CONTINUOUS-IN-TIME BUBBLING AND PROGRESS TOWARDS SOLITON RESOLUTION CONJECTURE FOR THE ENERGY-CRITICAL NONLINEAR HEAT FLOW

SHREY ARYAN

ABSTRACT. We show that any finite energy solution of the energy-critical nonlinear heat flow in dimensions $d \geq 3$ asymptotically resolves into a sum of possibly time-dependent solitons, a radiation term, and an error term that vanishes in the energy space. As a consequence, when the initial data has finite energy and is non-negative, we settle the Soliton Resolution Conjecture for all dimensions $d \geq 3$.

1. INTRODUCTION

1.1. **Problem Setting.** In this work, we study the long-term behavior of solutions to the energy-critical nonlinear heat flow in dimensions $d \geq 3$:

$$\partial_t u = \Delta u + |u|^{p-1} u \tag{1.1}$$
$$u(0, x) = u_0(x) \in \dot{H}^1(\mathbb{R}^d),$$

where $p := \frac{d+2}{d-2}$. This model arises as the negative gradient flow of the following nonlinear energy functional:

$$E(u) := \frac{1}{2} \int_{\mathbb{R}^d} |\nabla u|^2 \mathrm{d}x - \frac{1}{p+1} \int_{\mathbb{R}^d} |u(x)|^{p+1} \mathrm{d}x, \tag{1.2}$$

which appears naturally in the study of extremizers of the Sobolev inequality and, more generally, is connected to the Yamabe problem on the sphere via stereographic projection. The local well-posedness of (1.1) in $\dot{H}^1$-norm is classical and was initiated by Weissler in [Wei79, Wei80], with further contributions by Giga [Gig86], Ni–Sacks [NS85], and Brezis-Cazenave [BC96]. Observe that the solutions of (1.1) are invariant under translations and parabolic scaling,

$$u(t, x) \mapsto u_\lambda(t, x) := \lambda^{-\frac{d-2}{2}} u\left(t/\lambda^2, x/\lambda\right), \quad \lambda > 0.$$

Since the nonlinear energy is invariant under these symmetries, i.e., $E(u) = E(u_\lambda)$, the equation (1.1) is energy-critical. Testing (1.1) against $\partial_t u$ and integrating by parts we observe the formal energy identity

$$E(u(T)) + \int_0^T \|\partial_t u\|_{L^2}^2 \, \mathrm{d}t = E(u(0)), \tag{1.3}$$

for each $T > 0$. In particular, this implies that the nonlinear energy is non-increasing along the flow. Any function $W : \mathbb{R}^d \to \mathbb{R}$ solving the elliptic PDE

$$\Delta W + |W|^{p-1} W = 0 \tag{1.4}$$

is a stationary solution and will often be referred to as a bubble or soliton of (1.1).

1.2. **Statement of the Main result.** To state our main Theorem 1.6, we first define a notion of *scale* and *center* of a non-zero stationary solution. Let $S > 0$ in $\mathbb{R}^d$ be the sharp constant for the Sobolev inequality,

$$\|u\|_{L^{p+1}} \leq S\|\nabla u\|_{L^2}$$

for all $u \in \dot{H}^1$. Then observe that for any non-zero stationary solution $W \in \dot{H}^1$ we have

$$\int_{\mathbb{R}^d} |\nabla W|^2 \mathrm{d}x = \int_{\mathbb{R}^d} |W|^{p+1} \mathrm{d}x.$$

By the variational characterization of the Sobolev inequality, the best constant (or equality) is attained by a positive stationary solution W', i.e.,

$$\int_{\mathbb{R}^d} |W'|^{p+1} \mathrm{d}x = S^{p+1} \left(\int_{\mathbb{R}^d} |\nabla W'|^2 \mathrm{d}x \right)^{(p+1)/2} = \int_{\mathbb{R}^d} |\nabla W'|^2 \mathrm{d}x$$

implying that for any positive stationary solution we have

$$\int_{\mathbb{R}^d} |\nabla W'|^2 \mathrm{d}x = S^{-d}.$$

Therefore, for any sign-changing stationary solution, we have

$$\int_{\mathbb{R}^d} |\nabla W|^2 \mathrm{d}x = \int_{\mathbb{R}^d} |\nabla W^+|^2 \mathrm{d}x + \int_{\mathbb{R}^d} |\nabla W^-|^2 \mathrm{d}x > 2S^{-d}.$$

In particular, we deduce that any non-zero stationary solution $W \in \dot{H}^1$ satisfies

$$\|\nabla W\|_{L^2}^2 \geq S^{-d}.$$

Denote $\bar{E}_* := S^{-d}$ as the minimal energy of any non-zero stationary solution of (1.4), and in general, let $\bar{E}(u) := \|\nabla u\|_{L^2}^2$ for any $u \in \dot{H}^1$. Thus, given any non-zero stationary solution $W : \mathbb{R}^d \to \mathbb{R}$, we define its scale and center as follows:

Definition 1.1 (Scale of a stationary solution)**.** Let $\gamma_0 \in (0, \bar{E}_*/2)$. Then the scale associated to a non-trivial stationary solution W, denoted by $\lambda(W; \gamma_0)$, is defined by

$$\lambda(W; \gamma_0) := \inf\{\lambda \in (0, \infty) \mid \exists\, a \in \mathbb{R}^d \text{ such that } \bar{E}(W; B(a, \lambda)) \geq \bar{E}(W) - \gamma_0\}.$$

Definition 1.2 (Center of a stationary solution)**.** Let $\gamma_0 \in (0, \bar{E}_*/2)$ and let $\lambda(W; \gamma_0)$ be the scale of a non-zero stationary solution W. Then the center, denoted by $a(W; \gamma_0) \in \mathbb{R}^d$, is defined as

$$\bar{E}(W; B(a(W; \gamma_0), \lambda(W; \gamma_0))) \geq \bar{E}(W) - \gamma_0.$$

These quantities are well-defined as we will later prove in Lemma 2.1. Since our main result says that finite energy solutions of (1.1) eventually approach a sum of stationary solutions, it will be convenient to define their sum, which we will often refer to as a multi-bubble configuration.

Definition 1.3 (Multi-bubble configuration)**.** Let $K \in \{0, 1, 2, \dots\}$. A K-multi-bubble configuration is the sum

$$\mathbf{W}(x) = \sum_{j=1}^{K} W_j(x),$$

where $W_j : \mathbb{R}^d \to \mathbb{R}$ are smooth non-zero stationary solution. By convention if $K = 0$ then $\mathbf{W} \equiv 0$. To emphasize the dependence of $\mathbf{W}$ on the collection $\{W_j\}_{j=1}^K$, we will occasionally write $\mathbf{W} = \mathbf{W}(\vec{W})$, where $\vec{W} = (W_1, \dots, W_K)$.

Next, we quantify the distance of a function to some multi-bubble configuration.

Definition 1.4 (Localized distance to a multi-bubble configuration)**.** Given,

(1) some scales $\xi, \rho, \nu \in (0, \infty)$, such that $\xi \le \rho \le \nu$;
(2) a map $u : [0, T_+) \times B(y, \nu) \to \mathbb{R}$, where $T_+ > 0$ and $\gamma_0 \in (0, \bar{E}_*/2)$;
(3) a non-negative integer $K \in \mathbb{N}$ and non-zero stationary solutions $W_1, \ldots, W_K$ with centers $a(W_j; \gamma_0) \in B(y, \xi)$ and scales $\lambda(W_j; \gamma_0) \in (0, \infty)$ for each $j \in \{1, \ldots, K\}$;
(4) collection of radii $\vec{\nu} = (\nu, \nu_1, \ldots, \nu_K) \in (0, \infty)^{K+1}$ such that $B(a(W_j), \nu_j) \subset B(y, \xi)$ and smaller scales $\vec{\xi} = (\xi, \xi_1, \ldots, \xi_K) \in (0, \infty)^{K+1}$ such that $\xi_j < \lambda(W_j; \gamma_0)$ for each $j \in \{1, \ldots, K\}$.

Then the localized distance is defined as

$$\begin{aligned}\mathbf{d}_{\gamma_0}(u(t), \mathbf{W}; B(y, \rho); \vec{\nu}, \vec{\xi}) := \bar{E}\big(u - \mathbf{W}(\vec{W})); B(y, \rho)\big) + \bar{E}(u; B(y, \nu) \setminus B(y, \xi))\\ + \frac{\xi}{\rho} + \frac{\rho}{\nu} + \sum_{j \neq k} \left(\frac{\lambda(W_j)}{\lambda(W_k)} + \frac{\lambda(W_k)}{\lambda(W_j)} + \frac{|a(W_j) - a(W_k)|}{\lambda(W_j)} \right)^{-(d-2)/2}\\ + \sum_j \Big(\frac{\lambda(W_j)}{\operatorname{dist}(a(W_j), \partial B(y, \xi))} + \frac{\lambda(W_j)}{\nu_j} + \frac{\xi_j}{\lambda(W_j)} \Big)\\ + \sum_j \sum_{k \in \mathcal{I}_j} \frac{\xi_j}{\operatorname{dist}(a(W_k), \partial B(a(W_j), \nu_j))}.\end{aligned}$$

Minimizing over all the parameters in the above definition yields,

Definition 1.5 (Localized multi-bubble proximity function)**.** Given, $y \in \mathbb{R}^d$, $\rho \in (0, \infty)$, $u : [0, T_+) \times B(y, \rho) \to \mathbb{R}$, where $T_+ > 0$ and $\gamma_0 \in (0, \bar{E}_*/2)$, define

$$\boldsymbol{\delta}_{\gamma_0}(u(t); B(y, \rho)) := \inf_{\mathbf{W}, \vec{\nu}, \vec{\xi}} \mathbf{d}_{\gamma_0}(u, \mathbf{W}; B(y, \rho); \vec{\nu}, \vec{\xi})$$

where the infimum above is taken over all possible K-multi bubble configurations for any non-negative integer K, over all parameters $\vec{\nu} \in (0, \infty)^{K+1}$ and $\vec{\xi} \in (0, \infty)^{K+1}$ as in Definition 1.4. Since we will fix γ_0 later, we drop the subscript involving γ_0 in subsequent expressions.

With these definitions in hand, we state the main theorem in this paper.

Theorem 1.6 (Continuous Bubbling for NLH)**.** *Let $u(t)$ be a solution of* (1.1) *with initial data $u_0 \in \dot{H}^1$. Let $T_+ = T_+(u_0) \in (0, \infty]$ denote its maximal time of existence and assume that $u(t)$ has finite energy, i.e., $\sup_{t \in [0, T_+)} \bar{E}(u(t)) < \infty$. Then the following hold*

(i) *If $T_+ < \infty$, then there exists a finite energy map $u^* : \mathbb{R}^d \to \mathbb{R}$, an integer $K \ge 1$, and points $\{x^i\}_{i=1}^K \subset \mathbb{R}^d$ such that the following holds: let $t_n \to T_+$ be any time sequence. After passing to a subsequence (still denoted by t_n) we can associate to each $i \in \{1, \ldots, K\}$ an integer J_i, sequences $a^i_{j,n} \in \mathbb{R}^d$ and $\lambda^i_{j,n} \in (0, \infty)$ for each $j \in \{1, \ldots, J_i\}$, with $a^i_{j,n} \to x^i$, $\frac{\lambda^i_{j,n}}{\sqrt{T_+ - t_n}} \to 0$ as $n \to \infty$, and non-zero bubbles $W^i_1, \ldots, W^i_{J_i}$ such that*

$$\lim_{n \to \infty} \left(\frac{\lambda^i_{j,n}}{\lambda^i_{k,n}} + \frac{\lambda^i_{k,n}}{\lambda^i_{j,n}} + \frac{|a^i_{j,n} - a^i_{k,n}|}{\lambda^i_{j,n}} \right) = \infty \quad \textit{for all } j \neq k, \tag{1.5}$$

and

$$u(t_n) = u^* + \sum_{i=1}^K \sum_{j=1}^{J_i} W^i_j[a^i_{j,n}, \lambda^i_{j,n}] + o_{\dot{H}^1}(1), \tag{1.6}$$

where $W_j^i[a_{j,n}^i, \lambda_{j,n}^i](x) = (\lambda_{j,n}^i)^{-(d-2)/2} W_j((x - a_{j,n}^i)/\lambda_{j,n}^i)$ *and the error term* $o_{\dot{H}^1}(1) \to 0$ *strongly in* $\dot{H}^1$.

(ii) *If* $T_+ = \infty$, *then let* $t_n \to \infty$ *be any time sequence. After passing to a subsequence we can find an integer* $K \geq 0$, *sequences* $a_{j,n} \in \mathbb{R}^d$ *and* $\lambda_{j,n} \in (0, \infty)$ *for each* $j \in \{1, \ldots, K\}$, *with*

$$\lim_{n\to\infty} \frac{|a_{j,n}| + \lambda_{j,n}}{\sqrt{t_n}} = 0 \tag{1.7}$$

and non-zero bubbles $W_1, \ldots, W_K$, *so that*

$$\lim_{n\to\infty} \left(\frac{\lambda_{j,n}}{\lambda_{k,n}} + \frac{\lambda_{k,n}}{\lambda_{j,n}} + \frac{|a_{j,n} - a_{k,n}|}{\lambda_{j,n}} \right) = \infty \quad \textit{for all } j \neq k,$$

and

$$u(t_n) = \sum_{j=1}^{K} W_j[a_{j,n}, \lambda_{j,n}] + o_{\dot{H}^1}(1) \tag{1.8}$$

where $W_j[a_{j,n}, \lambda_{j,n}](x) = (\lambda_{j,n})^{-(d-2)/2} W_j((x - a_{j,n})/\lambda_{j,n})$ *and the error term* $o_{\dot{H}^1}(1) \to 0$ *strongly in* $\dot{H}^1$.

Note that the bubbles obtained in the above decomposition may depend on the sequence of times, which is similar to an issue encountered in [JLS25]; however, we can resolve this issue for (1.1) with a very reasonable assumption.

Corollary 1.7. *Let* $u(t)$ *be a solution of* (1.1) *with non-negative initial data* $u_0 \geq 0$ *and* $u_0 \in \dot{H}^1$. *Let* $T_+ = T_+(u_0) \in (0, \infty]$ *denote its maximal time of existence and assume that* $u(t)$ *has finite energy, i.e.,* $\sup_{t\in[0,T_+)} \bar{E}(u(t)) < \infty$. *Then the maps obtained in the decompositions* (1.6) *and* (1.8) *are unique and independent of the sequence of times.*

Theorem 1.6 is a consequence of the following localized bubbling result, in which we denote the ball centered at $x \in \mathbb{R}^d$ with radius $r > 0$ as $B(x, r) := \{y \in \mathbb{R}^d : |x - y| < r\}$.

Theorem 1.8 (Localized Bubbling for NLH)**.** *Let* $u(t)$ *be a solution of* (1.1) *with initial data* $u_0 \in \dot{H}^1$. *Let* $T_+ = T_+(u_0) \in (0, \infty]$ *denote its maximal time of existence and assume that* $u(t)$ *has finite energy, i.e.,* $\sup_{t\in[0,T_+)} \bar{E}(u(t)) < \infty$. *Then there exists* $\gamma_0 = \gamma_0(\sup_{t\in[0,T_+)} \bar{E}(u(t))) > 0$ *such that the following holds:*
(i) *If* $T_+ < \infty$, *then for any* $y \in \mathbb{R}^d$,

$$\lim_{t\to T_+} \boldsymbol{\delta}_{\gamma_0}\big(u(t); B(y, \sqrt{T_+ - t})\big) = 0.$$

Moreover, let $t_n \to T_+$ *be any sequence and let* $B(y_n, \rho_n)$ *be any sequence of balls such that* $B(y_n, R_n\rho_n) \subset B(y, \sqrt{T_+ - t_n})$ *for some sequence* $R_n \to \infty$. *Suppose* α_n, β_n *are sequences with* $\alpha_n \to 0$, $\beta_n \to \infty$, $\lim_{n\to\infty} \beta_n R_n^{-1} = 0$, *and*

$$\lim_{n\to\infty} \bar{E}\big(u(t_n); B(y_n, \beta_n\rho_n) \setminus B(y_n, \alpha_n\rho_n)\big) = 0.$$

Then,

$$\lim_{n\to\infty} \boldsymbol{\delta}_{\gamma_0}\big(u(t_n); B(y_n, \rho_n)\big) = 0.$$

(ii) *If* $T_+ = \infty$, *then for every* $y \in \mathbb{R}^d$,

$$\lim_{t\to\infty} \boldsymbol{\delta}_{\gamma_0}\big(u(t); B(y, \sqrt{t})\big) = 0.$$

Moreover, let $t_n \to \infty$ *be any sequence and let* $B(y_n, \nu_n)$ *be any sequence of balls such that* $B(y_n, R_n\nu_n) \subset B(y, \sqrt{t_n})$ *for some sequence* $R_n \to \infty$. *Suppose* α_n, β_n *are sequences with* $\alpha_n \to 0$, $\beta_n \to \infty$, $\lim_{n\to\infty} \beta_n R_n^{-1} = 0$ *and*

$$\lim_{n\to\infty} \bar{E}\big(u(t_n); B(y_n, \beta_n\rho_n) \setminus B(y_n, \alpha_n\rho_n)\big) = 0.$$

Then,

$$\lim_{n\to\infty} \boldsymbol{\delta}_{\gamma_0}\big(u(t_n); B(y_n, \rho_n)\big) = 0.$$

1.3. **Background and Motivation.** A fundamental problem in the analysis of nonlinear partial differential equations (PDEs) is describing the long-time behavior of their solutions. The Soliton Resolution Conjecture asserts that any finite-energy solution to a dispersive PDE asymptotically decomposes into a sum of decoupled solitons that are stationary solutions of the underlying equation, a radiation term that behaves like a solution to the linear flow, and an error term that vanishes in the natural energy norm. This conjecture arose from the numerical experiments of Fermi–Pasta–Ulam–Tsingou [FPU55] and Zabusky–Kruskal [ZK65], which provided evidence that it holds for the Korteweg-de Vries (KdV) equation. Since then, the problem has been extensively studied for the KdV equation as well as for several other integrable models.

Beyond integrable systems, analogues of the Soliton Resolution Conjecture have emerged across various areas of mathematics. In general relativity, the Final State Conjecture (cf. [Kla07]) predicts that generic solutions to Einstein's field equations asymptotically approach a finite number of stationary solutions or Kerr black holes moving apart from each other. In geometric analysis, Soliton Resolution arises naturally in the study of gradient flows associated with conformally invariant variational problems. For example, pioneering works of Struwe [Str85, Str94], Qing [Qin95], Qing–Tian [QT97], and Hong–Tian [HT04] have established bubbling or Soliton Resolution along a well-chosen sequence of times for the harmonic map and Yang–Mills heat flows.

Motivated by these parabolic works, in this paper we study the energy-critical nonlinear heat flow in dimension $d \geq 3$. Our main result, Theorem 1.6, establishes a continuous-in-time bubble-tree decomposition for all finite-energy solutions of (1.1). More precisely, any solution with uniformly bounded $\dot{H}^1$-norm decomposes into a sum of solitons that may vary along different time sequences, a radiation term that is asymptotically trivial or captured by a weak limit in $\dot{H}^1$, and an error term that vanishes in the energy space. Moreover, when the initial data is non-negative, Corollary 1.7 shows that Theorem 1.6 implies the Soliton Resolution Conjecture, since positive solitons have been classified and are unique up to the symmetries of the equation due to [Oba72, CGS89]. Therefore, Theorem 1.6 extends Struwe's classical compactness result [Str84], which establishes a similar decomposition only along a well-chosen sequence of times, while Corollary 1.7 provides the first instance of Soliton Resolution for a non-integrable PDE, beyond radial symmetry, and without restrictions on the size of the initial data.

To explain the significance of our result, we now review some key developments in the literature. In the integrable setting, where tools such as the inverse scattering transform are available, the conjecture is well understood for models including the KdV equation [ES83], the modified KdV equation [Sch06], the one-dimensional cubic nonlinear Schrödinger equation (NLS) [BJM18], the derivative NLS [JLPS19], and, more recently, the Calogero–Moser derivative NLS [KK24].

For non-integrable equations with radial symmetry, where the solitons do not move in space, the conjecture has been settled for the nonlinear wave equation [DKM12, DKM13, DKM23, DKMM22, JK17, CDKM22, JL23b, JL22], damped Klein–Gordon equation [BRS17, GZ23], equivariant self-dual Chern–Simons–Schrödinger equation [KKO22], equivariant harmonic map heat flow [JL23a], and energy-critical nonlinear heat flow [Ary24].

For non-integrable equations, without radial symmetry, Soliton Resolution is known in one dimension for the damped Klein-Gordon equation [CMY21], in the neighborhood of a few solitons for the energy-critical nonlinear heat flow and the damped Klein-Gordon equation [CMR17,IN23, Ish25], continuously in time for the harmonic map heat flow [JLS25] or along a sequence of times in dimensions $3 \leq d \leq 5$ for the energy-critical nonlinear wave equation [DJKM].
In contrast, establishing our main results requires working in any dimension $d \geq 3$, where solitons exhibit only weak decay and no longer enjoy radial symmetry, allowing them to translate in space and potentially behave pathologically (cf. [Din86, DPMPP11, DPMPP13]). Moreover, we impose no restrictions on the size of the initial data, which implies that the nonlinear energy is, in general, non-coercive, unlike the setting of [KM06, GR18]. We overcome these difficulties by introducing new ideas that are robust and adaptable to other nonlinear parabolic flows. In particular, our modified notion of collision intervals, introduced in Section 3, can be used to generalize the results of [JLS25] to higher-dimensional target manifolds.

1.4. **Proof Sketch.** The proofs of the main Theorems 1.6 and 1.8 build on the framework of [JLS25], but require addressing new difficulties that arise in the context of the energy-critical nonlinear heat flow. This includes:

- *Non-coercivity of the energy functional.* The lack of a definite sign for the energy functional (1.2), especially in non-radial settings, prevents the use of standard energy estimates (cf. [Ary24]). To overcome this, we develop new localized energy estimates and use profile decompositions to show that there is no concentration of energy outside the self-similar region, which is a key ingredient in our argument.
- *Absence of energy quantization.* Unlike the case of harmonic maps between the plane and the round two-sphere, solitons for (1.1) do not exhibit quantized energy, thereby preventing a direct application of the collision intervals from [JLS25]. Nevertheless, the existence of a uniform positive lower bound on the energy of any soliton allows us to define suitable collision intervals, which is sufficient to establish our main results.

We first sketch the proof of Theorem 1.8, which in turn is used to prove Theorem 1.6. The argument begins by contradiction. Thus, assume that there is a sequence of times along which the solution deviates from a multi-bubble configuration. Unfortunately, it is difficult to analyze this sequence, and so we give ourselves a bit of room and instead analyze a sequence of time intervals where the solution deviates from a multi-bubble configuration; these sequences of intervals are called collision intervals, for a precise definition, see (3.1).
Thus, consider $[a_n, b_n] \subset [0, T_+)$, a sequence of time intervals where near the endpoints a_n and b_n, $u(t)$ is close to some multi-bubble configuration while inside $[a_n, b_n]$, $u(t)$ deviates away from this multi-bubble configuration. We define K as the smallest non-negative real number such that, heuristically, $u(a_n)$ is close to a K-bubble configuration. Note that defining K is straightforward when the energy of each bubble is quantized, as in the case of harmonic maps from $\mathbb{S}^2$ to $\mathbb{S}^2$ since we could simply sum up the energies of each bubble arising in the limit when $n \to \infty$. However, in general, sign-changing stationary solutions could attain a continuum of energies, and thus we need to define K in an approximate sense; see Definition 3.1.
Next, the idea is to use the minimality of K to relate the length of the collision interval to the size of the largest bubble that loses its shape or comes into a collision. In other words, we show that there exist a sub-interval $[c_n, d_n] \subset [a_n, b_n]$ and a constant $C_1 > 0$ such that

$$|[c_n, d_n]| \geq C_1 \lambda_{\max,n}^2$$

where $\lambda_{\max,n}^2$ is the largest scale associated with a bubble that comes into a collision. An application of the elliptic bubbling Theorem 2.15 on the interval $[c_n, d_n]$ and a contradiction argument yield a constant $C_2 > 0$ such that

$$\inf_{t\in[c_n,d_n]} \lambda_{\max,n}\|\partial_t u(t)\|_{L^2} \geq C_2.$$

Combining the above two estimates with (1.3) gives

$$\infty > \int_0^{T_+} \|\partial_t u(t)\|_{L^2}^2 \,\mathrm{d}t \geq \sum_{n\in\mathbb{N}} \int_{c_n}^{d_n} \|\partial_t u(t)\|_{L^2}^2 \,\mathrm{d}t \gtrsim \sum_{n\in\mathbb{N}} 1 = \infty,$$

which is a contradiction, thus completing the proof of Theorem 1.6.

Now let $t_n \to T_+$ be any sequence of times. From Theorem 1.6, we see that $u(t_n)$ approaches K multi-bubble configuration on either $B(y, \sqrt{T_+ - t})$ when $T_+ < \infty$ or $B(y, \sqrt{t})$ when $T_+ = \infty$. In particular, the K multi-bubble configurations depend on n. To obtain a finite number of bubbles (independent of n) as in Theorem 1.8 that are asymptotically orthogonal in the sense of (1.5) and (1.7), we apply the Compactness Theorem 2.15 to each bubble obtained in the sequence of multi-bubble configurations arising from Theorem 1.6 and build a new bubble tree configuration by selecting bubbles such that (1.5) and (1.7) are satisfied. The resulting multi-bubble configuration then satisfies all the requirements of Theorem 1.8, thus completing the proof.

1.5. **Notation and Conventions.** We use the following conventions in this paper.

- We denote Strichartz spaces $L^p_t L^q_x$ where the subscripts indicate L^p integral in time and L^q integral in space. In general, we will use Sobolev spaces instead of L^p spaces.
- Some constants that will occur frequently include $p := \frac{d+2}{d-2}$, for $d \geq 3$ and $E_* := \|W\|^2_{\dot{H}^1}$ where W is a non-zero positive stationary solution of (1.4). Furthermore, the inequality $A \lesssim B$ means that $A \leq CB$ for some constant $C > 0$, while $A \simeq B$ means that $A \lesssim B$ and $B \lesssim A$.
- An open ball is defined as $B(x, r) = \{z : |z - x| < r\}$ while a parabolic ball $Q_r(x, t) := B(x, r) \times \big(t - r^2, t\big)$ for any $x \in \mathbb{R}^d$, $t > 0$, $r > 0$. For convenience, $Q_1 := B(0, 1) \times (-1, 0)$.
- We will often localize several quantities over the course of this paper. To simplify notation, first, we define the energy densities relevant to the energy-critical heat flow
$$\mathbf{e}(u) := \frac{|\nabla u|^2}{2} - \frac{|u|^{p+1}}{p+1}, \text{ and } \bar{\mathbf{e}}(u) := |\nabla u|^2,$$
where $u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$. Given $A \subset \mathbb{R}^d$, we measure these quantities localized to this region
$$E(u; A) := \int_A \mathbf{e}(u(t, x))\mathrm{d}x, \text{ and } \bar{E}(u; A) := \int_A \bar{\mathbf{e}}(u(t, x))\mathrm{d}x.$$
Sometimes the domain A might be time-dependent, in which case it is easier to localize using cut-off functions. To that end, given any $\phi \in C^\infty(\mathbb{R}^d)$ we define
$$E_\phi(u) := \int_{\mathbb{R}^d} \mathbf{e}(u(t, x))\phi^2(x)\mathrm{d}x, \text{ and } \bar{E}_\phi(u) := \int_{\mathbb{R}^d} \bar{\mathbf{e}}(u(t, x))\phi^2(x)\mathrm{d}x.$$
- A standard cut-off function will be denoted by $\chi \in C^\infty_c(\mathbb{R}^d)$ where $\chi \equiv 1$ on $B(0, 1)$ and $\chi \equiv 0$ outside $B(0, 2)$. The rescaling of χ, will be defined as $\chi_R(x) := \chi(x/R)$ for any $R > 0$.
- Given $\lambda > 0$, $z \in \mathbb{R}^d$ and a function $W : \mathbb{R}^d \to \mathbb{R}$, we define the rescaled function as
$$W[z, \lambda](x) := \frac{1}{\lambda^{\frac{d-2}{2}}} W\left(\frac{x - z}{\lambda}\right).$$

1.6. **Acknowledgments.** The author is grateful to Andrew Lawrie for proposing the problem and for many valuable discussions, Tobias Colding for his constant encouragement and invaluable advice, and Yvan Martel for insightful conversations. This work was partially supported by NSF DMS Grant 2405393.

2. Preliminaries

2.1. **Properties of Stationary Solutions.** In this section, we recall some standard properties of non-zero solutions to (1.4), show that the definitions of scale and center (see Definition 1.1 and 1.2) are well-defined, and establish some natural consequences of these definitions. Let $W : \mathbb{R}^d \to \mathbb{R}$ be a non-zero finite energy solution of (1.4). We will show that the definition of its scale $\lambda(W;\gamma_0)$ and center $a(W;\gamma_0)$ is well-defined.

Lemma 2.1 (Center and scale)**.** *Let $\gamma_0 \in (0, \bar{E}_*/2)$, let $W : \mathbb{R}^d \to \mathbb{R}$ be a non-zero stationary solution, let $\lambda(W) = \lambda(W, \gamma_0)$ be its scale from Definition 1.1 and let $a(W) = a(W, \gamma_0)$ be a choice of center from Definition 1.2. Then $\lambda(W)$ is uniquely defined and strictly positive, and $a(W)$ is well-defined. For all $(b, \mu) \in \mathbb{R}^d \times (0, \infty)$ we have*

$$\lambda\left(W[b,\mu]\right) = \lambda(W)\mu, \text{ and } \left|a\left(W[b,\mu]\right) - b - a(W)\mu\right| \leq 2\lambda(W)\mu. \tag{2.1}$$

Proof. Since $\bar{E}(W;B(0,R)) \to \bar{E}(W)$ as $R \to \infty$, it follows that the scale $\lambda(W)$ is well-defined. If $\lambda(W) = 0$, then there exists $a_n \in \mathbb{R}^d$ so that for $n \geq 1$ we have

$$\bar{E}(W;B(a_n,1/n)) \geq \bar{E}(W) - \gamma_0. \tag{2.2}$$

If $n \neq m$, the $B(a_n,1/n) \cap B(a_m,1/m) = \emptyset$. Indeed, otherwise

$$\bar{E}(W) \geq \bar{E}(W;B(a_n,1/n)) + \bar{E}(W;B(a_m,1/m)) \geq 2\bar{E}(W) - 2\gamma_0$$

whence $\bar{E}(W) \leq 2\gamma_0 < \bar{E}_*$ which contradicts that W is non-zero. Therefore, $\{a_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}^d$, and $a_n \to a_\infty$. Passing to the limit in (2.2) gives a contradiction. To see that the center $a(W)$ is well-defined, take $\lambda_n \to \lambda(W)$ and $a_n \in \mathbb{R}^d$ such that

$$\bar{E}(W;B(a_n,\lambda_n)) \geq \bar{E}(W) - \gamma_0.$$

As before, we conclude that no two disks $\{B(a_n,\lambda_n)\}_{n=1}^{\infty}$ can be disjoint. Thus, the sequence $a_n \in \mathbb{R}^d$ lies in a compact set and we may assume that $a_n \to a_\infty$ as $n \to \infty$, which is the desired center. We note that $\lambda(W)$ is uniquely defined, but $a(W)$ is defined only up to a distance of $2\lambda(W)$. The properties (2.1) are immediate from the definitions. □

Lemma 2.2 (Decay of stationary solutions)**.** *There exists $\gamma_0 \in (0, \bar{E}_*/2)$ with the following property. For any $0 < \gamma \leq \gamma_0$ and any non-zero stationary solution $W : \mathbb{R}^d \to \mathbb{R}$ the exterior energy decays at the following rate:*

$$\bar{E}(W;\mathbb{R}^d \setminus B(a(W;\gamma);R\lambda(W,\gamma))) \leq \frac{C}{R^{d-2}}$$

for all $R \geq 1$ with constant $C = C(d,W) > 0$.

Proof. Without loss of generality, assume that $a(W;\gamma) = 0$ and $\lambda(W;\gamma) = 1$. Then using Lemma 2.1 in [Pre24] we get precise asymptotics of $|\nabla W|$, which implies the desired estimate since

$$\bar{E}(W;\mathbb{R}^d \setminus B(0;R)) = \int_{B(0,R)^c} |\nabla W|^2 \mathrm{d}x \lesssim \int_R^\infty \frac{r^{d-1}}{(1+r)^{2d-2}} \mathrm{d}r \lesssim \frac{1}{R^{d-2}}.$$

□

Lemma 2.3 (Energy of multi-bubbles)**.** *Let* $(y_n, \rho_n, M) \in \mathbb{R}^d \times (0,\infty) \times \mathbb{N}$*. Let* $\{W_1, \dots, W_M\}$ *be a collection of non-zero stationary solutions, and for each* $j \in \{1, \dots, M\}$ *let* $(b_{n,j}, \mu_{n,j}) \in B(y_n, \rho_n) \times (0,\infty)$ *be sequences such that*

$$\lim_{n\to\infty}\left[\sum_{j\neq k}\left(\frac{\mu_{n,j}}{\mu_{n,k}} + \frac{\mu_{n,k}}{\mu_{n,j}} + \frac{|b_{n,j}-b_{n,k}|}{\mu_{n,j}}\right)^{-1} + \sum_{j=1}^{M}\frac{\mu_{n,j}}{\operatorname{dist}(b_{n,j}, \partial B(y_n,\rho_n))}\right] = 0. \tag{2.3}$$

Then

$$\lim_{n\to\infty}\bar{E}\Big(\mathbf{W}\big(W_1[b_{n,1},\mu_{n,1}], \dots, W_M[b_{n,M},\mu_{n,M}]\big); B(y_n,\rho_n)\Big) = \sum_{j=1}^{M}\bar{E}(W_j).$$

Proof. To simplify the notation within the proof, we use the shorthand $W_{n,j} = W_j[b_{n,j},\mu_{n,j}]$ for each $1 \le j \le M$. Expanding the energy, we obtain

$$\bar{E}(\mathbf{W}(W_{n,1},\dots,W_{n,M}); B(y_n,\rho_n)) = \sum_{j=1}^{M}\bar{E}(W_{n,j}; B(y_n,\rho_n)) + 2\sum_{j\neq k}\int_{B(y_n,\rho_n)}(\nabla W_{n,j}\cdot\nabla W_{n,k})\mathrm{d}x.$$

By the asymptotic orthogonality of the parameters in (2.3), Lemma 2.2 and the invariance of the $\dot{H}^1$ norm under translation and rescaling we get

$$\bar{E}(W_{n,j}; B(y_n,\rho_n)) = \bar{E}(W_{n,j}) + o_n(1) = \bar{E}(W_j) + o_n(1)$$

as $n\to\infty$. On the other hand, if $j\neq k$, then

$$\Big|\int_{B(y_n,\rho_n)}(\nabla W_{n,j}\cdot\nabla W_{n,k})\mathrm{d}x\Big| \le \int|\nabla W_{n,j}||\nabla W_{n,k}|\mathrm{d}x = o_n(1)$$

by (2.3). Combining the above two displays, we get the desired energy expansion. □

2.2. **Local well-posedness of the nonlinear heat flow.** Following the ideas of Struwe [Str85], we develop a local well-posedness theory that is well adapted to the bubbling analysis, which we will carry out later. We first define the function space on which we develop the local well-posedness theory in the space of finite energy solutions:

$$\mathcal{V}_\tau^{T,M} := \Big\{u : [\tau,T]\times\mathbb{R}^d \to \mathbb{R} \mid u \text{ is measurable, and } \int_\tau^T \|\partial_t u(t)\|_{L^2}^2 + \|\nabla^2 u(t)\|_{L^2}^2\,\mathrm{d}t < \infty,\ \sup_{t\in[\tau,T]}\bar{E}(u(t)) \le M\Big\}$$

For convenience, we denote $\mathcal{V}^{T,M} = \mathcal{V}_0^{T,M}$. The main theorem in this section is as follows:

Theorem 2.4 (Local well-posedness)**.** *Let* $u_0 \in \dot{H}^1$*. Then, there exist a maximal time of existence* $T_+ = T_+(u_0)$ *and a unique solution* $u \in \bigcap_{T<T_+}\mathcal{V}^{T,M}$ *to (1.1) with* $u(0) = u_0$ *for some* $M = M(u_0) > 0$ *that depends on the initial data* u_0*. The finite maximal time* $T_+ < \infty$ *is characterized by the existence of an integer* $L \ge 1$*, a number* $\varepsilon_0 > 0$*, and points* $\{x_\ell\}_{\ell=1}^L \subset \mathbb{R}^d$ *such that*

$$\limsup_{t\to T_+}\bar{E}(u(t); \mathcal{B}(x_\ell, R)) \ge \varepsilon_0, \quad \forall\ R > 0, \quad \forall 1 \le \ell \le L.$$

Assume that we are in the type-II regime, i.e.

$$\limsup_{t\to T_+}\bar{E}(u(t)) \le C_1 < +\infty$$

then the collection of bubbling points $\{x_\ell\}_{\ell=1}^L \subset \mathbb{R}^d$ *is finite. There exists a finite energy mapping* $u^* : \mathbb{R}^d \to \mathbb{R}$ *such that* $u(t) \rightharpoonup u^*$ *as* $t \to T_+$ *weakly in* $\dot{H}^1(\mathbb{R}^d)$ *and strongly in* $\dot{H}^1(\mathbb{R}^d \setminus \{x_\ell\}_{\ell=1}^L)$*.*

The nonlinear energy $E(u(t))$ is continuous and non-increasing as a function of $t \in [0, T_+)$, and for any $t_1 \le t_2 \in [0, T_+)$, there holds

$$E(u(t_2)) + \int_{t_1}^{t_2} \|\mathcal{T}(u(t))\|_{L^2}^2 \,\mathrm{d}t = E(u(t_1)).$$

In particular,

$$\int_0^{T_+} \|\mathcal{T}(u(t))\|_{L^2}^2 \,\mathrm{d}t \lesssim \sup_{t \in [0, T_+]} (\bar{E}(u(t)) + \bar{E}(u(t))^{2^*/2}) < +\infty.$$

In this section, we prove some preliminary estimates that will be needed in proving Theorem 2.4. We start with an estimate to control the nonlinear term.

Lemma 2.5. *For any $u \in \dot{H}^1(\mathbb{R}^d) \cap \dot{H}^2(\mathbb{R}^d)$ and any smooth cut-off function $\phi \in C_c^\infty(B(x, 2R))$, we have*

$$\int_{\mathbb{R}^d} |u|^{2p} \phi^2 dx \lesssim \bar{E}(u; B(x, 2R))^{\frac{4}{d-2}} \left(\int_{\operatorname{supp}\phi} |D^2 u|^2 \phi^2 \,\mathrm{d}x + \frac{1}{R^2} \bar{E}(u; B(x, 2R)) \right)$$

where the constant in the above inequality only depends on the dimension d.

Proof. By the Gagliardo–Nirenberg–Sobolev inequality, we have

$$\|u\|_{L^{2p}(\mathbb{R}^d)} \lesssim \|D^2 u\|_{L^2(\mathbb{R}^d)}^{\theta} \|\nabla u\|_{L^2(\mathbb{R}^d)}^{1-\theta}$$

where $\theta = \frac{d-2}{d+2}$. Localizing this estimate by using a smooth cut-off function $\phi \in C_c^\infty(B(x, 2R))$ with $\phi \equiv 1$ on $B(x, R)$ with $|\nabla \phi| \le CR^{-1}$ and using Hardy's inequality, we have

$$\begin{aligned}
\int_{\mathbb{R}^d} |u\phi|^{2p} &\lesssim \|D^2(u\phi)\|_{L^2(\mathbb{R}^d)}^2 \|\nabla(u\phi)\|_{L^2(\mathbb{R}^d)}^{2p(1-\theta)} \\
&\lesssim \bar{E}(u; B(x, 2R))^{\frac{4}{d-2}} \left(\int_{\operatorname{supp}\phi} |D^2 u|^2 \phi^2 + \frac{1}{R^2} \bar{E}(u; B(x, 2R)) \right)
\end{aligned}$$

□

Next, we will need some energy estimates to propagate energy at short time scales.

Lemma 2.6. *Let $u \in \mathcal{V}^{T,M}$ be a solution of (1.1). Consider $I \subset [0, T)$ and $\phi \in C_c^\infty(\mathbb{R}^d)$. Then, for any $t_1, t_2 \in I$ and $t_1 < t_2$ we have*

$$E_\phi(u(t_2)) - E_\phi(u(t_1)) = -\int_{t_1}^{t_2} \int_{\mathbb{R}^d} (\partial_t u)^2 \phi^2 - 2 \int_{t_1}^{t_2} \int_{\mathbb{R}^d} (\nabla u \cdot \nabla \phi) \phi \partial_t u, \tag{2.4}$$

$$\begin{aligned}
\bar{E}_\phi(u(t_2)) - \bar{E}_\phi(u(t_1)) = &-2 \int_{t_1}^{t_2} \int_{\mathbb{R}^d} (\partial_t u)^2 \phi^2 + 2 \int_{t_1}^{t_2} \int_{\mathbb{R}^d} |u|^{p-1} u (\partial_t u) \phi^2 \\
&- 4 \int_{t_1}^{t_2} \int_{\mathbb{R}^d} (\nabla u \cdot \nabla \phi) \phi \partial_t u.
\end{aligned} \tag{2.5}$$

Furthermore, we have the following estimates:

$$
\begin{aligned}
|E_\phi(u(t_2)) - E_\phi(u(t_1))| \leq & \int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t \\
& + 2\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |\nabla u|^2|\nabla\phi|^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}, \quad (2.6)
\end{aligned}
$$

$$
\begin{aligned}
\left|\bar{E}_\phi(u(t_2)) - \bar{E}_\phi(u(t_1))\right| \leq & \, 2\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t \\
& + \left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |u|^{2p}\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2} \\
& + 4\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |\nabla u|^2|\nabla\phi|^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}, \quad (2.7)
\end{aligned}
$$

$$
\begin{aligned}
\bar{E}_\phi(u(t_2)) - \bar{E}_\phi(u(t_1)) \leq & \, 4\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |\nabla u|^2|\nabla\phi|^2 \mathrm{d}x\mathrm{d}t \\
& + 2\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |u|^{2p}\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}, \quad (2.8)
\end{aligned}
$$

$$
\begin{aligned}
\bar{E}_\phi(u(t_2)) - \bar{E}_\phi(u(t_1)) \leq & \, 2\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |u|^{2p}\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2} \\
& + 4\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}\left(\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |\nabla u|^2|\nabla\phi|^2 \mathrm{d}x\mathrm{d}t\right)^{1/2}, \quad (2.9)
\end{aligned}
$$

$$
\bar{E}_\phi(u(t_2)) - \bar{E}_\phi(u(t_1)) \leq 4\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |u|^{2p}\phi^2 \mathrm{d}x\mathrm{d}t + 4\int_{t_1}^{t_2}\int_{\mathbb{R}^d} |\nabla u|^2|\nabla\phi|^2 \mathrm{d}x\mathrm{d}t. \quad (2.10)
$$

.

Proof. The first identity (2.4) follows from

$$
\begin{aligned}
\int_{t_1}^{t_2}\int_{\mathbb{R}^d} \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{e}(u(t))\phi^2 \mathrm{d}x\mathrm{d}t &= \int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\nabla u\cdot\nabla\partial_t u - |u|^{p-1}u\partial_t u)\phi^2 \mathrm{d}x\mathrm{d}t \\
&= \int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\mathrm{div}(\nabla u\partial_t u) - (\Delta u + |u|^{p-1}u)\partial_t u)\phi^2 \mathrm{d}x\mathrm{d}t \\
&= -\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 \mathrm{d}x\mathrm{d}t - 2\int_{t_1}^{t_2}\int_{\mathbb{R}^d} (\nabla u\cdot\nabla\phi)\phi\partial_t u \mathrm{d}x\mathrm{d}t.
\end{aligned}
$$

The identity (2.5) can be derived similarly. The remaining inequalities (2.6), (2.8), and (2.9) follow by applications of Cauchy-Schwarz and Young's inequality. □

Lemma 2.7. *For any solution $u \in \mathcal{V}^{T,M}$ of* (1.1) *we have*

$$
\int_0^T\int_{\mathbb{R}^d} |\partial_t u|^2 \ \mathrm{d}x\mathrm{d}t \leq E(u_0) - E(u(T)) \leq CM
$$

where $C > 0$ is a constant depending only on the dimension d.

Proof. Multiplying (1.1) by $\partial_t u$ and integrating by parts, we get

$$
E(u(T)) + \int_{I\times\mathbb{R}^d} |\partial_t u|^2 = E(u(0)).
$$

Thus, in the interval $I = [0,T]$ we have

$$\int_{I\times\mathbb{R}^d} |\partial_t u|^2 \leq E(u(0)) - E(u(T)) \leq CM$$

where $C > 0$ is a constant depending only on the dimension d. □

As a consequence, we have the following energy estimates

Corollary 2.8. *If u is a smooth solution to* (1.1) *on $B(x,2R)\times[0,T]$, then*

$$\bar{E}(u(T), B(x,R)) \leq \bar{E}(u(\tau), B(x,2R)) + \frac{4(T-\tau)}{R^2} M_{x,R} + CM_{x,2R}^{p-1}\int_\tau^T\int_{\mathbb{R}^d} |D^2u|^2\phi^2 \ \mathrm{d}x\mathrm{d}t + \frac{(T-\tau)}{R^2}M_{x,2R}^p,$$

where $C > 0$ is a dimension dependent constant and $M_{x,2R} := \sup_{t\in[0,T]} \bar{E}(u(t); B_{2R}(x))$.

Proof. Using (2.10) with $\phi \in C_c^\infty(B(x,2R))$, we get for any $\tau \in [0,T)$

$$\bar{E}_\phi(u(T)) \leq \bar{E}_\phi(u(\tau)) + CM_{x,2R}^{p-1}\int_\tau^T\int_{\mathbb{R}^d} |D^2u|^2\phi^2 \ \mathrm{d}x\mathrm{d}t + \frac{(T-\tau)}{R^2}M_{x,2R}^p + \frac{4M_{x,2R}}{R^2}(T-\tau).$$

where we used Lemma 2.5 to control the nonlinear terms and Lemma 2.7 to bound the tension. Removing the cut-off terms in the above estimate gives the desired inequality. □

Lemma 2.9. *There exists $\varepsilon_1 > 0$ such that the following holds. Let $u \in \mathcal{V}_\tau^{T,M}$ be a smooth solution of* (1.1) *on the domain $B(x,2R)\times[\tau,T]$. If*

$$\sup_{\tau\leq t\leq T} \bar{E}\left(u(t), B(x,2R)\right) < \varepsilon_1$$

then we have

$$\int_\tau^T\int_{\mathbb{R}^d} \left|D^2u\right|^2\phi^2\mathrm{d}x\mathrm{d}t + \int_\tau^T\int_{\mathbb{R}^d} |\partial_t u|^2\phi^2\mathrm{d}x\mathrm{d}t \leq C\varepsilon_1\left(1+\frac{T-\tau}{R^2}\right),$$

where the constant $C > 0$ depends only on the dimension d.

Proof. Suppose

$$\sup_{\tau\leq t\leq T} \bar{E}\left(u(t), B(x,2R)\right) < \varepsilon_1,$$

where ε_1 will be chosen later. Then from (2.6) we know that

$$\begin{aligned}\int_\tau^T\int_{\mathbb{R}^d} |\partial_t u|^2\phi^2 dxdt &\leq |E(u(\tau); B(x,2R))| + |E(u(T); B(x,2R))| + \frac{C}{R^2}\int_\tau^T\int_{B(x,2R)} |\nabla u|^2 \ \mathrm{d}x\mathrm{d}t \\ &\leq C\varepsilon_1\left(1+\frac{T-\tau}{R^2}\right),\end{aligned}$$

where the constant $C > 0$ in the above inequality depends on the dimension. Furthermore, integration by parts implies

$$\begin{aligned}\int_\tau^T\int_{B(x,2R)} \left|D^2u\right|^2\phi^2 \ \mathrm{d}x\mathrm{d}t &\leq C\int_\tau^T\int_{B(x,2R)} |\Delta u|^2\phi^2 \ \mathrm{d}x\mathrm{d}t + \frac{C}{R^2}\int_\tau^T\int_{\operatorname{supp}\phi} |\nabla u|^2 \ \mathrm{d}x\mathrm{d}t \\ &\leq C\int_\tau^T\int_{B(x,2R)} |\Delta u|^2\phi^2 \ \mathrm{d}x\mathrm{d}t + C\varepsilon_1\frac{T-\tau}{R^2}\end{aligned}$$

Testing the nonlinear heat equation (1.1) by $\phi^2\Delta u$ we get

$$\int_\tau^T\int_{B(x,2R)}\phi^2|\Delta u|^2\,\mathrm{d}x\mathrm{d}t\le C_1\int_\tau^T\int_{B(x,2R)}\phi^2\left|\partial_t u\right|^2\,\mathrm{d}x\mathrm{d}t+C_2\int_\tau^T\int_{B(x,2R)}\phi^2|u|^{2p}\,\mathrm{d}x\mathrm{d}t$$

where $C_1, C_2>0$ are some constants independent of d, M, R, τ and T. Then Lemma 2.5 implies that

$$\int_\tau^T\int_{B(x,2R)}\phi^2|u|^{2p}\,\mathrm{d}x\mathrm{d}t\le C\varepsilon_1^{p-1}\left(\int_\tau^T\int_{B(x,2R)}\left|D^2u\right|^2\phi^2\,\mathrm{d}x\mathrm{d}t+\frac{T-\tau}{R^2}\varepsilon_1\right).$$

Thus, we get

$$\int_\tau^T\int_{B(x,2R)}\left|D^2u\right|^2\phi^2\le C\varepsilon_1^{p-1}\int_\tau^T\int_{B(x,2R)}\left|D^2u\right|^2\phi^2+C\varepsilon_1^{p}\left(1+\frac{T-\tau}{R^2}\right)$$

Therefore, if $C\varepsilon_1^{p-1}\le 1/2$, then we can absorb the first term and complete the proof. □

Lemma 2.10. *There exists $\varepsilon_1>0$ such that the following holds. For a smooth solution u of (1.1) on $B(x,2R)\times\left[T-2\delta R^2,T\right]$ for some $\delta\in(0,1)$, if*

$$\sup_{T-2\delta R^2\le t\le T}\bar{E}\left(u(t),B(x,2R)\right)<\varepsilon_1$$

then, for any $t\in\left[T-\delta R^2,T\right]$, we have

$$\int_{B(x,R)}\left|D^2u(t)\right|^2\,\mathrm{d}x+\int_{B(x,R)}\left|\partial_t u(t)\right|^2\,\mathrm{d}x\le\frac{C\varepsilon_1}{R^2}\left(1+\frac{1}{\delta}\right).$$

Proof. Let ε_1 be the same as in Lemma 2.9 and suppose that u satisfies

$$\sup_{T-2\delta R^2\le t\le T}\bar{E}\left(u(t),B(x,2R)\right)<\varepsilon_1.$$

Suppose $\psi\in C_c^\infty\left(B(x,2R)\right)$ be a smooth cut-off function. Then differentiating (1.1) with respect to t yields

$$\partial_t^2u=\Delta\partial_t u+p|u|^{p-1}\partial_t u.$$

Multiplying the above display with $\psi^2\partial_t u$ and integrating over $\mathbb{R}^d$ gives

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^d}\frac12\left|\partial_t u\right|^2\psi^2\mathrm{d}x&=\int_{\mathbb{R}^d}\partial_t u\Delta\partial_t u\psi^2\mathrm{d}x+p\int_{\mathbb{R}^d}\psi^2|\partial_t u|^2|u|^{p-1}\mathrm{d}x\\&=-\int_{\mathbb{R}^d}\left|\nabla\partial_t u\right|^2\psi^2\mathrm{d}x-\int_{\mathbb{R}^d}\nabla\partial_t u\cdot\nabla\psi\partial_t u\psi\mathrm{d}x\\&\quad+p\int_{\mathbb{R}^d}\psi^2|\partial_t u|^2|u|^{p-1}\mathrm{d}x,\end{aligned}$$

where we use integration by parts for the last step. By applying Young's inequality to various terms, we have

$$\begin{aligned}\frac12\frac{d}{dt}\int_{\mathbb{R}^d}\left|\partial_t u\right|^2\psi^2\mathrm{d}x+\int_{\mathbb{R}^d}\left|\nabla\partial_t u\right|^2\psi^2\mathrm{d}x&\le C\int_{\mathbb{R}^d}\left|\partial_t u\right|^2|\nabla\psi|^2\mathrm{d}x+\frac12\int_{\mathbb{R}^d}\left|\nabla\partial_t u\right|^2\psi^2\,\mathrm{d}x\\&\quad+C\int_{\mathbb{R}^d}\left|\partial_t u\right|^2|u|^{p-1}\psi^2\mathrm{d}x,\end{aligned}$$

for some constant $C>0$ that depend on the dimension d. Therefore, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d}\left|\partial_t u\right|^2\psi^2\mathrm{d}x+\int_{\mathbb{R}^d}\left|\nabla\partial_t u\right|^2\psi^2\mathrm{d}x\le C\int_{\mathbb{R}^d}\left|\partial_t u\right|^2|\nabla\psi|^2\mathrm{d}x+C\int_{\mathbb{R}^d}\left|\partial_t u\right|^2|u|^{p-1}\psi^2\mathrm{d}x \tag{2.11}$$

Now, using Hölder's inequality, we estimate the last term of the above inequality to get

$$\int_{\mathbb{R}^d} |\partial_t u|^2 \, |u|^{p-1} \psi^2 \, \mathrm{d}x \leq \left(\int_{\operatorname{supp}\psi} |u|^{2p} \, \mathrm{d}x \right)^{\frac{2}{d+2}} \left(\int_{\mathbb{R}^d} |\partial_t u \psi|^{\frac{2(d+2)}{d}} \, \mathrm{d}x \right)^{\frac{d}{d+2}}$$

with exponents $\alpha = \frac{2p}{p-1} = \frac{d+2}{2}$ and $\beta = \frac{2p}{p+1} = \frac{d+2}{d}$. By Gagliardo–Nirenberg–Sobolev inequality, for any $f \in C_c^\infty(\mathbb{R}^d)$ we have

$$\|f\|_{L^{2\beta}(\mathbb{R}^d)} \leq C \, \|Df\|^\theta_{L^2(\mathbb{R}^d)} \, \|f\|^{1-\theta}_{L^2(\mathbb{R}^d)}$$

where $\theta = \frac{d}{d+2} = \frac{1}{\beta}$ and thus

$$\left(\int_{\mathbb{R}^d} |\partial_t u \psi|^{\frac{2(d+2)}{d}} \, \mathrm{d}x \right)^{\frac{d}{d+2}} \leq C \left(\int_{\mathbb{R}^d} (|\nabla \partial_t u|^2 \psi^2 + |\nabla \psi|^2 |\partial_t u|^2) \, \mathrm{d}x \right)^\theta \left(\int_{\mathbb{R}^d} |\partial_t u|^2 |\psi|^2 \, \mathrm{d}x \right)^{1-\theta}.$$

Hence, using Young's inequality, for any $\eta \in (0,1)$, we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} |\partial_t u|^2 \, |u|^{p-1} \psi^2 \, \mathrm{d}x \\
&\leq C \left(\int_{\operatorname{supp}\psi} |u|^{2p} \, \mathrm{d}x \right)^{\frac{2}{d+2}} \left(\int |\partial_t u|^2 |\psi|^2 \, \mathrm{d}x \right)^{\frac{2}{d+2}} \left(\int (|\nabla \partial_t u|^2 \psi^2 + |\nabla \psi|^2 |\partial_t u|^2) \, \mathrm{d}x \right)^{\frac{d}{d+2}} \\
&\leq \frac{C}{\eta} \int_{\operatorname{supp}\psi} |u|^{2p} \, \mathrm{d}x \cdot \int_{\mathbb{R}^d} |\partial_t u|^2 \, \psi^2 \, \mathrm{d}x + \eta \int_{\mathbb{R}^d} |\nabla \partial_t u|^2 \, \psi^2 \, \mathrm{d}x + \eta \int_{\mathbb{R}^d} |\partial_t u|^2 \, |\nabla \psi|^2 \, \mathrm{d}x
\end{aligned}$$

By taking sufficiently small η, we can absorb the integral $\int_{\mathbb{R}^d} |\nabla \partial_t u| \, \psi^2$ term from (2.11) to get the following differential inequality

$$\frac{d}{dt} \int_{\mathbb{R}^d} |\partial_t u|^2 \, \psi^2 \, \mathrm{d}x \leq C_1 \int_{\operatorname{supp}\psi} |u|^{2p} \, \mathrm{d}x \cdot \int_{\mathbb{R}^d} |\partial_t u|^2 \, \psi^2 \, \mathrm{d}x + C_2 \int_{\mathbb{R}^d} |\partial_t u|^2 \, |\nabla \psi|^2 \, \mathrm{d}x$$

for some dimension dependent constants $C_1, C_2 > 0$. If we solve this differential inequality on the interval $[t_0, t] \subset \left[T - 2\delta R^2, T\right]$, we obtain

$$\int_{\mathbb{R}^d} |\partial_t u(t)|^2 \, \psi^2 \, \mathrm{d}x \leq \exp\left(C_1 \int_{t_0}^t \int_{\operatorname{supp}\psi} |u|^{2p} \, \mathrm{d}x \right) \left(\int_{\mathbb{R}^d} |\partial_t u\,(t_0)|^2 \, \psi^2 \, \mathrm{d}x + C \int_{t_0}^t \int_{\mathbb{R}^d} |\partial_t u(t)|^2 \, |\nabla \psi|^2 \, \mathrm{d}x\mathrm{d}t \right)$$

Using Lemma 2.5 with $\phi \in C^\infty(B(x, 2R))$ be a smooth cut-off function such that $\phi \equiv 1$ on $B(x, 3R/2)$ along with Lemma 2.9, we get

$$\begin{aligned}
\int_{t_0}^t \int_{\operatorname{supp}\phi} |u|^{2p} \, \mathrm{d}x\mathrm{d}t &\leq \int_{t_0}^t \int_{B_{3R/2}} |u|^{2p} \, \mathrm{d}x\mathrm{d}t \leq \int_{t_0}^t \int_{\operatorname{supp}\phi} |u|^{2p} \phi^2 \, \mathrm{d}x\mathrm{d}t \\
&\leq \varepsilon_1^{\frac{4}{n-2}} \left(C\varepsilon_1 \left(1 + \frac{t - t_0}{R^2} \right) + \frac{(t - t_0)}{R^2} \varepsilon_1 \right) \leq C\varepsilon_1^p.
\end{aligned}$$

Therefore, if we pick $\varepsilon_1^p \ll 1$ small enough then it follows from that

$$\int_{\mathbb{R}^d} |\partial_t u(t)|^2 \, \psi^2 \, \mathrm{d}x \leq \int_{\mathbb{R}^d} |\partial_t u\,(t_0)|^2 \, \psi^2 \, \mathrm{d}x + C \int_{t_0}^t \int_{\mathbb{R}^d} |\partial_t u|^2 \, |\nabla \psi|^2 \, \mathrm{d}x\mathrm{d}t.$$

For any $t \in \left[T - \delta R^2, T\right]$, we can choose $t_0 \in \left[t - \delta R^2, t\right]$ such that

$$\int_{\mathbb{R}^d} |\partial_t u\,(t_0)|^2 \, \psi^2 \, \mathrm{d}x = \min_{s \in [t - \delta R^2, t]} \int_{\mathbb{R}^d} |\partial_t u(s)|^2 \, \psi^2 \, \mathrm{d}x.$$

Then from Lemma 2.9 implies that for $t \in \left[T-\delta R^2, T\right]$,

$$\begin{aligned}\int_{\mathbb{R}^d} |\partial_t u(t)|^2 \psi^2 \, \mathrm{d}x &\leq \inf_{s\in[t-\delta R^2,t]} \int_{\mathbb{R}^d} |\partial_t u(s)|^2 \psi^2 \, \mathrm{d}x + C\int_{t-\delta R^2}^t \int_{\mathbb{R}^d} |\partial_t u|^2 |\nabla \psi|^2 \, \mathrm{d}x\mathrm{d}t \\ &\leq \frac{1}{\delta R^2}\int_{t-\delta R^2}^t \int_{\mathbb{R}^d} |\partial_t u(s)|^2 \psi^2 \, \mathrm{d}x\mathrm{d}t + \frac{C}{R^2}\int_{t-\delta R^2}^t \int_{B(x,3R/2)} |\partial_t u|^2 \, \mathrm{d}x\mathrm{d}t \\ &\leq \frac{1}{\delta R^2}\left(1+\frac{\delta R^2}{R^2}\right)\varepsilon_1 + \frac{C}{R^2}\left(1+\frac{\delta R^2}{R^2}\right)\varepsilon_1 \leq \frac{C\varepsilon_1}{R^2}\left(1+\frac{1}{\delta}\right).\end{aligned}$$

Therefore, we have

$$\int_{B_R} |\partial_t u(t)|^2 \, \mathrm{d}x \leq \int_{\mathbb{R}^d} |\partial_t u(t)|^2 \psi^2 \, \mathrm{d}x \leq \frac{C\varepsilon_1}{R^2}\left(1+\frac{1}{\delta}\right), \quad \forall t \in [T-\delta R^2, T].$$

In order to estimate $\int_{\mathbb{R}^d} \left|D^2 u\right|^2 \psi^2 \, \mathrm{d}x$, we start with

$$\begin{aligned}\int_{\mathbb{R}^d} \left|D^2 u(t)\right|^2 \psi^2 \, \mathrm{d}x &\leq C\int_{\mathbb{R}^d} |\Delta u(t)|^2 \psi^2 \, \mathrm{d}x + \frac{C}{R^2}\int_{\operatorname{supp}\psi} |\nabla u(t)|^2 \, \mathrm{d}x \\ &\leq C\int_{\mathbb{R}^d} |\Delta u(t)|^2 \psi^2 \, \mathrm{d}x + \frac{C}{R^2}\varepsilon_1 \\ &\leq C\int_{\mathbb{R}^d} |\partial_t u(t)|^2 \psi^2 \, \mathrm{d}x + C\int_{\mathbb{R}^d} |u(t)|^{2p}\psi^2 \, \mathrm{d}x + \frac{C}{R^2}\varepsilon_1 \\ &\leq \frac{C\varepsilon_1}{R^2}\left(1+\frac{1}{\delta}\right) + C\int_{\mathbb{R}^d} |u(t)|^{2p}\psi^2 \, \mathrm{d}x.\end{aligned}$$

Then by Lemma 2.5 we have

$$\int_{\mathbb{R}^d} |u(t)|^{2p}\psi^2 \, \mathrm{d}x \leq C\varepsilon_1^{p-1}\left(\int_{\mathbb{R}^d} \left|D^2 u(t)\right|^2 \psi^2 \, \mathrm{d}x + \frac{\varepsilon_1}{R^2}\right)$$

and therefore, if we assume that ε_1^{p-1} is small enough, then we have

$$\int_{B_R} \left|D^2 u(t)\right|^2 \, \mathrm{d}x \leq \int_{\mathbb{R}^d} \left|D^2 u(t)\right|^2 \psi^2 \, \mathrm{d}x \leq \frac{C\varepsilon_1}{R^2}\left(1+\frac{1}{\delta}\right), \quad \forall t \in [T-\delta R^2, T].$$

□

Lemma 2.11. *There exists* $\varepsilon_1 > 0$ *such that if* u *is a smooth solution to* (1.1) *on* $B(x, 2R) \times \left[T-2\delta R^2, T\right]$ *and*

$$\sup_{T-2\delta R^2 \leq t \leq T} \bar{E}\left(u(t), B(x,2R)\right) < \varepsilon_1$$

then the Hölder norm of u *and its derivatives are uniformly bounded on* $B_R \times \left[T-\delta R^2, T\right]$ *in terms of* ε_1, δ, R. *In particular, the following estimate holds:*

$$\|D_t^l D_x^k u\|_{C^0(B(x,R)\times[T-\delta R^2,T])} \leq C R^{-(2l+k)}$$

for any $k, l \geq 0$ *where* C *depends on* $\delta, \varepsilon_1, k, l, d$.

Proof. From the previous Lemma 2.10, we know that the L^2 norm of $|D^2 u(t)|$ and $|\partial_t u(t)|$ on $B(x,R)$ are uniformly bounded for any $t \in \left[T-\delta R^2, T\right]$. By the Gagliardo–Nirenberg–Sobolev inequality we have that $|u(t)|^p \in L^2\left(B(x,R)\right)$. Since $|\partial_t u(t) - \Delta u(t)| \leq |u(t)|^p$ we deduce that $\|\partial_t u(t) - \Delta u(t)\|_{L^2(B(x,R))}$ is uniformly bounded on $\left[T-\delta R^2, T\right]$. Therefore,

$$|\partial_t u - \Delta u| \in L^2\left(B(x,R)\times\left[T-\delta R^2, T\right]\right).$$

Then higher regularity follows from a standard parabolic theory and a bootstrap argument. □

The previous results allow us to prove Theorem 2.4.

Proof of Theorem 2.4. We split the argument into two cases, one when u_0 is smooth and the other when $u_0 \in \dot{H}^1(\mathbb{R}^d)$.
First, consider the case when the initial map u_0 is smooth. By standard parabolic theory (cf. [BC96, CMR17, GR18]), we know that there is a solution u in $\mathcal{V}^{T,M}$ for some $T, M > 0$, where $M = M(u_0)$. Define $M_R := \sup_{x\in\mathbb{R}^d, 0\leq t<T} \bar{E}\left(u(t), B(x,R)\right)$. If $M_{x,R} < \varepsilon_1$ for some $R > 0$, then by Lemma 2.11 the Hölder norm of u and its derivative are uniformly bounded and hence u can be extended beyond the time T. Let $T > 0$ be the maximal time for which $M_R < \varepsilon_1$ holds for some R. Then there are some points x_i for which

$$\sup_{t\in[0,T]} \bar{E}\left(u(t), B\left(x_i, R\right)\right) \geq \varepsilon_1 \quad \text{for any } R > 0.$$

For any finite collection $\{x_i\}_{i=1}^k$ of such points and for any $R > 0$ for which $B(x_i, 2R)$ are disjoint balls, we can choose $t_i \in [\tau, T]$ such that

$$\int_{B(x_i,R)} |\nabla u(t_i)|^2 \mathrm{d}x \geq \bar{E}(u\left(t_i\right); B(x_i, R)) \geq \frac{\varepsilon_1}{2}$$

where $\tau = T - \delta R^2$ for $\delta > 0$ small enough. Then, since $u \in \mathcal{V}^{T,M}$ we have

$$\begin{aligned} M &\geq \sup_{[0,T]} \bar{E}(u(t)) \geq \sup_{[0,T]} \bar{E}(u(t); \cup_{i=1}^k B(x_i, R)) \\ &\geq \sum_{i=1}^k \sup_{[0,T]} \bar{E}(u(t); B(x_i, R)) \geq \sum_i \bar{E}(u(t_i); B(x_i, R)) \geq \frac{k\varepsilon_1}{2} \end{aligned}$$

which implies that $k \leq \frac{2M}{\varepsilon_1}$. Moreover, for any compact subset $K \subset \mathbb{R}^d \setminus \cup_{i=1}^k \{x_i\}$, there exists $R = R(K) > 0$ such that

$$\sup_{x\in K, 0\leq t<T} \bar{E}\left(u(t), B(x,R)\right) < \varepsilon_1.$$

Hence, by Lemma 2.11, we can extend our solution smoothly beyond the time T on K. Therefore, a smooth solution of (1.1) exists on $\mathbb{R}^d \times [0,T] \setminus \{(x_i, T)\}_{i=1,\ldots,k}$.
Now consider the general case when the initial map $u_0 \in \dot{H}^1(\mathbb{R}^d)$. Since $C_c^\infty(\mathbb{R}^d)$ is dense in $\dot{H}^1(\mathbb{R}^d)$, we may pick a sequence $u_{m0} \in C_c^\infty(\mathbb{R}^d)$ which converges to u_0 in $\dot{H}^1(\mathbb{R}^d)$ such that $\bar{E}(u_{m0}) \leq 2\bar{E}(u_0)$. For each smooth initial data u_{m0}, there is a solution u_m of (1.1) with initial data u_{m0} on $\mathcal{V}^{T_m,M}$ for some maximal values $T_m > 0$ and some $M = M(u_0)$ which is uniform in m by the fact that $\bar{E}(u_{m0}) \leq 2\bar{E}(u_0)$. Pick $R > 0$ such that $\sup_{x\in\mathbb{R}^d} \bar{E}(u_0; B(x, 2R)) \leq \frac{\varepsilon_1}{4}$. Then for m large enough, we have

$$\sup_{x\in\mathbb{R}^d} \bar{E}(u_{m0}; B(x, 2R)) \leq \frac{\varepsilon_1}{2}.$$

By Corollary 2.8, for $0 < T \leq \min(T_m, \varepsilon_1 R^2/(C'M))$, for some large constant $C' = C'(d) > 0$ we have

$$\begin{aligned}\sup_{x\in\mathbb{R}^d, 0\leq t<T} \bar{E}(u_m(t); B(x,R)) &\leq \sup_{x\in\mathbb{R}^d} \bar{E}(u_{m0}; B(x,2R)) + \frac{4T}{R^2} M_{x,2R} \\ &\quad + C\varepsilon_1^{p-1} \int_\tau^T \int_{\mathbb{R}^d} |D^2 u|^2 \phi^2 \, \mathrm{d}x\mathrm{d}t + \frac{T}{R^2}\varepsilon_1^p \\ &\leq \frac{\varepsilon_1}{2} + \frac{4T\varepsilon_1}{2R^2} + C'\varepsilon_1^p \left(1 + \frac{T}{R^2}\right) + \frac{T}{R^2}\varepsilon_1^p, \\ &\leq \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{2} \leq \varepsilon_1\end{aligned}$$

for $C' = C'(d) \gg 1$ and $\varepsilon_1 \ll 1$. Therefore by Lemma 2.11, the Hölder norm of u_m and its derivatives are uniformly bounded independent of m and we have $T_m \geq \frac{\varepsilon_1 R^2}{C'C_1}$ for all m because of the maximality of T_m. Moreover, by taking the limit of u_m as $m \to \infty$, we can have a solution $u \in \mathcal{V}^{T,M}$ of (1.1) with initial data u_0, which is indeed smooth on $\mathbb{R}^d \times [0,T)$ and satisfies

$$\sup_{0\leq t<T} \bar{E}(u(t)) \leq M.$$

Now let T_+ be the maximal time of existence of a smooth solution u to (1.1) on $\mathbb{R}^d \times [0,T)$. Then, using the above uniform energy bound and the same argument in the case when u_0 is smooth, we can show that the number of singular points is finite and bounded by a constant depending on M. □

As a consequence, we can prove the following stability result.

Lemma 2.12 (Short-time propagation of small energy). *Let $u(t)$ be a solution to (1.1) with initial data $u(0) = u_0 \in \dot{H}^1$. Let $T_+ = T_+(u_0)$ denote its maximal time of existence and assume that $\sup_{t\in[0,T_+)} \|u(t)\|_{\dot{H}^1} < \infty$. Let $0 < \sigma_n < \tau_n < T_+$ be two sequences of times such that $\sigma_n, \tau_n \to T_+$ as $n \to \infty$ and $\lim_{n\to\infty}(\tau_n - \sigma_n) = 0$. Let W be a stationary solution (possibly zero) and let $r_n > 0$ be a sequence such that $\lim_{n\to\infty}(\tau_n - \sigma_n) r_n^{-2} = 0$. If*

$$\lim_{n\to\infty} \bar{E}(u(\sigma_n) - W; B(0, 2r_n)) = 0,$$

then

$$\lim_{n\to\infty} \bar{E}(u(\tau_n) - W; B(0, r_n)) = 0. \tag{2.12}$$

Next, let $\varepsilon_n > 0$ be a sequence with $\varepsilon_n < r_n$ for all n and such that $\lim_{n\to\infty}(\tau_n - \sigma_n)\varepsilon_n^{-2} = 0$. Let $L \in \mathbb{N}$, $L \geq 1$, $\{x_\ell\}_{\ell=1}^L \subset \mathbb{R}^d$ such that the balls $B(x_\ell, \varepsilon_n)$ are disjoint and satisfy $B(x_\ell, \varepsilon_n) \subset B(0, r_n)$ for each $n \in \mathbb{N}$ and $\ell \in \{1, \ldots, L\}$. Moreover, $|x_\ell - x_m| \geq 5\varepsilon_n$ when $\ell \neq m$. If

$$\lim_{n\to\infty} \bar{E}(u(\sigma_n) - W; B(0, 2r_n) \setminus \cup_{\ell=1}^L B(x_\ell, \varepsilon_n/2)) = 0,$$

then

$$\lim_{n\to\infty} \bar{E}(u(\tau_n) - W; B(0, r_n) \setminus \cup_{\ell=1}^L B(x_\ell, \varepsilon_n)) = 0. \tag{2.13}$$

Proof. We prove (2.12). Set $v(t) := u(t) - W$. Then,

$$\partial_t v - \Delta v = |u|^{p-1}u - |W|^{p-1}W.$$

Then using the same idea as in (2.8) with a smooth cut-off function $\phi_n \in C_c^\infty(\mathbb{R}^d)$ supported on $B(0,2r_n)$ we get

$$\bar{E}_\phi(v(\tau_n)) \lesssim \bar{E}(v(\sigma_n);B(0,2r_n)) + \frac{(\tau_n-\sigma_n)^{1/2}}{r_n}\left(\int_{\sigma_n}^{\tau_n}\int_{\mathbb{R}^d}|\partial_t u|^2 \mathrm{d}x\,\mathrm{d}t\right)^{1/2} \tag{2.14}$$
$$+\left(\int_{\sigma_n}^{\tau_n}\int_{\mathbb{R}^d}|\partial_t u|^2 \mathrm{d}x\,\mathrm{d}t\right)^{1/2}\left(\int_{\sigma_n}^{\tau_n}\int_{B(0,2r_n)}|u|^{2p} \mathrm{d}x\,\mathrm{d}t\right)^{1/2}$$
$$+\left(\int_{\sigma_n}^{\tau_n}\int_{\mathbb{R}^d}|\partial_t u|^2 \mathrm{d}x\,\mathrm{d}t\right)^{1/2}\left(\int_{\sigma_n}^{\tau_n}\int_{B(0,2r_n)}|W|^{2p} \mathrm{d}x\,\mathrm{d}t\right)^{1/2}.$$

By Lemma 2.5 and 2.9 and using $\lim_{n\to\infty}(\tau_n-\sigma_n)r_n^{-2}=0$ we get

$$\int_{\sigma_n}^{\tau_n}\int_{B(0,2r_n)}|u|^{2p}\mathrm{d}x\,\mathrm{d}t \lesssim \left(1+\frac{\tau_n-\sigma_n}{r_n^2}\right) < \infty \text{ as } n\to\infty.$$

Next, using the decay of any stationary solution W from Lemma 2.1 in [Pre24] and $\lim_{n\to\infty}(\tau_n-\sigma_n)=0$ we get

$$\int_{\sigma_n}^{\tau_n}\int_{B(0,2r_n)}|W|^{2p}\mathrm{d}x\,\mathrm{d}t \lesssim (\tau_n-\sigma_n)\to 0, \text{ as } n\to\infty.$$

Finally, using the energy identity (1.3)

$$\int_{\sigma_n}^{\tau_n}\int_{\mathbb{R}^d}|\partial_t u|^2\mathrm{d}x\,\mathrm{d}t \le \int_{\sigma_n}^{T_+}\int_{\mathbb{R}^d}|\partial_t u|^2\mathrm{d}x\,\mathrm{d}t \to 0, \text{ as } n\to\infty.$$

This shows that all the error terms in (2.14) are asymptotically small, and thus the smallness of the energy $\bar{E}(v(\sigma_n);B(0,2r_n))$ can be transferred to the smallness of $\bar{E}(v(\tau_n);B(0,r_n))$ by using the fact that $\phi\equiv 1$ on $B(0,r_n)$.
The proof of (2.13) starts with (2.14) but uses a different cut-off function, which is supported on $B(0,2r_n)\setminus\cup_{\ell=1}^L B(x_\ell,\varepsilon_n/2)$ such that $\phi_n\equiv 1$ on the region $B(0,r_n)\setminus\cup_{\ell=1}^L B(x_\ell,\varepsilon_n)$, satisfying the bound $|\nabla\phi_n|\lesssim\varepsilon_n^{-1}$. Then, one can control the error terms following the same reasoning as above. □

2.3. **Concentration properties of the heat flow.** The goal of this section is to establish a crucial fact that energy cannot concentrate outside the self-similar scale, which is expected in type-II blowup scenario. Similar results are known for many other PDEs, for instance, energy-critical nonlinear wave equation [DJKM], wave maps [CTZ93, STZ92], and harmonic map heat flow [JLS25]. Due to the lack of finite speed of propagation, we cannot use the techniques developed for hyperbolic equations, while the lack of a coercive energy for the energy-critical heat flow (1.1) prevents us from using the arguments developed for the harmonic map heat flow when $T_+=\infty$.

Lemma 2.13 (No self-similar nonlinear energy concentration in the finite-time blowup case)**.** *Let $u(t)$ be a solution of (1.1) with initial data $u_0\in\dot{H}^1$ such that the maximal time of existence $T_+=T_+(u_0)<\infty$ and $\sup_{t\in[0,T_+)}\bar{E}(u(t))<\infty$. Let $x_0\in\mathcal{S}$ be a singular point as in (**??**) and let $r>0$ be sufficiently small such that $B(x_0,r)\cap(\mathcal{S}\setminus\{x_0\})=\emptyset$. Then*

$$\lim_{t\to T_+}\bar{E}(u(t);B(x_0,r)\setminus B(x_0,\alpha\sqrt{T_+-t}))=\bar{E}(u^*;B(x_0,r)) \tag{2.15}$$

for any $\alpha > 0$. Here, u^* denotes the weak limit of the flow, i.e. $u(t) \rightharpoonup u^*$ as $t \to T_+$. In particular, there exist $T_0 < T_+$ and functions $\nu, \xi : [T_0, T_+) \to (0, \infty)$ such that $\lim_{t\to T_+}(\nu(t) + \xi(t)) = 0$ and the following hold

$$\lim_{t\to T_+}\Big(\frac{\xi(t)}{\sqrt{T_+ - t}} + \frac{\sqrt{T_+ - t}}{\nu(t)}\Big) = 0, \quad \lim_{t\to T_+} E(u(t); B(x_0, \nu(t)) \setminus B(x_0, \xi(t))) = 0. \tag{2.16}$$

Proof. Consider a smooth radial cut-off function $\phi \in C_c^\infty(B(x_0, 2r_2) \setminus B(x_0, r_1/2))$ such that $\phi \equiv 1$ on $B(x_0, r_2) \setminus B(x_0, r_1)$ and $\phi \equiv 0$ outside $B(x_0, 2r_2) \setminus B(x_0, r_1/2)$ for any $0 < r_1 < r_2$. Using (2.5), Lemma 2.5 and Lemma 2.9, we see that for each $0 < s < \tau < T_+$ we have,

$$\begin{aligned}\Big|\bar{E}_\phi(u(\tau)) - \bar{E}_\phi(u(s))\Big| &\le \int_s^\tau \int_{\mathbb{R}^d} (\partial_t u)^2\phi^2 + \Big(\int_s^\tau \int_{\mathbb{R}^d} |u|^{2p}\phi^2\Big)^{1/2}\Big(\int_s^\tau \int_{\mathbb{R}^d} (\partial_t u)^2\phi^2\Big)^{1/2} \\ &\quad + 4\Big(\int_s^\tau \int_{\mathbb{R}^d} (\partial_t u)^2\phi^2\Big)^{1/2}\Big(\int_s^\tau \int_{\mathbb{R}^d} |\nabla u|^2|\nabla\phi|^2\Big)^{1/2} \\ &\lesssim \int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t + \sqrt{\int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t} + \frac{(T_+ - s)^{1/2}}{r_1}\sqrt{\int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t}.\end{aligned} \tag{2.17}$$

Let $s \to T_+$, then the above estimate implies that $\lim_{s\to T_+} \bar{E}_\phi(u(s))$ exists. Now observe that for some r' such that $0 < r' < \frac{r_1}{2} < r_1$ we have

$$\bar{E}_\phi(u(\tau)) - \bar{E}_\phi(u^*) = \int_{|x - x_0| \ge r'} (\mathbf{e}(u) - \mathbf{e}(u^*))\phi^2 \mathrm{d}x.$$

Since $u(t) \to u^*$ strongly in $\dot{H}^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \mathcal{S})$, the RHS in the above display tends to zero as $\tau \to T_+$. Thus, choosing $r_1 = \alpha(T_+ - s)^{1/2}$ and $r_2 = A(T_+ - s)^{1/2}$ in the definition of the cut-off function ϕ, where $0 < \alpha < A$ and sending $\tau \to T_+$ in (2.17) we get

$$\big|\bar{E}_\phi(u^*) - \bar{E}_\phi(u(s))\big| \lesssim \int_s^{T_+} \|\partial_t u(t)\|_{L^2}^2 \,\mathrm{d}t + \sqrt{\int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t} + \frac{1}{\alpha}\sqrt{\int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t}.$$

Therefore

$$\lim_{s\to T_+} \bar{E}_\phi(u(s)) = 0. \tag{2.18}$$

If, instead we set $r_1 = \alpha(T_+ - s)^{1/2}$ and $r_2 = r$ in the definition of the cut-off function ϕ where $r > 0$ is small enough such that $B(x_0, r)$ does not contain any other bubbling point, then we have

$$\big|\bar{E}_\phi(u^*) - \bar{E}_\phi(u(s))\big| \lesssim \int_s^{T_+} \|\partial_t u(t)\|_{L^2}^2 \,\mathrm{d}t + \sqrt{\int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t} + \frac{1}{\alpha}\sqrt{\int_s^{T_+} \|\partial_t u\|_{L^2}^2 \mathrm{d}t}.$$

Therefore, we have

$$\lim_{s\to T_+} \big|\bar{E}_\phi(u^*) - \bar{E}_\phi(u(s))\big| = 0. \tag{2.19}$$

Denote $A(s) = \{x \in \mathbb{R}^d : \alpha\sqrt{T_+ - s}/2 \le |x - x_0| \le \alpha\sqrt{T_+ - s}\}$ and $A(r) = \{x \in \mathbb{R}^d : r \le |x - x_0| \le 2r\}$ then if

$$\begin{aligned}
&\bar{E}_\phi(u(s)); B(x_0, r) \setminus B(x_0, \alpha\sqrt{T_+ - s})) - \bar{E}(u^*; B(x_0, r)) \\
&= \bar{E}_\phi(u(s)) - \bar{E}_\phi(u^*) - \int_{\{\phi \neq 1\}} \mathbf{e}(u(s))\phi^2 \mathrm{d}x + \int_{\{\phi \neq 1\}} \mathbf{e}(u^*)\phi^2 \mathrm{d}x - \bar{E}(u^*; B(x_0, \alpha\sqrt{T_+ - s})) \\
&= \bar{E}_\phi(u(s)) - \bar{E}_\phi(u^*) - \int_{A(s)} \mathbf{e}(u(s))\phi^2 \mathrm{d}x - \int_{A(r)} \mathbf{e}(u(s))\phi^2 \mathrm{d}x \\
&\quad + \int_{A(r)} \mathbf{e}(u^*)\phi^2 \mathrm{d}x + \int_{A(s)} \mathbf{e}(u^*)\phi^2 \mathrm{d}x - \bar{E}(u^*; B(x_0, \alpha\sqrt{T_+ - s})),
\end{aligned}$$

which implies that

$$\begin{aligned}
&\left|\bar{E}(u(s); B(x_0, r) \setminus B(x_0, \alpha\sqrt{T_+ - s})) - \bar{E}(u^*; B(x_0, r))\right| \\
&\lesssim \left|\bar{E}_\phi(u(s)) - \bar{E}_\phi(u^*)\right| + \bar{E}_{\pi'}(u(s); B(x_0, \alpha\sqrt{T_+ - s}) \setminus B(x_0, \alpha\sqrt{T_+ - s}/2)) \\
&\quad + \bar{E}(u^*; B(x_0, \alpha\sqrt{T_+ - s})) + \left|\int_{A(r)} \phi^2(\mathbf{e}(u^*) - \mathbf{e}(u(s)))\mathrm{d}x\right|.
\end{aligned}$$

By (2.19), (2.18), and strong convergence of $u(t)$ to u^* in $\dot{H}^1_{\mathrm{loc}}(\mathbb{R}^d \setminus \mathcal{S})$, we see that each term above tends to zero as $s \to T_+$. Thus,

$$\lim_{s \to T_+} \bar{E}(u(s); B(x_0, r) \setminus B(x_0, \alpha\sqrt{T_+ - s})) = \bar{E}(u^*; B(x_0, r)).$$

This completes the proof of (2.15). One can easily construct the curves ν and ξ such that the first equation in (2.16) holds. This, along with (2.15), implies the second equation in (2.16). □

To show the vanishing of the energy outside the self-similar scale in the global in time case, we follow the same argument as in the finite time blowup case.

Lemma 2.14 (Nonlinear energy dissipation in the global case)**.** *Let $u(t)$ be the solution to* (1.1) *with initial data $u_0 \in \dot{H}^1$, $T_+ = T_+(u_0) = \infty$ and finite energy $\sup_{t \ge 0} \bar{E}(u(t)) < \infty$. Then for any $y \in \mathbb{R}^d$ and any $\alpha > 0$ we have*

$$\lim_{t \to T_+} \bar{E}(u(t); \mathbb{R}^d \setminus B(y, \alpha\sqrt{t})) = 0.$$

Proof. Let $\varepsilon > 0$ be small enough. Then we can find $T_0 = T_0(\varepsilon) > 0$ such that

$$\Big(\int_{T_0}^\infty \int_0^\infty |\partial_t u|^2 \mathrm{d}x\mathrm{d}t\Big)^{1/2} \le \varepsilon.$$

Next, choose $T_1 \ge T_0$ so that for all $T \ge T_1$

$$\bar{E}(u(T_0); \mathbb{R}^d \setminus B(y, \alpha\sqrt{T}/4)) \le \varepsilon.$$

Fix any such $T \ge T_1$. Let $\phi(x) = 1 - \chi(|x - y|/\alpha\sqrt{T})$ where $\chi \in C_c^\infty(B(0, 2))$ is a smooth cut-off function. Then using (2.8), $|\nabla\phi|^2 \lesssim T^{-1}$, Lemma 2.5 and 2.9 we see that

$$\begin{aligned}
\bar{E}_\phi(u(T)) &\le \bar{E}_\phi(u(T_0)) + 2\left(\int_{T_0}^T \int_{\mathbb{R}^d} |u|^{2p}\phi^2\right)^{1/2} \left(\int_{T_0}^T \int_{\mathbb{R}^d} (\partial_t u)^2\phi^2\right)^{1/2} \\
&\quad + 4\left(\int_{T_0}^T \int_{\mathbb{R}^d} (\partial_t u)^2\phi^2\right)^{1/2} \left(\int_{T_0}^T \int_{\mathbb{R}^d} |\nabla u|^2 |\nabla\phi|^2\right)^{1/2} \\
&\le \varepsilon + C\varepsilon \le C_1\varepsilon,
\end{aligned}$$

for some constant C_1 that depends on $\alpha > 0$ and $\sup_{t\geq 0} \bar{E}(u(t)) < \infty$. Therefore, we get

$$\lim_{T\to T_+} \bar{E}_\phi(u(T)) = 0,$$

which implies the desired result. □

2.4. **Sequential Compactness.** In this section, we establish an elliptic compactness Theorem for Palais-Smale sequences for critical points associated to the equation (1.4). This result is quite classical with connections to concentration-compactness in analysis [Str84] and the Yamabe problem in differential geometry [BM10].

Theorem 2.15 (Elliptic Bubbling)**.** *Let $u_k : \mathbb{R}^d \to \mathbb{R}$ be a sequence of functions in $\dot{H}^1$ such that*

$$\limsup_{k\to\infty} \int_{\mathbb{R}^d} |\nabla u_k|^2 < \infty, \quad \lim_{k\to\infty} \rho_k \| \Delta u_k + |u_k|^{p-1} u_k \|_{L^2} = 0$$

for some sequence $\rho_k \in (0,\infty)$. Then given any sequence $y_k \in \mathbb{R}^d$, there exist a stationary solution $u_\infty \in \dot{H}^1$ (possibly trivial), an integer $m \in \mathbb{N}$, a constant $C > 0$, a sequence $R_k \to \infty$, a collection of elliptic solutions $W_1, \ldots, W_m$ each equipped with translation parameters $\{x_k^i\}_{j=1}^m \in B(y_k, C\rho_k)$ and scales $\{\lambda_k^j\}_{j=1}^m \in (0,\infty)$ such that

$$\begin{aligned}
&\lim_{k\to\infty} \bar{E}(u_k - u_\infty - \sum_{j=1}^m W_j[x_k^j, \lambda_k^j]; B(y_k, R_k\rho_k)) \\
&\quad + \sum_{j\neq j'} \left(\frac{\lambda_k^j}{\lambda_k^{j'}} + \frac{\lambda_k^{j'}}{\lambda_k^j} + \frac{|x_k^{j'} - x_k^j|^2}{\lambda_k^{j'}\lambda_k^j} \right)^{-1} + \sum_{j=1}^m \frac{\lambda_k^j}{\operatorname{dist}(x_k^j, \partial B(y_k, C\rho_k))} = 0.
\end{aligned} \tag{2.20}$$

Denote

$$\bar{\mathcal{S}} = \{x \in \mathbb{R}^d : \liminf_{k\to\infty} \lim_{r\to 0} \int_{B(x,r)\cap B(y_k, C\rho_k)} |u_k|^{p+1} \mathrm{d}x \geq \tilde{\varepsilon}\},$$

for some $\tilde{\varepsilon} = \tilde{\varepsilon}(n)$. Then $\bar{\mathcal{S}} = \{x^1, \ldots, x^l\}$, where $l \leq m$. Furthermore,

$$\begin{aligned}
&u_k(y_k + \rho_k \cdot) \rightharpoonup u_\infty \text{ weakly in } \dot{H}^1(B(0,C)) \\
&u_k(y_k + \rho_k \cdot) \to u_\infty \text{ strongly in } W^{2,2}_{\mathrm{loc}}(B(0,C) \setminus \bar{\mathcal{S}}).
\end{aligned} \tag{2.21}$$

For each $i \in \{1, \ldots, m\}$ there exists a finite set of points $\bar{\mathcal{S}}_i$, possibly empty and with $\operatorname{card}(\bar{\mathcal{S}}_i) \leq m$, such that

$$u_k(x_k^i + \lambda_k^i \cdot) \to W_j \text{ strongly in } W^{2,2}_{\mathrm{loc}}(\mathbb{R}^d \setminus \bar{\mathcal{S}}_j). \tag{2.22}$$

Finally, there exists a non-negative real number $K \geq 0$ so that

$$\lim_{k\to\infty} \bar{E}(u_k; B(y_k, R_k\rho_k)) = K\bar{E}_*. \tag{2.23}$$

Remark 2.16. The above Theorem 2.15 is similar in spirit to Theorem 1.1 in [Top04] for almost harmonic maps from $\mathbb{S}^2 \to \mathbb{S}^2$. The key difficulty in establishing the above theorem stems from the fact that the natural energy associated with (1.4) does not have a definite sign. Note that, unlike in the harmonic map case, we cannot expect to obtain L^∞ neck-estimates since $W^{2,2}(\mathbb{R}^d)$ does not embed into $C^0(\mathbb{R}^d)$ when $d \geq 4$. Lastly, observe that as a consequence of the above theorem, we have $\lim_{n\to\infty} \boldsymbol{\delta}(u_n; B(y_n, \widetilde{R}_n\rho_n)) = 0$ for any sequence $1 \ll \widetilde{R}_n \ll R_n$.

Proof. See Sections 2 and 3 in [Du13], where this argument has been carried out for $u_k \geq 0$ on bounded domains. However, the same argument can be repeated for sign–changing functions u_k on $\mathbb{R}^d$. The main difference is that the bubbles W_j arising from the blow-up argument are not necessarily positive solutions of (1.4). We briefly sketch the argument for the reader's convenience.

Step 1. Sequential Bubbling. By scaling and translational invariance, we can assume that $\rho_k = 1$ and $y_k = 0$. Denote the set of blowup points for the sequence

$$\mathcal{S} = \{x \in \mathbb{R}^d : \lim_{r\to 0} \liminf_{k\to\infty} \int_{B(x,r)\cap B(0,C)} |u_k|^{p+1} \mathrm{d}x \geq \tilde{\varepsilon}\},$$

where we will fix the constant $C > 0$ later and $\tilde{\varepsilon} = \tilde{\varepsilon}(n) > 0$ is a positive constant that appears in the ε-regularity Theorem 2.1 proved in [Du13], which says that $\int_{B(0,r)} |u|^{p+1} \leq \tilde{\varepsilon(n)}$ implies that $\int_{B(0,\delta)} |\nabla u|^2 \leq C_0$ for small $\delta \in (0,1)$ and some constant $C_0 > 0$. Since $\sup_{k\in\mathbb{N}} \int_{\mathbb{R}^d} |\nabla u_k|^2 < \infty$, by choosing $r > 0$ small enough such that $B(x^i, r) \cap B(x^j, r) = \emptyset$ for $i \neq j$, a standard covering argument implies that $\mathcal{S} = \{x^1, \ldots, x^N\}$ for some finite $N \in \mathbb{N}$, and $x^i \in \mathbb{R}^d$ for $1 \leq i \leq N$. We can choose C in the definition of $\mathcal{S}$ small enough such that $\mathcal{S}$ consists of a singleton, i.e., $\mathcal{S} = \{x^1\}$.

Step 1.1. Extracting the first bubble. Fix $x \in \overline{B(x^1, r) \cap B(0, C)}$ and let $r_k := r_k(x)$ be the unique radius depending on x such that

$$\int_{B(x,r_k)\cap B(0,C)} |u_k|^{p+1} \mathrm{d}x = \frac{\widetilde{\varepsilon}}{2}.$$

Let $x_k^1 \in \overline{B(x^1, r) \cap B(0, C)}$ be the point where $r_k(x)$ attains its minimum. Then define $\lambda_k^1 = r_k(x_k^1)$. Thus we have a blowup sequence, $\lambda_k^1 \to 0$ and $x_k^1 \to x^1$ as $k \to \infty$ such that

$$\int_{B(x_k^1,\lambda_k^1)} |u_k|^{p+1} \mathrm{d}x = \frac{\tilde{\varepsilon}}{2}.$$

Re-scaling the function u_k,

$$\tilde{u}_k(x) = (\lambda_k^1)^{2/(p-1)} u_k(\lambda_k^1 x + x_k^1)$$

and using the ε-regularity proved in Theorem 2.1 in [Du13] we see that since

$$\Delta \tilde{u}_k + |\tilde{u}_k|^{p-1} \tilde{u}_k = (\lambda_k^1)^{\frac{2}{p-1}} (\Delta u_k + |u_k|^{p-1} u_k),$$

the sequence $\tilde{u}_k \to W_1$ in $H^1_{\mathrm{loc}}(\mathbb{R}^d)$ where W_1 solves (1.4) either on $\mathbb{R}^d$ or $\mathbb{R}^d_+$ depending on whether x_k^1 lies in the interior of the domain $\overline{B(x^1, r) \cap B(0, C)}$ or on its boundary. The latter can be ruled out by showing

$$\frac{\lambda_k^1}{\mathrm{dist}(x_k^1, \partial B(0, C))} \to 0, \;\; k \to \infty$$

which can be done by a contradiction argument that involves assuming $\frac{\lambda_k^1}{\mathrm{dist}(x_k^1, \partial B(0,C))} \to c \in (0, \infty], \;\; k \to \infty$ and showing that this gives rise to a solution of (1.4) on the half-space which is known to be trivial by Pohozaev's identity. For more details, see page. 162, Section 3 in [Du13] or the proof of Proposition 2.1 in [Str84].

Step 1.2. Consider the re-normalized sequence

$$v_k(x) = u_k(x) - W_1[x_k^1, \lambda_k^1](x).$$

If v_k converges (up to subsequence) strongly to u_∞ in $\dot{H}^1(B(x_1,r)\cap B(0,C))$ then we are done. Otherwise, as in Step 1.1, we can find scales $\lambda_k^2 \to 0$ and centers $x_k^2 \to x^1$ such that

$$\int_{B(x_k^2,\lambda_k^2)} |v_k|^{p+1} = \frac{\tilde{\varepsilon_1}}{2} \tag{2.24}$$

for some constant $0 < \tilde{\varepsilon}_1 \leq \tilde{\varepsilon}$. We first claim that

$$\frac{\lambda_k^2}{\lambda_k^1} + \frac{|x_k^1 - x_k^2|}{\lambda_k^1 + \lambda_k^2} \to \infty, \text{ as } k \to \infty,$$

since otherwise there exists some constant $M > 0$ such that

$$\frac{\lambda_k^2}{\lambda_k^1} + \frac{|x_k^1 - x_k^2|}{\lambda_k^1 + \lambda_k^2} \leq M, \quad \text{ as } k \to \infty.$$

This, in turn, would imply that

$$\int_{B(x_k^2,\lambda_k^2)} |v_k|^{p+1}\mathrm{d}x \leq \int_{B((x_k^2-x_k^1)/\lambda_k^1,\lambda_k^2/\lambda_k^1)} |\tilde{u}_k - W_1|^{p+1}\mathrm{d}x \leq \int_{B(0,M)} |\tilde{u}_k - W_1|^{p+1}\mathrm{d}x \to 0$$

as $k \to \infty$ which contradicts the energy concentration in (2.24). The next subtle point here is to show that no energy is lost between the *neck*-region connecting the new bubble W_2 and the previous bubble W_1. This has been done in Section 4 [Du13] and therefore at the end of this step we get,

$$u_k - W_1[x_k^1,\lambda_k^1] - W_2[x_k^2,\lambda_k^2] \to 0$$

strongly in $\dot{H}^1(B(x^1,L\lambda_k^1)\cap B(x^1,L\lambda_k^2))$ for any $L > 0$.

Step 1.3. Iterate and conclude. One can then iterate this process finitely many times to extract the bubble tree as desired with asymptotically orthogonal parameters as in the second display in (2.20).

Step 2. Convergence results. The existence of the weak limit in (2.21) follows from the fact that u_k is a bounded sequence of $\dot{H}^1$ functions. The strong convergence in $W^{2,2}_{\mathrm{loc}}$ away from the blowup points follows from the ε-regularity result from Theorem 2.1 in [Du13]. Thus, u_∞ is a smooth stationary solution of (1.4) away from a finite set of points. Then the standard removable singularity theorem, see for instance [CGS89, Lemma 2.1], implies that u_∞ is a smooth solution of (1.4) on $\mathbb{R}^d$. The strong convergence in (2.22) follows from the definition of the blow-up parameters (x_k^i,λ_k^i) and ε-regularity from Theorem 2.1 in [Du13].

Step 3. Energy almost-quantization. The bubble tree convergence and the *no-neck* property established in Section 4 of [Du13] imply the energy identity

$$\lim_{k\to\infty} \bar{E}(u_k; B(0,R_k\rho_k)) = \sum_{j=1}^m \bar{E}(W_j).$$

Since $\bar{E}(W_j) \geq \bar{E}_*$ and we know that

$$\lim_{k\to\infty} \bar{E}(u_k; B(0,R_k\rho_k)) \geq m\bar{E}_*.$$

Furthermore, $\sum_{j=1}^m \bar{E}(W_j) \leq C_1$ since the sequence u_k has finite energy. Therefore, we can find a non-negative real number between m and less than or equal to $C_1/\bar{E}_*$ such that, up to passing to a subsequence, there exists a non-negative real number $K \geq m$ satisfying

$$\lim_{k\to\infty} \bar{E}(u_k; B(0,R_k)) = K\bar{E}_*$$

as desired. □

3. Analysis of Collision Intervals

For convenience, in this section, we let $u(t)$ be a solution of (1.1), with initial data $u_0 \in \dot{H}^1$, defined on the maximal time interval $I_+ = [0, T_+)$ where $T_+ < \infty$ in the finite time blow-up case and $T_+ = \infty$ in the global case. We will also assume that $C' = \sup_{t \in [0,T_+)} \bar{E}(u(t)) < \infty$. Let $0 < \gamma_0 \ll 1$ be such that Lemma 2.2 holds. We fix this choice of γ_0 and drop the subscript γ_0 from $\mathbf{d}_{\gamma_0}$ and $\boldsymbol{\delta}_{\gamma_0}$ and from the notation for the scale and center of a stationary solution W, in particular $\lambda(W) = \lambda(W; \gamma_0)$ and $a(W) = a(W; \gamma_0)$. Our goal in this section is to introduce the notion of collision intervals and show that if Theorem 1.6 fails, then these intervals have a nontrivial length.

Definition 3.1 (Collision Interval)**.** Let $K \in \mathbb{R}_+$ be the smallest positive real number with the following properties. There exist sequences of centers and scales $(y_n, \rho_n, \varepsilon_n) \in \mathbb{R}^d \times (0, \infty)^2$, sequences of times $\sigma_n, \tau_n \in (0, T_+)$ and small (but fixed) $\eta > 0$, satisfying $\varepsilon_n \to 0$, $0 < \sigma_n < \tau_n < T_+$, $\sigma_n, \tau_n \to T_+$, such that

1. $\boldsymbol{\delta}(u(\sigma_n); B(y_n, \rho_n)) \le \varepsilon_n$;
2. $\boldsymbol{\delta}(u(\tau_n); B(y_n, \rho_n)) \ge \eta$;
3. the interval $I_n := [\sigma_n, \tau_n]$ satisfies $|I_n| \le \varepsilon_n \rho_n^2$;
4. $K := \lim_{n \to \infty} \bar{E}(u(\sigma_n); B(y_n, \rho_n))/\bar{E}_*$.

Then the intervals $[\sigma_n, \tau_n]$ are called collision intervals associated to the energy level K and the parameters $(y_n, \rho_n, \varepsilon_n, \eta)$. We can conveniently package this information in the following notation $[\sigma_n, \tau_n] \in \mathcal{C}_K(y_n, \rho_n, \varepsilon_n, \eta)$.

Remark 3.2. By Definition 1.4 and item (1) in Definition 3.1, we can associate to each sequence of collision intervals $[\sigma_n, \tau_n] \in \mathcal{C}_K(y_n, \rho_n, \varepsilon_n, \eta)$ a sequence $(\xi_n, \nu_n) \in (0, \infty)^2$ with $\lim_{n \to \infty} \left(\frac{\xi_n}{\rho_n} + \frac{\rho_n}{\nu_n} \right) = 0$ such that

$$\lim_{n \to \infty} \bar{E}(u(\sigma_n); B(y_n, 2\nu_n) \setminus B(y_n, 2^{-1}\xi_n)) = 0. \tag{3.1}$$

Using item (3) in Definition 3.1 also allows us to assume that

$$|I_n| = \tau_n - \sigma_n \ll \xi_n^2. \tag{3.2}$$

Using Lemma 2.12 with (3.1) and (3.2), we get

$$\lim_{n \to \infty} \sup_{t \in [\sigma_n, \tau_n]} \bar{E}(u(t); B(y_n, \nu_n) \setminus B(y_n, \xi_n)) = 0. \tag{3.3}$$

The same argument works if we either enlarge ξ_n or shrink ν_n in the sense that we can replace (ξ_n, ν_n) by $(\widetilde{\xi}_n, \widetilde{\nu}_n)$ where $\xi_n \ll \widetilde{\xi}_n \ll \rho_n \ll \widetilde{\nu}_n \ll \nu_n$.

Lemma 3.3 (Existence of $K \ge 1$)**.** *If Theorem 1.6 is false, then K is well-defined with $K \ge 1$.*

Proof. Assume that Theorem 1.6 is false. Then there exist $\eta > 0$, sequences $\tau_n \to T_+$, $y_n \in \mathbb{R}^d$, $\rho_n \in (0, \infty)$ where $\rho_n \le \sqrt{T_+ - \tau_n}$ when $T_+ < \infty$ and $\rho_n \le \sqrt{\tau_n}$ when $T_+ = \infty$ and sequences $\alpha_n \to 0$ and $\beta_n \to \infty$ such that for all $n \in \mathbb{N}$ we have

$$\boldsymbol{\delta}(u(\tau_n); B(y_n, \rho_n)) \ge \eta, \ \lim_{n \to \infty} \bar{E}(u(\tau_n); B(y_n, \beta_n \rho_n) \setminus B(y_n, \alpha_n \rho_n)) = 0.$$

The existence of the sequences α_n and β_n follows from Lemma 2.13 or Lemma 2.14 when $\rho_n \simeq \sqrt{T_+ - \tau_n}$ or $\rho_n \simeq \sqrt{\tau_n}$.
Next, we can find a sequence σ_n such that $\sigma_n < \tau_n$, $\sigma_n, \tau_n \to T_+$, $|[\sigma_n, \tau_n]| \ll \rho_n^2$ and

$$\lim_{n \to \infty} \rho_n^2 \|\partial_t u(\sigma_n)\|_{L^2}^2 = 0.$$

To see this, assume to the contrary. Then there exist constants $c, c_0 > 0$ such that up to a subsequence we have

$$\rho_n^2 \|\partial_t u(t)\|_{L^2}^2 \geq c_0,$$

for all $t \in [\tau_n - c\rho_n^2, \tau_n]$. However, this yields a contradiction since $u(t)$ has finite energy, and therefore by the energy identity (1.3) we have

$$\infty > \int_0^{T_+} \int_{\mathbb{R}^d} |\partial_t u(t)|^2 \mathrm{d}x\mathrm{d}t \geq \sum_n \int_{\tau_n - c\rho_n^2}^{\tau_n} \int_{\mathbb{R}^d} |\partial_t u(t)|^2 \mathrm{d}x\mathrm{d}t \geq c_0 \sum_n \int_{\tau_n - c\rho_n^2}^{\tau_n} \rho_n^{-2}\,\mathrm{d}t = \infty.$$

Using (2.7), with $t_1 = \sigma_n$, $t_2 = \tau_n$, cut-off function $\phi \in C_c^\infty(B(y_n, \beta_n\rho_n) \setminus B(y_n, \alpha_n\rho_n))$ and showing that the error terms vanish as in the proof of Lemma 2.12, we get

$$\lim_{n\to\infty} \bar{E}(u(\sigma_n); B(y_n, 2^{-1}\beta_n\rho_n) \setminus B(y_n, 2\alpha_n\rho_n)) = 0. \tag{3.4}$$

Applying the sequential bubbling Theorem 2.15 to $u(\sigma_n)$, we obtain a bubble tree decomposition (2.20) along some subsequence of σ_n and for some sequence $R_n \to \infty$. Since energy vanishes in the neck region (3.4), we see that

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(\sigma_n); B(y_n, \rho_n)) = 0.$$

By Lemma 2.3 we can find $K \geq 0$ such that

$$K = \frac{\lim_{n\to\infty} \bar{E}(u(\sigma_n); B(y_n, \rho_n))}{\bar{E}^*}.$$

Thus, we have verified all the items in the Definition 3.1 for the interval $[\sigma_n, \tau_n]$, which shows that K is well defined and that $K \geq 0$.

To see that $K \geq 1$, we argue by contradiction. Suppose $K = 0$. Let ξ_n, ν_n be sequences as in Remark 3.2. Then, since $K = 0$, we get $\lim_{n\to\infty} \bar{E}(u(\sigma_n); B(y_n, \rho_n)) < \bar{E}_*$. This implies that $u(\sigma_n)$ cannot be close to any non-trivial multi-bubble configuration since any non-trivial bubble carries at least energy $\bar{E}_*$. Therefore, by item (1) in Definition 3.1, we get $\lim_{n\to\infty} \bar{E}(u(\sigma_n); B(y_n, \rho_n)) = 0$. Therefore, using $|[\sigma_n, \rho_n]| \ll \rho_n^2$, Lemma 2.12 and equation (3.4) we can propagate this smallness of energy at time $t = \sigma_n$ to time $t = \tau_n$ to get that

$$\bar{E}(u(\tau_n); B(y_n, \rho_n)) = o_n(1),$$

which contradicts item (2) in Definition 3.1. Thus $K \geq 1$. □

For the remainder of this section, assume that Theorem 1.6 is false. We will show that this implies a nontrivial lower bound on the length of the collision intervals. Let $K \geq 1$ be as in Lemma 3.3 and $[\sigma_n, \tau_n] \in \mathcal{C}_K(y_n, \rho_n, \varepsilon_n, \eta)$, where

$$y_n \in \mathbb{R}^d, \rho_n \in (0, \infty), \varepsilon_n \to 0, \eta > 0, 0 < \sigma_n < \tau_n < T_+, \sigma_n \to T_+, \tau_n \to T_+$$

are parameters that satisfy the requirements of Definition 3.1. We first prove a very general lower bound on the size of the intervals where the solution is initially close and later far from a multi-bubble configuration. We will call these *bad* intervals.

Lemma 3.4 (Lower bound on the length of *bad* intervals)**.** *There exists $\eta_0 > 0$ such that for all $\eta \in (0, \eta_0)$, there exist constants $\varepsilon, c_0 > 0$ such that the following holds: let $[\sigma, \tau] \subset [\sigma_n, \tau_n]$ be any subset such that*

$$\boldsymbol{\delta}(u(\sigma); B(y_n, \rho_n)) \leq \varepsilon, \quad \boldsymbol{\delta}(u(\tau); B(y_n, \rho_n)) \geq \eta,$$

let $\vec{W} = (W_1, \ldots, W_M)$ *be any collection of non-constant stationary solutions,* $\vec{\nu} = (\nu, \nu_1, \ldots, \nu_M) \in (0,\infty)^{M+1}$, $\vec{\xi} = (\xi, \xi_1, \ldots, \xi_M) \in (0,\infty)^{M+1}$ *any admissible vectors in the sense of Definition 1.4 such that*

$$\varepsilon \leq \mathbf{d}(u(\sigma), \mathbf{W}(\vec{W}); B(y_n, \rho_n); \vec{\nu}, \vec{\xi}) \leq 2\varepsilon.$$

Then

$$\tau - \sigma \geq c_0 \max_{j \in \{1,\ldots,M\}} \lambda(W_j)^2.$$

Remark 3.5 (Proof Sketch)**.** Since the proof of Lemma 3.4 is quite involved, we give a summary of the key ideas. As usual, we will argue by contradiction. Thus, there exists a sequence of intervals $[s_n, t_n] \subset [\sigma_n, \tau_n]$ such that $|[s_n, t_n]| \ll \lambda_{\max,n}^2$. The idea then is to contradict the minimality of $K \geq 1$ since size of the interval $[s_n, t_n]$ is too short compared to the scale $\lambda_{\max,n}$ implying that the collisions are captured on small balls $B(y_n', \rho_n') \subset B(y_n, \rho_n)$ with $\rho_n' \ll \rho_n$. As we do not see the large scales $\lambda_{\max,n}$ in these small balls $B(y_n', \rho_n')$, we deduce that these small balls must carry strictly smaller energy in the sense of the last item in Definition 3.1, which will contradict the minimality of K.
To make the above argument precise, it will be helpful to organize the bubbles that will arise when the localized distance $\mathbf{d}$ vanishes. To that end, we first distinguish the bubbles based on the size of their $\dot{H}^1$-interaction. In particular, if this interaction vanishes, then we say that the bubbles are asymptotically orthogonal.

Definition 3.6 (Asymptotic Orthogonality of Scales)**.** We say that two triples $(W_j, a_{j,n}, \lambda_{j,n})$ and $(W_{j'}, a_{j',n}, \lambda_{j',n})$ are *asymptotically orthogonal* if

$$\lim_{n\to\infty} \Big(\frac{\lambda_{j,n}}{\lambda_{j',n}} + \frac{\lambda_{j',n}}{\lambda_{j,n}} + \frac{|a_{j,n} - a_{j',n}|^2}{\lambda_{j,n}\lambda_{j',n}} \Big) = \infty, \tag{3.5}$$

where $W_j, W_{j'}$ are non-zero stationary solutions of (1.4), $a_{j,n}, a_{j',n} \in \mathbb{R}^d$ are sequences of points, and $\lambda_{j,n}, \lambda_{j',n} \in (0,\infty)$ are sequences of scales. We will use the short hand $(W_j, a_{j,n}, \lambda_{j,n}) \perp (W_{j'}, a_{j',n}, \lambda_{j',n})$ if the two triples are asymptotically orthogonal. See Proposition B.2 in [FG20] to understand the connection between (3.5) and the integral interaction between the bubbles W_j and $W_{j'}$ in the case when $W_j, W_{j'} \geq 0$.

Using the above notion of asymptotic orthogonality, we can organize a family of bubbles into a tree-like structure.

Definition 3.7 (Bubble Tree)**.** Given two collections of stationary solutions $\mathfrak{h}_1 = \{W_n\}_{n=1}^\infty$ and $\mathfrak{h}_2 = \{\widetilde{W}_n\}_{n=1}^\infty$, then $\mathfrak{h}_1 \prec \mathfrak{h}_2$ iff

$$\frac{\lambda(\widetilde{W}_n)}{\lambda(W_n)} \to \infty \ \text{ and } \ \exists C > 0 \text{ such that } B(a(W_n), \lambda(W_n)) \subset B(a(\widetilde{W}_n), C\lambda(\widetilde{W}_n)) \text{ for all } n \gg 1.$$

Then we say that $\mathfrak{h}_1$ is the parent and $\mathfrak{h}_2$ is its child. We will also allow for equality in the above relation by using the notation $\mathfrak{h}_1 \preceq \mathfrak{h}_2$. Given $M \in \mathbb{N}$, consider the collection $\{\mathfrak{h}_1, \ldots, \mathfrak{h}_M\}$ where $\mathfrak{h}_i = \{W_{k,i}\}_{k=1}^\infty$ and $W_{k,i}$ are stationary solutions. We define a *root* element $\mathfrak{h}_j$ as an element that is not a child of any parent $\mathfrak{h}_{j'}$ for $j' \in \{1, \ldots, M\}$. We define the collection of all root-indices as

$$\mathcal{R} := \{ j \in \{1, \ldots, M\} \mid \mathfrak{h}_j \text{ is a root } \}.$$

Finally, to each root $\mathfrak{h}_j$ we can define the *bubble tree* as the following collection $\mathcal{T}(j) := \{\mathfrak{h}_{j'} \mid \mathfrak{h}_{j'} \preceq \mathfrak{h}_j\}$ and $\mathcal{D}(j)$ as the set of all maximal elements (with respect to the partial order $\preceq$) of the pruned tree $\mathcal{T}(j) \setminus \{\mathfrak{h}_j\}$.

Proof of Lemma 3.4. Assume that Lemma 3.4 does not hold. Then there exists a sequence of intervals $[s_n, t_n] \subset [\sigma_n, \tau_n]$ such that

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(s_n); B(y_n, \rho_n)) = 0, \quad \lim_{n\to\infty} \boldsymbol{\delta}(u(t_n); B(y_n, \rho_n)) > 0, \tag{3.6}$$

a sequence of integers $M_n \geq 0$, sequences of M_n-bubble configurations $\mathbf{W}(\vec{W}_n)$, where $\vec{W}_n = (W_{1,n}, \ldots, W_{M_n,n})$ and sequences of vectors $\vec{\nu}_n = (\nu_n, \nu_{1,n}, \ldots, \nu_{M_n,n}) \in (0,\infty)^{M_n+1}$, $\vec{\xi}_n = (\xi_n, \xi_{1,n}, \ldots, \xi_{M_n,n}) \in (0,\infty)^{M_n+1}$ such that

$$\lim_{n\to\infty} \mathbf{d}(u(s_n), \mathbf{W}(\vec{W}_n); B(y_n, \rho_n); \vec{\nu}_n, \vec{\xi}_n) = 0, \tag{3.7}$$

and the largest scale $\lambda_{\max,n} := \max_{j=1,\ldots,M_n} \lambda(W_{j,n})$ satisfies $(t_n - s_n)^{1/2} \ll \lambda_{\max,n}$.
We can assume that $M_n = M$ is a fixed integer, by possibly passing to a subsequence. Consider the collection $\{\mathfrak{h}_1, \ldots, \mathfrak{h}_M\}$ where $\mathfrak{h}_j = \{W_{j,n}\}_{n=1}^{\infty}$ for $j \in \{1, \ldots, M\}$. Then construct a bubble tree as in Definition 3.7. By definition, for any $j, j' \in \mathcal{R}$ we can find a sequence $\widetilde{R}_n \to \infty$ such that up to a subsequence we have

$$B(a(W_{j,n}), 4R_n\lambda(W_{j,n})) \cap B(a(W_{j',n}), 4R_n\lambda(W_{j',n})) = \emptyset$$

for any sequence $R_n \leq \widetilde{R}_n$, where recall that $a(W_{j,n}), \lambda(W_{j,n})$ denotes the center and the scale of the stationary solution $W_{j,n}$. Then the decay estimate 2.2 implies that for any $j \in \mathcal{R}$ and any sequence $R_n \to \infty$ we have

$$\lim_{n\to\infty} \bar{E}(W_{j,n}; \mathbb{R}^d \setminus B(a(W_{j,n}); 4^{-1}R_n\lambda(W_{j,n}))) = 0,$$

which in turn combined with (3.7) yields

$$\lim_{n\to\infty} \bar{E}(u(s_n); B(y_n, \rho_n) \setminus \cup_{j\in\mathcal{R}} B(a(W_{j,n}), 4^{-1}R_n\lambda(W_{j,n}))) = 0. \tag{3.8}$$

Next, applying Theorem 2.15 to the sequence of stationary solutions $W_{j,n}$ and passing to a joint subsequence, we find a sequence $M_j \geq 0$ of non-negative integers, a sequence $\breve{R}_n \leq \widetilde{R}_n$ with $1 \ll \breve{R}_n \ll \xi_n\lambda_{\max,n}^{-1}$, stationary solutions $\mathcal{W}_{j,0}$, non-zero stationary solutions $\mathcal{W}_{j,k}$, scales $\Lambda_{j,k,n} \ll \lambda(W_{j,n})$ and points $p_{j,k,n} \in B(a(W_{j,n}), C\lambda(W_{j,n}))$ for each j and $k \in \{1, \ldots, M_j\}$, satisfying (2.21), (2.22), and

$$\begin{aligned}\lim_{n\to\infty} \bar{E}\Big(W_{j,n} - \mathcal{W}_{j,0}[a(W_{j,n}), \lambda_{j,n}] - \sum_{k=1}^{M_j} \mathcal{W}_{j,k}[p_{j,k,n}, \Lambda_{j,k,n}]); B(a(W_{j,n}), 4R_n\lambda(W_{j,n}))\Big)\\ + \sum_{k\neq k'} \Big(\frac{\Lambda_{j,k,n}}{\Lambda_{j,k',n}} + \frac{\Lambda_{j,k',n}}{\Lambda_{j,k,n}} + \frac{|p_{j,k,n} - p_{j,k',n}|^2}{\Lambda_{j,k,n}\Lambda_{j,k',n}}\Big)^{-1} + \sum_{k=1}^{M_j} \frac{\Lambda_{j,k,n}}{\operatorname{dist}(p_{j,k,n}, \partial B(a(W_{j,n}), C\lambda(W_{j,n}))} = 0.\end{aligned} \tag{3.9}$$

Here $C > 0$ is some finite constant, and R_n is a sequence, to be fixed below, such that $1 \ll R_n \leq \breve{R}_n$. To differentiate the weak limits $\mathcal{W}_{j,0}$ (which could be trivial) from the stationary solutions $\mathcal{W}_{j,k}$, we will call $\mathcal{W}_{j,0}$ body maps following the convention used in the harmonic map heat flow literature. Define the set of indices

$$\mathcal{J}_{\max} := \Big\{ j \in \{1, \ldots, M\} \mid C_j^{-1} \leq \frac{\lambda_{\max,n}}{\lambda(W_{j,n})} \leq C_j, \text{ for each } n \text{ for some } C_j > 1 \Big\}$$

and let $K_0 \geq 0$ be a real number such that

$$K_0 = \sum_{j\in\mathcal{J}_{\max}} \bar{E}(\mathcal{W}_{j,0}). \tag{3.10}$$

Then $0 \leq K_0 \leq K$ since we only collect the weak limits of each bubble $W_{j,n}$. Therefore it suffices to consider only two cases: $K_0 < K$ and $K_0 = K$.

Case 1: First, suppose that $K_0 = K$. Then $\mathcal{J}_{\max} = \mathcal{R} = \{1, \dots, M\}$ and $M_j = 0$. The idea is that if one of the above conditions does not hold, then there exists a bubble which will cost at least $\bar{E}_*$ amount of energy. More concretely, using (3.8)

$$\begin{aligned} K_0 \bar{E}_* = \sum_{j \in \mathcal{J}_{\max}} \bar{E}(\mathcal{W}_{j,0}) &\leq \sum_{j \in \mathcal{R}} \bar{E}(\mathcal{W}_{j,0}) \\ &\leq \sum_{j=1}^{M} \bar{E}(\mathcal{W}_{j,0}) + \sum_{j=1}^{M} \sum_{k=1}^{M_j} \bar{E}(\mathcal{W}_{j,k}) \\ &= \lim_{n \to \infty} \bar{E}(u(s_n); B(y_n, \rho_n)) = K \bar{E}_*. \end{aligned}$$

From the above expression it is clear that if $j_0 \in \mathcal{R} \setminus \mathcal{J}_{\max}$ then $\bar{E}(\mathcal{W}_{j_0,0}) \geq \bar{E}_*$ which contradicts $\lim_{n \to \infty} \bar{E}(u(s_n); B(y_n, \rho_n)) = K \bar{E}_*$. Therefore $\mathcal{J}_{\max} = \mathcal{R}$. By the same argument $\mathcal{R} = \{1, \dots, M\}$ and $M_j = 0$ for each $j \in \{1, \dots, M\}$. Therefore, for each $j \in \{1, \dots, M\}$

$$\lim_{n \to \infty} \bar{E}(W_{j,n} - \mathcal{W}_{j,0}[a(W_{j,n}), \lambda(W_{j,n})]; B(a(W_{j,n}),\ R_n \lambda(W_{j,n}))) = 0.$$

Fix a sequence $R_n \leq \breve{R}_n$ such that for each $j \in \{1, \dots, M\}$ we have $4 R_n \lambda_{\max,n} \leq \min_{j \in \{1,\dots,M\}} \nu_{j,n}$. Then since $\lambda(W_{j,n}) \simeq \lambda_{\max,n}$ for each $j \in \{1, \dots, M\}$ we can use (3.7) to get that

$$\lim_{n \to \infty} \bar{E}(u(s_n) - \mathcal{W}_{j,0}[a(W_{j,n}), \lambda(W_{j,n})]; B(a(W_{j,n}), 4 R_n \lambda_{\max,n}))) = 0.$$

Now we can use Lemma 2.12 with $(t_n - s_n)^{1/2} \ll \lambda_{\max,n}$ to propagate these estimates to time t_n for each $j \in \{1, \dots, M\}$ to get

$$\lim_{n \to \infty} \bar{E}(u(t_n) - \mathcal{W}_{j,0}[a(W_{j,n}), \lambda(W_{j,n})]; B(a(W_{j,n}), 4 R_n \lambda_{\max,n}))) = 0. \tag{3.11}$$

The same reasoning applied to (3.8) yields

$$\lim_{n \to \infty} \bar{E}(u(t_n) - \mathcal{W}_{j,0}[a(W_{j,n}), \lambda(W_{j,n})]; B(y_n, \rho_n) \setminus \cup_{j=1}^{M} B(a(W_{j,n}); R_n \lambda_{\max,n}))) = 0. \tag{3.12}$$

Using (3.11), (3.12), pairwise disjointness of distinct balls $B(a(W_{j,n}), R_n \lambda(W_{j,n}))$, asymptotic orthogonality of the triples $(\mathcal{W}_{j,0}, a(W_{j,n}), \lambda(W_{j,n}))$, and Remark 3.2, we get that

$$\lim_{n \to \infty} \boldsymbol{\delta}(u(t_n); B(y_n, \rho_n)) = 0,$$

which contradicts the second equation in (3.6).

Case 2: Next, consider the case $K_0 < K$. We show that this case leads to a contradiction with the minimality of K. Again we will need $R_n \to \infty$ such that $4 R_n \lambda_{\max,n} \leq \min\{\nu_{j,n}\}_{j \in \mathcal{J}_{\max}}$ and $R_n \leq \breve{R}_n$. We split the argument into several steps.

Step 1. We first show the existence of an integer $L \geq 1$, sequences $\{x_{\ell,n}\}_{\ell=1}^{L}$ with $x_{\ell,n} \in B(y_n, \xi_n)$ for each $n \in \mathbb{N}$ and each $\ell \in \{1, \dots, L\}$, and a sequence r_n satisfying the following properties

- $(t_n - s_n)^{1/2} \ll r_n \ll \lambda_{\max,n}$;
- the balls $B(x_{\ell,n}, r_n)$ are pairwise disjoint for $\ell \in \{1, \dots, L\}$ with

$$\lim_{n \to \infty} \frac{|x_{\ell,n} - x_{\ell',n}|}{r_n} = \infty \tag{3.13}$$

  for $\ell \neq \ell'$;
- on the union of all such balls, we approximately capture the missing energy, i.e.

$$K_1 = \frac{\lim_{n \to \infty} \bar{E}(u(s_n); \cup_{\ell=1}^{L} B(x_{\ell,n}, r_n))}{\bar{E}_*}, \tag{3.14}$$

where $K_1 = K - K_0$ and there is vanishing energy in the neck region, i.e., there exist sequences $\alpha_n \to 0, \beta_n \to \infty$ such that

$$\lim_{n\to\infty} \sum_{\ell=1}^{L} \bar{E}(u(s_n); B(x_{\ell,n}, \beta_n r_n) \setminus B(x_{\ell,n}, \alpha_n r_n))) = 0; \tag{3.15}$$

- and a sequence $\breve{\xi}_n$ such that

$$\xi_n \ll \breve{\xi}_n \ll \rho_n, \quad B(x_{\ell,n}, \beta_n r_n) \subset B(y_n, \breve{\xi}_n). \tag{3.16}$$

Step 1.1. We first construct the sequence of points $\mathcal{P} := \{\{x_{\ell,n}\}_{\ell=1}^{L}\}$ for some integer $L \geq 1$. The idea will be to do this inductively. Define our initial set $\mathcal{P}_0$ to consist of all points such that

- $a(\mathcal{W}_{j,n})$ with $j \in \mathcal{R} \setminus \mathcal{J}_{\max}$,
- $a(\mathcal{W}_{j,n})$ with $\mathfrak{h}_j \in \mathcal{D}(j_0)$ for some $j_0 \in \mathcal{J}_{\max,n}$, where note that $\mathcal{D}(j_0)$ is the collection of maximal elements in the pruned tree $\mathcal{T}(j_0) \setminus \mathfrak{h}_{j_0}$, and
- sequences $p_{j_0,k,n}$ associated to stationary solutions $\mathcal{W}_{j_0,k}\big(\frac{\cdot - p_{j_0,k,n}}{\Lambda_{j_0,k,n}}\big)$ for some $j_0 \in \mathcal{J}_{\max}$ that are
  - asymptotically orthogonal to every bubble in the collection $\mathfrak{h}_j \in \mathcal{D}(j_0)$,
  - and not children of any $\mathfrak{h}_j \in \mathcal{D}(j_0)$.

Enumerate the set of all such points $\mathcal{P}_0 = \{\{y_{\ell,n}\}_{\ell=1}^{L'}\}$ for some integer $L' \geq 1$. Observe that after possibly passing to a subsequence we have

$$\lim_{n\to\infty} \frac{(t_n - s_n)^{1/2}}{\mathrm{dist}(y_{\ell,n}, y_{\ell',n})} \in [0, \infty]$$

for any $\ell \neq \ell' \in \{1, \ldots, L'\}$. We add $y_{\ell_0,n}$ to our final collection $\mathcal{P}$ if

$$\lim_{n\to\infty} \frac{(t_n - s_n)^{1/2}}{\mathrm{dist}(y_{\ell_0,n}, y_{\ell,n})} = 0, \quad \forall \ell \in \{1, \ldots, L')\} \setminus \ell_0.$$

Otherwise, denote

$$\mathcal{D}(\ell_0) := \{\ell_0\} \cup \Big\{\ell : \lim_{n\to\infty} \frac{(t_n - s_n)^{1/2}}{\mathrm{dist}(y_{\ell_0,n}, y_{\ell,n})} > 0\Big\}.$$

Note that $\mathcal{D}(\ell_1) = \mathcal{D}(\ell_2)$ iff $\ell_2 \in \mathcal{D}(\ell_1)$. Define the barycenter

$$x_{\ell_0,n} := \sum_{\ell\in\mathcal{B}(\ell_0)} \frac{y_{\ell,n}}{|\mathcal{B}(\ell_0)|}.$$

We will include $x_{\ell_0,n} \in \mathcal{P}$. This finishes the construction of the set $\mathcal{P} = \{\{x_{\ell,n}\}_{\ell=1}^{L}\}$ for some integer $L \geq 1$ such that $\{x_{\ell,n}\} \subset B(y_n, \xi_n)$ for any $1 \leq \ell \leq L$.

Step 1.2. We choose the scale r_n such that

$$(t_n - s_n)^{1/2} \ll r_n \ll \lambda_{\max,n}, \quad \max\{R_n\lambda(W_{j,n}), \nu_{j,n}\} \ll r_n, \quad \forall j \notin \mathcal{J}_{\max}$$
$$\max(\Lambda_{j,k,n}, \xi_{j,n}) \ll r_n, \quad \forall\, (j,k) \in \mathcal{J}_{\max} \times \{1, \ldots, M_j\}$$

and such that the balls $B(x_{\ell,n}, r_n)$ satisfy (3.13). Note that $B(x_{\ell,n}, r_n)$ is asymptotically disjoint from $B(a(W_{j_0,n}), R_n\lambda_{\max,n})$ for any $j_0 \in \mathcal{J}_{\max}$ since $\lambda_{\max,n}^{-1}|a(W_{j_0,n}) - a(W_{j,n})| \to \infty$ for all $j \in \mathcal{R} \setminus \mathcal{J}_{\max}$ and $j_0 \in \mathcal{J}_{\max}$ and we choose the points $x_{\ell,n}$ coming from the centers $a(W_{j,n})$ for $j \in \mathcal{R} \setminus \mathcal{J}_{\max}$. This concludes the construction of the centers $\{x_{\ell,n}\}$ and scales r_n.

Step 1.3. It remains to verify (3.14), (3.15), and (3.16). The construction of the sequence $\breve{\xi}_n$ such that (3.16) holds follows from the construction of the scales r_n. For the other two estimates, observe that for any $j_0 \in \mathcal{J}_{\max}$, by definition $\{x_{\ell,n}\}_{\ell=1}^L$, the limit in (3.9), and the choice of r_n we have

$$\lim_{n\to\infty} \bar{E}(u(s_n) - \mathcal{W}_{j_0,0}[a(W_{j_0,n}), \lambda(W_{j_0,n})]; B(a(W_{j_0,n}), 4R_n\lambda_{\max,n}) \setminus \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) = 0. \tag{3.17}$$

Since $r_n \ll \lambda_{\max,n}$, the stationary solution

$$\lim_{n\to\infty} \bar{E}(\mathcal{W}_{j_0,0}[a(W_{j_0,n}), \lambda(W_{j_0,n})]; \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) = 0. \tag{3.18}$$

Equations (3.18), (3.17), (3.10), and (3.8) imply that

$$\lim_{n\to\infty} \bar{E}(u(s_n); B(y_n, \rho_n) \setminus \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) = \sum_{j\in\mathcal{J}_{\max}} \bar{E}(\mathcal{W}_{j,0}) = K_0\bar{E}_*. \tag{3.19}$$

Then

$$\begin{aligned}\lim_{n\to\infty} \bar{E}(u(s_n); \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) &= \lim_{n\to\infty} \bar{E}(u(s_n); B(y_n, \rho_n)) - \sum_{j\in\mathcal{J}_{\max}} \bar{E}(\mathcal{W}_{j,0}) \\ &= (K - K_0)\bar{E}_*,\end{aligned}$$

which verifies (3.14). The condition (3.15) follows from the construction of the set $\mathcal{P}$ and the choice of r_n.

Step 2. The key point of constructing the collection of balls $B(x_{\ell,n}, r_n)$ for $1 \le \ell \le L$, is that for large enough n, the function $u(t_n)$ deviates from a multi-bubble configuration on at least one of these balls. In other words, we will now show that there exists $1 \le \ell_1 \le L$ and $\eta_1 > 0$ such that (after possibly passing to a subsequence)

$$\boldsymbol{\delta}(u(t_n); B(x_{\ell_1,n}, r_n)) \ge \eta_1. \tag{3.20}$$

If not then for all $\ell \in \{1, \dots, L\}$ we have

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(t_n); B(x_{\ell,n}, r_n)) = 0. \tag{3.21}$$

We will argue that this implies that

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(t_n); B(y_n, \rho_n)) = 0,$$

which contradicts (3.6). First, since $(t_n - s_n)^{1/2} \ll r_n$ we can use Lemma 2.12 to propagate (3.15), (3.14), (3.19), and (3.17) up to time t_n to get

$$\begin{aligned}&\lim_{n\to\infty} \bar{E}(u(t_n); \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) = K_1\bar{E}_*, \\ &\lim_{n\to\infty} \bar{E}(u(t_n); B(y_n, \rho_n) \setminus \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) = K_0\bar{E}_*, \text{ and} \\ &\lim_{n\to\infty} \bar{E}(u(t_n) - \mathcal{W}_{j_0,0}[a(W_{j_0,n}), \lambda(W_{j_0,n})]; B(a(W_{j_0,n}), R_n\lambda_{\max,n}) \setminus \cup_{\ell=1}^L B(x_{\ell,n}, r_n)) = 0,\end{aligned} \tag{3.22}$$

where $K_1 = K - K_0$, $j_0 \in \mathcal{J}_{\max}$. Using again Lemma 2.12, (3.8), the construction of the sequences $\{x_{\ell,n}\}$ and r_n we have

$$\lim_{n\to\infty} \bar{E}(u(t_n); B(y_n, \rho_n) \setminus (\cup_{j\in\mathcal{J}_{\max}} B(a(W_{j,n}), R_n\lambda_{\max,n}) \cup \cup_{\ell=1}^L B(x_{\ell,n}, r_n))) = 0. \tag{3.23}$$

From (3.21), after passing to a joint subsequence in n, for each $\ell \in \{1, \dots, L\}$ we can find an integer $\widetilde{M}_\ell \ge 0$, a sequence of $\widetilde{M}_\ell$-bubble configurations $\mathbf{W}(\vec{\mathsf{W}}_{\ell,n})$, and sequences of vectors $\vec{\nu}_{\ell,n} = (\nu_{\ell,n}, \nu_{\ell,1,n}, \dots, \nu_{\ell,\widetilde{M}_\ell,n})$ and $\vec{\xi}_{\ell,n} = (\xi_{\ell,n}, \xi_{\ell,1,n}, \dots, \xi_{\ell,\widetilde{M}_\ell,n})$, so that

$$\lim_{n\to\infty} \mathbf{d}(u(t_n), \mathbf{W}(\vec{\mathsf{W}}_{\ell,n}); B(x_{\ell,n}, r_n); \vec{\nu}_{\ell,n}, \vec{\xi}_{\ell,n}) = 0. \tag{3.24}$$

Here note that $\mathbf{W}(\vec{\mathsf{W}}_{\ell,n}) = \sum_{j=1}^{\widetilde{M}_\ell} \mathsf{W}_{\ell,j,n}$ for some collection of stationary solutions $\mathsf{W}_{\ell,j,n}$. Consider collection of maps

$$\widetilde{\mathsf{W}}_n = ((\mathsf{W}_{\ell,j,n})_{\ell=1,j=1}^{L,\widetilde{M}_\ell}, (\mathcal{W}_{j,0}[a(W_{j,n}), \lambda(W_{j,n})])_{j\in\mathcal{J}_{\max}}).$$

Let $\mathbf{W}(\widetilde{\mathsf{W}}_n)$ denote the sum of all the maps in the above collection. For each $j \in \mathcal{J}_{\max}$ set $\nu_{j,n} := R_n$, $\xi_{j,n} = r_n$ and

$$\vec{\tilde{\nu}}_n := (\nu_n, (\nu_{\ell,n})_{\ell=1}^L, (\nu_{j,n})_{j\in\mathcal{J}_{\max}}), \quad \vec{\tilde{\xi}}_n := (\breve{\xi}_n, (\xi_{\ell,n})_{\ell=1}^L, (\xi_{j,n})_{j\in\mathcal{J}_{\max}}).$$

Then we claim that

$$\lim_{n\to\infty} \mathbf{d}\big(u(t_n), \mathbf{W}(\widetilde{\mathsf{W}}_n); B(y_n, \rho_n); \vec{\tilde{\nu}}_n, \vec{\tilde{\xi}}_n\big) = 0. \tag{3.25}$$

This follows from (3.24), asymptotic orthogonality of distinct triples $(\mathsf{W}_{\ell,k,n}, a(\mathsf{W}_{\ell,k,n}), \lambda(\mathsf{W}_{\ell,k,n}))$ and $(\mathsf{W}_{\ell',k',n}, a(\mathsf{W}_{\ell',k',n}), \lambda(\mathsf{W}_{\ell',k',n}))$ for $(\ell, k) \neq (\ell', k')$ since $B(x_{\ell,n}, r_n)$ are mutually disjoint, asymptotic orthogonality of triples

$$(\mathsf{W}_{\ell,k,n}, a(\mathsf{W}_{\ell,k,n}), \lambda(\mathsf{W}_{\ell,k,n})) \quad \text{and} \quad (\mathcal{W}_{j_0,0}[a(W_{j_0,n}), \lambda(W_{j_0,n})], a(W_{j_0,n}), \lambda(W_{j_0,n}))$$

for any $1 \le \ell \le L$ and $j_0 \in \mathcal{J}_{\max}$ since $r_n \ll \lambda_{\max,n}$ and

$$\lim_{n\to\infty} \bar{E}(\mathcal{W}_{j_0,0}[a(W_{j_0,n}), \lambda(W_{j_0,n})]; B(x_{\ell,n}, r_n)) = 0, \quad \forall j_0 \in \mathcal{J}_{\max}, \quad \forall \ell \in \{1, \ldots, L\},$$

$$\lim_{n\to\infty} \bar{E}(\mathsf{W}_{\ell,k,n}; B(y_n, \rho_n) \setminus B(x_{\ell,n}, r_n))) = 0, \quad \forall\, \ell \in \{1, \ldots, L\},\, k \in \{1, \ldots, \widetilde{M}_\ell\}.$$

These observations, together with (3.22), (3.23), and Remark 3.2 applied with scale $\breve{\xi}_n$, yields (3.25). This establishes (3.20).

Step 3. As a consequence of (3.20), we will show that there exists $\widetilde{\sigma}_n < t_n$ such that

$$t_n - \widetilde{\sigma}_n \ll r_n^2 \quad \text{and} \quad \lim_{n\to\infty} r_n \|\mathcal{T}(u(\widetilde{\sigma}_n))\|_{L^2} = 0. \tag{3.26}$$

This follows from the same contradiction argument as in the Proof of Lemma 3.3. Applying Theorem 2.15 and possibly passing to a subsequence, we have a bubble tree decomposition as in (2.20) for some sequence $\hat{R}_n \to \infty$. The estimate (3.15) can be propagated to time $\widetilde{\sigma}_n$ using (2.7) and the argument in Lemma 2.12 to get

$$\lim_{n\to\infty} \bar{E}\big(u(\widetilde{\sigma}_n); B(x_{\ell,n}, \beta_n r_n/2) \setminus B(x_{\ell,n}, 2\alpha_n r_n)\big) = 0.$$

Therefore, all the stationary solutions at scale r_n in Theorem 2.15 vanish, which implies that

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(\widetilde{\sigma}_n); B(x_{\ell_1,n}, r_n)) = 0. \tag{3.27}$$

By (2.23) we can find a real number $K' \ge 0$ such that,

$$K' \bar{E}_* = \lim_{n\to\infty} \bar{E}(u(\widetilde{\sigma}_n); B(x_{\ell_1,n}, r_n)). \tag{3.28}$$

The estimate (3.20) implies that $K' \ge 1$ since $t_n - \widetilde{\sigma}_n \ll r_n^2$.

Step 4. We will now show that $K' < K$ and that $[\widetilde{\sigma}_n, t_n] \in \mathcal{C}_{K'}(x_{\ell_1,n}, r_n, \varepsilon_{1,n}, \eta_1)$ for some sequence $\varepsilon_{1,n} \to 0$, which will contradict the minimality of K.

Step 4.1. We first show that $K' < K$. When $K_0 > 0$, i.e. $K_0 \ge 1$, then $K' < K$ since at least one bubble lives on the scale comparable to the maximum scale $\lambda_{\max,n}$, which is asymptotically larger than r_n, which contributes to an energy of at least $\bar{E}_*$. On the other hand, suppose $K_0 = 0$. Then $K \neq K_0$ since $K \ge 1$. If $K' = K$, then this implies that a part of the energy in $B(y_n, \rho_n)$ is successfully captured by the balls $B(x_{\ell_1,n}, r_n)$. However, since there is at least one index j_0 attaining the maximum scale, i.e., $\lambda(W_{j_0,n}) = \lambda_{\max,n}$, and $r_n \ll \lambda_{\max,n} = \lambda(W_{j_0,n})$, by Definition 1.1, we see that at least $\bar{E}_*/2$ energy must live outside the scale $B(x_{\ell_1,n}, r_n)$. Therefore,

$$K + o_n(1) \geq \bar{E}(u(\widetilde{\sigma}_n); B(y_n, \rho_n)) = \bar{E}(u(\widetilde{\sigma}_n); B(y_n, \rho_n) \setminus B(x_{\ell_1,n}, r_n)) + \bar{E}(u(\widetilde{\sigma}_n); B(x_{\ell_1,n}, r_n))$$
$$\geq \left(\frac{1}{2} + K'\right) \bar{E}_* - o_n(1).$$

Therefore, $K \geq K' + \frac{1}{2}$, contradicting $K' = K$. Thus $K' < K$.

Step 4.2. Next, we check the properties of Definition 3.1. Item (1) follows from (3.27), item (2) follows from (3.20), item (3) follows from (3.26) and item (4) follows from (3.28). Thus,

$$[\widetilde{\sigma}_n, t_n] \in \mathcal{C}_{K'}(x_{\ell_1,n}, r_n, \varepsilon_{1,n}, \eta_1)$$

which is a contradiction to the minimality of K, and therefore the proof is complete. □

By a standard continuity argument, we get the following Corollary of the above Lemma.

Corollary 3.8. *Let $\eta_0 > 0$ be as in Lemma 3.4, $\eta \in (0, \eta_0]$, and $[\sigma_n, \tau_n] \in \mathcal{C}_K(y_n, \rho_n, \varepsilon_n, \eta)$. Then, there exist $\varepsilon \in (0, \eta)$, $c_0 > 0$, $n_0 \in \mathbb{N}$, and $s_n \in (\sigma_n, \tau_n)$ such that for all $n \geq n_0$, the following conclusions hold. First,*

$$\boldsymbol{\delta}(u(s_n); B(y_n, \rho_n)) = \varepsilon.$$

Moreover, for each $n \geq n_0$ let $M_n \in \mathbb{N}$, and $\mathbf{W}(\vec{W}_n)$, where $\vec{W}_n = (W_1, \dots W_{M_n})$ be any sequence of M_n-bubble configurations, and let $\vec{\nu}_n = (\nu_n, \nu_{1,n}, \dots, \nu_{M,n}), \vec{\xi}_n = (\xi_n, \xi_{1,n}, \dots, \xi_{M,n}) \in (0, \infty)^{M+1}$ be any admissible sequences in the sense of Definition 1.4 such that

$$\varepsilon \leq \mathbf{d}(u(s_n), \mathbf{W}(\vec{W}_n); B(y_n, \rho_n), \vec{\nu}_n, \vec{\xi}_n) \leq 2\varepsilon$$

for each n. Define

$$\lambda_{\max,n} = \lambda_{\max}(s_n) := \max_{j=1,\dots,M_n} \lambda(W_{j,n}).$$

Then, $s_n + c_0 \lambda_{\max}(s_n)^2 \leq \tau_n$ and,

$$\boldsymbol{\delta}(u(t); B(y_n, \rho_n)) \geq \varepsilon, \quad \forall\, t \in [s_n, s_n + c_0 \lambda_{\max}(s_n)^2].$$

Proof. From Lemma 3.4, fix $\varepsilon, \eta_0 > 0$. Then we can define s_n by the first exit time

$$s_n := \inf\{t \in [\sigma_n, \tau_n] \mid \boldsymbol{\delta}(u(\tau); B(y_n, \rho_n)) \geq \varepsilon, \text{ for all } \tau \in [t, \tau_n]\}.$$

This is well-defined for all sufficiently large n. Then by continuity, $\boldsymbol{\delta}(u(s_n); B(y_n, \rho_n)) = \varepsilon$. Setting $\lambda_{\max}(s_n)$ and using Lemma 3.4 we see that for n large enough we have

$$s_n + c_0 \lambda_{\max}(s_n)^2 \leq \tau_n,$$

which completes the proof. □

4. Conclusion

In this section, we will prove Theorem 1.8 and use it to establish Theorem 1.6 and Corollary 1.7.

Proof of Theorem 1.8. The proof proceeds by a contradiction argument that we break into several steps.

Step 1. Setting up the contradiction hypothesis. If Theorem 1.8 fails then there exists a non-number $K \geq 1$, and parameters

$$y_n \in \mathbb{R}^d,\ \rho_n > 0,\ 0 < \sigma_n < \tau_n < T_+, \quad [\,\sigma_n, \tau_n\,] \in \mathcal{C}_K(y_n, \rho_n, \varepsilon_n, \eta),$$

with $\varepsilon_n \to 0$, $\sigma_n, \tau_n \to T_+$ such that

$$|\,\tau_n - \sigma_n\,| \le \varepsilon_n \rho_n^2, \quad \boldsymbol{\delta}\big(u(\sigma_n); B(y_n, \rho_n)\big) \le \varepsilon_n, \quad \boldsymbol{\delta}\big(u(\tau_n); B(y_n, \rho_n)\big) \ge \eta,$$

and $\bar{E}\big(u(\sigma_n); B(y_n, \rho_n)\big) = K\bar{E}_*$.

Step 2. Picking the first exit time inside each collision interval. By Corollary 3.8 there exist $\varepsilon \in (0, \eta)$, $c_0 > 0$, and times

$$s_n \in (\sigma_n, \tau_n), \quad \boldsymbol{\delta}\big(u(s_n); B(y_n, \rho_n)\big) = \varepsilon,$$

such that for $s_n + c_0\lambda_{\max,n}^2 \le \tau_n$ and for all $t \in [s_n, s_n + c_0\lambda_{\max,n}^2]$ we have

$$\boldsymbol{\delta}(u(t); B(y_n, \rho_n)) \ge \varepsilon \tag{4.1}$$

where $\lambda_{\max,n} := \lambda_{\max}(s_n)$.

Step 3. A quantitative lower bound on the $\|\partial_t u(t)\|_{L^2}^2$. We claim that there exists a constant $c_1 > 0$ such that for n large enough we have,

$$\lambda_{\max,n}^2 \,\|\partial_t u(t)\|_{L^2}^2 \;\ge\; c_1, \quad \forall t \in [s_n, s_n + c_0\lambda_{\max,n}^2]. \tag{4.2}$$

We will prove this by contradiction.

Step 3.1. Setting up the contradiction hypothesis. If (4.2) does not hold then there exists a sequence of times $t_n \in [s_n, s_n + c_0\lambda_{\max,n}^2]$ such that

$$\lambda_{\max,n}\,\|\partial_t u(t_n)\|_{L^2} \to 0$$

as $n \to \infty$. Using Theorem 2.15, we deduce that (up to a subsequence) there exists $R_n(x_n) \to \infty$ such that for any sequence $1 \ll \check{R}_n \ll R_n(x_n)$ we have

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(t_n); B(x_n, \check{R}_n\lambda_{\max,n})) = 0. \tag{4.3}$$

We will construct a set of points $x_{\ell,n}$ for $1 \le \ell \le L$ for some integer $L \ge 1$ and use (4.3) to conclude that

$$\lim_{n\to\infty} \boldsymbol{\delta}\big(u(t_n); B(y_n, \rho_n)\big) = 0. \tag{4.4}$$

which will contradict the lower bound $\boldsymbol{\delta}(u(t_n); B(y_n, \rho_n)) \ge \varepsilon$.

Step 3.2. Construction of the sequence $\{x_{\ell,n}\}_{\ell=1}^L$. We claim that there exist an integer $L \ge 1$, points $\{x_{\ell,n}\}$ for $1 \le \ell \le L$, $R \ge 2$ and a sequence $1 \ll \widetilde{R}_n \ll \lambda_{\max,n}^{-1}\xi_n$ such that

$$\bar{E}(u(s_n); B(y_n, \rho_n) \setminus \cup_{\ell=1}^L B(x_{\ell,n}, R\lambda_{\max,n})) \le \frac{\bar{E}_*}{4}, \tag{4.5}$$

$$B(x_{\ell,n}, \widetilde{R}_n\lambda_{\max,n}) \cap B(x_{\ell',n}, \widetilde{R}_n\lambda_{\max,n}) = \emptyset, \forall \ell \ne \ell' \in \{1, \dots, L\}, \tag{4.6}$$

where K is defined in Step 1, ξ_n comes from the multi-bubble configuration obtained at $t = s_n$, i.e., we consider multi-bubble configurations $\mathbf{W}(\vec{W}_n) = \sum_{j=1}^M W_{j,n}$ comprising of some fixed M number of bubbles, after possibly passing to a subsequence because our solution has finite energy with parameters $\vec{\nu}_n$ and $\vec{\xi}_n$ such that

$$\varepsilon \le \mathbf{d}(u(s_n), \mathbf{W}(\vec{W}_n); B(y_n, \rho_n); \vec{\nu}_n, \vec{\xi}_n) \le 2\varepsilon. \tag{4.7}$$

We define ξ_n, ν_n as the first components of the vectors $\vec{\xi}_n, \vec{\nu}_n$ respectively. Arguing as in Remark 3.2, we deduce that (3.2) and (3.3) hold. We will construct the sequence $\{x_{\ell,n}\}_{\ell=1}^L$ for some integer $L \in \mathbb{N}$ as follows. First up, to a subsequence we have that

$$L_{jk} := \lim_{n\to\infty} \frac{|a(W_{j,n}) - a(W_{k,n})|}{\lambda_{\max,n}} \in [0, \infty], \quad \forall j \ne k \in \{1, \dots, M\}.$$

Given an index $j \in \{1, \dots, M\}$, we collect all other indices for which L_{jk} is finite, i.e.,

$$\mathcal{L}(j) := \{j\} \cup \Big\{k \in \{1, \dots, M\} : L_{jk} < \infty\Big\}.$$

Observe that for $j \neq k$ either $\mathcal{L}(j) = \mathcal{L}(k)$ or $\mathcal{L}(j) \cap \mathcal{L}(k) = \emptyset$. Define the barycenter

$$x_{\mathcal{L}(j),n} := \sum_{i \in \mathcal{L}(j)} \frac{a(W_{i,n})}{|\mathcal{L}(j)|}.$$

Then our desired sequence of points $\{x_{\ell,n}\}_{\ell=1}^{L}$ is simply a collection of points $\{x_{\mathcal{L}(j),n}\}$ for each distinct index set $\mathcal{L}(j)$ with $L \leq M$.

Step 3.3 Verification of (4.5) and (4.6). Using Lemma 2.2, (4.7), and the definitions of $\mathbf{d}$ and $\lambda_{\max,n}$ there exists $R_1 \gg 1$ such that for $n \gg 1$ we have

$$E\Big(u(s_n); B(y_n, \rho_n) \setminus \bigcup_{j=1}^{M} B(a(\omega_{j,n}), R_1 \lambda_{\max,n})\Big) \leq \frac{\bar{E}_*}{4},$$

where K is defined in Step 1. Then, the definition of the points $x_{\ell,n}$ yields a sequence $1 \ll \widetilde{R}_n \ll \lambda_{\max,n}^{-1} \xi_n$ such that (4.6) holds.

Step 4. Vanishing of the distance in (4.4). Using the collection $\{x_{\ell,n}\}_{\ell=1}^{L}$ as the centers in (4.3) consider sequences $R_{\ell,n}$ such that for any $\breve{R}_n \leq R_{\ell,n}$ we have

$$\lim_{n \to \infty} \boldsymbol{\delta}(u(t_n); B(x_{\ell,n}, \breve{R}_n \lambda_{\max,n})) = 0, \quad \ell \in \{1, \dots, L\}. \tag{4.8}$$

This in particular implies that there exists a number $K_\ell \geq 0$ such that

$$\lim_{n \to \infty} \bar{E}(u(t_n); B(x_{\ell,n}, R_n \lambda_{\max,n})) = K_\ell \bar{E}_*$$

for each $\ell \in \{1, \dots, L\}$. Now consider $\widetilde{\xi}_n$ such that $\xi_n \ll \widetilde{\xi}_n \ll \rho_n$. Then, for each $\ell \in \{1, \dots, L\}$ we have

$$\lim_{n \to \infty} \frac{\lambda_{\max,n}}{\operatorname{dist}(x_{\ell,n}, \partial B(y_n, \widetilde{\xi}_n))} = 0, \tag{4.9}$$

as $x_{\ell,n} \in B(y_n, \xi_n)$ and $\lambda_{\max,n} \ll \xi_n$. Therefore, there exists a sequence $R_n \leq \min\{\widetilde{R}_n, \{R_{\ell,n}\}_{\ell=1}^{L}\}$ such that $B(x_{\ell,n}, R_n \lambda_{\max,n}) \subset B(y_n, \widetilde{\xi}_n)$ for each $\ell \in \{1, \dots, L\}$. Thus

$$\bar{E}(u(s_n); B(y_n, \rho_n) \setminus \cup_{\ell=1}^{L} B(x_{\ell,n}, R_n \lambda_{\max,n}/2)) \leq \frac{\bar{E}_*}{4}.$$

Propagating the above estimate using Lemma 2.12 we get

$$\lim_{n \to \infty} \bar{E}(u(t_n); B(y_n, \rho_n) \setminus \cup_{\ell=1}^{L} B(x_{\ell,n}, R_n \lambda_{\max,n})) \leq \frac{\bar{E}_*}{2}. \tag{4.10}$$

Using (4.3) for points in $\Omega_{n,L} := B(y_n, \rho_n) \setminus \cup_{\ell=1}^{L} B(x_{\ell,n}, R_n \lambda_{\max,n})$ we deduce that $u(t_n)$ cannot be close to a single bubble due to (4.10) and therefore

$$\lim_{n \to \infty} \bar{E}(u(t_n); \Omega_{n,L}) = 0.$$

We also know that (4.8) and the definition of the sequence R_n implies

$$\lim_{n \to \infty} \sum_{\ell=1}^{L} \boldsymbol{\delta}(u(t_n); B(x_{\ell,n}, R_n \lambda_{\max,n})) = 0. \tag{4.11}$$

Moreover, the balls $B(x_{\ell,n}, R_n\lambda_{\max,n})$ are disjoint by (4.6) and the choice of $R_n \le \widetilde{R}_n$. Combining (4.11), (4.9), the disjointness of the balls $B(x_{\ell,n}, R_n\lambda_{\max,n})$, (4.10), and Remark 3.2, we conclude that

$$\lim_{n\to\infty} \boldsymbol{\delta}(u(t_n); B(y_n, \rho_n)) = 0,$$

which contradicts (4.1). Thus (4.2) holds.
Step 5. Conclusion. By (4.2) we have

$$\begin{aligned}\int_0^{T_+} \|\partial_t u(t)\|_{L^2}^2 \,\mathrm{d}t &\ge \sum_n \int_{s_n}^{s_n+c_0\lambda_{\max}(s_n)^2} \|\partial_t u(t)\|_{L^2}^2 \,\mathrm{d}t \\ &\ge c_1 \sum_n \int_{s_n}^{s_n+c_0\lambda_{\max}(s_n)^2} \lambda_{\max}(s_n)^{-2} \,\mathrm{d}t \ge c_0 c_1 \sum_n 1 = \infty,\end{aligned}$$

which contradicts (1.3). Thus, we have proved Theorem 1.6. □

Proof of Theorem 1.6. We treat the finite–time blow–up case $T_+ < \infty$; the global case is analogous. Throughout we write $\rho(t) := \sqrt{T_+ - t}$.
Step 1. Reduction to small balls near the bubbling points. Theorem 2.4 furnishes the existence of the set $\{x_1, \dots, x_L\} \subset \mathbb{R}^d$ and a weak limit $u_* \in \dot{H}^1$. Choosing $0 < \rho_0 \ll 1$ so that the balls $B(x_\ell, 2\rho_0)$ are disjoint, we get from Lemmas 2.13 and 2.14 that

$$\lim_{t\to T_+} \bar{E}(u(t) - u_*;\, \mathbb{R}^d \backslash \cup_{\ell=1}^L B(x_\ell, \rho_0)) = 0 \quad \text{and} \quad \lim_{t\to T_+} \bar{E}(u(t) - u_*;\, B(x_\ell, \rho_0) \backslash B(x_\ell, \rho(t))\Big) = 0$$

for $1 \le \ell \le L$. Since $u_* \in \dot{H}^1$, we have $\bar{E}\big(u_*; B(x_\ell, \rho(t))\big) \to 0$ as $t \to T_+$. Hence it suffices to study $u(t)$ inside the shrinking balls $B(x_\ell, \rho(t))$.
Step 2. Bubbling at one blowup point. Fix one bubbling point and denote it by $y := x_\ell$. Theorem 1.8 gives

$$\lim_{t\to T_+} \boldsymbol{\delta}(u(t); B(y, \rho(t))) = 0.$$

Let $t_n \to T_+$ be an arbitrary sequence of times. Then there exist

- an integer $M_n \ge 1$, which is also finite since $u(t)$ is a finite energy solution;
- $\mathbf{W}(\vec{W}_n) = \sum_{j=1}^{M_n} W_{j,n}$, where $W_{j,n}$ are stationary solutions;
- and scales $\vec{\nu}_n = (\nu_{0,n}, \nu_{1,n}, \dots, \nu_{M_n,n})$ and $\vec{\xi}_n = (\xi_{0,n}, \xi_{1,n}, \dots, \xi_{M_n,n})$,

such that

$$\mathbf{d}(u(t_n), \mathbf{W}(\vec{W}_n), ; B(y, \rho(t_n)), \vec{\nu}_n, \vec{\xi}_n) \to 0 \quad \text{as } n \to \infty.$$

Upon passing to a subsequence, we may assume that $M_n = M$ for all n.
Step 2.1 Initial bubble tree construction. For every $j \in \{1, \dots, M\}$ the map $W_{j,n}$ is a stationary solution and therefore applying Theorem 2.15 we get for each fixed j

- an integer $M_j \ge 0$;
- a weak limit $\vartheta_{j,0}$;
- non-zero stationary solutions $\vartheta_{j,1}, \dots, \vartheta_{j,M_j}$ with center $p_{j,k,n} \in B(a(W_{j,n}), \lambda(W_{j,n}))$ and scales $\Lambda_{j,k,n} \ll \lambda(W_{j,n})$,

such that

$$\lim_{n\to\infty} \bar{E}\Big(W_{j,n} - \vartheta_{j,0}\big(\frac{\cdot - a(W_{j,})}{\lambda(W_{j,n})}\big) - \sum_{k=1}^{M_j} \vartheta_{j,k}\big(\frac{\cdot - p_{j,k,n}}{\Lambda_{j,k,n}}\big); B_{j,n}\Big) = 0,$$

where $B_{j,n} := B(a(W_{j,n}),\, R_n\lambda(W_{j,n}))$ for some $R_n \to \infty$, where the scales and centers satisfy

$$\lim_{n\to\infty} \sum_{k\neq k'} \Big(\frac{\Lambda_{j,k,n}}{\Lambda_{j,k',n}} + \frac{\Lambda_{j,k',n}}{\Lambda_{j,k,n}} + \frac{|p_{j,k,n} - p_{j,k',n}|^2}{\Lambda_{j,k,n}\Lambda_{j,k',n}} \Big)^{-1} = 0.$$

For convenience denote $\Lambda_{j,0,n} := \lambda(W_{j,n}), p_{j,0,n} := a(W_{j,n})$ so that in every bubble family indexed by (j,k) the index $k = 0$ corresponds to the original scale $\lambda(W_{j,n})$ and centre $a(W_{j,n})$.
Step 3. Refined bubble tree construction. By the construction in the previous step, we have found a family

$$\{(\vartheta_{j,k}, p_{j,k,n}, \Lambda_{j,k,n})\}_{j=1,k=0}^{j=M,k=M_j}$$

which looks promising, but unfortunately, might not be asymptotically orthogonal. However, we can follow the same argument as in the proof of Theorem 1 in [JLS25] to construct an asymptotic orthogonal family $(W_j, a_{j,n}, \lambda_{j,n})$. The idea is to analyze the bubble tree as in the proof of Lemma 3.4. Denote $\mathcal{R}$ to be set of root indices obtained after partially ordering the tree $\mathfrak{h}_j = \{W_{j,n}\}_{n=1}^{\infty}$ and for each $\mathfrak{h}_{j_0} \in \mathcal{R}$ consider the bubble tree $\mathcal{T}(j_0) := \{\mathfrak{h}_j \preceq \mathfrak{h}_{j_0}\}$. For some large constant $C' > 0$, $B(a(W_{j,n}), \lambda(W_{j,n})) \subset B(a(W_{j_0,n}), C'\lambda(W_{j_0,n}))$ and therefore the domain $B(a(W_{j_0,n}), C'\lambda(W_{j_0,n}))$ contains all the stationary solutions

$$\bigcup_{\mathfrak{h}_j \in \mathcal{T}(j_0)} \{(\vartheta_{j,k}, p_{j,k,n}, \Lambda_{j,k,n})\}_{k=0}^{M_j}.$$

We will refine this collection to obtain an asymptotic orthogonal family. To this end, define

$$\mathcal{K}(j,k) := \{(j,k)\} \cup \big\{(j',k') : (W_{j',k'}, p_{j',k',n}, \Lambda_{j',k',n}) \perp (W_{j,k}, p_{j,k,n}, \Lambda_{j,k,n})\big\}.$$

For each reference index $j_0 \in \mathcal{R}$ we examine every cluster $\mathcal{K}(j,k)$ attached to the preliminary list of triples $\big(\vartheta_{j,k}, p_{j,k,n}, \Lambda_{j,k,n}\big)$.

- *Case 1*: $\big|\mathcal{K}(j,k)\big| = 1$: we keep the lone triple $\big(\vartheta_{j,k}, p_{j,k,n}, \Lambda_{j,k,n}\big)$.
- *Case 2:* $\big|\mathcal{K}(j,k)\big| > 1$: discard all triples with first index in $\mathcal{K}(j,k)$ and or replace them by a single triple $\big(\Theta_{j,k}, p_{j,k,n}, \Lambda_{j,k,n}\big)$, where $\Theta_{j,k}$ is a stationary solution. The construction of this new bubble $\Theta_{j,k,n}$ uses Theorem 1.6 and Theorem 2.15, and therefore the argument from the proof of Theorem 1 in [JLS25] carries over to this setting as well.

Repeating this procedure for every root index $j_0 \in \mathcal{R}$ leaves a final family of triples that are pairwise asymptotically orthogonal and fulfill the conclusions of Theorem 1.6. □

Proof of Corollary 1.7. Since $u_0 \geq 0$, by the maximum principle $u(t) \geq 0$ for all $t \in [0, T_+)$. By [CGS89] all the nontrivial non-negative bubbles are classified and are up to scaling and translation of the form

$$W(x) = \left(1 + \frac{|x|^2}{d(d-2)}\right)^{-\frac{(d-2)}{2}}. \tag{4.12}$$

Modifying the definition of the localized distance 1.4 by only considering nonnegative bubbles, one can repeat the argument in Section 3, proof of Theorem 1.8 and Theorem 1.6 to deduce the (1.6) and (1.8) with the solitons being independent of the sequence of times. The key point in these lemmas is the application of the Elliptic Compactness Theorem 2.15, which produces nontrivial nonnegative bubbles given a sequence of finite energy nonnegative functions. This can be seen, for instance, by repeating the argument in the proof Theorem 2.15 under the additional assumption that the sequence $\{u_k\}$ is non-negative for all $k \in \mathbb{N}$. Then the bubbles extracted by the blowup argument in Step 1.1 will be nontrivial nonnegative solutions to (1.4). As a consequence, the bubbles obtained in the decompositions (1.6) or (1.8) are re-scalings of the function define in (4.12). Thus, with the above modifications, we can prove the

Soliton Resolution Conjecture for the energy-critical nonlinear heat flow with nonnegative initial data. □

References

[Ary24] Shrey Aryan. Soliton resolution for the energy-critical nonlinear heat equation in the radial case. *arXiv preprint arXiv:2405.06005*, 2024.

[BC96] Haïm Brezis and Thierry Cazenave. A nonlinear heat equation with singular initial data. *J. Anal. Math.*, 68:277–304, 1996.

[BJM18] Michael Borghese, Robert Jenkins, and Kenneth D. T.-R. McLaughlin. Long time asymptotic behavior of the focusing nonlinear Schrödinger equation. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 35(4):887–920, 2018.

[BM10] Simon Brendle and Fernando C Marques. Recent progress on the yamabe problem. *arXiv preprint arXiv:1010.4960*, 2010.

[BRS17] Nicolas Burq, Geneviève Raugel, and Wilhelm Schlag. Long time dynamics for damped Klein-Gordon equations. *Ann. Sci. Éc. Norm. Supér. (4)*, 50(6):1447–1498, 2017.

[CDKM22] Charles Collot, Thomas Duyckaerts, Carlos Kenig, and Frank Merle. Soliton resolution for the radial quadratic wave equation in six space dimensions. *arXiv preprint arXiv:2201.01848*, 2022.

[CGS89] Luis A. Caffarelli, Basilis Gidas, and Joel Spruck. Asymptotic symmetry and local behavior of semilinear elliptic equations with critical Sobolev growth. *Comm. Pure Appl. Math.*, 42(3):271–297, 1989.

[CMR17] Charles Collot, Frank Merle, and Pierre Raphaël. Dynamics near the ground state for the energy critical nonlinear heat equation in large dimensions. *Comm. Math. Phys.*, 352(1):215–285, 2017.

[CMY21] Raphaël Côte, Yvan Martel, and Xu Yuan. Long-time asymptotics of the one-dimensional damped nonlinear klein–gordon equation. *Arch. Ration. Mech. Anal.*, 239:1837–1874, 2021.

[CTZ93] Demetrios Christodoulou and A. Shadi Tahvildar-Zadeh. On the regularity of spherically symmetric wave maps. *Comm. Pure Appl. Math.*, 46(7):1041–1091, 1993.

[Din86] Wei Yue Ding. On a conformally invariant elliptic equation on $\mathbf{R}^n$. *Comm. Math. Phys.*, 107(2):331–335, 1986.

[DJKM] Thomas Duyckaerts, Hao Jia, Carlos Kenig, and Frank Merle. Soliton resolution along a sequence of times for the focusing energy critical wave equation. *Geom. Funct. Anal.*, 27.

[DKM12] Thomas Duyckaerts, Carlos Kenig, and Frank Merle. Profiles of bounded radial solutions of the focusing, energy-critical wave equation. *Geom. Funct. Anal.*, 22(3):639–698, 2012.

[DKM13] Thomas Duyckaerts, Carlos Kenig, and Frank Merle. Classification of radial solutions of the focusing, energy-critical wave equation. *Camb. J. Math.*, 1(1):75–144, 2013.

[DKM23] Thomas Duyckaerts, Carlos Kenig, and Frank Merle. Soliton resolution for the radial critical wave equation in all odd space dimensions. *Acta Math.*, 230(1):1–92, 2023.

[DKMM22] Thomas Duyckaerts, Carlos Kenig, Yvan Martel, and Frank Merle. Soliton resolution for critical co-rotational wave maps and radial cubic wave equation. *Comm. Math. Phys.*, 391(2):779–871, 2022.

[DPMPP11] Manuel Del Pino, Monica Musso, Frank Pacard, and Angela Pistoia. Large energy entire solutions for the yamabe equation. *J. Differential Equations*, 251(9):2568–2597, 2011.

[DPMPP13] Manuel Del Pino, Monica Musso, Frank Pacard, and Angela Pistoia. Torus action on $\mathbb{S}^n$ and sign-changing solutions for conformally invariant equations. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 12(1):209–237, 2013.

[Du13] Shi-Zhong Du. On partial regularity of the borderline solution of semilinear parabolic equation with critical growth. *Adv. Differential Equations*, 18(1-2):147–177, 2013.

[ES83] W. Eckhaus and P. Schuur. The emergence of solitons of the Korteweg-de Vries equation from arbitrary initial conditions. *Math. Methods Appl. Sci.*, 5(1):97–116, 1983.

[FG20] Alessio Figalli and Federico Glaudo. On the sharp stability of critical points of the sobolev inequality. *Arch. Ration. Mech. Anal.*, 237(1):201–258, 2020.

[FPU55] E. Fermi, J. Pasta, and S. Ulam. Los alamos report la-1940. 1955.

[Gig86] Yoshikazu Giga. Solutions for semilinear parabolic equations in L^p and regularity of weak solutions of the Navier–Stokes system. *J. Differential Equations*, 62(2):186–212, 1986.

[GR18] Stephen Gustafson and Dimitrios Roxanas. Global, decaying solutions of a focusing energy-critical heat equation in $\mathbb{R}^4$. *J. Differential Equations*, 264(9):5894–5927, 2018.

[GZ23] Jingyuan Gu and Lifeng Zhao. Soliton resolution for the energy critical damped wave equations in the radial case. *arXiv preprint arXiv:2401.04115*, 2023.

[HT04] Min-Chun Hong and Gang Tian. Asymptotical behaviour of the yang-mills flow and singular yang-mills connections. *Math. Ann.*, 330:441–472, 2004.

[IN23] Kenjiro Ishizuka and Kenji Nakanishi. Global dynamics around 2-solitons for the nonlinear damped klein-gordon equations. *Ann. PDE*, 9(1):2, 2023.

[Ish25] Kenjiro Ishizuka. Long-time asymptotics of 3-solitary waves for the damped nonlinear klein-gordon equation. *arXiv preprint arXiv:2503.14889*, 2025.

[JK17] Hao Jia and Carlos Kenig. Asymptotic decomposition for semilinear wave and equivariant wave map equations. *Amer. J. Math.*, 139(6):1521–1603, 2017.

[JL22] Jacek Jendrej and Andrew Lawrie. Soliton resolution for energy-critical wave maps in the equivariant case. *J. Amer. Math. Soc.*, 2022.

[JL23a] Jacek Jendrej and Andrew Lawrie. Bubble decomposition for the harmonic map heat flow in the equivariant case. *Calc. Var. Partial Differential Equations*, 62(9):Paper No. 264, 36, 2023.

[JL23b] Jacek Jendrej and Andrew Lawrie. Soliton resolution for the energy-critical nonlinear wave equation in the radial case. *Ann. PDE*, 9(2):Paper No. 18, 117, 2023.

[JLPS19] Robert Jenkins, Jiaqi Liu, Peter Perry, and Catherine Sulem. The derivative nonlinear schrödinger equation: Global well-posedness and soliton resolution. *arXiv preprint arXiv:1905.02866*, 2019.

[JLS25] Jacek Jendrej, Andrew Lawrie, and Wilhelm Schlag. Continuous in time bubble decomposition for the harmonic map heat flow. *Forum Math. Pi*, 13:Paper No. e4, 37, 2025.

[KK24] Taegyu Kim and Soonsik Kwon. Soliton resolution for calogero–moser derivative nonlinear schr\" odinger equation. *arXiv preprint arXiv:2408.12843*, 2024.

[KKO22] Kihyun Kim, Soonsik Kwon, and Sung-Jin Oh. Soliton resolution for equivariant self-dual chern-simons-schrödinger equation in weighted sobolev class. *arXiv preprint arXiv:2202.07314*, 2022.

[Kla07] Sergiu Klainerman. Mathematical challenges of general relativity. *Rend. Mat. Appl. (7)*, 27(2):105–122, 2007.

[KM06] Carlos E. Kenig and Frank Merle. Global well-posedness, scattering and blow-up for the energy-critical, focusing, non-linear Schrödinger equation in the radial case. *Invent. Math.*, 166(3):645–675, 2006.

[NS85] Wei-Ming Ni and Paul Sacks. Singular behavior in nonlinear parabolic equations. *Trans. Amer. Math. Soc.*, 287(2):657–671, 1985.

[Oba72] Morio Obata. Conformal changes of riemannian metrics on a euclidean sphere. *Differential geometry (in honor of Kentaro Yano)*, pages 347–353, 1972.

[Pre24] Bruno Premoselli. A priori estimates for finite-energy sign-changing blowing-up solutions of critical elliptic equations. *Int. Math. Res. Not. IMRN*, (6):5212–5273, 2024.

[Qin95] Jie Qing. On singularities of the heat flow for harmonic maps from surfaces into spheres. *Comm. Anal. Geom.*, 3(1-2):297–315, 1995.

[QT97] Jie Qing and Gang Tian. Bubbling of the heat flows for harmonic maps from surfaces. *Comm. Pure Appl. Math.*, 50(4):295–310, 1997.

[Sch06] Peter Cornelis Schuur. *Asymptotic analysis of soliton problems: an inverse scattering approach*, volume 1232. Springer, 2006.

[Str84] Michael Struwe. A global compactness result for elliptic boundary value problems involving limiting nonlinearities. *Math. Z.*, 187(4):511–517, 1984.

[Str85] Michael Struwe. On the evolution of harmonic mappings of Riemannian surfaces. *Comment. Math. Helv.*, 60(4):558–581, 1985.

[Str94] Michael Struwe. The Yang-Mills flow in four dimensions. *Calc. Var. Partial Differential Equations*, 2(2):123–150, 1994.

[STZ92] J. Shatah and A. Tahvildar-Zadeh. Regularity of harmonic maps from the Minkowski space into rotationally symmetric manifolds. *Comm. Pure Appl. Math.*, 45(8):947–971, 1992.

[Top04] Peter Topping. Repulsion and quantization in almost-harmonic maps, and asymptotics of the harmonic map flow. *Ann. of Math. (2)*, 159(2):465–534, 2004.

[Wei79] Fred B. Weissler. Semilinear evolution equations in Banach spaces. *J. Functional Analysis*, 32(3):277–296, 1979.

[Wei80] Fred B. Weissler. Local existence and nonexistence for semilinear parabolic equations in L^p. *Indiana Univ. Math. J.*, 29(1):79–102, 1980.

[ZK65] N. J. Zabusky and M. D. Kruskal. Interaction of solitons in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 15:240–243, Aug 1965.

Shrey Aryan

Department of Mathematics, Massachusetts Institute of Technology
77 Massachusetts Avenue Cambridge, MA 02139-4307, USA
shrey183@mit.edu