ignition condition with parameters $c_1^\flat$, $\frac{1}{2}\min(1, c_2(c_1^\flat/2)) = c_2^\flat(c_1^\flat)$, $2C_2(c_1^\flat/2) = C_2^\flat(c_1^\flat)$.

By the induction hypothesis, we have

$$\int_{t_1^\flat}^{t_2^\flat} \mathbf{d}^\flat(t)\mathrm{d}t \leq C_1^\flat \sup_{t_1^\flat \leq t \leq t_2^\flat} (\lambda_K(t)\mathbf{d}^\flat(t)),$$

where

$$\mathbf{d}^\flat(t) := \sqrt{\sum_{i \in \mathcal{S}\setminus\{1\}} \frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=2}^K \big(a_i^+(t)^2 + a_i^-(t)^2\big)}.$$

The bound (B.23) and $c_2^\flat(c_{1,\min}^\flat) \leq 1$ imply $\mathbf{d}_{\mathrm{par}}(t) \leq \sqrt{2}\mathbf{d}^\flat(t)$ for all $t \in [t_1^\flat, t_2^\flat]$, and the desired bound follows. □

Next, we prove that the modulation parameters satisfy the ignition condition.

Lemma B.9 *For any $c_1 \in (0, c_1^*]$ there exist $\eta_0, c_2, C_2 > 0$ such that the following is true. Let $\boldsymbol{u}$ be a solution of (1.1) for $D = 4$, let $\mathbf{d}$ be defined by (1.6), I a time interval such that $\mathbf{d}(t) \le \eta_0$ for all $t \in I$, and let $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=1}^K$ be the modulation parameters as in Lemma 5.12. Then $(\iota_j, \lambda_j, a_j^-, a_j^+)_{j=1}^K$ satisfy the ignition condition with parameters (c_1, c_2, C_2) on I.*

Proof Assume first that (B.16) and (B.17) hold. We have $|a_j^+(t_0)| \ge |a_j^-(t_0)|$ or $|a_j^+(t_0)| \le |a_j^-(t_0)|$. We will show that the former implies (B.21), and the latter implies (B.20). Since the two cases are analogous, we only consider the first one. To fix ideas, assume $a_j^+(t_0) > 0$, the case $a_j^+(t_0) < 0$ being analogous, so that

$$a_j^+(t_0) \ge \frac{\sqrt{c_1}}{\sqrt{2}}\mathbf{d}_{\mathrm{par}}(t_0) \quad \Rightarrow \quad \widetilde{a}_j^+(t_0) \ge \frac{2\sqrt{c_1}}{3}\mathbf{d}_{\mathrm{par}}(t_0),$$

where the last inequality follows from (B.5).

Set

$$t_3 := \sup\left\{t \in [t_0, t_2] : \widetilde{a}_j^+(t) \ge \frac{\sqrt{c_1}}{4}\mathbf{d}_{\mathrm{par}}(t) \text{ for all } t \in [t_0, t_3]\right\}.$$

If η_0 is small enough, then (B.6) yields

$$\big(\widetilde{a}_j^+\big)'(t) \ge \frac{\kappa}{2\lambda_j(t)}\widetilde{a}_j^+(t), \qquad \text{for all } t \in [t_0, t_3].$$

Integrating, and using again the inequality defining t_3, we obtain

$$\int_{t_0}^{t_3} \mathbf{d}_{\mathrm{par}}(t)\mathrm{d}t \lesssim \mathbf{d}_{\mathrm{par}}(t_3) \sup_{t\in[t_0,t_3]} \lambda_j(t). \tag{B.24}$$

Since $|\lambda_k'(t)| \lesssim \mathbf{d}(t) \lesssim \mathbf{d}_{\mathrm{par}}(t)$, for all $k \ge j$ we obtain

$$\sup_{t\in[t_0,t_3]} \lambda_k(t) \le (1 + c_4) \inf_{t\in[t_0,t_3]} \lambda_k(t), \tag{B.25}$$

where c_4 can be made arbitrarily small upon adjusting η_0. Similarly, for $k > j$ (B.6) yields

$$\big|\big(\widetilde{a}_k^\pm\big)'(t)\big| \lesssim \mathbf{d}_{\mathrm{par}}(t)/\lambda_k(t),$$

so (B.24) together with (B.25) yield

$$|\widetilde{a}_k^\pm(t_3) - \widetilde{a}_k^\pm(t_0)| \le c_4\mathbf{d}_{\mathrm{par}}(t_3), \qquad \text{for all } k > j.$$

Also, (B.6) yields

$$\max\left(0, \frac{\mathrm{d}}{\mathrm{d}t}|\widetilde{a}_j^-(t)|\right) \lesssim \mathbf{d}_{\mathrm{par}}(t)^2/\lambda_j(t),$$

thus using again (B.24) and (B.25) with $k = j$, we have

$$|\widetilde{a}_j^-(t_3)| - |\widetilde{a}_j^-(t_0)| \leq c_4 \mathbf{d}_{\mathrm{par}}(t_3).$$

Set

$$\mathbf{d}^\sharp(t)^2 := \sum_{i\in\mathcal{S}, i<j} \frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=1}^{j-1} a_i^\pm(t)^2,$$
$$\mathbf{d}^\flat(t)^2 := \sum_{i\in\mathcal{S}, i\geq j} \frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=j+1}^{K} a_i^\pm(t)^2 + a_j^-(t)^2.$$

From the bounds above, we obtain $\mathbf{d}^\flat(t_3)^2 \leq \mathbf{d}^\flat(t_0)^2 + c_4\mathbf{d}_{\mathrm{par}}(t_3)^2$, with c_4 small. By (B.16), we have $\mathbf{d}^\sharp(t_3)^2 \leq c_2 \mathbf{d}_{\mathrm{par}}(t_3)$. Since $\widetilde{a}_j^+$ is increasing on $[t_0, t_3]$, we obtain

$$\begin{aligned}\mathbf{d}_{\mathrm{par}}(t_3)^2 &= \mathbf{d}^\sharp(t_3)^2 + \mathbf{d}^\flat(t_3)^2 + a_j^+(t_3)^2 \leq c_2\mathbf{d}_{\mathrm{par}}(t_3)^2 + \mathbf{d}^\flat(t_0)^2 + c_4\mathbf{d}_{\mathrm{par}}(t_3)^2 + a_j^+(t_3)^2 \\ &\leq (c_2 + c_4)\mathbf{d}_{\mathrm{par}}(t_3)^2 + \big(1 + 9/(4c_1)\big)a_j^+(t_3)^2.\end{aligned}$$

If c_1, c_2 and c_4 are small enough, this implies $a_j^+(t_3) \geq \frac{\sqrt{c_1}}{2}\mathbf{d}_{\mathrm{par}}(t_3)$, thus $t_3 = t_2$ and (B.24) yields (B.21).

Assume now that (B.18) and (B.19) hold. We will prove that $\beta_{j-1}(t_0) \geq 0$ implies (B.21). An analogous argument would show that $\beta_{j-1}(t_0) \leq 0$ implies (B.20).

Set

$$t_3 := \sup\left\{t \in [t_0, t_2] : \xi_{j-1}(t)/\lambda_j(t) \geq \frac{c_1}{4}\mathbf{d}_{\mathrm{par}}(t)^2\right\}.$$

Then, choosing c_2 in (B.18) small enough, (B.10) yields

$$\beta_{j-1}'(t) \geq 8\lambda_j(t)^{-1}, \qquad \text{for all } t \in [t_0, t_3].$$

For $0 < x \ll 1$, set $\Phi(x) := \sqrt{-x\log x}$. 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 $c_3 > 0$ to be determined, consider the auxiliary function

$$\phi(t) := \beta_{j-1}(t) + c_3\Phi\big(\xi_{j-1}(t)/\lambda_j(t)\big).$$

The Chain Rule gives

$$\phi'(t) = \beta'_{j-1}(t) + c_3 \frac{\xi'_{j-1}(t)}{\lambda_j(t)} \Phi'\Big(\frac{\xi_{j-1}(t)}{\lambda_j(t)}\Big).$$

By (B.8), we have $|\xi'_{j-1}(t)| \lesssim (\xi_{j-1}(t)/\lambda_j(t))^{\frac{1}{2}} \log(-\xi_{j-1}(t)/\lambda_j(t))^{-1/2}$, hence we can choose c_3 such that

$$\phi'(t) \geq 4\lambda_j(t)^{-1}, \qquad \text{for all } t \in [t_0, t_3]. \tag{B.26}$$

If we consider $\widetilde{\phi}(t) := \beta_{j-1}(t) + \frac{c_3}{2}\Phi\big(\xi_{j-1}(t)/\lambda_j(t)\big)$ instead of ϕ, then the computation above shows that $\widetilde{\phi}$ is increasing. We have $\widetilde{\phi}(t_0) \geq 0$, so $\widetilde{\phi}(t) \geq 0$ for all $t \in [t_0, t_3]$, implying

$$\mathbf{d}(t) \lesssim \sqrt{\xi_{j-1}(t)/\lambda_j(t)} \lesssim \phi(t)/\sqrt{-\log \phi(t)}. \tag{B.27}$$

The bound (B.26) yields

$$\big(\lambda_j(t)\phi(t)^2/\sqrt{-\log \phi(t)}\big)' \gtrsim \phi(t)/\sqrt{-\log \phi(t)}.$$

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

$$\int_{t_0}^{t_3} \phi(t)/\sqrt{-\log \phi(t)}\mathrm{d}t \lesssim \lambda_j(t_3)\phi(t_3)^2/\sqrt{-\log \phi(t_3)} \lesssim \mathbf{d}(t_3)^2\sqrt{-\log \mathbf{d}(t_3)}\lambda_j(t_3).$$

Thus, (B.27) yields

$$\int_{t_0}^{t_3} \mathbf{d}(t)\mathrm{d}t \lesssim \mathbf{d}(t_3)^2\sqrt{-\log \mathbf{d}(t_3)}\lambda_j(t_3). \tag{B.28}$$

The argument from the first part of the proof yields (B.25), for all $k \geq j$. Also, (B.6) gives $\big|\big(\widetilde{a}_j^{\pm}\big)'(t)\big| \lesssim \lambda_j(t)^{-1}\mathbf{d}(t)$, thus using again (B.27) and (B.25) we get, for all $k \geq j$,

$$|\widetilde{a}_k^{\pm}(t_3) - \widetilde{a}_k^{\pm}(t_0)| \leq c_4 \mathbf{d}_{\mathrm{par}}(t_3)$$

with c_4 small, since the right hand side of (B.28) is $\ll \mathbf{d}(t_3)\lambda_j(t_3)$ if η_0 is small. Finally, (B.9) and $\beta_{j-1}(t) \geq 0$ imply, again using (B.28),

$$\frac{\xi_{j-1}(t_3)}{\lambda_j(t_3)} \geq (1 - c_4)\frac{\xi_{j-1}(t_0)}{\lambda_j(t_0)} \quad \Rightarrow \quad \frac{\lambda_{j-1}(t_3)}{\lambda_j(t_3)} \geq (1 - c_4)\frac{\lambda_{j-1}(t_0)}{\lambda_j(t_0)}, \tag{B.29}$$

where in the last step we use (B.4). Set

$$\mathbf{d}^\sharp(t)^2 := \sum_{i\in\mathcal{S}, i<j-1} \frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=1}^{j-1} a_i^\pm(t)^2,$$
$$\mathbf{d}^\flat(t)^2 := \sum_{i\in\mathcal{S}, i\geq j} \frac{\lambda_i(t)}{\lambda_{i+1}(t)} + \sum_{i=j}^{K} a_i^\pm(t)^2.$$

From the bounds above, we obtain $\mathbf{d}^\flat(t_3)^2 \leq \mathbf{d}^\flat(t_0)^2 + c_4\mathbf{d}_{\mathrm{par}}(t_3)^2$, with c_4 small. By (B.16), we have $\mathbf{d}^\sharp(t_3)^2 \leq c_2\mathbf{d}_{\mathrm{par}}(t_3)$. Applying (B.29), we obtain

$$\begin{aligned}\mathbf{d}_{\mathrm{par}}(t_3)^2 &= \mathbf{d}^\sharp(t_3)^2 + \mathbf{d}^\flat(t_3)^2 + \lambda_{j-1}(t_3)/\lambda_j(t_3)\\ &\leq c_2\mathbf{d}_{\mathrm{par}}(t_3)^2 + \mathbf{d}^\flat(t_0)^2 + c_4\mathbf{d}_{\mathrm{par}}(t_3)^2 + \lambda_{j-1}(t_3)/\lambda_j(t_3)\\ &\leq (c_2+c_4)\mathbf{d}_{\mathrm{par}}(t_3)^2 + \big((1-c_4)^{-1} + c_1^{-1}\big)\lambda_{j-1}(t_3)/\lambda_j(t_3).\end{aligned}$$

If c_1, c_2 and c_4 are small enough, this implies $\xi_{j-1}(t_3)/\lambda_j(t_3) \geq \frac{c_1}{2}\mathbf{d}_{\mathrm{par}}(t_3)^2$, thus $t_3 = t_2$ and (B.28) yields (B.21). □

Proof of Lemma B.3 It suffices to prove (B.15), and (B.14) will follow by the same argument as in the proof of Lemma 6.5.

We claim that there exist inreasing functions $c_2, \eta_0 : (0, c_1^*] \to (0, \infty)$, and a decreasing function $C_2 : (0, c_1^*] \to (0, \infty)$, such that for all c_1 the modulation parameters satisfy the ignition condition with parameters $(c_1, c_2(c_1), C_2(c_1))$ on any time interval on which $\mathbf{d}(t) \leq \eta_0(c_1)$. Indeed, take a strictly decreasing sequence of values $c_1^{(n)}$ converging to zero, and let

$$c_2^{(n)} := c_2(c_1^{(n)}), \quad \eta_0^{(n)} := \eta_0(c_1^{(n)}), \quad C_2^{(n)} := C_2(c_1^{(n)})$$

by Lemma B.9. By Remark B.5, one can always decrease $\eta_0^{(n)}$ and $c_2^{(n)}$, and increase $C_2^{(n)}$, so we can assume that $\eta_0^{(n)}$ and $c_2^{(n)}$ are decreasing sequences, and $C_2^{(n)}$ is an increasing sequence. We now set

$$c_2(c_1) := c_2(c_1^{(n)}),\ \eta_0(c_1) := \eta_0(c_1^{(n)}),\ C_2(c_1) := C_2(c_1^{(n)}), \qquad \text{for all } c_1 \in (c_1^{(n+1)}, c_1^{(n)}].$$

The conclusion follows from Lemma B.8 and the fact that $\mathbf{d}_{\mathrm{par}}(t) \simeq \mathbf{d}(t)$ for all $t \in I$, see (B.3). □

References

1. Bahouri, H., Chemin, J.-Y., Danchin, R.: Fourier Analysis and Nonlinear Partial Differential Equations, volume 343 of Grundlehren der mathematischen Wissenschaften. Springer (2011)

2. Bahouri, H., Gérard, P.: High frequency approximation of solutions to critical nonlinear wave equations. Amer. J. Math. **121**, 131–175 (1999)
3. Bahouri, H., Shatah, J.: Decay estimates for the critical semilinear wave equation. Ann. Inst. H. Poincaré Anal. Non Linéaire **15**(6), 783–789 (1998)
4. Beceanu, M.: A center-stable manifold for the energy-critical wave equation in $\mathbb{R}^3$ in the symmetric setting. J. Hyperbolic Differ. Equ. **11**(3), 437–476 (2014)
5. Bizoń, P., Chmaj, T., Tabor, Z.: On blowup for semilinear wave equations with a focusing nonlinearity. Nonlinearity **17**(6), 2187–2201 (2004)
6. Brezis, H., Coron, J.M.: Convergence of solutions of H-systems or how to blow bubbles. Arch. Ration. Mech. Anal. **89**, 21–56 (1985)
7. Bulut, A., Czubak, M., Li, D., Pavlović, N., Zhang, X.: Stability and unconditional uniqueness of solutions for energy critical wave equations in high dimensions. Comm. Part. Diff. Eq. **38**(4), 575–607 (2013)
8. Chow, S.-N., Hale, J.K.: Methods of Bifurcation Theory, volume 251 of Grundlehren der mathematischen Wissenschaften. Springer, (1982)
9. Christodoulou, D., Tahvildar-Zadeh, A.S.: On the regularity of spherically symmetric wave maps. Comm. Pure Appl. Math. **46**(7), 1041–1091 (1993)
10. Collot, C., Duyckaerts, T., Kenig, C., Merle, F.: Soliton resolution for the radial quadratic wave equation in six space dimensions. arXiv E-Prints, arXiv:2201.01848 (2022)
11. Côte, R.: On the soliton resolution for equivariant wave maps to the sphere. Comm. Pure Appl. Math. **68**(11), 1946–2004 (2015)
12. Cote, R., Kenig, C., Lawrie, A., Schlag, W.: Profiles for the radial focusing 4d energy-critical wave equation. Comm. Math. Phys. **357**(3), 934–1008 (2018)
13. Côte, R., Kenig, C., Schlag, W.: Energy partition for the linear radial wave equation. Math. Ann. **358**(3–4), 573–607 (2014)
14. Côte, R., Kenig, C.E., Lawrie, A., Schlag, W.: Characterization of large energy solutions of the equivariant wave map problem: I (erratum). arXiv E-Prints arXiv:1209.3682, (2019)
15. Donninger, R.: Strichartz estimates in similarity coordinates and stable blowup for the critical wave equation. Duke Math. J. **166**(9), 1627–1683 (2017)
16. Duyckaerts, T., Jia, H., Kenig, C., Merle, F.: Soliton resolution along a sequence of times for the focusing energy critical wave equation. Geom. Funct. Anal. **27**(4), 798–862 (2017)
17. Duyckaerts, T., Jia, H., Kenig, C., Merle, F.: 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. Duyckaerts, T., Kenig, C., Martel, Y., Merle, F.: Soliton resolution for critical co-rotational wave maps and radial cubic wave equation. Comm. Math. Phys. **391**, 779–871 (2022)
19. Duyckaerts, T., Kenig, C., Merle, F.: Erratum: Profiles of bounded radial solutions of the focusing, energy-critical wave equation. arXiv e-prints, arXiv:1201.4986 01 (2015)
20. Duyckaerts, T., Kenig, C., Merle, F.: Scattering profile for global solutions of the energy-critical wave equation. J. Eur. Math. Soc. **21**(7), 2117–2162 (2019)
21. Duyckaerts, T., Kenig, C.E., Merle, F.: 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. Duyckaerts, T., Kenig, C.E., Merle, F.: Profiles of bounded radial solutions of the focusing, energy-critical wave equation. Geom. Funct. Anal. **22**(3), 639–698 (2012)
23. Duyckaerts, T., Kenig, C.E., Merle, F.: Classification of the radial solutions of the focusing, energy-critical wave equation. Camb. J. Math. **1**(1), 75–144 (2013)
24. Duyckaerts, T., Kenig, C.E., Merle, F.: Exterior energy bounds for the critical wave equation close to the ground state. Comm. Math. Phys. **379**, 1113–1175 (2020)
25. Duyckaerts, T., Kenig, C.E., Merle, F.: Decay estimates for nonradiative solutions of the energy-critical focusing wave equation. J. Geom. Anal., (2021)
26. Duyckaerts, T., Kenig, C.E., Merle, F.: Soliton resolution for the radial critical wave equation in all odd space dimensions. Acta Math., to appear
27. Duyckaerts, T., Merle, F.: Dynamics of threshold solutions for energy-critical wave equation. Int. Math. Res. Pap, IMRP (2008)
28. Duyckaerts, T., Merle, F.: Dynamic of threshold solutions for energy-critical NLS. Geom. Funct. Anal. **18**(6), 1787–1840 (2009)
29. Eckhaus, P., Schuur, W.: The emergence of solitons of the korteweg-de vries equation from arbitrary initial conditions. Math. Methods Appl. Sci. **5**(1), 97–116 (1983)

30. Fermi, E., Pasta, J., Ulam, S.: S. Los Alamos Report LA-1940, (1955)
31. Gérard, P.: Description du défaut de compacité de l'injection de Sobolev. ESAIM Control Optim. Calc. Var. **3**, 213–233 (1998)
32. Ginibre, J., Velo, G.: Scattering theory in the energy space for a class of nonlinear wave equations. Comm. Math. Phys. **123**(4), 535–573 (1989)
33. Ginibre, J., Velo, G.: Generalized Strichartz inequalities for the wave equation. J. Funct. Anal. **133**, 50–68 (1995)
34. Grillakis, M.: Regularity and asymptotic behavior of the wave equation with a critical nonlinearity. Ann. of Math. **2**(132), 485–509 (1990)
35. Hillairet, M., Raphaël, P.: Smooth type II blow up solutions to the four dimensional energy critical wave equation. Anal. PDE **5**(4), 777–829 (2012)
36. Jendrej, J.: Construction of two-bubble solutions for the energy-critical NLS. Anal. PDE **10**(8), 1923–1959 (2017)
37. Jendrej, J.: Construction of two-bubble solutions for energy-critical wave equations. Amer. J. Math. **141**(1), 55–118 (2019)
38. Jendrej, J., Lawrie, A.: An asymptotic expansion of two-bubble wave maps in high equivariant classes. Anal. PDE **15**(2), 327–403 (2022)
39. Jendrej, J., Lawrie, A.: Uniqueness of two-bubble wave maps in high equivariance classes. Comm. Pure Appl. Math. **76**(8), 1608–1656 (2023)
40. Jendrej, J., Lawrie, A.: Two-bubble dynamics for threshold solutions to the wave maps equation. Invent. Math. **213**(3), 1249–1325 (2018)
41. Jendrej, J., Lawrie, A.: Soliton resolution for energy–critical wave maps in the equivariant case. to appear in J. Amer. Math. Soc., https://doi.org/10.1090/jams/1012
42. Jia, H., Kenig, C.: Asymptotic decomposition for semilinear wave and equivariant wave map equations. Amer. J. Math. **139**(6), 1521–1603 (2017)
43. Jörgens, K.: Das Anfangswertproblem im Grossen für eine Klasse nichtlinearer Wellengleichungen. Math. Z. **77**, 295–308 (1961)
44. Kapitanski, L.V.: Global and unique weak solutions of nonlinear wave equations. Math. Res. Lett. **1**, 211–223 (1994)
45. Kenig, C., Merle, F.: Global well-posedness, scattering and blow-up for the energy-critical focusing non-linear wave equation. Acta Math. **201**(2), 147–212 (2008)
46. Krieger, J., Nakanishi, K., Schlag, W.: Global dynamics away from the ground state for the energy-critical nonlinear wave equation. Amer. J. Math. **135**(4), 935–965 (2013)
47. Krieger, J., Nakanishi, K., Schlag, W.: Center-stable manifold of the ground state in the energy space for the critical wave equation. Math. Ann. **361**(1–2), 1–50 (2015)
48. Krieger, J., Schlag, W.: On the focusing critical semi-linear wave equation. Amer. J. Math. **129**(3), 843–913 (2007)
49. Krieger, J., Schlag, W.: Full range of blow up exponents for the quintic wave equation in three dimensions. J. Math Pures Appl. **101**(6), 873–900 (2014)
50. Krieger, J., Schlag, W., Tataru, D.: Slow blow-up solutions for the $H^1(\mathbb{R}^3)$ critical focusing semilinear wave equation. Duke Math. J. **147**(1), 1–53 (2009)
51. Levine, H.A.: Instability and nonexistence of global solutions to nonlinear wave equations of the form $Pu_{tt} = -Au + \mathcal{F}(u)$. Trans. Amer. Math. Soc. **192**, 1–21 (1974)
52. Lindblad, H., Sogge, C.D.: On existence and scattering with minimal regularity for semilinear wave equations. J. Funct. Anal. **130**(2), 357–426 (1995)
53. Lions, P.-L.: 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)
54. Lions, P.-L.: The concentration-compactness principle in the Calculus of Variations. The limit case, part II. Rev. Mat. Iberoam. **1**, 145–201 (1985)
55. Martel, Y., Merle, F.: Description of two soliton collision for the quartic gKdV equation. Ann. of Math. (2) **174**(2), 757–857 (2011)
56. Martel, Y., Merle, F.: Inelastic interaction of nearly equal solitons for the quartic gKdV equation. Invent. Math. **183**(3), 563–648 (2011)
57. Martel, Y., Merle, F.: Inelasticity of soliton collisions for the 5D energy critical wave equation. Invent. Math. **214**(3), 1267–1363 (2018)
58. Merle, F., Vega, L.: 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)

59. Morawetz, C.S., Strauss, W.A.: Decay and scattering of solutions of a nonlinear relativistic wave equation. Comm. Pure Appl. Math. **25**(1), 1–31 (1972)
60. Morawetz, C.S.: Time decay for the nonlinear Klein-Gordon equations. Proc. Roy. Soc. London Ser. A **306**, 291–296 (1968)
61. Muscalu, C., Schlag, W.: Classical and Multilinear Harmonic Analysis, vol. I. Cambridge University Press (2013)
62. Nakanishi, K., Schlag, W.: Global dynamics above the ground state energy for the focusing nonlinear Klein-Gordon equation. J. Differential Equations **250**(5), 2299–2333 (2011)
63. Nakanishi, K., Schlag, W.: 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)
64. Pecher, H.: Ein nichtlinearer Interpolationssatz und seine Anwendung auf nichtlineare Wellengleichungen. Math. Z. **161**(1), 9–40 (1978)
65. Pecher, H.: Scattering for semilinear wave equations with small data in three space dimensions. Math. Z. **198**(2), 277–289 (1988)
66. Qing, J.: On singularities of the heat flow for harmonic maps from surfaces into spheres. Comm. Anal. Geom., 3(297–316), (1995)
67. Qing, J., Tian, G.: Bubbling of the heat flows for harmonic maps from surfaces. Comm. Pure Appl. Math. **50**(4), 295–310 (1997)
68. Raphaël, P., Szeftel, J.: 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)
69. Rauch. J.: I. The u^5 Klein-Gordon equation. II. Anomalous singularities for semilinear wave equations. In Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. I (Paris, 1978/1979), number 53 in Res. Notes in Math., pages 335–364. Pitman, Boston, Mass.-London, (1981)
70. Rodnianski, I., Sterbenz, J.: On the formation of singularities in the critical $O(3)$ σ-model. Ann. of Math. **2**(172), 187–242 (2010)
71. Rodriguez, C.: Profiles for the radial focusing energy-critical wave equation in odd dimensions. Adv. Differential Equations **21**(5/6), 505–570 (2016)
72. Schuur, P.C.: Asymptotic analysis of soliton problems, an inverse scattering approach. Lecture Notes in Math- ematics. Springer-Verlag, Berlin (1986)
73. Segal, I.E.: The global Cauchy problem for a relativistic scalar field with power interaction. Bull. Soc. Math. France **91**, 129–135 (1963)
74. Segal, I.E.: Non-linear semi-groups. Ann. of Math. **2**(78), 339–364 (1963)
75. Segur, H., Ablowitz, M.J.: Asymptotic solutions and conservation laws for the nonlinear Schrödinger equation I. J. Math. Phys. **17**(5), 710–713 (1976)
76. Shatah, J., Struwe, M.: Regularity results for nonlinear wave equations. Ann. of Math. (2) **138**(3), 503–518 (1993)
77. Shatah, J., Struwe, M.: Well-posedness in the energy space for semilinear wave equations with critical growth. Int. Math. Res. Not. IMRN, (7), (1994)
78. Shatah, J., Struwe, M.: Geometric wave equations. Courant Lecture notes in Mathematics, New York University, Courant Institute of Mathematical Sciences, New York. American Mathematical Society, Providence RI, (1998)
79. Shatah, J., Tahvildar-Zadeh, A.: Regularity of harmonic maps from the Minkowski space into rotationally symmetric manifolds. Comm. Pure Appl. Math. **45**(8), 947–971 (1992)
80. Sogge, C.D.: Lectures on non-linear wave equations, second edition International Press, Boston, MA (2008)
81. Strauss, W.A.: Nonlinear wave equations, volume 73 of CBMS Regional Conference Series in Mathematics. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, (1989)
82. Strauss, W.A.: Decay and asymptotics for $\Box u = F(u)$. J. Funct. Anal. **2**, 409–457 (1968)
83. Struwe, M.: On the evolution of harmonic mappings of riemannian surfaces. Comment. Math. Helv. **60**(4), 558–581 (1985)
84. Struwe, M.: Globally regular solutions to the u^5 Klein-Gordon equation. Ann. Scuola Norm. Pisa Cl. Sci. **15**, 495–513 (1988)
85. Tao, T.: A (concentration-)compact attractor for high-dimensional non-linear Schrödinger equations. Dyn. Partial Differ. Equ. **4**(1), 1–53 (2007)

86. Topping, P.: Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow. Ann. of Math. (2) **159**(2), 465–534 (2004)
87. Topping, P.M.: Rigidity in the harmonic map heat flow. J. Diff. Geom. **45**(3), 593–610 (1997)
88. Zabusky, N.J., Kruskal, M.D.: Interaction of "solitons" in a collisionless plasma and the recurrence of initial states. Phys. Rev. Lett. **15**(6), 240 (1965)