NONEXISTENCE OF TYPE II BLOWUPS FOR AN ENERGY CRITICAL NONLINEAR HEAT EQUATION

KELEI WANG[†] AND JUNCHENG WEI[*]

ABSTRACT. We consider the energy critical semilinear heat equation

$$\begin{cases} \partial_t u - \Delta u = |u|^{\frac{4}{n-2}} u & \text{in } \mathbb{R}^n \times (0, T), \\ u(x, 0) = u_0(x), \end{cases}$$

where $n \geq 3$, $u_0 \in L^\infty(\mathbb{R}^n)$, and $T \in \mathbb{R}^+$ is the first blow up time. We prove that if $n \geq 7$ and $u_0 \geq 0$, then any blowup must be of Type I, i.e.,

$$\|u(\cdot, t)\|_{L^\infty(\mathbb{R}^n)} \leq C(T-t)^{-\frac{n-2}{4}}.$$

A similar result holds for interior blowups in the Cauchy-Dirichlet problem. The proof is built on several key ingredients: first we perform tangent flow analysis and study bubbling formation in this process; next we give a second order bubbling analysis in the multiplicity one case, where we use a reverse inner-outer gluing mechanism; finally, in the higher multiplicity case (bubbling tower/cluster), we develop Schoen's Harnack inequality and obtain next order estimates in Pohozaev identities for critical parabolic flows.

CONTENTS

1. INTRODUCTION

In this paper we consider the blowup problem for the nonlinear heat equation

$$\partial_t u - \Delta u = |u|^{p-1}u \tag{1.1}$$

where $p > 1$.

1.1. **Cauchy problem.** We first consider the Cauchy problem

$$\begin{cases} \partial_t u - \Delta u = |u|^{p-1}u & \text{in } \mathbb{R}^n \times (0, T), \\ u(x, 0) = u_0(x). \end{cases} \tag{1.2}$$

Here $u_0 \in L^\infty(\mathbb{R}^n)$, and $T \in \mathbb{R}^+$ is the first blow up time, that is, $u \in L^\infty(\mathbb{R}^n \times [0, t])$ for any $t < T$, while

$$\lim_{t \to T} \|u(t)\|_{L^\infty(\mathbb{R}^n)} = +\infty.$$

Problem (1.2) is one of classical nonlinear parabolic equations which has been studied extensively in recent decades. See the monograph by Quitter and Souplet [75] for backgrounds and references therein. Following the seminal work of Fujita [33], it is well-known that finite time blow-up must occur when the initial datum is sufficiently large (in some suitable sense). An important and fundamental question is the classification of the blowup near the first blowup time T. Since the equation (1.1) is invariant under the scaling

$$u^\lambda(x, t) := \lambda^{\frac{2}{p-1}} u(\lambda x, \lambda^2 t), \tag{1.3}$$

the blow up is divided into two types: it is **Type I** if there exists a constant C such that for any $t < T$,

$$\|u(\cdot, t)\|_{L^\infty(\mathbb{R}^n)} \leq C(T - t)^{-\frac{1}{p-1}}, \tag{1.4}$$

otherwise it is called **Type II**.

When p is subcritical (i.e. $1 < p < +\infty$ for $n = 1, 2$, or $1 < p < \frac{n+2}{n-2}$ for $n \geq 3$), in classical works of Giga-Kohn [37] (in the case of $u_0 \geq 0$) and Giga-Matsui-Sasayama [39] (in the case of sign-changing u_0), it is shown that any blow up is of Type I.

In contrast, less is known about the energy **critical** case, i.e. $n \geq 3$ and $p = \frac{n+2}{n-2}$, see e.g. [75, Open Problem 2.8 in Appendix I]. In this paper we establish the following

Theorem 1.1. *If $n \geq 7$, $p = \frac{n+2}{n-2}$ and $u_0 \geq 0$, then any blow up to (1.2) is of Type I.*

Some remarks on Theorem 1.1 are in order. First we remark that the dimension restriction in Theorem 1.1 is **optimal**: when $n \leq 6$ and $p = \frac{n+2}{n-2}$, Type II blowup solutions to (1.2) do exist. This was first predicted and proved formally in the pioneering work of Filippas-Herrero-Velázquez in [30] via method of matched asymptotics. The first rigorous proof of Type II blowup solutions in the radial setting is given by Schweyer in [79] for $n = 4$. (For nonradial setting and multiply blowups in dimension $n = 4$ we refer to [23].) For the remaining dimensions rigorous construction of Type II blowups are established recently: for $n = 3$, see [22], and for $n = 5$, we refer to [19, 41]. For $n = 6$, see Harada [42]. We should mention that when $p = \frac{n+2}{n-2}$ all the type II blowup solutions to (1.2) constructed so far are sign-changing. It is an open question if there are Type II blowups for positive solutions in low dimensions $n = 3, 4, 5, 6$.

Second we remark that the exponent restriction in Theorem 1.1 is also **optimal**: when $p > \frac{n+2}{n-2}$ many types of Type II blowup solutions have been found. The first example was discovered in the radial setting by Herrero and Velazquez in [43] for $p > p_{JL}$ where p_{JL} is the Joseph-Lundgreen exponent,

$$p_{JL} = \begin{cases} 1 + \frac{4}{n-4-2\sqrt{n-1}} & \text{if } n \geq 11 \\ +\infty, & \text{if } n \leq 10. \end{cases}$$

See also Mizoguchi [69] for the case of a ball, Seki [80] for the case of $p = p_{JL}$, and Collot [10] for the case of general domains with the restriction that $p > p_{JL}$ and p is an odd integer. A new and interesting anisotropic Type II blow-up solutions is also constructed recently for $p > p_{JL}(n-1), n \geq 14$ by Collot, Merle and Raphael in [13]. In the intermediate exponent regime $\frac{n+2}{n-2} < p < p_{JL}$, Matano and Merle [63, 64] proved that no type II blow-up is present for radial solutions (under some extra technical conditions). However type II blow-up also exists in this intermediate regime. In [21], the authors successfully constructed non-radial type II blow-up solution in the Matano-Merle range $p = \frac{n+1}{n-3} \in \left(\frac{n+2}{n-2}, p_{JL}(n)\right)$. Another type of non-radial Type blow-up with shrinking spheres is recently found for $n \geq 5, p = 4$ in [18].

Theorem 1.1 is the first instance of classification in the critical exponent case for general initial datum. In the pioneering work of Filippas-Herrero-Velázquez [30] it is shown that blowup is type I if the initial data is positive and radially decreasing, when $n \geq 3$. For initial datum with low energy, we mention that in [11], Collot, Merle and Raphael proved a result similar to Theorem 1.1 under the condition that $\|u_0 - W\|_{\dot{H}^1(\mathbb{R}^n)} \ll 1$, where W is a positive Aubin-Talenti solution satisfying (1.9), i.e. positive steady state of (1.2). More precisely assume that $\|u_0 - W\|_{\dot{H}^1(\mathbb{R}^n)} \ll 1$ and the dimension $n \geq 7$, then there is a trichotomy to solution to (1.2): it either

dissipates to zero, or approaches to a rescaled Aubin-Talenti solution, or blows up in finite time and the blow-up is of Type I. Note that in Theorem 1.1, we have assumed neither decaying nor energy condition.

1.2. **Outline of proof.** The proof of Theorem 1.1 consists of the following five steps.

(1) First we perform tangent flow analysis at a possible blow up point, see Part 4. This is in the same spirit of Giga-Kohn [36, 37, 38], but we rewrite it in the form which is more familiar in geometric measure theory such as the tangent cone analysis/ blowing up analysis/tangent flow analysis used in the study of minimal surfaces, mean curvature flows and many other geometric variational problems.

(2) Bubbles may appear during this blow up procedure. A general theory on the energy concentration phenomena in (1.1) is then developed in Part 1, where we mainly follow the treatment (on harmonic map heat flows) in Lin-Wang [55, 56, 57], see also their monograph [58].

(3) We then perform a refined analysis of this bubbling phenomena, first in the case of one bubble (*multiplicity one case*), see Part 2. Here we mainly use *the inner-outer gluing mechanism* developed by the second author with Davila, del Pino and Musso in [15], [16], [19], [20], [73], [22] (see also [17] for a survey);

(4) In Part 3, we combine the analysis in Part 1 and Part 2 to establish the refined analysis in the general case (*higher multiplicity*), where in particular we will exclude the **bubble towering** formation in this energy concentration phenomena. Arguments used here are motivated by those used in the study of Yamabe problems through pioneering work of Schoen [78], including secondary blow ups, construction of Green functions, next order expansion of Pohozaev identities, see for example Schoen [77, 78], Kuhri-Marques-Schoen [46], Li-Zhu [50], Li-Zhang [48]. In fact, the Pohozaev identity enters the proof twice: for the second one it is used in the same way as in the Yamabe problem, but for the first one it is used in combination with *the reverse inner-outer gluing method* to give a next order expansion of this identity.

(5) Finally, the refined analysis in Part 3 and Part 2 are applied to the first time singularity problem. Results in Part 3 are used to exclude blow ups with **bubble clustering**, while results in Part 2 are used to exclude blow ups with only one bubble.

In [87, 88] we have performed *second order estimates* for stable solutions of singularly perturbed Allen-Cahn equation and as a consequence we have obtained $C^{2,\alpha}$ estimates for interfaces. The present work can be considered a natural generalization of the second order estimates for elliptic equations to nonlinear parabolic equations which are more involved. The ideas presented in this paper may be useful in studying singularities occurring in geometric flows.

1.3. **Cauchy-Dirichlet problem.** Our method can also be applied to the initial-boundary value problem

$$
\begin{cases}
\partial_t u - \Delta u = |u|^{p-1}u & \text{in } \Omega^T := \Omega \times (0, T), \\
u = 0 & \text{on } \partial\Omega \times (0, T), \\
u(x, 0) = u_0(x).
\end{cases}
\tag{1.5}
$$

Here $\Omega \subset \mathbb{R}^n$ is a bounded domain with C^2 boundary, $u_0 \in L^q(\Omega)$ for some $q \geq n(p-1)/2$. (The exponent $\frac{n(p-1)}{2}$ is optimal. See [70].) By [75, Theorem 15.2] or [7] and [90], there exists a $T > 0$ such that $u \in L^\infty(\overline{\Omega} \times (0, t])$ for any $t < T$. Here we assume u blows up in finite time, and T is the first blow up time, that is,

$$
\lim_{t \to T} \|u(\cdot, t)\|_{L^\infty(\Omega)} = +\infty.
$$

For this Cauchy-Dirichlet problem we prove

Theorem 1.2. *If $n \geq 7$, $p = \frac{n+2}{n-2}$ and $u_0 \geq 0$, the first time singularity in the interior must be Type I, that is, for any $\Omega' \Subset \Omega$, there exists a constant C such that*

$$
\|u(\cdot, t)\|_{L^\infty(\Omega')} \leq C(T - t)^{-\frac{1}{p-1}}, \quad \forall 0 < t < T.
$$

The proof of this theorem is similar to the one for Theorem 1.1. It is also generally believed that there is no boundary blow up. This, however, is only known in some special cases, e.g. when Ω is convex (cf. [38, Theorem 5.3], [40]).

Once we have this Type I blow up bound, it will be interesting to know if the set of blow up points enjoy further regularities, and if the blow up profiles satisfy the uniform, refined estimates as in the subcritical case, see for example Liu [60], Filippas-Kohn [31], Filippas-Liu [32], Velázquez [83, 84, 85], Merle-Zaag [65, 66, 67, 68], Zaag [94, 95, 96] and Kammerer-Merle-Zaag [29].

Our proof also implies that the energy collapses and the L^{p+1} norm blows up at T, cf. Giga [35], Zaag [93] and Du [27].

Corollary 1.3. *Under the assumptions of Theorem 1.2, if there exists an interior blow up point, then*

$$
\lim_{t \to T} \int_\Omega \left[\frac{1}{2}|\nabla u(x, t)|^2 - \frac{1}{p+1}u(x, t)^{p+1} \right] dx = -\infty,
\tag{1.6}
$$

and

$$
\lim_{t \to T} \int_\Omega u(x, t)^{p+1} dx = +\infty.
\tag{1.7}
$$

In fact, we prove that

$$
\lim_{t \to T} \int_0^t \int_\Omega |\partial_t u(x, t)|^2 dx = +\infty.
\tag{1.8}
$$

Then (1.6) follows from the standard energy identity for u. A local version of (1.8) and (1.6) also hold for the solution of the Cauchy problem (1.2).

We do not know if the $L^2(\Omega)$ norm of $\nabla u(t)$ blows up as $t \to T$. We conjecture that the blow up must be complete, i.e. the solution cannot be extended beyond T

as a weak solution, cf. Baras-Cohen [3], Galaktinov-Vazquez [34], Martel [62] and [75, Section 27].

1.4. **List of notations and conventions used throughout the paper.**

- The open ball in $\mathbb{R}^n$ is denoted by $B_r(x)$, and by B_r if the center is the origin 0.
- The parabolic cylinder is $Q_r(x,t) := B_r(x) \times (t-r^2, t+r^2)$, the forward parabolic cylinder is $Q_r^+(x,t) := B_r(x) \times (t, t+r^2)$, and the backward parabolic cylinder is $Q_r^-(x,t) := B_r(x) \times (t-r^2, t)$. If the center is the origin $(0,0)$, it will not be written down explicitly.
- The parabolic distance is
$$\operatorname{dist}_P((x,t),(y,s)) := \max\{|x-y|, \sqrt{|t-s|}\}.$$
Hölder, Lipschitz continuous functions with respect to this distance is defined as usual.
- Given a domain $\Omega \subset \mathbb{R}^n$, $H^1(\Omega)$ is the Sobolev space endowed with the norm
$$\left[\int_\Omega \left(|\nabla u|^2 + u^2\right) dx\right]^{1/2}.$$
Given an interval $\mathcal{I} \subset \mathbb{R}$, we use $L_t^2 H_x^1$ to denote the space $L^2(\mathcal{I}; H^1)$.
- A bubble is an entire solution of the stationary equation
$$-\Delta W = |W|^{\frac{4}{n-2}} W \quad \text{in } \mathbb{R}^n \tag{1.9}$$
with finite energy
$$\int_{\mathbb{R}^n} |\nabla W|^2 = \int_{\mathbb{R}^n} |W|^{\frac{2n}{n-2}} < +\infty.$$
By the translational and scaling invariance, if W is a bubble, so is
$$W_{\xi,\lambda}(x) := \lambda^{\frac{n-2}{2}} W\left(\frac{x-\xi}{\lambda}\right), \quad \xi \in \mathbb{R}^n, \ \lambda \in \mathbb{R}^+.$$
- By Caffarelli-Gidas-Spruck [8], all entire positive solutions to the stationary equation (1.9) are given by Aubin-Talenti bubbles
$$W_{\xi,\lambda}(x) := \left(\frac{\lambda}{\lambda^2 + \frac{|x-\xi|^2}{n(n-2)}}\right)^{\frac{n-2}{2}}, \qquad \lambda > 0, \quad \xi \in \mathbb{R}^n. \tag{1.10}$$
They have finite energy, which are always equal to
$$\Lambda := \int_{\mathbb{R}^n} |\nabla W_{\xi,\lambda}|^2 = \int_{\mathbb{R}^n} W_{\xi,\lambda}^{p+1}. \tag{1.11}$$
For simplicity, denote $W := W_{0,1}$.
- We use C (large) and c (small) to denote universal constants. They could be different from line to line. Given two quantities A and B, if $A \leq CB$ for some universal constant C, we simply write $A \lesssim B$. If the constant C depends on a quantity D, it will be written as $C(D)$ or $\lesssim_D$.

Part 1. **Energy concentration behavior**

2. Setting

In this part we assume that $p > 1$ (which may not necessarily be the critical exponent), and that the solutions could be sign-changing. Here we study the energy concentration behavior of the nonlinear heat equation (1.1).

First we need to define what we call as a solution.

Definition 2.1 (Suitable weak solution). *A function u is a suitable weak solution of* (1.1) *in Q_1, if $\partial_t u, \nabla u \in L^2(Q_1)$, $u \in L^{p+1}(Q_1)$, and*

- *u satisfies* (1.1) *in the weak sense, that is, for any $\eta \in C_0^\infty(Q_1)$,*

$$\int_{Q_1} \left[\partial_t u \eta + \nabla u \cdot \nabla \eta - |u|^{p-1} u \eta \right] = 0; \tag{2.1}$$

- *u satisfies the localized energy inequality: for any $\eta \in C_0^\infty(Q_1)$,*

$$\int_{Q_1} \left[\left(\frac{|\nabla u|^2}{2} - \frac{|u|^{p+1}}{p+1} \right) \partial_t \eta^2 - |\partial_t u|^2 \eta^2 - 2\eta \partial_t u \nabla u \cdot \nabla \eta \right] \geq 0; \tag{2.2}$$

- *u satisfies the stationary condition: for any $Y \in C_0^\infty(Q_1, \mathbb{R}^n)$,*

$$\int_{Q_1} \left[\left(\frac{|\nabla u|^2}{2} - \frac{|u|^{p+1}}{p+1} \right) divY - DY(\nabla u, \nabla u) + \partial_t u \nabla u \cdot Y \right] = 0. \tag{2.3}$$

A smooth solution satisfies all of these conditions. (In fact, (2.2) will become an equality, which is just the standard energy identity.) But a suitable weak solution need not to be smooth everywhere.

Definition 2.2. *Given a weak solution u of* (1.1), *a point (x, t) is a regular point of u if there exists an $r > 0$ such that $u \in C^\infty(Q_r(x, t))$, otherwise it is a singular point of u. The corresponding sets are denoted by $\mathcal{R}(u)$ and $\mathcal{S}(u)$.*

By definition, $\mathcal{R}(u)$ is open and $\mathcal{S}(u)$ is closed.

In this part, u_i denotes a sequence of suitable weak solutions to (1.1) in Q_1, satisfying a uniform energy bound

$$\limsup_{i \to +\infty} \int_{Q_1} \left[|\nabla u_i|^2 + |u_i|^{p+1} + |\partial_t u_i|^2 \right] dxdt < +\infty. \tag{2.4}$$

The integral bound on $\partial_t u_i$ in (2.4) can be deduced by substituting the bounds on the first two integrands into (2.2), if we shrink the domain Q_1 a little.

We will state the results about the energy concentration behavior after introducing some necessary notations. The main results in this part are

(1) Theorem 4.2, where we establish basic properties about this energy concentration behavior;
(2) Theorem 6.1 about the case $2\frac{p+1}{p-1}$ is not an integer, where we prove a strong convergence result;
(3) Theorem 8.1 about the case $2\frac{p+1}{p-1}$ is an integer, where we show that the limiting problem is a generalized Brakke's flow;

(4) Theorem 9.1, which is about the energy quantization result in the critical case $p = \frac{n+2}{n-2}$.

Our treatment in this part mainly follows the work of Lin and Wang [55], [56], [57]. See also their monograph [58].

There are still many problems remaining open about this energy concentration behavior, such as the energy quantization result in the general case (cf. Lin-Riviere [53] for the corresponding results for harmonic maps, and Naber-Valtorta [74] for Yang-Mills fields), Brakke type regularity result for the limiting problem (cf. Brakke [6], Kasai-Tonegawa [45]). But we will content with such a preliminary analysis of this energy concentration phenomena, because the main goal of this part is to providing a setting for later parts.

2.1. **List of notations and conventions used in this part.**

- For any $\lambda > 0$ and each set $A \subset \mathbb{R}^n \times \mathbb{R}$, we use λA to denote the parabolic scaling of A, which is the set

$$\left\{ (\lambda x, \lambda^2 t) : (x, t) \in A \right\}.$$

- Given a set $A \subset \mathbb{R}^n \times \mathbb{R}$, its time slices are $A_t := A \cap (\mathbb{R}^n \times \{t\})$.
- Denote $m := \frac{2(p+1)}{p-1}$. Note that $m > 2$, and $p = \frac{m+2}{m-2}$. If m is an integer, then p is the Sobolev critical exponent in dimension m.
- For $s \geq 0$, $\mathcal{P}^s$ denotes the s-dimensional parabolic Hausdorff measure. The dimension of a subset of $\mathbb{R}^n \times \mathbb{R}$ is always understood to be the parabolic Hausdorff dimension.

3. MONOTONICITY FORMULA AND ε-REGULARITY

In this section, u denotes a fixed suitable weak solution of (1.1) in Q_1. Here we recall the monotonicity formula and ε-regularity theorems. We also derive a Morrey space estimate, which will be used below in Section 5 to preform the tangent flow analysis. Our use of Morrey space estimates follows closely Chou-Du-Zheng [9] and Du [25, 27], which is different from the ones used in Blatt-Struwe [5] and Souplet [81].

Fix a function $\psi \in C_0^\infty(B_1)$ such that $0 \leq \psi \leq 1$, $\psi \equiv 1$ in $B_{3/4}$ and $|\nabla \psi| + |\nabla^2 \psi| \leq 100$. For $t > 0$ and $x \in \mathbb{R}^n$, let

$$G(x, t) := (4\pi t)^{-\frac{n}{2}} e^{-\frac{|x|^2}{4t}}$$

be the standard heat kernel on $\mathbb{R}^n$.

Take a large constant C and a small constant c. For each $(x, t) \in B_{3/4} \times (-3/4, 1]$ and $s \in (0, 1/4)$, define

$$
\begin{aligned}
\Theta_s(x, t; u) \;\; := \;\; & s^{\frac{p+1}{p-1}} \int_{B_1} \left[\frac{|\nabla u(y, t - s)|^2}{2} - \frac{|u(y, t - s)|^{p+1}}{p+1} \right] G(y - x, s) \psi(y)^2 dy \\
& + \;\; \frac{1}{2(p-1)} s^{\frac{2}{p-1}} \int_{B_1} u(y, t - s)^2 G(y - x, s) \psi(y)^2 dy + C e^{-cs^{-1}}.
\end{aligned}
$$

Remark 3.1. *Because u and ∇u are only integrable in space-time, rigorously we should integrate one more time in s, e.g. to consider the quantity*

$$\overline{\Theta}_s(x,t;u) := s^{-1} \int_s^{2s} \Theta_\tau(x,t;u) d\tau.$$

However, to simplify notations we will not use this quantity.

The following is a localized version of the monotonicity formula of Giga and Kohn, see [36, 38] and [39].

Proposition 3.2 (Localized monotonicity formula). *If C is universally large and c is universally small, then for any $(x,t) \in B_{3/4} \times (-3/4, 1]$ and $0 < s_1 < s_2 < 1/4$,*

$$\begin{aligned}
&\Theta_{s_2}(x,t) - \Theta_{s_1}(x,t) \\
&\geq \int_{s_1}^{s_2} \tau^{\frac{2}{p-1}-1} \int_{B_1} \left| (t-\tau)\partial_t u(y,t-\tau) + \frac{u(y,t-\tau)}{p-1} + \frac{y}{2} \cdot \nabla u(y,t-\tau) \right|^2 \\
&\qquad \times G(y-x,\tau)\psi(y)^2 dy d\tau.
\end{aligned}$$

This almost monotonicity allows us to define

$$\Theta(x,t) := \lim_{s \to 0^+} \Theta_s(x,t).$$

For each $s > 0$, from the definition we see $s^{-1} \int_s^{2s} \Theta_\tau(x,t) d\tau$ is a continuous function of (x,t). Combining this fact with the monotonicity formula, we get

Lemma 3.3. $\Theta(x,t)$ *is upper semi-continuous in* (x,t).

The following Morrey space bound is essentially Giga-Kohn's [37, Proposition 2.2] or [38, Proposition 3.1].

Proposition 3.4. *For any $(x,t) \in Q_{1/2}$ and $r \in (0, 1/4)$,*

$$\begin{aligned}
&r^{m-2-n} \int_{Q_r^-(x,t-r^2)} \left(|\nabla u|^2 + |u|^{p+1} \right) + r^{m-n} \int_{Q_r^-(x,t-r^2)} |\partial_t u|^2 \\
&\leq C \max \left\{ \Theta_{4r^2}(x,t), \Theta_{4r^2}(x,t)^{\frac{2}{p+1}} \right\}.
\end{aligned} \tag{3.1}$$

Combining this estimate with the monotonicity formula, we obtain a Morrey space estimate of u.

Corollary 3.5 (Morrey space bound). *There exists a universal constant M depending on $\int_{Q_1} [|\nabla u|^2 + |u|^{p+1}]$ such that for any $(x,t) \in Q_{1/2}$ and $r \in (0, 1/4)$,*

$$r^{m-2-n} \int_{Q_r(x,t)} \left(|\nabla u|^2 + |u|^{p+1} \right) + r^{m-n} \int_{Q_r(x,t)} |\partial_t u|^2 \leq M. \tag{3.2}$$

These Morrey space estimates are invariant under the scaling (1.3).
Combining (3.1) and (3.2), we also get

Corollary 3.6 (Change of base point). *For any $\varepsilon > 0$, there exist $0 < \delta(\varepsilon) \ll \theta(\varepsilon) \ll 1$ and $C(\varepsilon)$ so that the following holds. For any $(x,t) \in Q_{1/2}$, $r < 1/4$ and $(y,s) \in Q_{\delta r}(x,t)$,*

$$\Theta_{\theta r^2}(y,s) \le C \max \left\{ \Theta_{r^2}(x,t), \Theta_{r^2}(x,t)^{\frac{2}{p+1}} \right\} + \varepsilon.$$

Next we recall two standard ε-regularity theorems. The first one is a reformulation of Du [27, Theorem 3.1].

Theorem 3.7 (ε-regularity I). *There exist three universal constants ε_*, δ_* and c_* so that the following holds. If*

$$\Theta_{r^2}(x,t) \le \varepsilon_*,$$

then

$$\sup_{Q_{\delta_* r}(x,t)} |u| \le c_* r^{-\frac{2}{p-1}}.$$

Once we have an L^∞ bound on u, higher order regularity follows by applying standard parabolic estimates to (1.1). As a consequence, we get

Corollary 3.8. $\mathcal{R}(u) = \{\Theta = 0\}$ *and* $\mathcal{S}(u) = \{\Theta \ge \varepsilon_*\}$.

The second ε-regularity theorem is Chou-Du-Zheng [9, Theorem 2] (see also Du [27, Proposition 4.1]).

Theorem 3.9 (ε-regularity II). *After decreasing ε_*, δ_*, and enlarging c_* further, the following holds. There exists a constant $\theta_* > 0$ such that, if*

$$\int_{Q_r^-(x,t-\theta_* r^2)} \left(|\nabla u|^2 + |u|^{p+1} \right) \le \varepsilon_* r^{n+2-m},$$

then

$$\sup_{Q_{\delta_* r}(x,t)} |u| \le c_* r^{-\frac{2}{p-1}}.$$

A standard covering argument then gives

Corollary 3.10. *If u is a suitable weak solution of (1.1) in Q_1, then*

$$\mathcal{P}^{n+2-m} \left(\mathcal{S}(u) \right) = 0.$$

In particular, $\dim\left(\mathcal{S}(u)\right) \le n + 2 - m$.

4. Defect measures

From now on until Section 9, we will be concerned with the energy concentration behavior in (1.1). In this section we define the defect measure and prove some basic properties about it.

4.1. **Definition and basic properties.** By (2.4), after passing to a subsequence, we may assume u_i converges weakly to a limit u_∞ in $L^{p+1}(Q_1)$, ∇u_i and $\partial_t u_i$ converges weakly to ∇u_∞ and $\partial_t u_\infty$ respectively in $L^2(Q_1)$.

By Sobolev embedding theorem and an interpolation argument, for any $1 \le q < p + 1$, u_i converges to u_∞ strongly in $L^q_{loc}(Q_1)$. As a consequence, u_∞ is a weak solution of (1.1).

There exist three Radon measures μ_1, μ_2 and ν such that

$$\begin{cases} |u_i|^{p+1}dxdt \rightharpoonup |u_\infty|^{p+1}dxdt + \mu_1, \\ |\nabla u_i|^2 dxdt \rightharpoonup |\nabla u_\infty|^2 dxdt + \mu_2, \\ |\partial_t u_i|^2 dxdt \rightharpoonup |\partial_t u_\infty|^2 dxdt + \nu. \end{cases}$$

These measures, called *defect measures*, characterize the failure of corresponding strong convergence.

We also let

$$\nabla u_i \otimes \nabla u_i dxdt \rightharpoonup \nabla u_\infty \otimes \nabla u_\infty dxdt + \mathcal{T}d\mu_2,$$

where $\mathcal{T}$ is a matrix valued μ_2-measurable functions. Furthermore, $\mathcal{T}$ is symmetric, semi-positive definite and its trace equals 1 μ_2-a.e.

Lemma 4.1 (Energy partition). $\mu_1 = \mu_2$.

Proof. For any $\eta \in C_0^\infty(Q_1)$, multiplying the equation (1.1) by $u_i \eta^2$ and integrating by parts, we obtain

$$\int_{Q_1} \left[-u_i^2 \eta \partial_t \eta + |\nabla u_i^2| \eta^2 - |u_i|^{p+1} \eta^2 + 2\eta u_i \nabla u_i \nabla \eta \right] = 0. \tag{4.1}$$

Letting $i \to \infty$, and noting that u_∞ also satisfies (4.1), we obtain

$$\int_{Q_1} \eta^2 d\left(\mu_1 - \mu_2\right) = 0.$$

Since this holds for any $\eta \in C_0^\infty(Q_1)$, $\mu_1 = \mu_2$ as Radon measures. $\qquad\square$

By this lemma, we can write μ_1 and μ_2 just as μ. Denote $\Sigma := \mathrm{spt}(\mu)$, and define

$$\Sigma^* := \left\{ (x,t): \ \forall r > 0, \ \limsup_{i \to \infty} r^{m-2-n} \int_{Q_r^-(x,t)} \left(|\nabla u_i|^2 + |u_i|^{p+1} \right) \ge \varepsilon_*/2 \right\}$$

to be *the blow up locus*.

Now we state some basic properties on this energy concentration phenomena.

Theorem 4.2. *Suppose u_i is a sequence of suitable weak solutions of (1.1) in Q_1, satisfying (2.4). Define μ, ν, Σ and Σ^* as above. Then the following holds.*

(1) *(Morrey space bound) For any $(x,t) \in Q_{1/2}$ and $r \in (0, 1/4)$,*

$$\begin{cases} \mu(Q_r(x,t)) + \displaystyle\int_{Q_r(x,t)} \left(|\nabla u_\infty|^2 + |u_\infty|^{p+1} \right) \le Mr^{n+2-m}, \\ \nu(Q_r(x,t)) + \displaystyle\int_{Q_r(x,t)} |\partial_t u_\infty|^2 \le Mr^{n-m}. \end{cases} \tag{4.2}$$

Here M is the constant in Corollary 3.5.

(2) *The blow up locus Σ^* is closed.*
(3) *(Smooth convergence) u_i converges to u_∞ in $C^\infty_{loc}(Q_1 \setminus \Sigma^*)$. As a consequence,*
  - *$u_\infty \in C^\infty(Q_1 \setminus \Sigma^*)$, that is, $\mathcal{S}(u_\infty) \subset \Sigma^*$.*
  - *$\Sigma \subset \Sigma^*$ and $spt(\nu) \subset \Sigma^*$.*
(4) *(Measure estimate of the blow up locus) For any $(x,t) \in \Sigma^* \cap Q_{1/2}$ and $0 < r < 1/2$,*

$$\mathcal{P}^{n+2-m}(\Sigma^* \cap Q_r(x,t)) \leq \frac{C}{\varepsilon_*} r^{n+2-m}.$$

(5) *(Lower density bound) For μ-a.e. $(x,t) \in \Sigma \cap Q_{1/2}$ and $r \in (0, 1/2)$,*

$$\mu(Q_r(x,t)) \geq \mu(Q^-_{r/2}(x, t - \theta_* r^2)) \geq \varepsilon_* r^{n+2-m}. \tag{4.3}$$

(6) $\mathcal{P}^{n+2-m}\left(\Sigma^* \setminus \Sigma\right) = 0.$
(7) *There exists a measurable function θ on Σ such that*

$$\mu = \theta(x,t)\mathcal{P}^{n+2-m}\lfloor_\Sigma.$$

Moreover,

$$\frac{\varepsilon_*}{C} \leq \theta \leq C \qquad \mathcal{P}^{n+2-m} - a.e. \ in \ \Sigma.$$

Before presenting the proof, we first note the following Federer-Ziemer type result [28]. It can be proved by a Vitali covering argument.

Lemma 4.3. (1) *For $\mathcal{P}^{n+2-m}$ a.e. $(x,t) \in Q_{1/2}$,*

$$\lim_{r \to 0} r^{m-2-n} \int_{Q_r(x,t)} \left(|\nabla u_\infty|^2 + |u_\infty|^{p+1} \right) = 0.$$

(2) *For $\mathcal{P}^{n-m}$ a.e. $(x,t) \in Q_{1/2}$,*

$$\lim_{r \to 0} r^{m-n} \int_{Q_r(x,t)} |\partial_t u_\infty|^2 = 0.$$

Proof of Theorem 4.2. (1) This Morrey space bound follows directly by passing to the limit in (3.1) (for u_i).
(2) By definition, for any $(x,t) \notin \Sigma^*$, there exists an $r > 0$ such that for all i large,

$$\int_{Q^-_r(x,t)} \left(|\nabla u_i|^2 + |u_i|^{p+1} \right) < \varepsilon_* r^{n+2-m}.$$

By Theorem 3.9 and standard parabolic regularity theory, u_i are uniformly bound in $C^k(Q_{\delta_* r}(x,t))$ for each $k \in \mathbb{N}$. By the weak convergence of u_i and Arzela-Ascolli theorem, they converge to u_∞ in $C^\infty(Q_{\delta_* r}(x,t))$.
(3) This follows directly from the previous point.
(4) This follows from a standard Vitali type covering argument, by utilizing the lower density bound coming from the definition of Σ^*, that is, for any $(x,t) \in \Sigma^* \cap Q_{1/2}$ and $0 < r < 1/2$,

$$\mu(Q_r(x,t)) + \int_{Q_r(x,t)} \left(|\nabla u_\infty|^2 + |u_\infty|^{p+1} \right) \geq \frac{\varepsilon_*}{2} r^{n+2-m}. \tag{4.4}$$

(5) This follows by combining (4.4) and Lemma 4.3.
(6) This follows from a Vitali type covering argument, by using (4.4) and Lemma 4.3.
(7) This follows from differentiation theorem for measures, and an application of (4.2) and (4.3). □

4.2. **Definition of the mean curvature.** In this subsection we define a mean curvature type term for defect measures.

Lemma 4.4. *There exists an $\mathbb{R}^n$ valued function $\mathbf{H} \in L^2(Q_1, d\mu)$ such that*

$$\partial_t u_i \nabla u_i dxdt \rightharpoonup \partial_t u_\infty \nabla u_\infty dxdt + \frac{1}{m}\mathbf{H}d\mu$$

weakly as Radon measures.

Proof. By Cauchy-Schwarz inequality,

$$\int_{Q_1} |\partial_t u_i||\nabla u_i| \leq \left(\int_{Q_1} |\partial_t u_i|^2\right)^{1/2} \left(\int_{Q_1} |\nabla u_i|^2\right)^{1/2} \tag{4.5}$$

are bounded as $i \to +\infty$. Therefore we may assume

$$\partial_t u_i \nabla u_i dxdt \rightharpoonup \xi \quad \text{weakly as vector valued Radon measures,}$$

where ξ is an $\mathbb{R}^n$-valued Radon measure.

Take the Radon-Nikodym decomposition of ξ with respect to the Lebesgue measure, $\xi = \xi^a + \xi^s$, where ξ^a is the absolute continuous part, and ξ^s is the singular part.

By Point (3) in Theorem 4.2,

$$\xi = \partial_t u_\infty \nabla u_\infty dxdt \quad \text{outside } \Sigma^*.$$

In view of Point (4) in Theorem 4.2, this is just the absolutely continuous part ξ^a. In other words,

$$\xi^a = \partial_t u_\infty \nabla u_\infty dxdt. \tag{4.6}$$

On the other hand, we can also define

$$\mathbf{H}_i := \begin{cases} \frac{\partial_t u_i \nabla u_i}{|\nabla u_i|^2}, & \text{if } |\nabla u_i| \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

It satisfies

$$\int_{Q_1} |\mathbf{H}_i|^2 |\nabla u_i|^2 dxdt \leq \int_{Q_1} |\partial_t u_i|^2 dxdt.$$

By Hutchinson [44], we may assume $(\mathbf{H}_i, |\nabla u_i|^2 dxdt)$ converges to $(\widetilde{\mathbf{H}}, |\nabla u_\infty|^2 dxdt + \mu)$ weakly as measure-functions pairs. By Fatou lemma,

$$\int_{Q_1} |\widetilde{\mathbf{H}}|^2 |\nabla u_\infty|^2 dxdt + \int_{Q_1} |\widetilde{\mathbf{H}}|^2 d\mu < +\infty. \tag{4.7}$$

Since ξ is the weak limit of $\partial_t u_i \nabla u_i dxdt$, we have

$$\xi = \widetilde{\mathbf{H}} \left(|\nabla u_\infty|^2 dxdt + \mu\right). \tag{4.8}$$

By (4.6),

$$\widetilde{\mathbf{H}} = \begin{cases} \frac{\partial_t u_\infty \nabla u_\infty}{|\nabla u_\infty|^2}, & \text{if } |\nabla u_\infty| \neq 0 \\ 0, & \text{otherwise,} \end{cases}$$

a.e. with respect to the Lebesgue measure in Q_1. (In fact, this holds everywhere in $Q_1 \setminus \Sigma^*$.) Hence in view of (4.6), we get

$$\xi^s = \widetilde{\mathbf{H}} d\mu.$$

In particular, $\widetilde{\mathbf{H}}$ is the Radon-Nikodym derivative $d\xi^s/d\mu$. The proof is complete by defining $\mathbf{H} := m\widetilde{\mathbf{H}}$. □

Similar to (4.5), we obtain

Corollary 4.5. *For each $Q_r(x,t) \subset Q_1$,*

$$\frac{1}{m^2} \int_{Q_r(x,t)} |\mathbf{H}|^2 d\mu \leq \nu(Q_r(x,t))\mu(Q_r(x,t)).$$

A more precise estimate will be given in Lemma 8.3 below.

Passing to the limit in (2.2) gives the limiting energy inequality:

$$\begin{aligned}
0 \leq & \int_{Q_1} \left[\left(\frac{|\nabla u_\infty|^2}{2} - \frac{|u_\infty|^{p+1}}{p+1} \right) \partial_t \eta^2 - |\partial_t u_\infty|^2 \eta^2 - 2\eta \partial_t u_\infty \nabla u_\infty \cdot \nabla \eta \right] \\
& + \frac{1}{m} \int_{Q_1} \partial_t \eta^2 d\mu - \int_{Q_1} \eta^2 d\nu - \frac{1}{m} \int_{Q_1} \nabla \eta^2 \cdot \mathbf{H} d\mu.
\end{aligned} \tag{4.9}$$

Passing to the limit in (2.3) gives the limiting stationary condition: for any $Y \in C_0^\infty(Q_1, \mathbb{R}^n)$,

$$\begin{aligned}
0 = & \int_{Q_1} \left[\left(\frac{|\nabla u_\infty|^2}{2} - \frac{|u_\infty|^{p+1}}{p+1} \right) \operatorname{div} Y - DY(\nabla u_\infty, \nabla u_\infty) + \partial_t u_\infty \nabla u_\infty \cdot Y \right] \\
& + \frac{1}{m} \int_{Q_1} \left[(I - m\mathcal{T}) \cdot DY + \mathbf{H} \cdot Y \right] d\mu.
\end{aligned} \tag{4.10}$$

4.3. **Limiting monotonicity formula.** In this section we establish the monotonicity formula for the limit (u_∞, μ). For this purpose, we need first to define the time slices of μ.

Lemma 4.6. *There exists a family of Radon measures μ_t on B_1 (defined for a.e. $t \in (-1,1)$) such that*

$$\mu = \mu_t dt. \tag{4.11}$$

Proof. Denote the projection onto the time axis by π. For each $r < 1$, let μ^r be the restriction of μ to $B_r \times (-1,1)$.

Take a function $\eta_r \in C_0^\infty(B_1)$, $0 \leq \eta_r \leq 1$ and $\eta_r \equiv 1$ in B_r. The limiting energy inequality (4.9) implies that

$$\int_{B_1} \left(\frac{1}{2} |\nabla u_\infty(x,t)|^2 - \frac{1}{p+1} |u_\infty(x,t)|^{p+1} \right) \eta_r(x)^2 dx + \frac{1}{m} \pi_\sharp(\eta_r \mu)$$

is a BV function on $(-1,1)$. Hence there exists a function $\overline{e}_r(t) \in L^1(-1,1)$ such that

$$\pi_\sharp(\eta_r\mu) = \overline{e}_r(t)dt,$$

Because

$$0 \le \pi_\sharp\mu^r \le \pi_\sharp(\eta_r\mu),$$

we find another function $e_r(t) \in L^1(-1,1)$ such that

$$\pi_\sharp\mu^r = e^r(t)dt.$$

By the disintegration theorem, there exists a family of probability measure $\overline{\mu}_t$ on $(-1,1)$ such that

$$\mu^r = \overline{\mu}_t d(\pi_\sharp\mu^r).$$

By defining

$$\mu_t := e_r(t)\overline{\mu}_t,$$

we get (4.11). $\hfill\square$

Remark 4.7. *Unlike harmonic map heat flows, because the energy density for* (1.1) *is sign-changing, we do not know if*

$$\left[\frac{1}{2}|\nabla u_\infty(x,t)|^2 - \frac{1}{p+1}|u_\infty(x,t)|^{p+1}\right]dx + \mu_t$$

is a well-defined measure for all t. *Similarly, we also do not know if there is an estimate on the Hausdorff measure of* Σ_t^*.

Define $\Theta_s(x,t;u_\infty,\mu)$ to be

$$s^{\frac{p+1}{p-1}}\int_{B_1}\left[\frac{|\nabla u_\infty(y,t-s)|^2}{2} - \frac{|u_\infty(y,t-s)|^{p+1}}{p+1}\right]G(y-x,s)\psi(y)^2dy$$

$$+ \frac{1}{2(p-1)}s^{\frac{2}{p-1}}\int_{B_1}u_\infty(y,t-s)^2G(y-x,s)\psi(y)^2dy$$

$$+ \frac{1}{m}s^{\frac{p+1}{p-1}}\int_{\mathbb{R}^n}G(y-x,s)\psi(y)^2d\mu_{t-s}(y) + Ce^{-cs^{-1}}.$$

As in Remark 3.1, rigourously we should integrate one more time in s.

Passing to the limit in the monotonicity formula for u_i, we obtain

Proposition 4.8 (Limiting, localized monotonicity formula). *For any* $(x,t) \in Q_{1/2}$ *and a.a.* $0 < s_1 < s_2 < 1/4$,

$$\Theta_{s_2}(x,t;u_\infty,\mu) - \Theta_{s_1}(x,t;u_\infty,\mu)$$

$$\ge \int_{s_1}^{s_2}\tau^{\frac{2}{p-1}-1}\int_{B_1}\left|(t-\tau)\partial_t u_\infty(y,t-\tau) + \frac{u_\infty(y,t-\tau)}{p-1} + \frac{y}{2}\cdot\nabla u_\infty(y,t-\tau)\right|^2$$

$$\times G(y-x,\tau)\psi(y)^2dyd\tau$$

$$+ \int_{s_1}^{s_2}\int_{B_1}\tau^{\frac{2}{p-1}-1}(t-\tau)^2G(y-x,\tau)\psi(y)^2d\nu(y,\tau)$$

$$+ \int_{s_1}^{s_2}\int_{B_1}\tau^{\frac{2}{p-1}-1}\frac{|y|^2}{4}G(y-x,\tau)\psi(y)^2d\mu(y,\tau)$$

$$+ \quad \frac{1}{m} \int_{s_1}^{s_2} \int_{B_1} \tau^{\frac{2}{p-1}-1}(t-\tau)y \cdot \mathbf{H}(y,\tau)G(y-x,\tau)\psi(y)^2 d\mu(y,\tau).$$

5. Tangent flow analysis, I

In this section we perform the tangent flow analysis for (u_∞, μ).

For any $(x,t) \in \Sigma^*$ and a sequence $\lambda_i \to 0$, define the blowing up sequence

$$\begin{cases} u_\infty^{\lambda_i}(y,s) := \lambda_i^{\frac{2}{p-1}} u_\infty(x+\lambda_i y, t+\lambda_i^2 s), \\ \mu^{\lambda_i}(A) := \lambda_i^{m-2-n}\mu(\lambda_i A), \quad \nu^{\lambda_i}(A) := \lambda_i^{m-n}\nu(\lambda_i A), \quad \text{for any } A \subset \mathbb{R}^n \times \mathbb{R}. \end{cases}$$

By scaling (4.2), we see that with the same constant M in Corollary 3.5, for any $Q_R \subset \mathbb{R}^n \times \mathbb{R}$, we have

$$\begin{cases} \mu^{\lambda_i}(Q_R) + \int_{Q_R} \left(|\nabla u_\infty^{\lambda_i}|^2 + |u_\infty^{\lambda_i}|^{p+1}\right) \leq MR^{n+2-m}, \\ \nu^{\lambda_i}(Q_R) + \int_{Q_R} |\partial_t u_\infty^{\lambda_i}|^2 \leq MR^{n-m}. \end{cases} \tag{5.1}$$

Therefore there exists a subsequence (still denoted by λ_i) such that $u_\infty^{\lambda_i} \rightharpoonup u_\infty^0$ weakly in $L_t^2 H_{x,loc}^1(\mathbb{R}^n \times \mathbb{R}) \cap L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R})$, and

$$\begin{cases} |\nabla u_\infty^{\lambda_i}|^2 dxdt + \mu^{\lambda_i} \rightharpoonup |\nabla u_\infty^0|^2 dxdt + \mu^0, \\ |\partial_t u_\infty^{\lambda_i}|^2 dxdt + \nu^{\lambda_i} \rightharpoonup |\partial_t u_\infty^0|^2 dxdt + \nu^0 \end{cases}$$

weakly as Radon measures on any compact set of $\mathbb{R}^n \times \mathbb{R}$.

Remark 5.1. *By Lemma 4.3, we can avoid a set of zero $\mathcal{P}^{n+2-m}$ measure so that*

$$\lim_{r \to 0} \left[r^{m-2-n} \int_{Q_r(x,t)} \left(|\nabla u_\infty|^2 + |u_\infty|^{p+1}\right) + r^{m-n} \int_{Q_r(x,t)} |\partial_t u_\infty|^2 \right] = 0. \tag{5.2}$$

Under this assumption, $u_\infty^{\lambda_i}$ converges to 0 strongly in $L_t^2 H_{x,loc}^1(\mathbb{R}^n \times \mathbb{R}) \cap L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R})$, $\partial_t u_\infty^{\lambda_i}$ converges to 0 strongly in $L_{loc}^2(\mathbb{R}^n \times \mathbb{R})$, so

$$\mu^{\lambda_i} \rightharpoonup \mu^0, \quad \nu^{\lambda_i} \rightharpoonup \nu^0.$$

This special choice will be used in Section 6.

Because μ is the defect measure coming from u_i, we can take a further subsequence of u_i so that, by defining

$$u_i^{\lambda_i}(y,s) := \lambda_i^{\frac{2}{p-1}} u_i\left(x+\lambda_i y, t+\lambda_i^2 s\right),$$

we have

$$\begin{cases} |\nabla u_i^{\lambda_i}|^2 dyds \rightharpoonup |\nabla u_\infty^0|^2 dxdt + \mu^0, \\ |\partial_t u_i^{\lambda_i}|^2 dyds \rightharpoonup |\partial_t u_\infty^0|^2 dxdt + \nu^0. \end{cases}$$

As a consequence, all results in Section 4 hold for u_∞^0, μ^0 and ν^0. In particular,

(1) by Lemma 4.6, there exists a family of Radon measures μ_t^0 on $\mathbb{R}^n$ (for a.a. $t \in \mathbb{R}$) such that

$$\mu^0 = \mu_t^0 dt; \tag{5.3}$$

(2) there exists an $\mathbf{H}^0 \in L^2_{loc}(\mathbb{R}^n \times \mathbb{R}, d\mu^0)$ such that

$$\partial_t u_i^{\lambda_i} \nabla u_i^{\lambda_i} dxdt \rightharpoonup \partial_t u_\infty^0 \nabla u_\infty^0 dxdt + \frac{1}{m}\mathbf{H}^0 d\mu^0;$$

(3) for any $(x,t) \in \mathbb{R}^n \times \mathbb{R}$ and $s > 0$, we define $\Theta_s(x,t; u_\infty^0, \mu^0)$ to be

$$s^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla u_\infty^0(y,t-s)|^2}{2} - \frac{|u_\infty^0(y,t-s)|^{p+1}}{p+1} \right] G(y-x,s)dy$$

$$+ \frac{1}{2(p-1)} s^{\frac{2}{p-1}} \int_{\mathbb{R}^n} u_\infty^0(y,t-s)^2 G(y-x,s)dy + \frac{1}{m} s^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} G(y-x,s)d\mu^0_{t-s}(y),$$

which is still non-decreasing in $s > 0$.

For any $(x,t) \in \mathbb{R}^n \times \mathbb{R}$, we still define

$$\Theta(x,t; u_\infty^0, \mu^0) := \lim_{s \to 0} \Theta_s(x,t; u_\infty^0, \mu^0).$$

By the scaling invariance of Θ, we obtain

$$\Theta_s(0,0; u_\infty^0, \mu^0) = \Theta(x,t; u_\infty, \mu), \quad \forall s > 0. \tag{5.4}$$

Then an application of Proposition 4.8 gives

Lemma 5.2.   (1) *The function u_∞^0 is backwardly self-similar in the sense that*

$$u_\infty^0(\lambda x, \lambda^2 t) = \lambda^{-\frac{2}{p-1}} u_\infty(x,t), \quad \forall (x,t) \in \mathbb{R}^n \times \mathbb{R}^-, \ \lambda > 0.$$

(2) *The measure μ^0 is backwardly self-similar in the sense that*

$$\mu^0(\lambda A) = \lambda^{n+2-m} \mu^0(A), \quad \forall \lambda > 0, \ A \subset \mathbb{R}^n \times \mathbb{R}^-.$$

By (5.4), following the proof of Lin-Wang [58, Lemma 8.3.3], we obtain

Lemma 5.3. *For any $(x,t) \in \mathbb{R}^n \times \mathbb{R}$,*

$$\Theta(x,t; u_\infty^0, \mu^0) \leq \Theta(0,0; u_\infty^0, \mu^0).$$

Moreover, if the equality is attained at (x,t), then u_∞^0 and μ^0 are translational invariant in the (x,t)-direction.

By this lemma,

$$\mathcal{L}(u_\infty^0, \mu^0) := \left\{ (x,t) : \ \Theta(x,t; u_\infty^0, \mu^0) = \Theta(0,0; u_\infty^0, \mu^0) \right\}$$

is a linear subspace of $\mathbb{R}^n \times \mathbb{R}$, which is called *the invariant subspace of (u_∞^0, μ^0).*

Definition 5.4. *The invariant dimension of (u_∞^0, μ^0) is*

$$\begin{cases} k+2, & \text{if } \mathcal{L}_{\mu_0} = \mathbb{R}^k \times \mathbb{R} \\ k, & \text{otherwise.} \end{cases}$$

Using these notations we can define a stratification of Σ^*, and give a dimension estimate on these stratifications as in White [91], see also [58, Section 8.3].

6. The case m is not an integer

In this section, we use Marstrand theorem ([61], see also [59, Theorem 1.3.12]) to study the case when m is not an integer. This is similar to the elliptic case studied in Du [26] and the authors [86]. The main result in this case is

Theorem 6.1. *Suppose m is not an integer. Then under the assumptions of Theorem 4.2, u_i converges strongly in $L^{p+1}_{loc}(Q_1)$ and ∇u_i converges strongly to ∇u_∞ in $L^2_{loc}(Q_1)$. In other words, $\mu = 0$. As a consequence, u_∞ is a suitable weak solution of* (1.1) *in Q_1.*

Here we do not claim the strong convergence of $\partial_t u_i$.

To prove this theorem, we use the tangent flow analysis in the previous section, but only performed at a point (x, t) satisfying the condition (5.2). Hence by Remark 5.1, $u^0_\infty = 0$ and

$$\mu^\lambda \rightharpoonup \mu^0, \quad \nu^\lambda \rightharpoonup \nu^0.$$

By (4.3) in Theorem 4.2, μ^0 is not identically zero in $\mathbb{R}^n \times \mathbb{R}^-$. Then by the self-similarity of μ^0, there exists a point $(x_0, -1) \in \operatorname{spt}(\mu^0)$. We blow up μ^0 again at $(x_0, -1)$ as in Section 5, producing a tangent measure μ^1 to μ^0 at this point.

Lemma 6.2. *The measure μ^1 is static, that is,*

$$\partial_t \mu^1 = 0 \quad \text{in the distributional sense.}$$

Proof. By Lemma 5.2 and the scaling invariance of Θ_s, for any $\lambda > 0$,

$$\Theta_{\lambda^2 s}(\lambda x_1, -\lambda^2; \mu^0) = \Theta_s(x_1, -1; \mu^0).$$

Letting $s \to 0$, we obtain

$$\Theta(\lambda x_1, -\lambda^2; \mu^0) = \Theta(x_1, -1; \mu^0).$$

After blowing up to μ^1, this equality implies that

$$\Theta(0, t; \mu^1) = \Theta(0, 0; \mu^1), \quad \forall t \in \mathbb{R}.$$

By Lemma 5.3, μ^1 is invariant under translations in the time direction. $\square$

Corollary 6.3. $\nu^1 = 0$ *and* $\mathbf{H}^1 = 0$ μ^1*-a.e..*

Proof. This follows from the energy identity for μ^1,

$$\frac{d}{dt} \int_{\mathbb{R}^n} \eta \, d\mu^1_t = -m \int \eta \, d\nu^1 - \int_{\mathbb{R}^n} \nabla \eta \cdot \mathbf{H}^1 d\mu^1_t.$$

Note that $\mathbf{H}^1$ can be controlled by ν^1 as in Corollary 4.5. $\square$

By the previous lemma, we can view μ^1 as a Radon measure on $\mathbb{R}^n$. Now the stationary condition (4.10) reads as

$$\int_{\mathbb{R}^n} (I - m\mathcal{T}) \cdot DY \, d\mu^1 = 0, \quad \forall Y \in C^\infty_0(\mathbb{R}^n, \mathbb{R}^n). \tag{6.1}$$

Following Moser [72], this is called a stationary measure. Similar to [72, Lemma 2.1], we have

Lemma 6.4. *For any $x \in \mathbb{R}^n$,*

$$\Theta_r(x; \mu^1) := r^{m-n} \mu_1(B_r(x))$$

is non-decreasing in $r > 0$.

As a consequence,

$$\Theta(x; \mu^1) := \lim_{r \to 0} r^{m-n} \mu_1(B_r(x))$$

exists. By (4.2), (4.3) and Lemma 4.3,

$$\frac{\varepsilon_*}{2} \le \Theta(x; \mu^1) \le C, \quad \mu^1 - \text{a.e. in } \mathbb{R}^n.$$

By Marstrand theorem, m must be an integer. In other words, if m is not an integer, then μ^1, and consequently, μ^0 and μ, must be trivial. This finishes the proof of Theorem 6.1.

7. Partial regularity of suitable weak solutions

In this section, as an application of Theorem 6.1, we prove the following partial regularity result for a fixed, suitable weak solution. This is almost the same with the one obtained in [9], with a small improvement.

Theorem 7.1. *Suppose u is a suitable weak solution of* (1.1) *in Q_1. Then*
- *If $1 < p < \frac{n+2}{n-2}$, u is smooth, that is, $\mathcal{S}(u) = \emptyset$.*
- *If there exists an integer k with $3 \le k < n$ such that $\frac{k+3}{k-1} < p < \frac{k+2}{k-2}$, then $dim_{\mathcal{P}}(\mathcal{S}(u)) \le n - k + 2$.*
- *If m is an integer, then $\mathcal{P}^{n-m+2}(\mathcal{S}(u)) = 0$.*

We have already obtained a dimension bound on the singular set $\mathcal{S}(u)$ in Corollary 3.10. Here we need only to improve this bound when m is not an integer. This follows from a standard dimension reduction argument, cf. [54], [86].

Suppose $(x, t) \in \mathcal{S}(u) \cap Q_{1/2}$. By Theorem 3.9, for any $r \in (0, 1/2)$,

$$\int_{Q_r^-(x,t)} \left(|\nabla u|^2 + |u|^{p+1} \right) \ge \varepsilon_* r^{n+2-m}. \tag{7.1}$$

For $\lambda \to 0$, define the blowing up sequence

$$u^\lambda(y, s) := \lambda^{\frac{2}{p-1}} u(x + \lambda y, t + \lambda^2 s).$$

As in Section 5, we can take a subsequence so that $u^\lambda \rightharpoonup u^0$ weakly in $L_t^2 H_{x,loc}^1(\mathbb{R}^n \times \mathbb{R}) \cap L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R})$, and

$$\begin{cases} |\nabla u^\lambda|^2 dxdt \rightharpoonup |\nabla u^0|^2 dxdt + \mu^0, \\ |\partial_t u^\lambda|^2 dxdt \rightharpoonup |\partial_t u^0|^2 dxdt + \nu^0 \end{cases}$$

weakly as Radon measures on any compact set of $\mathbb{R}^n \times \mathbb{R}$.

Furthermore, because m is not an integer, by Theorem 6.1, we deduce that $\mu^0 = 0$ and u^λ in fact converges strongly. Because tangent flows are always nontrivial, $u^0 \neq 0$.

By Lemma 5.2, u^0 is backwardly self-similar. For these solutions, the following Liouville theorem was established by Giga-Kohn in [36].

Theorem 7.2. *When $p \le (n+2)/(n-2)$, if u is a backward self-similar solution of* (1.1) *in $\mathbb{R}^n \times \mathbb{R}^-$, satisfying*

$$\sup_{x \in \mathbb{R}^n} \int_{Q_1^-(x,-1)} \left(|\nabla u|^2 + |u|^{p+1} \right) < +\infty, \tag{7.2}$$

then either $u \equiv 0$, or $u \equiv \pm(p-1)^{-\frac{1}{p-1}}(-t)^{-\frac{1}{p-1}}$.

Here the original $L^\infty(\mathbb{R}^n)$ assumption in [36, Theorem 1] is replaced by an integral one (7.2). With this condition, we are still able to do the same computation in [36] to deduce the Pohozaev identity [36, Proposition 1]. For u^0, the condition (7.2) follows by scaling (3.2) and passing to the limit.

By this theorem, if $p < (n+2)/(n-2)$ is subcritical, because u^0 also satisfies

$$\int_{Q_1} |u^0|^{p+1} < +\infty,$$

we must have $u^0 = 0$. This contradiction implies that there is no singular point of u. In other words, u is smooth.

If $p > (n+2)/(n-2)$ is supcritical, by Point (3) in Theorem 4.2, the strong convergence of u^λ implies that if λ is small enough, then $\lambda^{-1}(\mathcal{S}(u) - (x,t))$ is contained in a small neighborhood of $\mathcal{S}(u^0)$. Since u^0 is backwardly self-similar (see Lemma 5.2), applying White's stratification theorem in [91], we conclude the proof of Theorem 7.1.

8. THE CASE m IS AN INTEGER

In this section we continue the analysis of energy concentration behavior, now under the assumption that m is an integer. For this case we prove

Theorem 8.1. *If m is an integer, then (u_∞, μ) is a generalized Brakke's flow in the following sense: For any $\varphi \in C_0^\infty(B_1; \mathbb{R}^+)$ and $-1 < t_1 < t_2 < 1$,*

$$\left[\int_{B_1} \left(\frac{1}{2} |\nabla u_\infty(t_2)|^2 - \frac{1}{p+1} |u_\infty(t_2)|^{p+1} \right) \varphi dx + \frac{1}{m} \int_{B_1} \varphi d\mu_{t_2} \right]$$
$$- \left[\int_{B_1} \left(\frac{1}{2} |\nabla u_\infty(t_1)|^2 - \frac{1}{p+1} |u_\infty(t_1)|^{p+1} \right) \varphi dx + \frac{1}{m} \int_{B_1} \varphi d\mu_{t_1} \right]$$
$$\le -\int_{t_1}^{t_2} \int_{B_1} \left[|\partial_t u_\infty|^2 \varphi + \partial_t u_\infty \nabla u_\infty \cdot \nabla \varphi \right] dx dt \tag{8.1}$$
$$- \frac{1}{m} \int_{t_1}^{t_2} \int_{\Sigma_t} \left[|\mathbf{H}_t|^2 \varphi - \nabla \varphi \cdot \mathbf{H}_t \right] d\mu_t dt.$$

By Lemma 4.6, for a.e. $t \in (-1,1)$ and any $Y \in C_0^\infty(B_1, \mathbb{R}^n)$,

$$0 = \int_{B_1} \left[\left(\frac{|\nabla u_\infty(x,t)|^2}{2} - \frac{|u_\infty(x,t)|^{p+1}}{p+1} \right) \text{div} Y - DY \left(\nabla u_\infty(x,t), \nabla u_\infty(x,t) \right) \right] dx$$

$$+ \quad \int_{B_1} \partial_t u_\infty(x,t) \nabla u_\infty(x,t) \cdot Y \, dx \tag{8.2}$$

$$+ \quad \int_{B_1} \left[\left(\frac{1}{m} I - \mathcal{T}(x,t) \right) \cdot DY + \frac{1}{m} \mathbf{H}(x,t) \cdot Y \right] d\mu_t(x).$$

In view of the lower density bound in (4.3), as in [1] or [52] (see also [72] or [2, Proposition 3.1]), we deduce that μ_t is countably $(n-m)$-rectifiable. In other words, Σ_t is countably $(n-m)$-rectifiable, and

$$I - m\mathcal{T} = T_x \Sigma_t, \quad \mathcal{H}^{n-m} \text{ a.e. on } \Sigma_t, \tag{8.3}$$

where $T_x \Sigma_t$ is the weak tangent space (identified with the projection map onto it) of Σ_t at x.

Similar to Lin-Wang [58, Lemma 9.2.2], we get

Lemma 8.2. *For a.e. $t \in (-1,1)$,*

$$\mathbf{H}(x,t) \perp T_x \Sigma_t, \quad \text{for } \mathcal{H}^{n-m} \text{ a.e. } x \in \Sigma_t.$$

Similar to Lin-Wang [58, Lemma 9.2.7], we also get

Lemma 8.3. *For any $\eta \in C_0^\infty(Q_1)$,*

$$\int_{Q_1} \eta^2 |\mathbf{H}(x,t)|^2 d\mu_t(x) dt \leq \frac{1}{m} \int_{Q_1} \eta^2 d\nu.$$

Plugging this estimate into (4.9), we obtain (8.1). This finishes the proof of Theorem 8.1.

9. The case $p = (n+2)/(n-2)$

In this section we assume $p = (n+2)/(n-2)$ is the Sobolev critical exponent, and all of the solutions u_i are *smooth*. Note that now $m = 2(p+1)/(p-1) = n$. The main result of this section is about the quantization of energy.

Theorem 9.1. *For any $(x,t) \in \Sigma$, there exist finitely many bubbles W^k, $k = 1, \cdots, N$, such that*

$$\Theta(x,t) = \frac{1}{n \left(4\pi\right)^{n/2}} \sum_{k=1}^{N} \int_{\mathbb{R}^n} |\nabla W^k|^2.$$

Furthermore, if all solutions are positive, then there exists an $N \in \mathbb{N}$ such that

$$\Theta(x,t) = N \frac{\Lambda}{n \left(4\pi\right)^{n/2}}.$$

We first prove a local version of this proposition in Subsection 9.1, then prove this theorem in Subsection 9.2. During this course, some further properties of defect measures will also be established in Subsection 9.2 and Subsection 9.3, in particular, for applications in Part 2, Part 3 and Part 4, the special case of positive solutions will be discussed.

9.1. **A local quantization result.** In this subsection we prove the following

Lemma 9.2. *Given a constant $M > 0$, suppose a sequence of smooth solutions u_i to* (1.1) *in Q_1 satisfies*

$$\begin{cases} |\nabla u_i|^2 dxdt \rightharpoonup M\delta_0 \otimes dt, \\ |u_i|^{p+1} dxdt \rightharpoonup M\delta_0 \otimes dt, \\ \displaystyle\int_{Q_1} |\partial_t u_i|^2 dxdt \to 0. \end{cases}$$

Then there exist finitely many bubbles W^k such that

$$M = \sum_k \int_{\mathbb{R}^n} |\nabla W^k|^2 dx.$$

First let us present some immediate consequences of the assumptions in this lemma.

- An application of the ε-regularity theorem (Theorem 3.9) implies that

$$u_i \to 0 \quad \text{in } C^\infty_{loc}\left((B_1 \setminus \{0\}) \times (-1, 1)\right). \tag{9.1}$$

- By the $L^2(Q_1)$ bound on $\partial_t u_i$, we deduce that $u_i \to 0$ in $C_{loc}(-1, 1; L^2_{loc}(B_1))$.
- By Fatou lemma,

$$\int_{-1}^1 \lim_{i \to +\infty} \left(\int_{B_1} |\partial_t u_i(x, t)|^2 dx \right) dt \leq \lim_{i \to +\infty} \int_{Q_1} |\partial_t u_i|^2 dxdt = 0.$$

Hence for a.e. $t \in (-1, 1)$,

$$\lim_{i \to +\infty} \int_{B_1} |\partial_t u_i(x, t)|^2 dx = 0. \tag{9.2}$$

The following lemma describes the energy concentration behavior for a.e. time slice of u_i, under the assumptions in Lemma 9.2.

Lemma 9.3. *For a.e. $t \in (-1, 1)$,*

$$\begin{cases} |\nabla u_i(x, t)|^2 dx \rightharpoonup M\delta_0, \\ |u_i(x, t)|^{p+1} dx \rightharpoonup M\delta_0. \end{cases}$$

Proof. For any $\eta \in C_0^\infty(B_1)$, by the energy identity for u_i, the function

$$E_{\eta,i}(t) := \int_{B_1} \left(\frac{|\nabla u_i(x, t)|^2}{2} - \frac{|u_i(x, t)|^{p+1}}{p+1} \right) \eta(x)^2 dx$$

is uniformly bounded in $BV_{loc}(-1, 1)$. After passing to a subsequence, they converge in $L^1_{loc}(-1, 1)$ and a.e. in $(-1, 1)$ to the limit $(M/n)\eta(0)^2$.

Another consequence of this uniform BV bound is, $E_{\eta,i}$ are uniformly bounded in any compact set of $(-1, 1)$. If (9.2) holds at t, then

$$\int_{B_1} \left(|\nabla u_i(t)|^2 - |u_i(t)|^{p+1} \right) \eta^2 = - \int_{B_1} \left[\partial_t u_i(t) u_i(t) \eta^2 + 2\eta u_i(t) \nabla u_i(t) \cdot \nabla \eta \right]$$

are also bounded as $i \to +\infty$. Therefore

$$\limsup_{i \to +\infty} \int_{B_1} \left[|\nabla u_i(t)|^2 + |u_i(t)|^{p+1} \right] dx < +\infty. \tag{9.3}$$

Because $u_i(t) \to 0$ in $L^2(B_1)$, with the help of (9.1), we find a nonnegative constant $M(t)$ such that

$$|\nabla u_i(x,t)|^2 dx \rightharpoonup M(t)\delta_0, \quad |u_i(x,t)|^{p+1} dx \rightharpoonup M(t)\delta_0.$$

Then using the above a.e. convergence of $E_{\eta,i}(t)$, and letting η vary in $C_0^\infty(B_1)$, we deduce that $M(t) = M$. $\square$

For those t satisfying (9.2), we have the uniform $H^1(B_1)$ bound (9.3), hence by Struwe's global compactness theorem ([82]), the following bubble tree convergence holds for $u_i(t)$.

Proposition 9.4 (Bubble tree convergence). *There exist $N(t)$ points $\xi_{ik}^*(t)$, positive constants $\lambda_{ik}^*(t)$, $k = 1, \cdots, N(t)$, all converging to 0 as $i \to +\infty$, and $N(t)$ bubbles W^k, such that*

$$u_i(x,t) = \sum_{k=1}^{N(t)} W_{\xi_{ik}^*(t), \lambda_{ik}^*(t)}^k(x) + o_i(1),$$

where $o_i(1)$ are measured in $H^1(B_1)$.
  As a consequence,

$$\int_{B_1} |\nabla u_i(x,t)|^2 dx = \sum_{k=1}^{N(t)} \int_{\mathbb{R}^n} |\nabla W^k|^2 + o_i(1). \tag{9.4}$$

Remark 9.5. *If u_i are positive, then all bubbles constructed in this proposition are positive, see [24, Section 3.2]. In view of the Liouville theorem of Caffarelli-Gidas-Spruck [8], we can take all of these W^k to be the standard Aubin-Talenti bubble W. As a consequence, (9.4) reads as*

$$\int_{B_1} |\nabla u_i(x,t)|^2 dx = N(t)\Lambda + o_i(1). \tag{9.5}$$

Remark 9.6. *About bubble tree convergence, there are two phenomena that we will investigate more closely latter.*
  Bubble towering: *Two bubbles at $\xi_{ik}^*(t)$ and $\xi_{i\ell}^*(t)$ are towering if*

$$\limsup_{i \to +\infty} \frac{|\xi_{ik}^*(t) - \xi_{i\ell}^*(t)|}{\max\{\lambda_{ik}^*(t), \lambda_{i\ell}^*(t)\}} < +\infty.$$

  In this case, either $\frac{\lambda_{ik}^*(t)}{\lambda_{i\ell}^*(t)} \to +\infty$ or $\frac{\lambda_{i\ell}^*(t)}{\lambda_{ik}^*(t)} \to +\infty$. Hence these two bubbles are located at almost the same point (with respect to the bubble scales), but the height of one bubble is far larger than the other one's.
  Bubble clustering: *If for some $k \neq \ell$,*

$$\lim_{i \to +\infty} |\xi_{ik}^*(t) - \xi_{i\ell}^*(t)| = 0$$

but

$$\lim_{i \to +\infty} \frac{|\xi_{ik}^*(t) - \xi_{i\ell}^*(t)|}{\max\{\lambda_{ik}^*(t), \lambda_{i\ell}^*(t)\}} = +\infty,$$

we say these bubbles are clustering.

Using the terminology introduced in the study of Yamabe problem (see Schoen [77]), if there is no bubble clustering and towering, the blow up is called isolated and simple.

Combining Lemma 9.3 with Proposition 9.4, we conclude the proof of Lemma 9.2. Furthermore, if all solutions are positive, by Remark 9.5, there exists an $N \in \mathbb{N}$ such that

$$M = N\Lambda, \quad \text{and} \quad N(t) = N \quad \text{a.e. in } (-1, 1). \tag{9.6}$$

9.2. **Proof of Theorem 9.1.** Under the critical exponent assumption, the tangent flow analysis in Section 5 can give more information.

First, by noting that $m = n$, in the definition of the blowing up sequence, we have

$$\nu^\lambda(A) := \nu\left((x, t) + \lambda A\right), \quad \forall A \subset \mathbb{R}^n \times \mathbb{R}.$$

In the same way, for any $R > 0$,

$$\int_{Q_R} |\partial_t u_\infty^\lambda|^2 = \int_{Q_{\lambda R}} |\partial_t u_\infty|^2 \to 0 \quad \text{as } \lambda \to 0. \tag{9.7}$$

Therefore by defining

$$Q := \lim_{\lambda \to 0} \nu\left(Q_\lambda(x, t)\right), \tag{9.8}$$

we get

$$\nu^0 = Q\delta_{(0,0)}. \tag{9.9}$$

There is an atom of ν at (x, t) if and only if $Q > 0$.

By Corollary 4.5, we find that

$$\mathbf{H}^0 d\mu = \mathbf{H}^0(0, 0)\delta_{(0,0)}. \tag{9.10}$$

Recall that u_∞ is backwardly self-similar (Lemma 5.2). By combining Theorem 7.2 with (9.7), we deduce that

$$u_\infty^0 \equiv 0 \quad \text{in } \mathbb{R}^n \times \mathbb{R}.$$

Hence in view of (9.9) and (9.10), the energy identity (4.9) for (u_∞^0, μ^0) reads as

$$\frac{1}{m} \int_{Q_1} \partial_t \eta d\mu^0 = Q\eta(0, 0) + \frac{1}{m} \nabla \eta(0, 0) \cdot \mathbf{H}^0(0, 0), \quad \forall \eta \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}). \tag{9.11}$$

In particular,

$$\partial_t \mu^0 = 0 \quad \text{in the distributional sense in } (\mathbb{R}^n \times \mathbb{R}) \setminus \{(0, 0)\}.$$

Combining this fact with the backward self-similarity of μ^0 (see Lemma 5.2), we deduce that there exists a constant $M \geq Q$ such that

$$\mu^0 = M\delta_0 \otimes dt\lfloor_{\mathbb{R}^-} + (M - Q)\,\delta_0 \otimes dt\lfloor_{\mathbb{R}^+}. \tag{9.12}$$

By choosing $\eta(x, t) = \varphi(x)\psi(t)x$ in (9.11), where $\varphi \in C_0^\infty(\mathbb{R}^n)$ and $\psi \in C_0^\infty(\mathbb{R})$, we also deduce that

$$\mathbf{H}^0(0, 0) = 0. \tag{9.13}$$

Now we come to the proof of Theorem 9.1.

Proof. As in Section 5, there exist two sequences of solutions to (1.1) (with $p = (n+2)/(n-2)$) satisfying the assumptions in Lemma 9.2 in $Q_1(0, -2)$ and $Q_1(0, 2)$. Therefore there exist two groups of finitely many bubbles, $\{W^k\}$ and $\{W^\ell\}$, such that

$$M = \sum_k \int_{\mathbb{R}^n} |\nabla W^k|^2 dx, \quad M - Q = \sum_\ell \int_{\mathbb{R}^n} |\nabla W^\ell|^2 dx. \tag{9.14}$$

Furthermore, if all solutions are positive, then there exist $N_1, N_2 \in \mathbb{N}$ such that

$$M = N_1 \Lambda, \quad M - Q = N_2 \Lambda. \tag{9.15}$$

By the weak convergence of $|\nabla u_\infty^\lambda|^2 dx dt + \mu^\lambda$ etc., we get

$$
\begin{aligned}
\Theta(x, t; u_\infty, \mu) &= \lim_{\lambda \to 0} \int_1^2 \Theta_{\lambda^2 s}(x, t; u_\infty, \mu) ds \\
&= \lim_{\lambda \to 0} \int_1^2 \Theta_s(0, 0; u_\infty^\lambda, \mu^\lambda) ds \\
&= \frac{1}{n} \int_1^2 s^{\frac{n}{2}} \left[\int_{\mathbb{R}^n} G(y, 1) d\mu^0_{-s}(y) dy \right] ds \\
&= \frac{M}{n (4\pi)^{n/2}}.
\end{aligned}
$$

Substituting (9.14) into this equality, we conclude the proof. $\qquad\square$

A consequence of Theorem 9.1 is the following relation between Θ and θ.

Corollary 9.7. *For μ-a.e. (x, t),*

$$\Theta(x, t; u_\infty, \mu) = \frac{1}{n (4\pi)^{\frac{n}{2}}} \theta(x, t),$$

and consequently, there exist finitely many bubbles W^k, $k = 1, \cdots, N$, such that

$$\theta(x, t) = \sum_{k=1}^N \int_{\mathbb{R}^n} |\nabla W^k|^2.$$

Moreover, if all solutions are positive, there exist an $N(x, t) \in \mathbb{N}$ such that

$$\theta(x, t) = N(x, t)\Lambda \quad \mu - a.e..$$

Proof. For those (x, t) satisfying the condition (5.2), we have

$$\theta(x, t) = \lim_{\lambda \to 0} \frac{\mu(Q_\lambda(x, t))}{2\lambda^2} = \frac{1}{2} \mu^0(Q_1).$$

We can also assume there is no atom for ν at (x, t) by avoiding a at most countable set. Then by the above structure theory of tangent flows and Theorem 9.1, we get

$$\mu^0(Q_1) = 2M = 2n (4\pi)^{\frac{n}{2}} \Theta(x, t). \qquad\square$$

9.3. **Further properties of defect measures.** Some consequences follow from the above structure result on tangent flows. The first one is

Lemma 9.8. *For each t, Σ_t^* is isolated.*

Proof. Assume by the contrary, there exists a sequence of points $(x_j, t) \in \Sigma_t^*$ converging to a limit point (x, t). Because Σ^* is closed, $(x, t) \in \Sigma^*$. Denote

$$\lambda_j := |x_j - x| \to 0.$$

Define the blow up sequence $(u_\infty^{\lambda_j}, \mu^{\lambda_j})$ with respect to the bast point (x, t) as before. Assume without loss of generality that

$$\frac{x_j - x}{\lambda_j} \to x_\infty \in \partial B_1.$$

By the upper semi-continuity of Θ, we get

$$\Theta(x_\infty, 0; \mu^0) \geq \limsup_{j \to +\infty} \Theta(x_j, t; u_\infty, \mu) \geq \varepsilon_*. \tag{9.16}$$

On the other hand, the tangent flow analysis shows that $\mathrm{spt}(\mu^0) = \{0\} \times \mathbb{R}$. This fact, together with the ε-regularity theorem and the fact that $u_\infty^0 = 0$, implies that

- $u_\infty^{\lambda_j} \to 0$ in $C_{loc}^\infty((\mathbb{R}^n \setminus \{0\}) \times \mathbb{R})$,
- and for all λ_j small, $\mathrm{spt}(\mu^{\lambda_j})$ is contained in $B_{1/2} \times \mathbb{R}$.

Therefore,

$$\Theta(x_\infty, 0; \mu^0) = \lim_{j \to +\infty} \Theta\left(\frac{x_j - x}{\lambda_j}, 0; u_\infty^{\lambda_j}, \mu^{\lambda_j}\right) = 0.$$

This is a contradiction with (9.16). In other words, there does not exist converging sequences in Σ_t^*. $\qquad\square$

The next one is about the form of the stress-energy tensor $\mathcal{T}$.

Lemma 9.9. $\mathcal{T} = I/n$ μ-*a.e..*

Proof. This is (8.3) in this special case. Here we explain briefly a direct proof. First, because $\mathcal{T}$ is μ-measurable, it is approximately continuous μ-a.e.. Choose a point (x, t) so that $\mathcal{T}$ is approximate continuous and Lemma 4.3 holds at this point. Take a sequence $\lambda \to 0$ and define the tangent flow at this point as before, denoted by μ^0. Let

$$\mathcal{T}^\lambda(y, s) := \mathcal{T}(x + \lambda y, t + \lambda^2 s).$$

By the approximate continuity of $\mathcal{T}$ at (x, t), we see

$$\mathcal{T}^\lambda d\mu^\lambda \rightharpoonup \mathcal{T}(x, t) d\mu^0.$$

Then we obtain the stationary condition for the tangent flow,

$$\int_{\mathbb{R}^n \times \mathbb{R}} [I - n\mathcal{T}(x, t)] \cdot DY \, d\mu^0 = 0, \quad \forall Y \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}, \mathbb{R}^n).$$

In view of the form of μ^0 in (9.12), substituting suitable Y as test functions into this identity we deduce that $\mathcal{T}(x, t) = I/n$. $\qquad\square$

The following lemma is a rigourous statement that zero dimensional objects should have zero mean curvatures.

Lemma 9.10. $\mathbf{H} = 0$ μ-a.e..

Proof. We will show that, for a.e. $t \in (-,1,1)$,

$$\mathbf{H} = 0 \quad \mu_t - \text{everywhere.} \tag{9.17}$$

For a.e. t,

$$\int_{B_1} \left[|\nabla u_\infty(x,t)|^2 + |u_\infty(x,t)|^{p+1} + |\partial_t u_\infty(x,t)|^2 \right] dx < +\infty, \tag{9.18}$$

and by Lemma 4.6, μ_t is a well defined Radon measure.

By the form of $\mathcal{T}$, now the stationary condition (4.10) reads as

$$
\begin{aligned}
0 &= \int_{Q_1} \left[\left(\frac{|\nabla u_\infty|^2}{2} - \frac{|u_\infty|^{p+1}}{p+1} \right) \mathrm{div} Y - DY(\nabla u_\infty, \nabla u_\infty) + \partial_t u_\infty \nabla u_\infty \cdot Y \right] \\
&+ \frac{1}{m} \int_{Q_1} \mathbf{H} \cdot Y d\mu, \qquad \forall Y \in C_0^\infty(Q_1, \mathbb{R}^n).
\end{aligned} \tag{9.19}
$$

Hence for a.e. t, we have the stationary condition

$$
\begin{aligned}
0 &= \int_{B_1} \left[\left(\frac{|\nabla u_\infty|^2}{2} - \frac{|u_\infty|^{p+1}}{p+1} \right) \mathrm{div} X - DX(\nabla u_\infty, \nabla u_\infty) + \partial_t u_\infty \nabla u_\infty \cdot X \right] \\
&+ \frac{1}{m} \int_{B_1} \mathbf{H} \cdot X d\mu_t, \qquad \forall X \in C_0^\infty(B_1, \mathbb{R}^n).
\end{aligned} \tag{9.20}
$$

Now we claim that

Claim. $u_\infty(t) \in W_{loc}^{2,2}(B_1) \cap L_{loc}^{2p}(B_1)$.

Once we have this claim, by integration by parts, we can show that $u_\infty(t)$ satisfies the stationary condition

$$0 = \int_{B_1} \left[\left(\frac{|\nabla u_\infty|^2}{2} - \frac{|u_\infty|^{p+1}}{p+1} \right) \mathrm{div} X - DX(\nabla u_\infty, \nabla u_\infty) + \partial_t u_\infty \nabla u_\infty \cdot X \right].$$

Combining this identity with (9.20) we get

$$\int_{B_1} \mathbf{H} \cdot X d\mu_t = 0, \qquad \forall X \in C_0^\infty(B_1, \mathbb{R}^n),$$

from which (9.17) follows.

Proof of the claim. By Lemma 9.8, Σ_t^* is isolated. Since u_∞ is smooth outside Σ^*, we only need to show that for each $\xi \in \Sigma_t^*$, there exists a ball $B_r(\xi)$ such that $u_\infty(t) \in W^{2,2}(B_r(\xi)) \cap L^{2p}(B_r(\xi))$.

We take a sufficiently small σ and choose this ball so that

$$\int_{B_r(\xi)} |u_\infty(t)|^{\frac{2n}{n-2}} \le \sigma.$$

Take a standard cut-off function η in $B_r(\xi)$ such that $\eta \equiv 1$ in $B_{r/2}(\xi)$. A direct calculation gives

$$-\Delta \left(u_\infty(t)\eta \right) = |u_\infty(t)|^{\frac{n-2}{2}} u_\infty \eta + f_\eta,$$

where $f_\eta \in L^2(B_r(\xi))$. Then by the $W^{2,2}$ estimate for Laplacian operator and an application of Hölder inequality, if σ is sufficiently small, we obtain

$$\|u_\infty(t)\eta\|_{W^{2,2}(B_r(\xi))} \lesssim \|u(t)\|_{L^2(B_r(\xi))} + \|f_\eta\|_{L^2(B_r(\xi))}.$$

The L^{2p} estimate on $u_\infty(t)$ follows by combining this estimate with the equation for $u_\infty(t)$. The proof of the claim is complete. $\qquad\square$

Corollary 9.11. u_∞ *satisfies the stationary condition* (2.3).

We do not know if the energy inequality (4.9) can be decoupled in the same way.
Finally, for positive solutions, we note the following fact as a consequence of (9.15):

- either $Q = 0$, which implies that there is no atom at (x, t),
- or $Q \geq \Lambda$, which implies that there is an atom in ν at (x, t) and its mass is at least Λ.

In conclusion, we get

Lemma 9.12. *If all solutions are positive, then the mass of each atom in ν is at least Λ. Consequently, there are at most finitely many atoms in ν.*

The atoms of ν correspond to the singular point of Σ. Microscopically, such an atom comes from scalings of a connecting orbit (or a terrace of connecting orbits)

$$\begin{cases} \partial_t u - \Delta u = u^p, & \text{in } \mathbb{R}^n \times \mathbb{R}, \\ \displaystyle\int_{\mathbb{R}^n} \left[\frac{1}{2}|\nabla u|^2 - \frac{1}{p+1}u^{p+1} \right]\Big|_{+\infty}^{-\infty} = Q. \end{cases}$$

However, we do not know if Σ is smooth outside this singular set.

Part 2. **Energy concentration with only one bubble**

10. Setting

In this part, $p = (n + 2)/(n - 2)$ is the Sobolev critical exponent, u_i denotes a sequence of *smooth, positive solutions* of (1.1) in Q_1, satisfying the following three assumptions:

(II.a) Weak limit: u_i converges weakly to u_∞ in $L^{p+1}(Q_1)$, and ∇u_i converges weakly to ∇u_∞ in $L^2(Q_1)$. Here u_∞ is a *smooth solution* of (1.1) in Q_1.

(II.b) Energy concentration set: weakly as Radon measures

$$\begin{cases} |\nabla u_i|^2 dxdt \rightharpoonup |\nabla u_\infty|^2 dxdt + \Lambda\delta_0 \otimes dt, \\ u_i^{p+1} dxdt \rightharpoonup u_\infty^{p+1} dxdt + \Lambda\delta_0 \otimes dt. \end{cases}$$

(II.c) Convergence of time derivatives: as $i \to \infty$, $\partial_t u_i$ converges to $\partial_t u_\infty$ strongly in $L^2(Q_1)$.

The assumption **(II.b)** says there is only one bubble. From the above assumptions, it is also seen that u_i converges to u_∞ in $C(-1, 1; L^2(B_1))$.

The main result of this part is the following theorem, which can be viewed as a weak form of Schoen's Harnack inequality for Yamabe problem (see [77], [47]). As in Yamabe problem, this will be used to prove that there is no bubble towering.

Theorem 10.1. *Under the above assumptions, we must have $u_\infty \equiv 0$.*

In the following, we also assume there exists a constant L such that for all i,

$$\sup_{-1 < t < 1} \int_{B_1} |\partial_t u_i(x, t)|^2 dx \leq L. \tag{10.1}$$

This assumption is in fact a consequence of **(II.a-II.c)**, see Section 24 in Part 3 for the proof.

The proof of Theorem 10.1 uses mainly a reverse version of *the inner-outer gluing mechanism*.

(1) In Section 11, we describe the blow up profile of u_i, i.e. the form of bubbles when energy concentration phenomena appears, see Proposition 11.1. This is the main order term of u_i, and it provides us the starting point for the decomposition in the next step.

(2) In Section 12, we take two decompositions for u_i: the first one is an orthogonal decomposition where we decompose u_i into a standard bubble (which is the main order term) and an error function (which is the next order term); the second one is the inner-outer decomposition, where we divide the error function into two further parts, the first one (the inner part) is localized near the bubble, and the second one (the outer part) is on the original scale.

(3) In Section 13, we establish an estimate for the inner problem, where we mainly use the nondegenaracy of the bubble (see Section A). Roughly speaking, this estimate reads as

$$\mathcal{I} \leq A\mathcal{O} + \text{higher order terms from scaling parameters etc.},$$
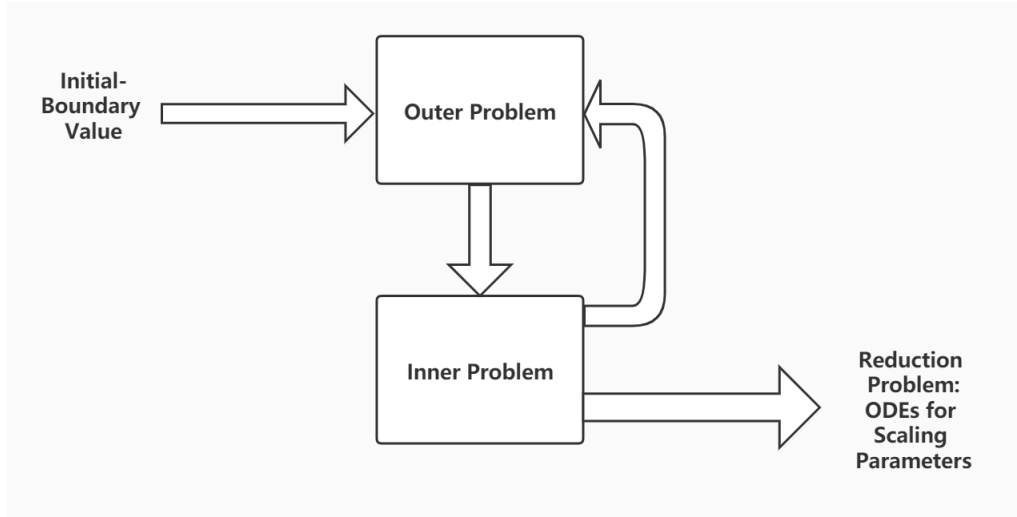
FIGURE 1. Inner-outer gluing mechanism

where $\mathcal{I}$ is a quantity measuring the inner component, $\mathcal{O}$ is a quantity measuring the outer component, and A is a constant.

(4) In Section 14, we establish an estimate for the outer problem. This estimate reads as

$$\begin{aligned} \mathcal{O} \quad \leq \quad & B\mathcal{I} + \text{effect from initial-boundary value} \\ + \quad & \text{higher order terms from scaling parameters etc.,} \end{aligned}$$

where B is a constant.

The *inner-outer gluing mechanism* works thanks to the fact that

$$AB < 1.$$

This follows from a fast decay estimate away from the bubble domains, where we mainly rely on a Gaussian bound on heat kernels associated to a parabolic operator with small Hardy term, see Moschini-Tesei [71].

(5) In Section 15, we combine these two estimates on inner and outer problems to establish an Harnack inequality for the scaling parameter. This gives a uniform in time control on the height of bubbles.

(6) In Section 16, by the estimate in Section 15, we improve the estimates on inner and outer problems.

(7) In Section 17, we improve further the estimates on the error function to an optimal one, such as uniform L^∞ estimate, first order gradient Hölder estimate and Schauder estimate.

(8) In Section 18, with the help of these estimates on the error function, we are able to linearize the Pohozaev identity. This is because a Pohozaev identity holds for u, and u is approximated by the bubble at main order, which also satisfies a Pohozaev identity, then the next order term in the Pohozaev identity of u gives further information.

(9) In Section 19, we use this linearization of Pohozaev identity to establish a weak form of Schoen's Harnack inequality, which also finishes the proof of Theorem 10.1.

(10) In Section 20, we use this weak form of Schoen's Harnack inequality to exclude a special case of bubble clustering, where we use the same idea of constructing Green functions in the study of Yamabe problems. This result will be used in Section 30 of Part 4.

10.1. **List of notations and conventions used in this part.**

- Given $\theta \in [0, 1)$, $\alpha \geq 0$ and $k \in \mathbb{N}$, let $\mathcal{X}_\alpha^{k+\theta}$ be the space of functions $\phi \in C^{k,\theta}(\mathbb{R}^n)$ with the weighted $C^{k,\theta}$ norm

$$\|\phi\|_{\alpha, k+\theta} := \sup_{x \in \mathbb{R}^n} \sum_{\ell=0}^{k-1} (1 + |x|)^{\ell + \alpha} |\nabla^\ell \phi(x)| + \sup_{x \in \mathbb{R}^n} (1 + |x|)^{k+\alpha} \|\nabla^k \phi\|_{C^\theta(B_1(x))}.$$

  If $\theta = 0$ and $k = 0$, this space is written as $\mathcal{X}_\alpha$.

- Throughout this part, an even function $\eta \in C_0^\infty(-2, 2)$ will be fixed, which satisfies $\eta \equiv 1$ in $(-1, 1)$ and $|\eta'| + |\eta''| \leq 10$. For any $R > 0$, denote

$$\eta_R(y) := \eta\left(\frac{|y|}{R}\right).$$

- $\alpha := \frac{n-2}{2}$.
- $\bar{p} := \min\{p, 2\}$.
- Two large constants $K \gg L \gg 1$ will be chosen in Section 12.
- In this part, unless otherwise stated, it is always assumed that $n \geq 7$, which implies $\alpha > 2$.

11. BLOW UP PROFILE

In this section we only assume $n \geq 3$. Here we prove

Proposition 11.1 (Blow up profile). *For any $t \in [-81/100, 81/100]$, there exists a unique maxima point of $u_i(\cdot, t)$ in the interior of $B_1(0)$. Denote this point by $\xi_i^*(t)$ and let $\lambda_i^*(t) := u_i(\xi_i^*(t), t)^{-\frac{2}{n-2}}$.*
As $i \to +\infty$,

$$\lambda_i^*(t) \to 0, \quad \xi_i^*(t) \to 0, \quad \text{uniformly in } C([-81/100, 81/100]),$$

and the function

$$u_i^t(y, s) := \lambda_i^*(t)^{\frac{n-2}{2}} u_i\left(\xi_i^*(t) + \lambda_i^*(t)y, t + \lambda_i^*(t)^2 s\right),$$

converges to $W(y)$ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.

Let us first present some preliminary results, which are needed for the proof of this proposition. The first one is a direct consequence of Theorem 3.9 (with the help of **(II.b)**).

Corollary 11.2. *As $i \to +\infty$, u_i are uniformly bounded in $C_{loc}^\infty(Q_1 \setminus \Sigma)$.*

We also note that there is no concentration for $\partial_t u_i$.

Lemma 11.3. *For any $\sigma > 0$, there exists an $r(\sigma) > 0$ such that for any $(x,t) \in Q_{9/10}$,*

$$\limsup_{i \to +\infty} \int_{Q_{r(\sigma)}(x,t)} |\partial_t u_i|^2 < \sigma.$$

Proof. Assume by the contrary, there exists a $\sigma > 0$, a sequence of points $(x_i, t_i) \in Q_{9/10}$ and a sequence of $r_i \to 0$ such that

$$\int_{Q_{r_i}(x_i, t_i)} |\partial_t u_i|^2 \geq \sigma. \tag{11.1}$$

Without loss of generality, assume $(x_i, t_i) \to (x_\infty, t_\infty)$. For any $r > 0$ fixed, by **(II.c)**, if i is large enough,

$$\int_{Q_r(x_\infty, t_\infty)} |\partial_t u_\infty|^2 \geq \int_{Q_{r_i}(x_i, t_i)} |\partial_t u_i|^2 \geq \sigma.$$

This is a contradiction with the fact that $\partial_t u_\infty \in L^2(Q_1)$. $\square$

Next is a result about the energy concentration behavior of each time slice $u_i(t)$.

Lemma 11.4. *For any $\delta > 0$, there exists an $r(\delta) \in (0,1)$ such that for any $t \in [-81/100, 81/100]$,*

$$\limsup_{i \to +\infty} \left(\left| \int_{B_{r(\delta)}} |\nabla u_i(x,t)|^2 dx - \Lambda \right| + \left| \int_{B_{r(\delta)}} u_i(x,t)^{p+1} dx - \Lambda \right| \right) < \delta. \tag{11.2}$$

Proof. Take a sufficiently small $\sigma > 0$, and then choose $r(\sigma)$ according to Lemma 11.3. Without loss of generality, assume $t = 0$. By the scaling invariance of energy, in the following we need only to prove the corresponding result for the rescaling of u_i,

$$\widetilde{u}_i(x,t) := r(\sigma)^{\frac{n-2}{2}} u_i \left(r(\sigma)x, r(\sigma)^2 t \right).$$

However, in order not to complicate notations, we still use u_i to denote $\widetilde{u}_i$. In particular, now u_i also satisfyies

$$\int_{Q_1} |\partial_t u_i|^2 \leq \sigma. \tag{11.3}$$

For each $r \in (0,1)$, recall that $\eta_r(x) := \eta(|x|/r)$ is a standard cut-off function in B_{2r}. Because u_i is smooth, we have the standard localized energy identity

$$\frac{d}{dt} \int_{B_1} \left[\frac{1}{2} |\nabla u_i(x,t)|^2 - \frac{1}{p+1} u_i(x,t)^{p+1} \right] \eta_r(x)^2 dx \tag{11.4}$$

$$= - \int_{B_1} |\partial_t u_i(x,t)|^2 \eta_r(x)^2 dx + 2 \int_{B_1} \eta_r(x) \partial_t u_i(x,t) \nabla u_i(x,t) \cdot \nabla \eta_r(x) dx.$$

Integrating this identity and applying Cauchy-Schwarz inequality, we get

$$\operatorname*{osc}_{t \in (-1,1)} \int_{B_1} \left[\frac{1}{2} |\nabla u_i(x,t)|^2 - \frac{1}{p+1} u_i(x,t)^{p+1} \right] \eta_r(x)^2 \lesssim \left(\int_{Q_1} |\partial_t u_i|^2 \right)^{1/2}$$

$$\lesssim \sigma^{1/2}, \tag{11.5}$$

where we have used (11.3).

By (11.5), **(II.b)** and the smoothness of u_∞, we get

$$\int_{B_1} \left[\frac{1}{2} |\nabla u_i(x,t)|^2 - \frac{1}{p+1} u_i(x,t)^{p+1} \right] \eta_r(x)^2 dx dt \tag{11.6}$$

$$= \frac{\Lambda}{n} + O\left(\sigma^{1/2}\right) + O\left(r^n\right) + o_i(1), \qquad \forall t \in (-1,1).$$

Next, for each $t \in (-1,1)$, multiplying (1.1) by $u_i(t)\eta_r^2$ and integrating on B_1, we get

$$\int_{B_1} \left[|\nabla u_i(x,t)|^2 - u_i(x,t)^{p+1} \right] \eta_r(x)^2 dx \tag{11.7}$$

$$= \int_{B_1} \left[\partial_t u_i(x,t) u_i(x,t) \eta_r(x)^2 - 2\eta_r(x) u_i(x,t) \nabla \eta_r(x) \cdot \nabla u_i(x,t) \right] dx.$$

Concerning the right hand side of this equation, the following estimates hold.

- By Cauchy-Schwarz inequality and (10.1), we have

$$\left| \int_{B_1} \partial_t u_i(x,t) u_i(x,t) \eta_r(x)^2 dx \right|$$

$$\lesssim \left(\int_{B_r} \partial_t u_i(x,t)^2 dx \right)^{\frac{1}{2}} \left(\int_{B_r} u_i(x,t)^2 dx \right)^{\frac{1}{2}}$$

$$\lesssim L^{1/2} \left[\left(\int_{B_r} u_\infty(x,t)^2 dx \right)^{\frac{1}{2}} + o_i(1) \right],$$

  where in the second inequality we have used the Lipschitz assumption (10.1) and the uniform convergence of $u_i(t)$ in $C_{loc}(-1,1;L^2(B_1))$.
- Still by the uniform convergence of $u_i(t)$ in $C_{loc}(-1,1;L^2(B_1))$, we obtain

$$\int_{B_1} \eta_r(x) u_i(x,t) \nabla \eta_r(x) \cdot \nabla u_i(x,t) dx$$

$$= -\frac{1}{2} \int_{B_1} u_i(x,t)^2 \Delta \eta_r(x)^2 dx$$

$$= -\frac{1}{2} \int_{B_1} u_\infty(x,t)^2 \Delta \eta_r(x)^2 dx + o_i(1).$$

Substituting these two estimates into (11.7) and noting the smoothness of u_∞, we get

$$\int_{B_1} \left[|\nabla u_i(x,t)|^2 - u_i(x,t)^{p+1} \right] \eta_r(x)^2 dx = o_r(1) + o_i(1). \tag{11.8}$$

Plugging this relation into (11.6), we obtain

$$\begin{cases} \int_{B_1} |\nabla u_i(x,t)|^2 \eta_r(x)^2 = \Lambda + O\left(\sigma^{1/2}\right) + o_r(1) + o_i(1), \\ \int_{B_1} u_i(x,t)^{p+1} \eta_r(x)^2 = \Lambda + O\left(\sigma^{1/2}\right) + o_r(1) + o_i(1). \end{cases} \tag{11.9}$$

Then (11.2) follows by noting that $\eta_r \leq \chi_{B_r} \leq \eta_{2r}$ and choosing a suitable r.   □

Some consequences follow directly from this lemma.

(1) There exists a constant $C(u_\infty, L)$ such that

$$\int_{B_{4/5}} \left[|\nabla u_i(t)|^2 + u_i(t)^{p+1}\right] \leq C(u_\infty, L), \quad \forall t \in [-81/100, 81/100]. \qquad (11.10)$$

(2) By the convergence of $u_i(t)$ in $C^\infty_{loc}(B_1 \setminus \{0\})$, we deduce that for any $t \in [-81/100, 81/100]$,

$$|\nabla u_i(t)|^2 dx \rightharpoonup |\nabla u_\infty|^2 dx + \Lambda \delta_0, \quad u_i(t)^{p+1} dx \rightharpoonup u_\infty^{p+1} dx + \Lambda \delta_0. \qquad (11.11)$$

Equation (11.11) implies that for any $t \in [-81/100, 81/100]$,

$$\sup_{B_1} u_i(x, t) \to +\infty \quad \text{as} \quad i \to +\infty.$$

Because $u_i(t)$ are bounded in $C_{loc}(B_1 \setminus \{0\})$, this supremum must be a maxima, and it is attained near 0. Take one such maximal point $\xi_i^*(t)$, and define $\lambda_i^*(t)$, u_i^t as in the statement of Proposition 11.1. By definition,

$$u_i^t(0, 0) = \max_{y \in B_{\lambda_i^*(t)^{-1}/2}} u_i^t(y, 0) = 1. \qquad (11.12)$$

Scaling (10.1) gives

$$\lim_{i \to +\infty} \int_{B_R} \partial_t u_i^t(y, s)^2 dy = 0, \quad \text{for any } R > 0, \ s \in \mathbb{R}. \qquad (11.13)$$

By scaling (11.2), we see for any $R > 0$ and $s \in \mathbb{R}$,

$$\begin{cases} \limsup_{i \to +\infty} \int_{B_R} |\nabla u_i^t(y, s)|^2 dy \leq \Lambda, \\ \limsup_{i \to +\infty} \int_{B_R} u_i^t(y, s)^{p+1} dy \leq \Lambda. \end{cases} \qquad (11.14)$$

Then by Struwe's global compactness theorem (see [82]), we deduce that for any $s \in \mathbb{R}$, $u_i^t(\cdot, s)$ converges strongly in $W^{1,2}_{loc}(\mathbb{R}^n)$ and $L^{p+1}_{loc}(\mathbb{R}^n)$. Denote its limit by u_∞^t. By (11.13) and (11.14), u_∞^t is a nonnegative solution of (1.9). In view of its H^1 regularity, it must be smooth. Hence for any $(y, s) \in \mathbb{R}^n \times \mathbb{R}$, there exists an $r > 0$ such that

$$\int_{Q_r^-(y,s)} \left[|\nabla u_\infty^t|^2 + \left(u_\infty^t\right)^{p+1}\right] \leq \frac{\varepsilon_*}{2} r^2.$$

By the above strong convergence of u_i^t, we get

$$\limsup_{i \to +\infty} \int_{Q_r^-(y,s)} \left[|\nabla u_i^t|^2 + \left(u_i^t\right)^{p+1}\right] < \varepsilon_* r^2.$$

Therefore Theorem 3.7 is applicable, which implies that u_i^t are uniformly bounded in $C^\infty(Q_{\delta_* r}(y, s))$. In conclusion, u_i^t are uniformly bounded in $C^\infty_{loc}(\mathbb{R}^n \times \mathbb{R})$. Hence

it converges to u_∞^t in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$. Then we can let $i \to +\infty$ in (11.12), which leads to

$$u_\infty^t(0) = \max_{y \in \mathbb{R}^n} u_\infty^t(y) = 1.$$

By the Liouville theorem of Caffarelli-Gidas-Spruck [8], we get

$$u_\infty^t \equiv W \quad \text{in } \mathbb{R}^n.$$

Next we give a scaling invariant estimate for u_i away from $\xi_i^*(t)$.

Lemma 11.5. *For any R large enough, there exist two constants $\sigma(R) \ll 1$ and $C(R) > 0$ such that, outside $B_{R\lambda_i^*(t)}(\xi_i^*(t))$ we have*

$$\begin{cases} u_i(x,t) \le \sigma(R)|x - \xi_i^*(t)|^{-\frac{n-2}{2}} + C(R), \\ |\nabla u_i(x,t)| \le \sigma(R)|x - \xi_i^*(t)|^{-\frac{n}{2}} + C(R), \\ |\nabla^2 u_i(x,t)| + |\partial_t u_i(x,t)| \le \sigma(R)|x - \xi_i^*(t)|^{-\frac{n+2}{2}} + C(R). \end{cases} \tag{11.15}$$

Proof. For any $\delta > 0$, by the smooth convergence of u_i^t obtained above, there exists an $R(\delta) \gg 1$ such that for any $t \in [-81/100, 81/100]$,

$$\begin{cases} \displaystyle\fint_{B_{R(\delta)\lambda_i^*(t)}(\xi_i^*(t))} |\nabla u_i(x,t)|^2 \ge \Lambda - \delta/4, \\ \displaystyle\fint_{B_{R(\delta)\lambda_i^*(t)}(\xi_i^*(t))} u_i(x,t)^{p+1} \ge \Lambda - \delta/4. \end{cases} \tag{11.16}$$

By Lemma 11.4, there exists an $r(\delta) \ll 1$ such that for any $t \in [-81/100, 81/100]$,

$$\begin{cases} \displaystyle\fint_{B_{r(\delta)}} |\nabla u_i(x,t)|^2 \le \Lambda + \delta/4, \\ \displaystyle\fint_{B_{r(\delta)}} u_i(x,t)^{p+1} \le \Lambda + \delta/4. \end{cases} \tag{11.17}$$

Combining (11.16) with (11.17), we get

$$\sup_{|t| \le 81/100} \fint_{B_{r(\delta)} \setminus B_{R(\delta)\lambda_i^*(t)}(\xi_i^*(t))} \left[|\nabla u_i(x,t)|^2 + u_i(x,t)^{p+1} \right] dx \le \delta. \tag{11.18}$$

By Corollary 11.2, there exists a constant $C(\delta)$ such that

$$u_i(x,t) + |\nabla u_i(x,t)| + |\nabla^2 u_i(x,t)| + |\partial_t u_i(x,t)| \le C(\delta), \quad \text{if } x \notin B_{r(\delta)}(\xi_i^*(t)).$$

It remains to prove (11.15) when $x \in B_{r(\delta)} \setminus B_{R(\delta)\lambda_i^*(t)}(\xi_i^*(t))$. We argue by contradiction, so assume there exists a sequence (x_i, t_i), with $t_i \in [-81/100, 81/100]$ and $x_i \in B_{r(\delta)}(0) \setminus B_{R(\delta)\lambda_i^*(t_i)}(\xi_i^*(t_i))$, but for some σ (to be determined below),

$$u_i(x_i, t_i) \ge \sigma |x_i - \xi_i^*(t_i)|^{-\frac{n-2}{2}}. \tag{11.19}$$

Let $r_i := |x_i - \xi_i^*(t_i)|$ and

$$\widetilde{u}_i(y,s) := r_i^{\frac{n-2}{2}} u_i \left(\xi_i^*(t_i) + r_i y, t_i + r_i^2 s \right).$$

Denote

$$\widetilde{x}_i := \frac{x_i - \xi_i^*(t_i)}{r_i}.$$

It lies on ∂B_1. Assume it converges to a limit point $\widetilde{x}_\infty \in \partial B_1$.

Scaling (11.16) and (11.2) leads to

$$\begin{cases} \displaystyle\fint_{B_{R(\delta)^{-1}}} |\nabla \widetilde{u}_i(y,0)|^2 \geq \Lambda - \delta/4, \\[2mm] \displaystyle\fint_{B_{R(\delta)^{-1}}} \widetilde{u}_i(y,0)^{p+1} \geq \Lambda - \delta/4 \end{cases} \tag{11.20}$$

and for any $R > 0$ and $s \in \mathbb{R}$,

$$\begin{cases} \displaystyle\fint_{B_R(0)} |\nabla \widetilde{u}_i(y,s)|^2 \leq \Lambda + \delta/4, \\[2mm] \displaystyle\fint_{B_R(0)} \widetilde{u}_i(y,s)^{p+1} \leq \Lambda + \delta/4. \end{cases} \tag{11.21}$$

Then by noting that for any s, $\partial_s \widetilde{u}_i(\cdot, s)$ converges to 0 strongly in $L^2_{loc}(\mathbb{R}^n)$, we can apply Lemma 11.4 to find an $\rho(\delta) > 0$ such that for any $s \in [-1, 1]$,

$$\left| \fint_{B_{\rho(\delta)}} |\nabla \widetilde{u}_i(y,s)|^2 dy - \Lambda \right| + \left| \fint_{B_{\rho(\delta)}} \widetilde{u}_i(y,s)^{p+1} dy - \Lambda \right| < \delta.$$

Combining this estimate with (11.21), we get

$$\int_{-1}^{0} \fint_{B_{1/2}(\widetilde{x}_i)} \left(|\nabla \widetilde{u}_i|^2 + \widetilde{u}_i^{p+1} \right) \leq C\delta.$$

By choosing $C\delta < \varepsilon_*$, we can apply Theorem 3.7 to deduce that

$$\widetilde{u}_i(\widetilde{x}_i, 0) + |\nabla \widetilde{u}_i(\widetilde{x}_i, 0)| + |\nabla^2 \widetilde{u}_i(\widetilde{x}_i, 0)| + |\partial_s \widetilde{u}_i(\widetilde{x}_i, 0)| \leq c(\delta),$$

where $c(\delta)$ is a small constant depending on δ. Scaling back to u_i we get a contradiction with (11.19). In other words, (11.15) must hold in $B_{r(\delta)} \setminus B_{R(\delta)\lambda_i^*(t)}(\xi_i^*(t))$. $\qquad\square$

Combining Proposition 11.1 and Lemma 11.5. we get

Corollary 11.6. *For each $t \in [-81/100, 81/100]$, $\xi_i^*(t)$ is the unique maximal point of $u_i(t)$ in B_1.*

This completes the proof of Proposition 11.1.

12. Orthogonal and inner-outer decompositions

For simplicity of notations, from here to Section 16 we are concerned with a fixed solution u_i with sufficiently large index i, and denote it by u. In this section we define the orthogonal decomposition and inner-outer decomposition for u.

12.1. **Orthogonal decomposition.** From now on, until Section 19, a constant $K \gg 1$ will be fixed. In the following we will use the notations about bubbles and kernels to their linearized equations, $W_{\xi,\lambda}$ and $Z_{i,\xi,\lambda}$, see Appendix A.

Proposition 12.1 (Orthogonal condition). *For any $t \in [-81/100, 81/100]$, there exists a unique $(a(t), \xi(t), \lambda(t)) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^+$ with*

$$\frac{|\xi(t) - \xi^*(t)|}{\lambda(t)} + \left|\frac{\lambda(t)}{\lambda^*(t)} - 1\right| + \left|\frac{a(t)}{\lambda(t)}\right| = o(1), \tag{12.1}$$

such that for each $i = 0, \cdots, n+1$,

$$\int_{B_1} \left[u(x,t) - W_{\xi(t),\lambda(t)}(x) - a(t)Z_{0,\xi(t),\lambda(t)}(x)\right] \eta_K\left(\frac{x - \xi(t)}{\lambda(t)}\right) Z_{i,\xi(t),\lambda(t)}(x)dx = 0. \tag{12.2}$$

Proof. For these t, set

$$\widetilde{u}(y) := \lambda^*(t)^{\frac{n-2}{2}} u(\xi^*(t) + \lambda^*(t)y, t).$$

Define a smooth map F from $B_{1/2}^{n+2} \subset \mathbb{R}^{n+2}$ into $\mathbb{R}^{n+2}$ as

$$F(\xi, \lambda, a) := \left(\int_{\mathbb{R}^n} \left[\widetilde{u}(y) - W_{\xi,1+\lambda}(y) - aZ_{0,\xi,1+\lambda}(y)\right] \eta_K\left(\frac{y - \xi}{1 + \lambda}\right) Z_{i,\xi,1+\lambda}(y)dy\right)_{i=0}^{n+1}.$$

The task is reduced to find a solution of $F(\xi, \lambda, a) = 0$.

By Proposition 11.1, $|F(0,0,0)| \ll 1$. By a direct calculation and applying Proposition 11.1 once again, we find a fixed small constant ρ_* such that, for any $(\xi, \lambda, a) \in B_{\rho_*}^{n+2}$, $DF(\xi, \lambda, a)$ is diagonally dominated, and hence invertible. Then the inverse function theorem implies the existence and uniqueness of the solution $F = 0$ in $B_{\rho_*}^{n+2}$. $\qquad\square$

In the above, in order to show that $DF(\xi, \lambda, a)$ is diagonally dominated, we need $n \geq 5$ to make sure $Z_{n+1} \in L^2(\mathbb{R}^n)$, and we also need the $n \geq 7$ assumption because at first we only know that u decays like $|y|^{-(n-2)/2}$. (This term appears in the integral containing $\nabla\eta$.)

For later purpose, we notice a Lipschitz bound on the parameters $(a(t), \xi(t), \lambda(t))$.

Lemma 12.2. *For any $t \in [-81/100, 81/100]$,*

$$|a'(t)| + |\xi'(t)| + |\lambda'(t)| \lesssim \left(\int_{B_1} \partial_t u(x,t)^2 dx\right)^{1/2}. \tag{12.3}$$

Proof. For each $i = 1, \cdots, n$, because $Z_i\eta_K$ is orthogonal to Z_0 in $L^2(\mathbb{R}^n)$, the orthogonal condition (12.2) reads as

$$\int_{B_1} \left[u(x,t) - W_{\xi(t),\lambda(t)}(x)\right] \eta_K\left(\frac{x - \xi(t)}{\lambda(t)}\right) Z_{i,\xi(t),\lambda(t)}(x)dx = 0.$$

Differentiating this equality in t, by the L^2 orthogonal relations between different Z_i, we obtain

$$\left[\int_{B_1} \eta_K\left(\frac{x - \xi(t)}{\lambda(t)}\right) Z_{i,\xi(t),\lambda(t)}(x)^2 dx\right] \xi_i'(t)$$

$$= \int_{B_1} \partial_t u(x,t) \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i,\xi(t),\lambda(t)}(x) dx \qquad (12.4)$$

$$+ \int_{B_1} \left[u(x,t) - W_{\xi(t),\lambda(t)}(x) \right] \frac{\partial}{\partial t} \left[\eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i,\xi(t),\lambda(t)}(x) \right] dx.$$

First, by the definition of $Z_{i,\xi(t),\lambda(t)}$, we have

$$\int_{B_1} \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i,\xi(t),\lambda(t)}(x)^2 dx = \int_{\mathbb{R}^n} \eta_K(y) Z_i(y)^2 dy \geq c.$$

Next, by Cauchy-Schwarz inequality and (10.1), we get

$$\left| \int_{B_1} \partial_t u(x,t) \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i,\xi(t),\lambda(t)}(x) dx \right|$$

$$\leq \left(\int_{B_1} \partial_t u(x,t)^2 dx \right)^{1/2} \left(\int_{B_1} \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right)^2 Z_{i,\xi(t),\lambda(t)}(x)^2 dx \right)^{1/2}$$

$$\lesssim \left(\int_{B_1} \partial_t u(x,t)^2 dx \right)^{1/2}.$$

Finally, because in $B_{K\lambda(t)}(\xi(t))$,

$$\begin{cases} \left| u(x,t) - W_{\xi(t),\lambda(t)}(x) \right| = o(1) \lambda(t)^{-\frac{n-2}{2}} \left(1 + \frac{|x - \xi(t)|}{\lambda(t)} \right)^{-\frac{n-2}{2}}, \\ \left| \frac{\partial}{\partial t} \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i,\xi(t),\lambda(t)}(x) \right| \lesssim \frac{|\xi'(t)| + |\lambda'(t)|}{\lambda(t)^{\frac{n+2}{2}}} \left(1 + \frac{|x - \xi(t)|}{\lambda(t)} \right)^{1-n}, \\ \left| \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) \frac{\partial}{\partial t} Z_{i,\xi(t),\lambda(t)}(x) \right| \lesssim \frac{|\xi'(t)| + |\lambda'(t)|}{\lambda(t)^{\frac{n+2}{2}}} \left(1 + \frac{|x - \xi(t)|}{\lambda(t)} \right)^{1-n}, \end{cases}$$

we get

$$\left| \int_{B_1} \left[u(x,t) - W_{\xi(t),\lambda(t)}(x) \right] \frac{\partial}{\partial t} \left[\eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i,\xi(t),\lambda(t)}(x) \right] dx \right|$$

$$\ll \lambda(t)^{-n} \left(|\xi'(t)| + |\lambda'(t)| \right) \int_{B_1} \left(1 + \frac{|x - \xi(t)|}{\lambda(t)} \right)^{2 - \frac{3n}{2}}$$

$$\lesssim |\xi'(t)| + |\lambda'(t)|.$$

Plugging these three estimates into (12.4), we obtain

$$|\xi'(t)| \lesssim \left(\int_{B_1} \partial_t u(x,t)^2 dx \right)^{1/2} + o \left(|\xi'(t)| + |\lambda'(t)| \right). \qquad (12.5)$$

In the same way, we get

$$|\lambda'(t)| \lesssim \left(\int_{B_1} \partial_t u(x,t)^2 dx \right)^{1/2} + o \left(|\xi'(t)| + |\lambda'(t)| + |a'(t)| \right) \qquad (12.6)$$

and

$$|a'(t)| \lesssim \left(\int_{B_1} \partial_t u(x,t)^2 dx \right)^{1/2} + o\left(|\xi'(t)| + |\lambda'(t)| + |a'(t)| \right). \tag{12.7}$$

Adding these three estimates together we get (12.3). $\qquad\square$

12.2. Inner-outer decomposition.

In the following we denote the error function

$$\phi(x,t) := u(x,t) - W_{\xi(t),\lambda(t)}(x) - a(t)Z_{0,\xi(t),\lambda(t)}(x),$$

$$W_*(x,t) := W_{\xi(t),\lambda(t)}(x), \quad Z_{i,*}(x,t) := Z_{i,\xi(t),\lambda(t)}(x)$$

and $Z_* := (Z_{0,*}, Z_{1,*}, \cdots, Z_{n+1,*})$.

Combining Lemma 11.5 and Proposition 12.1, we obtain

Proposition 12.3. *In B_1,*

$$\begin{cases} |\phi(x,t)| = o\left((\lambda(t) + |x - \xi_*(t)|)^{-\frac{n-2}{2}} \right) + O(1), \\[2mm] |\nabla\phi(x,t)| = o\left((\lambda(t) + |x - \xi_*(t)|)^{-\frac{n}{2}} \right) + O(1), \\[2mm] |\nabla^2\phi(x,t)| + |\phi_t(x,t)| = o\left((\lambda(t) + |x - \xi_*(t)|)^{-\frac{n+2}{2}} \right) + O(1). \end{cases}$$

The error function ϕ satisfies

$$\partial_t\phi - \Delta\phi = pW_*^{p-1}\phi + \mathcal{N} + \left(-a' + \mu_0\frac{a}{\lambda^2}, \xi', \lambda' \right) \cdot Z_* - a\partial_t Z_{0,*}, \tag{12.8}$$

where $\{\}'$ denotes differentiation in t, and the nonlinear term is defined by

$$\mathcal{N} := (W_* + \phi + aZ_{0,*})^p - W_*^p - pW_*^{p-1}(\phi + aZ_{0,*}).$$

Now take a further decomposition of ϕ, *the inner-outer decomposition*, as follows. Keep K as the large constant used in Proposition 12.1. Take another constant L satisfying $1 \ll L \ll K$. Denote

$$\eta_{in}(x,t) := \eta\left(\frac{x - \xi(t)}{K\lambda(t)} \right), \quad \eta_{out}(x,t) := \eta\left(\frac{x - \xi(t)}{L\lambda(t)} \right). \tag{12.9}$$

Set

$$\phi_{in}(x,t) := \phi(x,t)\eta_K(x,t), \quad \phi_{out}(x,t) := \phi(x,t)\left[1 - \eta_{out}(x,t) \right].$$

For later purpose, we introduce two quantities:

$$\mathcal{I}(t) := [L\lambda(t)]^\alpha \sup_{B_{2L\lambda(t)}(\xi(t))\setminus B_{L\lambda(t)}(\xi(t))} |\phi| + [L\lambda(t)]^{1+\alpha} \sup_{B_{2L\lambda(t)}(\xi(t))\setminus B_{L\lambda(t)}(\xi(t))} |\nabla\phi|$$

and

$$\mathcal{O}(t) := [K\lambda(t)]^\alpha \sup_{B_{2K\lambda(t)}(\xi(t))\setminus B_{K\lambda(t)}(\xi(t))} |\phi| + [K\lambda(t)]^{1+\alpha} \sup_{B_{2K\lambda(t)}(\xi(t))\setminus B_{K\lambda(t)}(\xi(t))} |\nabla\phi|.$$

By Proposition 12.3, for any $t \in [-81/100, 81/100]$,

$$\mathcal{I}(t) + \mathcal{O}(t) \ll 1.$$

Starting from this smallness, we will improve it to an explicit bound in the following sections.

13. Inner problem

In this section we give an $C^{1,\theta}$ estimate on the inner component ϕ_{in}, see Proposition 13.4.

Define a new coordinate system around $\xi(t)$ by

$$y := \frac{x - \xi(t)}{\lambda(t)}, \qquad \tau = \tau(t).$$

Here the new time variable τ is determined by the relation

$$\tau'(t) = \lambda(t)^{-2}, \qquad \tau(0) = 0.$$

Because there is a one to one correspondence between τ and t, in the following we will not distinguish between them. For example, we will use $\mathcal{O}(\tau)$ instead of the notation $\mathcal{O}(t(\tau))$.

It is convenient to write ϕ in these new coordinates. By denoting

$$\varphi(y, \tau) := \lambda(t)^{\frac{n-2}{2}} \phi(x, t),$$

we get

$$
\begin{aligned}
\partial_\tau \varphi - \Delta \varphi &= pW^{p-1}\varphi + \left(-\frac{\dot{a}}{\lambda} + \mu_0 \frac{a}{\lambda}, \frac{\dot{\xi}}{\lambda}, \frac{\dot{\lambda}}{\lambda}\right) \cdot Z \\
&+ \mathcal{N} + \frac{\dot{\xi}}{\lambda} \cdot \nabla\varphi + \frac{\dot{\lambda}}{\lambda}\left(y \cdot \nabla\varphi + \frac{n-2}{2}\varphi\right) \\
&+ \frac{a}{\lambda}\left[\frac{\dot{\xi}}{\lambda} \cdot \nabla Z_0 + \frac{\dot{\lambda}}{\lambda}\left(y \cdot \nabla Z_0 + \frac{n}{2}Z_0\right)\right].
\end{aligned}
\tag{13.1}
$$

Here $\dot{\{\}}$ denotes differentiation in τ, and by abusing notations,

$$\mathcal{N} := \left(W + \varphi + \frac{a}{\lambda}Z_0\right)^p - W^p - pW^{p-1}\left(\varphi + \frac{a}{\lambda}Z_0\right).$$

We will need the following pointwise estimate on $\mathcal{N}$. Recall that $\bar{p} = \min\{p, 2\}$.

Lemma 13.1. *The following point-wise inequality holds:*

$$|\mathcal{N}| \lesssim |\varphi|^{\bar{p}} + \left|\frac{a}{\lambda}\right|^{\bar{p}} Z_0^{\bar{p}}.$$

Proof. There exists a $\vartheta \in [0, 1]$ such that

$$
\begin{aligned}
&\left(W + \varphi + \frac{a}{\lambda}Z_0\right)^p - W^p - pW^{p-1}\left(\varphi + \frac{a}{\lambda}Z_0\right) \\
&= p\left\{\left[W + \vartheta\left(\varphi + \frac{a}{\lambda}Z_0\right)\right]^{p-1} - W^{p-1}\right\}\left(\varphi + \frac{a}{\lambda}Z_0\right).
\end{aligned}
$$

Since both W and $W + \varphi + \frac{a}{\lambda}Z_0 = u$ are bounded by a universal constant, and $\varphi + \frac{a}{\lambda}Z_0$ is small, by considering the two cases $W \geq |\varphi + \frac{a}{\lambda}Z_0|$ and $W < |\varphi + \frac{a}{\lambda}Z_0|$ separately, we get

$$|\mathcal{N}| \lesssim \left|\varphi + \frac{a}{\lambda}Z_0\right|^{\bar{p}} \lesssim |\varphi|^{\bar{p}} + \left|\frac{a}{\lambda}\right|^{\bar{p}} Z_0^{\bar{p}}. \qquad \square$$

By scaling the estimates in Proposition 12.3, we obtain

Proposition 13.2. *For each τ,*

$$K^{\alpha} \sup_{y \in B_{2K}} \left[|\varphi(y,\tau)| + K|\nabla\varphi(y,\tau)| + K^2 \left(|\nabla^2\varphi(y,\tau)| + |\partial_{\tau}\varphi(y,\tau)| \right) \right] = o(1).$$

In (y,τ) coordinates, the inner error function ϕ_{in} is

$$\varphi_K(y,\tau) := \varphi(y,\tau)\eta_K(y).$$

For each τ, the support of $\varphi_K(\cdot,\tau)$ is contained in B_K. By (12.2), for each $i = 0, \cdots, n+1$,

$$\int_{\mathbb{R}^n} \varphi_K(y,\tau)Z_i(y)dy = 0. \tag{13.2}$$

The equation for φ_K is

$$\partial_{\tau}\varphi_K - \Delta_y\varphi_K = pW^{p-1}\varphi_K + \lambda^{-1}\left(-\dot{a} + \mu_0 a, \dot{\xi}, \dot{\lambda}\right) \cdot Z + E_K. \tag{13.3}$$

Here

$$
\begin{aligned}
E_K \quad := \quad & -2\nabla\varphi \cdot \nabla\eta_K - \varphi\Delta\eta_K + \mathcal{N}\eta_K \\
+ \quad & \lambda^{-1}\left(-\dot{a} + \mu_0 a, \dot{\xi}, \dot{\lambda}\right) \cdot Z\left(\eta_K - 1\right) \\
+ \quad & \lambda^{-1}\left[\dot{\lambda}\left(\frac{n-2}{2}\varphi + y \cdot \nabla_y\varphi\right) + \dot{\xi} \cdot \nabla_y\varphi\right]\eta_K \\
+ \quad & \frac{a}{\lambda}\left[\frac{\dot{\xi}}{\lambda} \cdot \nabla Z_0 + \frac{\dot{\lambda}}{\lambda}\left(y \cdot \nabla Z_0 + \frac{n}{2}Z_0\right)\right]\eta_K.
\end{aligned}
$$

The following estimate holds for E_K.

Lemma 13.3. *For any τ, we have*

$$
\begin{aligned}
\|E_K(\tau)\|_{2+\alpha} \quad \lesssim \quad & \mathcal{O}(\tau) + K^{2+\alpha-\bar{p}\alpha}\|\varphi_K(\tau)\|_{\alpha}^{\bar{p}} + \left|\frac{a(\tau)}{\lambda(\tau)}\right|^{\bar{p}} \\
+ \quad & K^{3+\alpha-n}\left|\frac{\dot{\xi}(\tau)}{\lambda(\tau)}\right| + K^{4+\alpha-n}\left|\frac{\dot{\lambda}(\tau)}{\lambda(\tau)}\right| + e^{-cK}\left|\frac{\dot{a}(\tau) - \mu_0 a(\tau)}{\lambda(\tau)}\right|.
\end{aligned}
$$

Proof. (1) Because $\alpha = (n-2)/2$, by the definition of $\mathcal{O}$, we have

$$K^{2+\alpha}\sup_y |2\nabla\varphi(y,\tau) \cdot \nabla\eta_K(y) + \varphi(y,\tau)\Delta\eta_K(y)| \lesssim \mathcal{O}(\tau).$$

(2) By Lemma 13.1, we obtain

$$
\begin{aligned}
|\mathcal{N}\eta_K| \quad \lesssim \quad & |\varphi|^{\bar{p}}\eta_K + \left|\frac{a}{\lambda}\right|^{\bar{p}} Z_0^{\bar{p}} \\
= \quad & |\varphi|^{\bar{p}}\left(\eta_K - \eta_K^{\bar{p}}\right) + |\varphi_K|^{\bar{p}} + \left|\frac{a}{\lambda}\right|^{\bar{p}} Z_0^{\bar{p}}.
\end{aligned}
$$

First, because $\mathcal{O}(\tau) \ll 1$,

$$\sup_{y \in \mathbb{R}^n} K^{2+\alpha}|\varphi(y,\tau)|^{\bar{p}}\left[\eta_K(y) - \eta_K(y)^{\bar{p}}\right] \lesssim K^{2+\alpha}\mathcal{O}(\tau)^{\bar{p}} \ll \mathcal{O}(\tau).$$

Secondly, because the support of φ_K is contained in B_K,

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} |\varphi_K(y, \tau)|^{\bar{p}} \lesssim K^{2+\alpha - \bar{p}\alpha} \|\varphi_K(\tau)\|_\alpha^{\bar{p}}.$$

Finally, by the exponential decay of Z_0 at infinity,

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} \left| \frac{a(\tau)}{\lambda(\tau)} \right|^{\bar{p}} |Z_0(y)|^{\bar{p}} \lesssim \left| \frac{a(\tau)}{\lambda(\tau)} \right|^{\bar{p}}.$$

(3) For $i = 1, \cdots, n$, $Z_i(y) = O\left(|y|^{1-n}\right)$. Hence

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} \left| \frac{\dot{\xi}(\tau)}{\lambda(\tau)} \cdot Z(y) \left[\eta_K(y) - 1 \right] \right| \lesssim K^{3+\alpha-n} \left| \frac{\dot{\xi}(\tau)}{\lambda(\tau)} \right|.$$

(4) As in the previous case, because $Z_{n+1}(y) = O\left(|y|^{2-n}\right)$, we get

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} \left| \frac{\dot{\lambda}(\tau)}{\lambda(\tau)} Z_{n+1}(y) \left[\eta_K(y) - 1 \right] \right| \lesssim K^{4+\alpha-n} \left| \frac{\dot{\lambda}(\tau)}{\lambda(\tau)} \right|.$$

(5) By the exponential decay of Z_0, we get

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} \left| \frac{\dot{a}(\tau) - \mu_0 a(\tau)}{\lambda(\tau)} Z_0(y) \left[\eta_K(y) - 1 \right] \right| \lesssim e^{-cK} \frac{|\dot{a}(\tau) - \mu_0 a(\tau)|}{\lambda(\tau)}.$$

(6) By Proposition 13.2, we have

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} \left| \frac{\dot{\lambda}(\tau)}{\lambda(\tau)} \left[\frac{n-2}{2} \varphi(y) + y \cdot \nabla_y \varphi(y) \right] \eta_K(y) \right| \ll K^{4+\alpha-n} \frac{|\dot{\lambda}(\tau)|}{\lambda(\tau)}.$$

(7) As in the previous case,

$$\sup_{y \in \mathbb{R}^n} (1 + |y|)^{2+\alpha} \left| \frac{a(\tau)}{\lambda(\tau)} \left[\frac{\dot{\xi}(\tau)}{\lambda(\tau)} \cdot \nabla Z_0(y) + \frac{\dot{\lambda}(\tau)}{\lambda(\tau)} \left(y \cdot \nabla Z_0(y) + \frac{n}{2} Z_0(y) \right) \right] \eta_K(y) \right|$$

$$\ll K^{4+\alpha-n} \left(\frac{|\dot{\xi}(\tau)|}{\lambda(\tau)} + \frac{|\dot{\lambda}(\tau)|}{\lambda(\tau)} \right).$$

Putting these estimates together we finish the proof. $\qquad\square$

Because $n \geq 7$ and $\alpha = (n-2)/2$,

$$4 + \alpha - n < 0.$$

Furthermore, by noting that those terms like $\|\varphi_K(\tau)\|_\alpha$ are small, this lemma can be restated as

$$\|E_K(\tau)\|_{2+\alpha} \lesssim o\left(\|\varphi_K(\tau)\|_\alpha + \left| \frac{\dot{\xi}(\tau)}{\lambda(\tau)} \right| + \left| \frac{\dot{\lambda}(\tau)}{\lambda(\tau)} \right| + \left| \frac{a(\tau)}{\lambda(\tau)} \right| + \left| \frac{\dot{a}(\tau) - \mu_0 a(\tau)}{\lambda(\tau)} \right| \right)$$

$$+ \mathcal{O}(\tau). \tag{13.4}$$

In particular, the main order term in E_K is $\mathcal{O}(\tau)$, which comes from the outer component.

Now we prove our main estimate in this section.

Proposition 13.4 ($C^{1,\theta}$ estimates). • *For any τ,*

$$\left|\frac{\dot{\xi}(\tau)}{\lambda(\tau)}\right| + \left|\frac{\dot{\lambda}(\tau)}{\lambda(\tau)}\right| + \left|\frac{\dot{a}(\tau) - \mu_0 a(\tau)}{\lambda(\tau)}\right| \lesssim \mathcal{O}(\tau) + K^{2+\alpha-\bar{p}\alpha}\|\varphi_K(\tau)\|_\alpha^{\bar{p}} + \left|\frac{a(\tau)}{\lambda(\tau)}\right|^{\bar{p}}. \quad (13.5)$$

• *For any $\sigma > 0$, there exist two constants C (universal) and $T(\sigma, K)$ such that for any $\tau_1 < \tau_2$,*

$$\|\varphi_K\|_{C^{(1+\theta)/2}(\tau_1, \tau_2; \mathcal{X}_\alpha^{1+\theta})} \leq \sigma\|\varphi_K(\tau_1 - T(\sigma, K))\|_{\mathcal{X}_\alpha} \quad (13.6)$$

$$+ C\left\|\left(\mathcal{O}, \left|\frac{a}{\lambda}\right|^{\bar{p}}\right)\right\|_{L^\infty(\tau_1 - T(\sigma, K), \tau_2)}.$$

Proof. Take a decomposition $\varphi_K = \varphi_{K,1} + \varphi_{K,2}$, where $\varphi_{K,1}$ is the solution of (A.3) with initial value (at $\tau_1 - T(\sigma)$, with $T(\sigma, K)$ to be determined below) φ_K, and $\varphi_{K,2}$ is the solution of (A.8) with non-homogeneous term E_K.

To apply Lemma A.4, take an $\alpha' > n/2$. Because φ_K is supported in B_K, we have

$$\|\varphi_K(\tau_1 - T(\sigma, K))\|_{\alpha'} \leq (2K)^{\alpha'-\alpha} \|\varphi_K(\tau_1 - T(\sigma, K))\|_\alpha. \quad (13.7)$$

By Lemma A.4, there exists a $T(\sigma, K) > 0$ such that

$$\|\varphi_{K,1}\|_{C^{(1+\theta)/2}(\tau_1, \tau_2; C^{1+\theta}(B_{2K}))} \leq \sigma (4K)^{-1-\alpha} (2K)^{\alpha-\alpha'}. \quad (13.8)$$

For $\varphi_{K,2}$, in view of the estimate in Lemma 13.3, a direct application of Lemma A.5 gives

$$\|\varphi_{K,2}\|_{C^{(1+\theta)/2}(\tau_1, \tau_2; \mathcal{X}_\alpha^{1+\theta})} \leq C(\alpha)\left\|\left(\mathcal{O}, \left|\frac{a}{\lambda}\right|^{\bar{p}}\right)\right\|_{L^\infty(\tau_1 - T(\sigma, K), \tau_2)}. \quad (13.9)$$

Adding (13.8) and (13.9), by using again the fact that $\varphi_K(\tau)$ is supported in B_K, we get (13.6). □

For applications in Section 18, we give another estimate on the parameters $\dot{\lambda}$ etc. in terms of the L^∞ norm of φ.

Lemma 13.5. *For each τ,*

$$\left|\frac{\dot{a}(\tau) - \mu_0 a(\tau)}{\lambda(\tau)}\right| + \left|\frac{\dot{\xi}(\tau)}{\lambda(\tau)}\right| + \left|\frac{\dot{\lambda}(\tau)}{\lambda(\tau)}\right| \lesssim \sup_{B_{2K}\setminus B_K} |\varphi(\tau)| + \sup_{B_{2K}} |\varphi(\tau)|^{\bar{p}} + \left|\frac{a(\tau)}{\lambda(\tau)}\right|^{\bar{p}}. \quad (13.10)$$

Proof. For each $j = 0, \cdots, n+1$, multiplying (13.3) by Z_j, utilizing the orthogonal condition (13.2), we get

$$\begin{cases} \dfrac{\dot{a}(\tau) - \mu_0 a(\tau)}{\lambda(\tau)} = \displaystyle\int_{\mathbb{R}^n} Z_0(y) E_K(y, \tau) dy, \\[2ex] \dfrac{\dot{\xi}_j(\tau)}{\lambda(\tau)} = -\displaystyle\int_{\mathbb{R}^n} Z_j(y) E_K(y, \tau) dy, \quad j = 1, \cdots, n, \\[2ex] \dfrac{\dot{\lambda}(\tau)}{\lambda(\tau)} = -\displaystyle\int_{\mathbb{R}^n} Z_{n+1}(y) E_K(y, \tau) dy. \end{cases}$$

An integration by parts gives

$$\left| \int_{\mathbb{R}^n} (2\nabla\varphi \cdot \nabla\eta_K + \varphi\Delta\eta_K) Z_j \right| = \left| \int_{\mathbb{R}^n} \varphi \left(2\nabla\eta_K \cdot \nabla Z_j + \Delta\eta_K Z_j\right) \right|$$
$$\lesssim \sup_{B_{2K}\setminus B_K} |\varphi|. \tag{13.11}$$

By the at most $|y|^{2-n}$ decay of Z_j, Lemma 13.1 and Proposition 13.2, we find the contribution from other terms in E_K is of the order

$$O\left(\sup_{y\in B_{2K}} |\varphi(y,\tau)|^{\bar p} + \left|\frac{a(\tau)}{\lambda(\tau)}\right|^{\bar p}\right) + o\left(\left|\frac{\dot\xi(\tau)}{\lambda(\tau)}\right| + \left|\frac{\dot\lambda(\tau)}{\lambda(\tau)}\right| + \left|\frac{\dot a(\tau) - \mu_0 a(\tau)}{\lambda(\tau)}\right|\right).$$

Adding these estimates for $\frac{\dot a(\tau)-\mu_0 a(\tau)}{\lambda(\tau)}$, $\frac{\dot\xi_j(\tau)}{\lambda(\tau)}$ and $\frac{\dot\lambda(\tau)}{\lambda(\tau)}$ together, we get (13.10). $\qquad\square$

Finally, we will need the following backward estimate on a/λ.

Lemma 13.6. *For any $\tau_1 < \tau_2$, we have*

$$\frac{|a(\tau_1)|}{\lambda(\tau_1)} \lesssim e^{-\frac{\mu_0}{2}(\tau_2-\tau_1)}\frac{|a(\tau_2)|}{\lambda(\tau_2)} + \int_{\tau_1}^{\tau_2} e^{-\frac{\mu_0}{2}(\tau_2-s)} \left[\mathcal{O}(s) + \|\varphi_K(s)\|_\alpha^{\bar p}\right] ds. \tag{13.12}$$

Proof. By a direct differentiation and then applying (13.5), we obtain

$$\frac{d}{d\tau}\frac{a(\tau)}{\lambda(\tau)} = \frac{\dot a(\tau)}{\lambda(\tau)} - \frac{\dot\lambda(\tau)}{\lambda(\tau)}\frac{a(\tau)}{\lambda(\tau)}$$
$$= \mu_0\frac{a(\tau)}{\lambda(\tau)} + O\left(\mathcal{O}(\tau) + \|\varphi_K(\tau)\|_\alpha^{\bar p} + \left|\frac{a(\tau)}{\lambda(\tau)}\right|^{\bar p}\right)$$
$$= [\mu_0 + o(1)]\frac{a(\tau)}{\lambda(\tau)} + O\left(\mathcal{O}(\tau) + \|\varphi_K(\tau)\|_\alpha^{\bar p}\right).$$

Integrating this differential inequality gives (13.12). $\qquad\square$

14. Outer problem

In this section we establish an estimate on the outer component ϕ_{out}. Our main goal is to obtain an estimate of $\mathcal{O}$ in terms of $\mathcal{I}$ and the parameters λ' etc..

14.1. Decomposition of the outer equation. First we need to take a further decomposition of ϕ_{out}. Recall that ϕ_{out} satisfies

$$\begin{aligned}
\partial_t\phi_{out} - \Delta\phi_{out} &= (u^p - W_*^p)(1 - \eta_{out}) \\
&\quad - \phi\left(\partial_t\eta_{out} - \Delta\eta_{out}\right) + 2\nabla\phi \cdot \nabla\eta_{out} \\
&\quad + \left(-a' + \mu_0\frac{a}{\lambda^2}, \xi', \lambda'\right) Z_* (1 - \eta_{out}) \\
&\quad - a\partial_t Z_0^* (1 - \eta_{out}).
\end{aligned} \tag{14.1}$$

Introducing

$$
\begin{cases}
V_* := \begin{cases} \frac{u^p - W_*^p}{u - W_*} \chi_{B_{L\lambda(t)}(\xi(t))^c}, & \text{if } u \neq W_*, \\ p u^{p-1} \chi_{B_{L\lambda(t)}(\xi(t))^c}, & \text{otherwise}, \end{cases} \\
F_2 := -\phi \left(\partial_t \eta_{out} - \Delta \eta_{out} \right) + 2\nabla\phi \cdot \nabla\eta_{out}, \\
F_3 := \lambda' Z_{n+1,*} \left(1 - \eta_{out} \right), \\
F_4 := \sum_{j=1}^{n} \xi_j' Z_{j,*} \left(1 - \eta_{out} \right), \\
F_5 := \left[a V_* Z_0^* - \left(a' - \mu_0 \frac{a}{\lambda^2} \right) Z_0^* - a \partial_t Z_0^* \right] \left(1 - \eta_{out} \right),
\end{cases}
$$

then (14.1) is written as

$$
\partial_t \phi_{out} - \Delta \phi_{out} = V_* \phi_{out} + F_2 + F_3 + F_4 + F_5. \tag{14.2}
$$

Corresponding to this decomposition of the right hand side, we have $\phi_{out} = \phi_1 + \phi_2 + \phi_3 + \phi_4 + \phi_5$, with these five functions solving the following five equations respectively: first,

$$
\begin{cases} \partial_t \phi_1 - \Delta \phi_1 = V_* \phi_1 & \text{in } Q_{9/10}, \\ \phi_1 = \phi_{out}, & \text{on } \partial^p Q_{9/10}; \end{cases}
$$

next,

$$
\begin{cases} \partial_t \phi_2 - \Delta \phi_2 = V_* \phi_2 + F_2 & \text{in } Q_{9/10}, \\ \phi_2 = 0, & \text{on } \partial^p Q_{9/10}; \end{cases}
$$

and ϕ_3, ϕ_4, ϕ_5 solve the same equation with ϕ_2, but with the right hand side term replaced by F_3, F_4 and F_5 respectively.

The following estimates hold for V_* and F_2-F_5.

(i) Because $u > 0$ and $W_* > 0$, by convexity of the function $x \mapsto x^p$ and Lemma 11.5, for any $\delta > 0$ and $C > 0$, there exists a constant $R(\delta) > 0$ such that, if $L \geq R(\delta)$, then

$$
|V_*| \leq p \left(u^{p-1} + W_*^{p-1} \right) \leq \frac{\delta}{|x - \xi(t)|^2} + C. \tag{14.3}
$$

(ii) By the definition of η_{out} in (12.9), we get

$$
\begin{aligned}
& \left| \phi \left(\partial_t \eta_{out} - \Delta \eta_{out} \right) \right| \\
\lesssim{}& |\phi| \left(\frac{|\lambda'(t)|}{\lambda(t)} + \frac{|\xi'(t)|}{L\lambda(t)} + \frac{1}{L^2 \lambda(t)^2} \right) \chi_{B_{2L\lambda(t)}(\xi(t)) \backslash B_{L\lambda(t)}(\xi(t))} \\
\lesssim{}& \frac{\mathcal{I}(t)}{L^\alpha \lambda(t)^\alpha} \left(\frac{|\lambda'(t)|}{\lambda(t)} + \frac{|\xi'(t)|}{L\lambda(t)} + \frac{1}{L^2 \lambda(t)^2} \right) \chi_{B_{2L\lambda(t)}(\xi(t)) \backslash B_{L\lambda(t)}(\xi(t))}.
\end{aligned} \tag{14.4}
$$

In the same way, we get

$$
\begin{aligned}
2|\nabla\phi||\nabla\eta_{out}| &\lesssim \frac{|\nabla\phi|}{L\lambda(t)} \chi_{B_{2L\lambda(t)}(\xi(t)) \backslash B_{L\lambda(t)}(\xi(t))} \\
&\lesssim \frac{\mathcal{I}(t)}{L^{2+\alpha} \lambda(t)^{2+\alpha}} \chi_{B_{2L\lambda(t)}(\xi(t)) \backslash B_{L\lambda(t)}(\xi(t))}.
\end{aligned} \tag{14.5}
$$

Combining (14.4) and (14.5), we obtain

$$|F_2| \lesssim \frac{\mathcal{I}(t)}{L^\alpha \lambda(t)^\alpha} \left[\frac{|\lambda'(t)|}{\lambda(t)} + \frac{|\xi'(t)|}{L\lambda(t)} + \frac{1}{L^2\lambda(t)^2} \right] \chi_{B_{2L\lambda(t)}(\xi(t)) \setminus B_{L\lambda(t)}(\xi(t))}. \qquad (14.6)$$

(iii) By the decay of Z_{n+1} at infinity, we get

$$|F_3| \lesssim \lambda(t)^{\frac{n-4}{2}} |\lambda'(t)| |x - \xi(t)|^{2-n} \chi_{B_{L\lambda(t)}(\xi(t))^c}. \qquad (14.7)$$

(iv) By the decay of $Z_1, \cdots, Z_n$ at infinity, we get

$$|F_4| \lesssim \lambda(t)^{\frac{n-2}{2}} |\xi'(t)| |x - \xi(t)|^{1-n} \chi_{B_{L\lambda(t)}(\xi(t))^c}. \qquad (14.8)$$

(v) By the decay of Z_0 at infinity, we get

$$\left| \left(a' - \mu_0 \frac{a}{\lambda^2} \right) Z_0^* (1 - \eta_{out}) \right| \lesssim \left| a'(t) - \mu_0 \frac{a(t)}{\lambda(t)^2} \right| \lambda(t)^{-\frac{n}{2}} e^{-c\frac{|x-\xi(t)|}{\lambda(t)}} \chi_{B_{L\lambda(t)}(\xi(t))^c}. \qquad (14.9)$$

Similarly, we have

$$\left| a\partial_t Z_0^* (1 - \eta_{out}) \right| \lesssim |a(t)| \lambda(t)^{-\frac{n+2}{2}} \left(|\lambda'(t)| + |\xi'(t)| \right) e^{-c\frac{|x-\xi(t)|}{\lambda(t)}} \chi_{B_{L\lambda(t)}(\xi(t))^c}. \qquad (14.10)$$

Finally, by (14.3), we obtain

$$\begin{aligned} |a V_* Z_0^* (1 - \eta_{out})| &\lesssim \frac{|a(t)|}{|x - \xi(t)|^2} \lambda(t)^{-\frac{n}{2}} e^{-c\frac{|x-\xi(t)|}{\lambda(t)}} \chi_{B_{L\lambda(t)}(\xi(t))^c} \qquad (14.11) \\ &\lesssim \frac{|a(t)|}{\lambda(t)^2} \lambda(t)^{-\frac{n}{2}} e^{-c\frac{|x-\xi(t)|}{\lambda(t)}} \chi_{B_{L\lambda(t)}(\xi(t))^c}. \end{aligned}$$

Combining (14.9)-(14.11), we get

$$\begin{aligned} |F_5| &\lesssim \left[\frac{|a(t)|}{\lambda(t)^2} + \left| a'(t) - \mu_0 \frac{a(t)}{\lambda(t)^2} \right| + \frac{|a(t)|}{\lambda(t)} \left(|\lambda'(t)| + |\xi'(t)| \right) \right] \qquad (14.12) \\ &\quad \times \lambda(t)^{-\frac{n}{2}} e^{-c\frac{|x-\xi(t)|}{\lambda(t)}} \chi_{B_{L\lambda(t)}(\xi(t))^c}. \end{aligned}$$

For each $i = 1, \cdots, 5$, let

$$\overline{\phi}_i(x, t) := |\phi_i (x - \xi(t), t)|. \qquad (14.13)$$

By the Kato inequality, we obtain

$$\partial_t \overline{\phi}_1 - \Delta \overline{\phi}_1 + \xi'(t) \cdot \nabla \overline{\phi}_1 \leq |V_*| \overline{\phi}_1 \qquad (14.14)$$

and for $i = 2, 3, 4, 5$,

$$\partial_t \overline{\phi}_i - \Delta \overline{\phi}_i + \xi'(t) \cdot \nabla \overline{\phi}_i \leq |V_*| \overline{\phi}_i + |F_i|. \qquad (14.15)$$

14.2. **Estimates for the outer equation.** In this subsection, we establish some pointwise estimates on $\overline{\phi}_1, \cdots, \overline{\phi}_5$.

Denote the Dirichlet heat kernel for the operator $\mathcal{H} := \partial_t - \Delta - (\delta|x|^{-2} + C) + \xi'(t) \cdot \nabla$ in Q_1 by $G(x, t; y, s)$ $(t > s)$.

Let

$$\gamma := \frac{n-2}{2} - \sqrt{\frac{(n-2)^2}{4} - 4\delta}.$$

Then $w(x) := |x|^{-\gamma}$ is a weak solution of the elliptic equation

$$-\Delta w(x) = \frac{\delta}{|x|^2} w(x).$$

Lemma 14.1 (Estimate on $\overline{\phi}_1$). $\overline{\phi}_1(x, t) \lesssim |x|^{-\gamma}$ in $Q_{8/9}$.

Proof. Set $\varphi_1 := \overline{\phi}_1/w$. It is a sub-solution to the parabolic equation

$$\partial_t \varphi = w^{-2} \text{div} \left(w^2 \nabla \varphi \right) - \xi'(t) \cdot \nabla \varphi + \left[C + \gamma \xi'(t) \cdot \frac{x}{|x|^2} \right] \varphi. \tag{14.16}$$

We note the following three facts about the coefficients of this equation.

(1) Because $\gamma \ll 1$, by [71, Theorem 3.1], the volume doubling property and the scaling invariant Poincare inequality holds on $\mathbb{R}^n$ with the weighted measure $w^2 dx$. By [76, Theorem 5.2.3], for some $q > 2$ (independent of γ), there is a local Sobolev embedding from $W^{1,2}(B_1, w^2 dx)$ into $L^q(B_1, w^2 dx)$.
(2) By the Lipschitz hypothesis (10.1), $\xi'(t) \in L^\infty(-1, 1)$.
(3) Similarly, the zeroth order term $C + \gamma \xi'(t) \cdot \frac{x}{|x|^2} \in L^\infty(-1, 1; L^{n-3\gamma}(B_1, w^2 dx))$.

Then by standard De Giorgi-Nash-Moser estimate (cf. [71, Section 3] and [76, Chapter 5]), we deduce that φ_1 is bounded in $Q_{8/9}$. (The lower order terms in this parabolic operator do not affect this argument by noting their higher integrability.) $\square$

In fact, Moser's Harnack inequality holds for positive solutions of (14.16). This is similar to [71, Theorem 3.5]. Then by [76, Section 5.4.7] (see also [71, Theorem 4.3]), the heat kernel G satisfies a Gaussian bound

$$G(x, t; y, s) \leq C (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1 + \frac{\sqrt{t-s}}{|x|} \right)^\gamma \left(1 + \frac{\sqrt{t-s}}{|y|} \right)^\gamma. \tag{14.17}$$

Hereafter we fix two constants $\beta > 1$ (but sufficiently close to 1) and $\mu \in (0, 1)$ (to be determined below). We choose μ so that

$$\left(\frac{n-2}{2} - 2\gamma \right) (1 - \mu) > 2.$$

which is denoted by $2 + 2\kappa$. This inequality is guaranteed by the assumption that $n \geq 7$.

Lemma 14.2 (Estimate on $\overline{\phi}_2$). *If $K\lambda(t)/2 \leq |x| \leq 4K\lambda(t)$, then*

$$\overline{\phi}_2(x, t) \lesssim \left[K^{-\left(\frac{n-2}{2} - \gamma\right)(\beta-1)} + \frac{L^{\frac{n-2}{2} - \gamma}}{K^{\frac{n-2}{2} - (2\beta-1)\gamma}} \right] [K\lambda(t)]^{-\frac{n-2}{2}} \|\mathcal{I}\|_{L^\infty(t-\lambda(t)^{2\mu}, t)}$$

$$+ \quad \lambda(t)^{-\gamma(1-\mu)-\frac{n-2}{2}\mu}.$$

Proof. By the heat kernel representation formula, for any (x,t) we have

$$\overline{\phi}_2(x,t) \le \int_{-81/100}^{t} \int_{B_{2L\lambda(s)} \setminus B_{L\lambda(s)}} G(x,t;y,s) \frac{\mathcal{I}(s)}{L^\alpha \lambda(s)^\alpha} \left(\frac{|\lambda'(s)|}{\lambda(s)} + \frac{|\xi'(s)|}{L\lambda(s)} + \frac{1}{L^2\lambda(s)^2} \right).$$

Divide this integral into three parts, the first part I being on $(-81/100, t - \lambda(t)^{2\mu})$, the second part II involving the integral on $(t - \lambda(t)^{2\mu}, t - K^{2\beta}\lambda(t)^2)$, and the third part III involving the integral on $(t - K^{2\beta}\lambda(t)^2, t)$.

Estimate of I. By (13.5),

$$|\lambda'(s)| + |\xi'(s)| \ll \lambda(s)^{-1}. \tag{14.18}$$

Hence

$$\left\| \frac{\mathcal{I}(s)}{L^\alpha \lambda(s)^\alpha} \left(\frac{|\lambda'(s)|}{\lambda(s)} + \frac{|\xi'(s)|}{L\lambda(s)} + \frac{1}{L^2\lambda(s)^2} \right) \chi_{B_{2L\lambda(s)} \setminus B_{L\lambda(s)}} \right\|_{L^{\frac{2n}{n+2}}(B_1)} \lesssim \mathcal{I}(s).$$

By the Guassian bound on $G(x,y;t,s)$ in (14.17),

$$\|G(x,t;\cdot,s)\|_{L^{\frac{2n}{n-2}}(B_1)} \lesssim (t-s)^{-\frac{n+2}{4}} \left(1 + \frac{\sqrt{t-s}}{|x|} \right)^\gamma. \tag{14.19}$$

Then by Hölder inequality we get

$$\mathrm{I} \;\lesssim\; |x|^{-\gamma} \int_{-81/100}^{t - \lambda(t)^{2\mu}} (t-s)^{-\frac{n+2}{4}+\frac{\gamma}{2}} \mathcal{I}(s) ds \tag{14.20}$$

$$\lesssim\; K^{-\gamma} \lambda(t)^{-(1-\mu)\gamma - \frac{n-2}{2}\mu},$$

where in the last step we used only the estimate $|\mathcal{I}(s)| \le C$.

Estimate of II. This case differs from the previous one only in the last step. Now we have

$$\mathrm{II} \;\lesssim\; |x|^{-\gamma} \int_{t-\lambda(t)^{2\mu}}^{t-K^{2\beta}\lambda(t)^2} (t-s)^{-\frac{n+2}{4}+\frac{\gamma}{2}} \mathcal{I}(s)$$

$$\lesssim\; K^{-\frac{n-2}{2}\beta+(\beta-1)\gamma} \lambda(t)^{-\frac{n-2}{2}} \|\mathcal{I}\|_{L^\infty(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)}.$$

Estimate of III. Still by the heat kernel representation formula, III is bounded by

$$\int_{t-K^{2\beta}\lambda(t)^2}^{t} \int_{B_{2L\lambda(s)} \setminus B_{L\lambda(s)}} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1 + \frac{\sqrt{t-s}}{|x|} \right)^\gamma \left(1 + \frac{\sqrt{t-s}}{|y|} \right)^\gamma$$

$$\times \frac{\mathcal{I}(s)}{L^\alpha \lambda(s)^\alpha} \left(\frac{|\lambda'(s)|}{\lambda(s)} + \frac{|\xi'(s)|}{L\lambda(s)} + \frac{1}{L^2\lambda(s)^2} \right) dy ds$$

$$= \int_{t-K^{2\beta}\lambda(t)^2}^{t} \frac{\mathcal{I}(s)}{L^\alpha \lambda(s)^\alpha} \left(\frac{|\lambda'(s)|}{\lambda(s)} + \frac{|\xi'(s)|}{L\lambda(s)} + \frac{1}{L^2\lambda(s)^2} \right) \left(1 + \frac{\sqrt{t-s}}{|x|} \right)^\gamma$$

$$\times \left[\int_{B_{2L\lambda(s)} \setminus B_{L\lambda(s)}} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1 + \frac{\sqrt{t-s}}{|y|} \right)^\gamma dy \right] ds.$$

For $s \in [t - K^{2\beta}\lambda(t)^2, t]$, by Proposition 11.1 we have

$$\lambda(s) \sim \lambda(t). \tag{14.21}$$

Thus by noting that $K\lambda(t)/2 \leq |x| \leq 4K\lambda(t)$ and $K \gg L$, we have

$$|x| \gg |y| \quad \text{for any } y \in B_{2L\lambda(s)}.$$

Therefore in the integral above we can replace $e^{-c\frac{|x-y|^2}{t-s}}$ by $e^{-c\frac{|x|^2}{t-s}}$.

Then in view of (14.18), III is controlled by

$$K^{(\beta-1)\gamma}\|\mathcal{I}\|_{L^\infty(t-K^{2\beta}\lambda(t)^2,t)}L^{\frac{n-2}{2}}\lambda(t)^{\frac{n-2}{2}}$$

$$\times \int_{t-K^{2\beta}\lambda(t)^2}^t (t-s)^{-\frac{n}{2}}e^{-c\frac{|x|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{L\lambda(t)}\right)^\gamma ds$$

$$\lesssim K^{(\beta-1)\gamma}\|\mathcal{I}\|_{L^\infty(t-K^{2\beta}\lambda(t)^2,t)}L^{\frac{n-2}{2}}\lambda(t)^{\frac{n-2}{2}}$$

$$\left(1+\frac{K^\beta}{L}\right)^\gamma \int_{t-K^{2\beta}\lambda(t)^2}^t (t-s)^{-\frac{n}{2}}e^{-c\frac{|x|^2}{t-s}}ds$$

$$\lesssim K^{(2\beta-1)\gamma}L^{\frac{n-2}{2}-\gamma}\|\mathcal{I}\|_{L^\infty(t-K^{2\beta}\lambda(t)^2,t)}\lambda(t)^{\frac{n-2}{2}}|x|^{2-n}$$

$$\lesssim \frac{L^{\frac{n-2}{2}-\gamma}}{K^{\frac{n-2}{2}-(2\beta-1)\gamma}}\left[K\lambda(t)\right]^{-\frac{n-2}{2}}\|\mathcal{I}\|_{L^\infty(t-K^2\lambda(t)^2,t)}.$$

Adding up the estimates for I, II and II, we finish the proof. $\qquad\square$

Lemma 14.3 (Estimate on $\overline{\phi}_3$). *For $K\lambda(t)/2 \leq |x| \leq 4K\lambda(t)$,*

$$\overline{\phi}_3(x,t) \lesssim \lambda(t)^{-\frac{n-2}{2}\mu+2(\mu-1)\gamma} + K^{-\frac{n-2}{2}(\beta-1)+2(\beta-1)\gamma}\left[K\lambda(t)\right]^{-\frac{n-2}{2}}\|\lambda\lambda'\|_{L^\infty(t-\lambda(t)^{2\mu},t)}.$$

Proof. By the heat kernel representation,

$$\overline{\phi}_3(x,t) \lesssim \int_{-81/100}^t \int_{B_{L\lambda(s)}^c}(t-s)^{-\frac{n}{2}}e^{-c\frac{|x-y|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{|x|}\right)^\gamma\left(1+\frac{\sqrt{t-s}}{|y|}\right)^\gamma$$

$$\times \lambda(s)^{\frac{n-4}{2}}|\lambda'(s)||y|^{2-n}dyds$$

$$= \int_{-81/100}^t \left(1+\frac{\sqrt{t-s}}{|x|}\right)^\gamma \lambda(s)^{\frac{n-4}{2}}|\lambda'(s)|$$

$$\times \left[\int_{B_{L\lambda(s)}^c}(t-s)^{-\frac{n}{2}}e^{-c\frac{|x-y|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{|y|}\right)^\gamma|y|^{2-n}dy\right]ds.$$

We still divide this integral into three parts, I being on the interval $(-81/100, t-\lambda(t)^{2\mu})$, II being on the interval $(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)$, and III on $(t-K^{2\beta}\lambda(t)^2, t)$.

Estimate of I. Direct calculation gives

$$\mathcal{P} := \int_{B_{L\lambda(s)}^c}(t-s)^{-\frac{n}{2}}e^{-c\frac{|x-y|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{|y|}\right)^\gamma|y|^{2-n}dy$$

$$\lesssim (t-s)^{-\frac{n-2}{2}}\left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^\gamma \int_{\frac{L\lambda(s)}{\sqrt{t-s}}}^{+\infty}e^{-c\left(\frac{|x|}{\sqrt{t-s}}-r\right)^2}rdr.$$

Recall that $K\lambda(t)/2 \leq |x| \leq 4K\lambda(t)$, $t - s \geq K^{2\beta}\lambda(t)^2$. There are two cases.

- If $L\lambda(s) \geq 6K\lambda(t)$,

$$\mathcal{P} \lesssim (t-s)^{-\frac{n-2}{2}} \left[1 + \frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}}.$$

- If $L\lambda(s) \leq 6K\lambda(t)$,

$$\begin{aligned}
\mathcal{P} &\lesssim (t-s)^{-\frac{n-2}{2}} \left[1 + \frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} \\
&\lesssim (t-s)^{-\frac{n-2-\gamma}{2}} L^{-\gamma}\lambda(s)^{-\gamma}.
\end{aligned}$$

Hence

$$\begin{aligned}
\mathrm{I} &\lesssim |x|^{-\gamma} \int_{(-81/100, t-\lambda(t)^{2\mu})\cap\{L\lambda(s)\geq 6K\lambda(t)\}} (t-s)^{-\frac{n-2-\gamma}{2}} \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| \\
&\qquad\qquad \times \left[1 + \frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}} \qquad\qquad (14.22) \\
&+ L^{-\gamma}|x|^{-\gamma} \int_{(-81/100, t-\lambda(t)^{2\mu})\cap\{L\lambda(s)\leq 6K\lambda(t)\}} (t-s)^{-\frac{n-2}{2}+\gamma} \lambda(s)^{\frac{n-4}{2}-\gamma}|\lambda'(s)|.
\end{aligned}$$

Because

$$\begin{aligned}
&(t-s)^{-\frac{n-2-\gamma}{2}} \lambda(s)^{\frac{n-6}{2}} \left[1 + \frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}} \\
&= (t-s)^{-\frac{n+2-2\gamma}{4}} \left[\frac{\lambda(s)}{\sqrt{t-s}}\right]^{\frac{n-6}{2}} \left[1 + \frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}} \\
&\lesssim L^{-\frac{n-6}{2}-\gamma} \lambda(s)^{-\gamma} (t-s)^{-\frac{n+2}{4}+\gamma},
\end{aligned}$$

by noting that we have assumed $L\lambda(s) \geq 6K\lambda(t)$, the first integral is controlled by

$$\begin{aligned}
&L^{-\frac{n-6}{2}} [K\lambda(t)]^{-2\gamma} \int_{-81/100}^{t-\lambda(t)^{2\mu}} (t-s)^{-\frac{n+2}{4}+\gamma} |\lambda(s)\lambda'(s)| \\
&\lesssim L^{-\frac{n-6}{2}} K^{-2\gamma} \lambda(t)^{-\frac{n-2}{2}\mu+2(\mu-1)\gamma}.
\end{aligned}$$

In the last step we used only the estimate $|\lambda\lambda'| \leq C$.

For the second integral, by the estimate $|\lambda\lambda'| \leq C$ again, we obtain

$$\begin{aligned}
&L^{-\gamma}|x|^{-\gamma} \int_{(-81/100, t-\lambda(t)^{2\mu})\cap\{L\lambda(s)\leq 6K\lambda(t)\}} (t-s)^{-\frac{n-2}{2}+\gamma} \lambda(s)^{\frac{n-4}{2}-\gamma}|\lambda'(s)| \\
&\lesssim L^{-\frac{n-6}{2}} K^{\frac{n-6}{2}-2\gamma} \lambda(t)^{\frac{n-6}{2}-2\gamma} \int_{-81/100}^{t-\lambda(t)^{2\mu}} (t-s)^{-\frac{n-2}{2}+\gamma} \\
&\lesssim L^{-\frac{n-6}{2}} K^{\frac{n-6}{2}-2\gamma} \lambda(t)^{-(n-4)\mu+\frac{n}{2}-3+2(\mu-1)\gamma}.
\end{aligned}$$

Because

$$-(n-4)\mu + \frac{n}{2} - 3 + 2(\mu-1)\gamma > -\frac{n-2}{2}\mu + 2(\mu-1)\gamma,$$

combining these two estimates, we get

$$\mathrm{I} \lesssim L^{-\frac{n-6}{2}} K^{-2\gamma} \lambda(t)^{-\frac{n-2}{2}\mu+2(\mu-1)\gamma}.$$

Estimate of II. As in the previous case, we have

$$
\begin{aligned}
\mathrm{II} &\lesssim |x|^{-\gamma} \int_{(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)\cap\{L\lambda(s)\geq 6K\lambda(t)\}} (t-s)^{-\frac{n-2-\gamma}{2}} \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| \\
&\qquad\qquad\qquad\times \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}} \\
&\quad+ L^{-\gamma}|x|^{-\gamma} \int_{(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)\cap\{L\lambda(s)\leq 6K\lambda(t)\}} (t-s)^{-\frac{n-2}{2}+\gamma} \lambda(s)^{\frac{n-4}{2}-\gamma} |\lambda'(s)| \\
&\lesssim L^{-\frac{n-6}{2}} [K\lambda(t)]^{-2\gamma} \int_{t-\lambda(t)^{2\mu}}^{t-K^{2\beta}\lambda(t)^2} (t-s)^{-\frac{n+2}{4}+\gamma} |\lambda(s)\lambda'(s)| \\
&\quad+ L^{-\frac{n-6}{2}} K^{\frac{n-6}{2}-2\gamma} \lambda(t)^{\frac{n-6}{2}-2\gamma} \int_{t-\lambda(t)^{2\mu}}^{t-K^{2\beta}\lambda(t)^2} (t-s)^{-\frac{n-2}{2}+\gamma} |\lambda(s)\lambda'(s)| \\
&\lesssim L^{-\frac{n-6}{2}} \left[K^{-\frac{n-2}{2}\beta+2(\beta-1)\gamma} + K^{-\frac{n-2}{2}\beta-\left(\frac{n-6}{2}-2\gamma\right)(\beta-1)}\right] \\
&\qquad \times \lambda(t)^{-\frac{n-2}{2}} \|\lambda\lambda'\|_{L^{\infty}(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)} \\
&\lesssim K^{-\frac{n-2}{2}\beta+2(\beta-1)\gamma} \lambda(t)^{-\frac{n-2}{2}} \|\lambda\lambda'\|_{L^{\infty}(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)}.
\end{aligned}
$$

Estimate of III. By (14.21),

$$
\begin{aligned}
\mathrm{III} &\lesssim \int_{t-K^{2\beta}\lambda(t)^2}^{t} \int_{B_1} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1+\frac{\sqrt{t-s}}{|x|}\right)^{\gamma} \left(1+\frac{\sqrt{t-s}}{|y|}\right)^{\gamma} \\
&\qquad\qquad \times \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| |y|^{2-n} \\
&= \int_{t-K^{2\beta}\lambda(t)^2}^{t} \left(1+\frac{\sqrt{t-s}}{|x|}\right)^{\gamma} \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| \\
&\qquad\qquad \times \left[\underbrace{\int_{B_1} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1+\frac{\sqrt{t-s}}{|y|}\right)^{\gamma} |y|^{2-n} dy}_{\text{Lemma B.1}}\right] ds \\
&\lesssim \int_{t-K^{2\beta}\lambda(t)^2}^{t} \left(|x|+\sqrt{t-s}\right)^{2-n} \left(1+\frac{\sqrt{t-s}}{|x|}\right)^{2\gamma} \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| ds \\
&\lesssim [K\lambda(t)]^{2-n} K^{2(\beta-1)\gamma} \int_{t-K^{2\beta}\lambda(t)^2}^{t} \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| ds \\
&\lesssim K^{2-n+2\beta+2(\beta-1)\gamma} \lambda(t)^{-\frac{n-2}{2}} \|\lambda\lambda'\|_{L^{\infty}(t-K^{2\beta}\lambda(t)^2, t)}.
\end{aligned}
$$

Because $n > 6$ and β is sufficiently close to 1, we have

$$2 - n + 2\beta + 2(\beta-1)\gamma < -\frac{n-2}{2}\beta + 2(\beta-1)\gamma.$$

Adding up estimates for I, II and III we finish the proof. $\qquad\square$

Lemma 14.4 (Estimates on $\overline{\phi}_4$). *For $K\lambda(t)/2 \le |x| \le 4K\lambda(t)$,*

$$\overline{\phi}_4(x,t) \lesssim \lambda(t)^{-\frac{n-2}{2}\mu+2(\mu-1)\gamma} + K^{-\frac{n-2}{2}(\beta-1)+2(\beta-1)\gamma} \left[K\lambda(t)\right]^{-\frac{n-2}{2}} \|\lambda\xi'\|_{L^\infty(t-\lambda(t)^{2\mu},t)}.$$

Proof. We still have the heat kernel representation

$$\begin{aligned}
\overline{\phi}_4(x,t) &\lesssim \int_{-81/100}^t \int_{B_{L\lambda(s)}^c} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1+\frac{\sqrt{t-s}}{|x|}\right)^\gamma \left(1+\frac{\sqrt{t-s}}{|y|}\right)^\gamma \\
&\qquad\qquad \times \lambda(s)^{\frac{n-2}{2}} |\xi'(s)||y|^{1-n} dy\, ds \\
&= \int_{-81/100}^t \left(1+\frac{\sqrt{t-s}}{|x|}\right)^\gamma \lambda(s)^{\frac{n-2}{2}}|\xi'(s)| \\
&\qquad \times \left[\int_{B_{L\lambda(s)}^c} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1+\frac{\sqrt{t-s}}{|y|}\right)^\gamma |y|^{1-n} dy\right] ds.
\end{aligned}$$

We still divide this integral into three parts, I being on the interval $(-81/100, t-\lambda(t)^{2\mu})$, II being on the interval $(t-\lambda(t)^{2\mu}, t-K^{2\beta}\lambda(t)^2)$, and III on the interval $(t-K^{2\beta}\lambda(t)^2, t)$.

Estimate of I. Direct calculation gives

$$\begin{aligned}
\mathcal{P} &:= \int_{B_{L\lambda(s)}^c} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1+\frac{\sqrt{t-s}}{|y|}\right)^\gamma |y|^{1-n} dy \\
&\lesssim (t-s)^{-\frac{n-1}{2}} \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^\gamma \int_{\frac{L\lambda(s)}{\sqrt{t-s}}}^{+\infty} e^{-c\left(\frac{|x|}{\sqrt{t-s}}-r\right)^2} dr.
\end{aligned}$$

Recall that $K\lambda(t)/2 \le |x| \le 4K\lambda(t)$ and $t-s \ge K^{2\beta}\lambda(t)^2$. There are two cases.

- If $L\lambda(s) \ge 6K\lambda(t)$,

$$\mathcal{P} \lesssim (t-s)^{-\frac{n-1}{2}} \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^\gamma e^{-c\frac{L^2\lambda(s)^2}{t-s}}.$$

- If $L\lambda(s) \le 6K\lambda(t)$,

$$\mathcal{P} \lesssim (t-s)^{-\frac{n-1-\gamma}{2}} L^{-\gamma}\lambda(s)^{-\gamma}.$$

Plugging these estimates into the formula for I, we obtain

$$\begin{aligned}
\mathrm{I} &\lesssim |x|^{-\gamma} \int_{(-81/100, t-\lambda(t)^{2\mu})\cap\{L\lambda(s)\ge 6K\lambda(t)\}} (t-s)^{-\frac{n-1-\gamma}{2}} \lambda(s)^{\frac{n-2}{2}}|\xi'(s)| \\
&\qquad\qquad \times \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^\gamma e^{-c\frac{L^2\lambda(s)^2}{t-s}} \qquad\qquad\qquad (14.23) \\
&+ L^{-\gamma}|x|^{-\gamma} \int_{(-81/100, t-\lambda(t)^{2\mu})\cap\{L\lambda(s)\le 6K\lambda(t)\}} (t-s)^{-\frac{n-1}{2}+\gamma} \lambda(s)^{\frac{n-2}{2}-\gamma}|\xi'(s)|.
\end{aligned}$$

Because

$$(t-s)^{-\frac{n-1-\gamma}{2}} \lambda(s)^{\frac{n-4}{2}} \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^\gamma e^{-c\frac{L^2\lambda(s)^2}{t-s}}$$

$$= (t-s)^{-\frac{n+2}{4}+\frac{\gamma}{2}} \left[\frac{\lambda(s)}{\sqrt{t-s}}\right]^{\frac{n-4}{2}} \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}}$$

$$\lesssim L^{-\gamma}\lambda(s)^{-\gamma}(t-s)^{-\frac{n+2}{4}+\gamma},$$

the first integral is controlled by

$$[K\lambda(t)]^{-2\gamma} \int_{-81/100}^{t-\lambda(t)^{2\mu}} (t-s)^{-\frac{n+2}{4}+\gamma}|\lambda(s)\xi'(s)|$$

$$\lesssim K^{-2\gamma}\lambda(t)^{-\frac{n-2}{2}\mu+2(\mu-1)\gamma}.$$

In the last step we used only the estimate $|\lambda\xi'| \leq C$.

Similarly, for the second integral, we have

$$L^{-\gamma}|x|^{-\gamma} \int_{(-81/100,t-\lambda(t)^{2\mu})\cap\{L\lambda(s)\leq 4K\lambda(t)\}} (t-s)^{-\frac{n-1}{2}+\gamma}\lambda(s)^{\frac{n-2}{2}-\gamma}|\xi'(s)|$$

$$\lesssim L^{-\frac{n-4}{2}}K^{\frac{n-4}{2}-2\gamma}\lambda(t)^{\frac{n-4}{2}-2\gamma} \int_{-81/100}^{t-\lambda(t)^{2\mu}} (t-s)^{-\frac{n-1}{2}+\gamma}$$

$$\lesssim L^{-\frac{n-4}{2}}K^{\frac{n-4}{2}-2\gamma}\lambda(t)^{\frac{n-4}{2}-(n-3)\mu+2(\mu-1)\gamma}.$$

Because

$$-\frac{n-2}{2}\mu+2(\mu-1)\gamma < \frac{n-4}{2}-(n-3)\mu+2(\mu-1)\gamma,$$

we get

$$\mathrm{I} \lesssim \lambda(t)^{-\frac{n-2}{2}\mu+2(\mu-1)\gamma}.$$

Estimate of II. We have

$$\mathrm{II} \lesssim |x|^{-\gamma} \int_{(t-\lambda(t)^{2\mu},t-K^{2\beta}\lambda(t)^2)\cap\{L\lambda(s)\geq 6K\lambda(t)\}} (t-s)^{-\frac{n-1-\gamma}{2}}\lambda(s)^{\frac{n-2}{2}}|\xi'(s)|$$

$$\times \left[1+\frac{\sqrt{t-s}}{L\lambda(s)}\right]^{\gamma} e^{-c\frac{L^2\lambda(s)^2}{t-s}}$$

$$+ L^{-\gamma}|x|^{-\gamma} \int_{(t-\lambda(t)^{2\mu},t-K^{2\beta}\lambda(t)^2)\cap\{L\lambda(s)\leq 6K\lambda(t)\}} (t-s)^{-\frac{n-1}{2}+\gamma}\lambda(s)^{\frac{n-2}{2}-\gamma}|\xi'(s)|$$

$$\lesssim [K\lambda(t)]^{-2\gamma} \int_{t-\lambda(t)^{2\mu}}^{t-K^{2\beta}\lambda(t)^2} (t-s)^{-\frac{n+2}{4}+\gamma}|\lambda(s)\xi'(s)|$$

$$+ L^{-\frac{n-4}{2}}K^{\frac{n-4}{2}-2\gamma}\lambda(t)^{\frac{n-4}{2}-2\gamma} \int_{t-\lambda(t)^{2\mu}}^{t-K^{2\beta}\lambda(t)^2} (t-s)^{-\frac{n-1}{2}+\gamma}|\lambda(s)\xi'(s)|$$

$$\lesssim \left[K^{-\frac{n-2}{2}\beta+2(\beta-1)\gamma} + L^{-\frac{n-4}{2}}K^{-\beta(n-3-2\gamma)+\frac{n-4}{2}-2\gamma}\right]$$

$$\times \lambda(t)^{-\frac{n-2}{2}}\|\lambda\xi'\|_{L^\infty(t-\lambda(t)^{2\mu},t-K^{2\beta}\lambda(t)^2)}$$

$$\lesssim K^{-\frac{n-2}{2}\beta+2(\beta-1)\gamma}\lambda(t)^{-\frac{n-2}{2}}\|\lambda\xi'\|_{L^\infty(t-\lambda(t)^{2\mu},t-K^{2\beta}\lambda(t)^2)}.$$

Estimate of III. As in the $\overline{\phi}_3$ case, we have

$$
\begin{aligned}
\text{II} \;&=\; \int_{t-K^{2\beta}\lambda(t)^2}^{t}\left(1+\frac{\sqrt{t-s}}{|x|}\right)^{\gamma}\lambda(s)^{\frac{n-2}{2}}|\xi'(s)| \\
&\qquad\times\left[\int_{B_1}(t-s)^{-\frac{n}{2}}\,e^{-c\frac{|x-y|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{|y|}\right)^{\gamma}|y|^{1-n}dy\right]ds \\
&\lesssim\; |x|^{1-n}\int_{t-K^{2\beta}\lambda(t)^2}^{t}\left(1+\frac{\sqrt{t-s}}{|x|}\right)^{2\gamma}\lambda(s)^{\frac{n-2}{2}}|\xi'(s)|ds \\
&\lesssim\; K^{1-n+2\beta+2(\beta-1)\gamma}\lambda(t)^{-\frac{n-2}{2}}\|\lambda\xi'\|_{L^\infty(t-K^{2\beta}\lambda(t)^2,t)}.
\end{aligned}
$$

Because

$$
1-n+2\beta+2(\beta-1)\gamma<-\frac{n-2}{2}\beta+2(\beta-1)\gamma,
$$

adding up estimates for I, II and III we finish the proof. $\qquad\square$

Lemma 14.5 (Estimates on $\overline{\phi}_5$). *For $K\lambda(t)/2\le |x|\le 4K\lambda(t)$,*

$$
\begin{aligned}
\overline{\phi}_5(x,t) \;&\lesssim\; \left[K^{-\left(\frac{n-2}{2}-\gamma\right)(\beta-1)}+K^{\frac{n-2}{2}+2\beta+2(\beta-1)\gamma}e^{-cK}+\frac{K^{\frac{n-2}{2}+2\beta+(\beta-1)\gamma}}{e^{cL}}\right] \\
&\qquad\times [K\lambda(t)]^{-\frac{n-2}{2}}\left\|\left(\frac{a}{\lambda},\lambda a'-\mu_0\frac{a}{\lambda},\lambda\lambda',\lambda\xi'\right)\right\|_{L^\infty(t-\lambda(t)^{2\mu},t)} \\
&\;+\; \lambda(t)^{-\frac{n-2}{2}\mu+(\mu-1)\gamma}.
\end{aligned}
$$

Proof. By the heat kernel representation, we have

$$
\begin{aligned}
\overline{\phi}_5(x,t) \;&\le\; \int_{-81/100}^{t}\int_{B^c_{L\lambda(s)}}G(x,y;t,s)\lambda(s)^{-\frac{n+2}{2}}e^{-c\frac{|y|}{\lambda(s)}} \\
&\qquad\times\left[\frac{|a(s)|}{\lambda(s)}+\left|\lambda(s)a'(s)-\mu_0\frac{a(s)}{\lambda(s)}\right|+\frac{|a(s)|}{\lambda(s)}\left(|\lambda(s)\lambda'(s)|+|\lambda(s)\xi'(s)|\right)\right].
\end{aligned}
$$

We still divide this integral into three parts, I being on the interval $(-81/100,t-\lambda(t)^{2\mu})$, II being on the interval $(t-\lambda(t)^{2\mu},t-K^{2\beta}\lambda(t)^2)$, and III on $(t-K^{2\beta}\lambda(t)^2,t)$.

Estimate of I. For I, just using the estimate

$$
\left\|\left(\frac{a}{\lambda},\lambda a'-\mu_0\frac{a}{\lambda},\frac{a}{\lambda}\left(|\lambda\lambda'|+|\lambda\xi'|\right)\right)\right\|_{L^\infty(-1,t)}\le C,
$$

by the operator bound on G (see (14.19)) and Hölder inequality, we get

$$
\begin{aligned}
\text{I} \;&\lesssim\; \int_{-81/100}^{t-\lambda(t)^{2\mu}}\int_{B_1}G(x,t;y,s)\lambda(s)^{-\frac{n+2}{2}}e^{-c\frac{|y|}{\lambda(s)}}dyds \\
&\lesssim\; \int_{-81/100}^{t-\lambda(t)^{2\mu}}(t-s)^{-\frac{n+2}{4}}\left(1+\frac{\sqrt{t-s}}{|x|}\right)^{\gamma}ds \\
&\lesssim\; \lambda(t)^{-\frac{n-2}{2}\mu+(\mu-1)\gamma}.
\end{aligned}
$$

Estimate of II. For II, in the same way we obtain

$$
\text{II} \lesssim \left\| \left(\frac{a}{\lambda}, \lambda a' - \mu_0 \frac{a}{\lambda}, \frac{a}{\lambda} \left(|\lambda \lambda'| + |\lambda \xi'| \right) \right) \right\|_{L^\infty(t - \lambda(t)^{2\mu}, t - K^{2\beta}\lambda(t)^2)}
$$

$$
\times \int_{t - \lambda(t)^{2\mu}}^{t - K^{2\beta}\lambda(t)^2} \int_{B_1} G(x, y; t, s) \lambda(s)^{-\frac{n+2}{2}} e^{-c\frac{|y|}{\lambda(s)}} \, dy \, ds
$$

$$
\lesssim \left\| \left(\frac{a}{\lambda}, \lambda a' - \mu_0 \frac{a}{\lambda}, \lambda \lambda', \lambda \xi' \right) \right\|_{L^\infty(t - \lambda(t)^{2\mu}, t - K^{2\beta}\lambda(t)^2)}
$$

$$
\times \int_{-1}^{t - K^{2\beta}\lambda(t)^2} (t - s)^{-\frac{n+2}{4}} \left(1 + \frac{\sqrt{t - s}}{|x|} \right)^\gamma ds
$$

$$
\lesssim K^{-\frac{n-2}{2}\beta + (\beta - 1)\gamma} \lambda(t)^{-\frac{n-2}{2}} \left\| \left(\frac{a}{\lambda}, \lambda a' - \mu_0 \frac{a}{\lambda}, \lambda \lambda', \lambda \xi' \right) \right\|_{L^\infty(t - \lambda(t)^{2\mu}, t - K^{2\beta}\lambda(t)^2)}.
$$

Estimate of III. In this case, first note that

$$
\int_{t - K^{2\beta}\lambda(t)^2}^{t} \int_{B_{L\lambda(s)}^c} G(x, t; y, s) \lambda(s)^{-\frac{n+2}{2}} e^{-c\frac{|y|}{\lambda(s)}} \, dy \, ds
$$

$$
\lesssim \int_{t - K^{2\beta}\lambda(t)^2}^{t} \int_{B_{L\lambda(s)}^c} (t - s)^{-\frac{n}{2}} e^{-c\frac{|x - y|^2}{t - s}} \left(1 + \frac{\sqrt{t - s}}{|x|} \right)^\gamma \left(1 + \frac{\sqrt{t - s}}{|y|} \right)^\gamma
$$

$$
\times \lambda(s)^{-\frac{n+2}{2}} e^{-c\frac{|y|}{\lambda(s)}} \, dy \, ds
$$

$$
\lesssim \int_{t - K^{2\beta}\lambda(t)^2}^{t} \left(1 + \frac{\sqrt{t - s}}{|x|} \right)^\gamma \lambda(s)^{-\frac{n+2}{2}}
$$

$$
\times \left[\underbrace{\int_{B_{L\lambda(s)}^c} (t - s)^{-\frac{n}{2}} e^{-c\frac{|x - y|^2}{t - s}} \left(1 + \frac{\sqrt{t - s}}{|y|} \right)^\gamma e^{-c\frac{|y|}{\lambda(s)}} \, dy}_{\text{Lemma B.2}} \right] ds
$$

$$
\lesssim \int_{t - K^{2\beta}\lambda(t)^2}^{t} \left(1 + \frac{\sqrt{t - s}}{|x|} \right)^\gamma \lambda(s)^{-\frac{n+2}{2}}
$$

$$
\times \left[e^{-c\frac{|x|}{\lambda(s)}} \left(1 + \frac{\sqrt{t - s}}{|x|} \right)^\gamma + \left(1 + \frac{\lambda(s)}{\sqrt{t - s}} \right)^{n+\gamma} e^{-cL - c\frac{|x|^2}{t - s}} \right] ds
$$

$$
\lesssim \left[K^{2\beta + 2(\beta - 1)\gamma} e^{-cK} + K^{2\beta + (\beta - 1)\gamma} e^{-cL} \right] \lambda(t)^{-\frac{n-2}{2}},
$$

where we have used (14.21) to deduce that last inequality.

Then similar to the estimate of II, we get

$$
\text{III} \lesssim \left[K^{2\beta + 2(\beta - 1)\gamma} e^{-cK} + K^{2\beta + (\beta - 1)\gamma} e^{-cL} \right] \lambda(t)^{-\frac{n-2}{2}}
$$

$$
\times \left\| \left(\frac{a}{\lambda}, \lambda a' - \mu_0 \frac{a}{\lambda}, \lambda \lambda', \lambda \xi' \right) \right\|_{L^\infty(t - K^{2\beta}\lambda(t)^2, t)}.
$$

Adding up estimates for I, II and III we finish the proof.   □

14.3. **Estimate of $\mathcal{O}(t)$.** As an application of the pointwise estimates on $\overline{\phi}_1, \cdots, \overline{\phi}_5$ obtained in the previous subsection, we give an estimate on $\mathcal{O}(t)$.

Proposition 14.6. *There exist two constants $C(K,L)$ and $\sigma(K,L) \ll 1$ (depending on K and L) such that for any $t \in (-81/100, 81/100)$,*

$$\mathcal{O}(t) \leq C(K,L)\lambda(t)^{2+2\kappa} + \sigma(K,L)\left\|\left(\mathcal{I}, \frac{a}{\lambda}, \lambda a' - \mu_0\frac{a}{\lambda}, \lambda\lambda', \lambda\xi'\right)\right\|_{L^\infty(t-\lambda(t)^{2\mu}, t)}.$$
(14.24)

Proof. First by the five lemmas in the previous subsection, we obtain

$$[K\lambda(t)]^{\frac{n-2}{2}} \sup_{B_{2K\lambda(t)}(\xi(t))\setminus B_{K\lambda(t)}(\xi(t))} |\phi(x,t)| \tag{14.25}$$

$$\leq [K\lambda(t)]^{\frac{n-2}{2}} \sup_{B_{2K\lambda(t)}\setminus B_{K\lambda(t)}} \left[\overline{\phi}_1(x,t) + \overline{\phi}_2(x,t) + \overline{\phi}_3(x,t) + \overline{\phi}_4(x,t) + \overline{\phi}_5(x,t)\right]$$

$$\leq C(K,L)\lambda(t)^{2+2\kappa} + \sigma(K,L)\left\|\left(\mathcal{I}, \frac{a}{\lambda}, \lambda a' - \mu_0\frac{a}{\lambda}, \lambda\lambda', \lambda\xi'\right)\right\|_{L^\infty(t-\lambda(t)^{2\mu}, t)}.$$

Next, by a rescaling and standard interior gradient estimates for heat equation, we get

$$[K\lambda(t)]^{\frac{n}{2}} \sup_{B_{2K\lambda(t)}(\xi(t))\setminus B_{K\lambda(t)}(\xi(t))} |\nabla\phi_1(x,t)|$$

$$\lesssim [K\lambda(t)]^{\frac{n-2}{2}} \sup_{\left(B_{4K\lambda(t)(\xi(t))}\setminus B_{K\lambda(t)/2(\xi(t))}\right)\times(t-K^2\lambda(t)^2, t)} |\phi_1(y,s)|$$

$$\lesssim K^{\frac{n-2}{2}-\gamma}\lambda(t)^{\frac{n-2}{2}-\gamma}.$$

The same estimate holds for ϕ_2, if we note that $F_2 \equiv 0$ in $B_{4K\lambda(t)(\xi(t))} \setminus B_{K\lambda(t)/2(\xi(t))}$.
In the same way, for ϕ_3 we have

$$[K\lambda(t)]^{\frac{n}{2}} \sup_{B_{2K\lambda(t)}(\xi(t))\setminus B_{K\lambda(t)}(\xi(t))} |\nabla\phi_3(x,t)|$$

$$\lesssim [K\lambda(t)]^{\frac{n-2}{2}} \sup_{\left(B_{4K\lambda(t)(\xi(t))}\setminus B_{K\lambda(t)/2(\xi(t))}\right)\times(t-K^2\lambda(t)^2, t)} \left[|\phi_3(y,s)| + (K\lambda(t))^2 |F_3(y,s)|\right]$$

$$\lesssim K^{-\frac{n-2}{2}(\beta-1)+2(\beta-1)\gamma}\|\lambda\lambda'\|_{L^\infty(t-\lambda(t)^{2\mu}, t)} + K^{\frac{n-2}{2}}\lambda(t)^{2+2\kappa}$$

$$+ K^{-\frac{n-6}{2}}\|\lambda\lambda'\|_{L^\infty(t-K^2\lambda(t)^2, t)}.$$

The same estimates hold for ϕ_4 and ϕ_5.

Putting these gradient estimates together and arguing as in (14.25), we get the estimate on the remaining terms in $\mathcal{O}(t)$. □

15. A HARNACK INEQUALITY FOR λ

In this section, we combine the estimates in the previous two sections to establish an Harnack inequality for λ.

For any $t \in (-81/100, 81/100)$, denote

$$\mathcal{D}(t) := \|\phi_{in}(t)\|_{\alpha;1+\theta} + |\lambda(t)\xi'(t)| + |\lambda(t)\lambda'(t)| + \left|\lambda(t)a'(t) - \mu_0 \frac{a(t)}{\lambda(t)}\right| + \left|\frac{a(t)}{\lambda(t)}\right|.$$

For $r \in (0, 81/100)$, let

$$g(r) := \sup_{-r < t < r} \mathcal{D}(t)$$

and

$$M(r) := \sup_{-r < t < r} \lambda(t)^2, \quad m(r) := \inf_{-r < t < r} \lambda(t)^2.$$

By Proposition 13.4 and Lemma 13.6, there exist three universal constants C, $T \gg 1$ and $\sigma \ll 1$ such that

$$\mathcal{D}(t) \leq \sigma \sup_{[t-T\lambda(t)^2, t+T\lambda(t)^2]} \mathcal{D}(s) + C \sup_{[t-T\lambda(t)^2, t+T\lambda(t)^2]} \mathcal{O}(s). \tag{15.1}$$

By Proposition 14.6,

$$\mathcal{O}(t) \leq C(K, L)\lambda(t)^{2+2\kappa} + \sigma(K, L) \sup_{[t-\lambda(t)^{2\mu}, t]} \mathcal{D}(s). \tag{15.2}$$

Combining these two estimates we get

$$\mathcal{D}(t) \leq \frac{1}{2} \sup_{[t-2\lambda(t)^{2\mu}, t+2K\lambda(t)^2]} \mathcal{D}(s) + C \sup_{[t-2\lambda(t)^{2\mu}, t+2K\lambda(t)^2]} \lambda(s)^{2+2\kappa}. \tag{15.3}$$

For any $r \in (0, 1)$, taking supremum over $t \in (-r + 2M(r)^\mu, r - 2M(r)^\mu)$ in (15.3), we obtain

$$g\left(r - 2M(r)^\mu\right) \leq \frac{1}{2}g(r) + CM(r)^{1+\kappa}. \tag{15.4}$$

For any $r_1 < r_2$, an iteration of this inequality from r_2 to r_1 in $\lfloor (r_2 - r_1)/M(r_2)^\mu \rfloor$ steps leads to

$$g(r_1) \leq g(r_2)e^{-c(r_2-r_1)M(r_2)^{-\mu}} + CM(r_2)^{1+\kappa}.$$

For any $r \in (0, 1)$, $M(r) \ll 1$, so

$$e^{-cM(r)^{-\mu/2}} \lesssim M(r)^{1+\kappa}.$$

Hence by choosing $r_2 = r$ and $r_1 = r - M(r)^{\mu/2}$, we get

$$g\left(r - M(r)^{\mu/2}\right) \lesssim M(r)^{1+\kappa}.$$

By this estimate, integrating $\lambda\lambda'$ on $[-r + M(r)^{\mu/2}, r - M(r)^{\mu/2}]$, we find a constant C_H such that

$$M\left(r - M(r)^{\mu/2}\right) \leq m\left(r - M(r)^{\mu/2}\right) + C_H M(r)^{1+\kappa}. \tag{15.5}$$

Lemma 15.1. *There exists an $r_H \in [1/2, 3/4]$ such that*

$$M\left(r_H - M(r_H)^{\mu/2}\right) \geq 2C_H M(r_H)^{1+\kappa}.$$

Proof. Assume that for any $r \in [1/2, 3/4]$,

$$M\left(r - M(r)^{\mu/2}\right) \leq 2C_H M(r)^{1+\kappa}. \tag{15.6}$$

Set

$$r_0 := 3/4, \quad a_0 := M(r_0)$$

and for $k \in \mathbb{N}$,

$$r_{k+1} := r_k - a_k^{\mu/2}, \quad a_{k+1} := M(r_{k+1}).$$

Then (15.6) says

$$a_{k+1} \leq 2C_H a_k^{1+\kappa}.$$

By our assumption, $a_0 \ll 1$. An induction gives

$$a_k \leq a_0^{k+1}, \quad r_k \geq r_0 - \sum_{k=0}^{+\infty} a_0^{(k+1)\frac{\mu}{2}} \geq 1/2.$$

As a consequence, $M(1/2) = 0$. This is impossible. $\qquad\square$

For this r_H, (15.5) can be written as a Harnack inequality

$$m(r_H) \geq [1 - C_H M(r_H)^\kappa] M(r_H) \geq \frac{1}{2} M(r_H). \tag{15.7}$$

After a scaling $u(x,t) \mapsto r_H^{(n-2)/2} u(r_H x, r_H^2 t)$, from here to Section 18 we work in the following setting:
- we denote

$$\varepsilon := \lambda(0), \tag{15.8}$$

  and after a translation in the spatial direction, we assume $\xi(0) = 0$;
- by the above Harnack inequality (15.7), for any $t \in [-1, 1]$,

$$\lambda(t) = \varepsilon + O\left(\varepsilon^{1+\kappa}\right) \tag{15.9}$$

  and

$$\|\phi_{in}(t)\| + |\lambda(t)\lambda'(t)| + |\lambda(t)\xi'(t)| + \left|\lambda(t)a'(t) - \mu_0 \frac{a(t)}{\lambda(t)}\right| + \left|\frac{a(t)}{\lambda(t)}\right| \lesssim \varepsilon^{1+\kappa}. \tag{15.10}$$

- integrating ξ' and using the above estimate, we obtain

$$|\xi(t)| \lesssim \varepsilon^\kappa, \quad \text{for any } t \in [-1, 1]. \tag{15.11}$$

16. INNER PROBLEM AND OUTER PROBLEM AGAIN

In this section, under the assumptions (15.8)-(15.10), we prove

Proposition 16.1. *For any $\gamma > 0$, there exists a constant $C(\gamma)$ such that*

$$g\left(\frac{8}{9}\right) \leq C(\gamma)\varepsilon^{\frac{n-2}{2} - \gamma}. \tag{16.1}$$

This gives an improvement on the estimate of scaling parameters etc.. A more precise estimate on ϕ also follows:

Proposition 16.2. *For any $\gamma > 0$, there exists a constant $C(\gamma) > 0$ such that*

$$|\phi(x,t)| \leq C(\gamma)\left(\varepsilon + |x - \xi(t)|\right)^{-\gamma} \quad in \quad Q_{7/8}. \tag{16.2}$$

To prove these propositions, estimates on $\overline{\phi}_2$, $\overline{\phi}_3$, $\overline{\phi}_4$ and $\overline{\phi}_5$ in Section 14 are not sufficient. We need to upgrade them.

Lemma 16.3 (Estimate on $\overline{\phi}_2$). *If* $|x| \geq K\lambda(t)/2$, *then*

$$\overline{\phi}_2(x,t) \lesssim \left[K^{-\left(\frac{n-2}{2}-\gamma\right)\beta}\frac{\varepsilon^{\gamma}}{|x|^{\gamma}} + L^{\frac{n-2}{2}-\gamma}K^{\beta\gamma}\frac{\varepsilon^{n-2}}{|x|^{n-2}} \right] \varepsilon^{-\frac{n-2}{2}}\|\mathcal{I}\|_{L^{\infty}(-1,t)}. \tag{16.3}$$

Proof. As in the proof of Lemma 14.2, we have the heat kernel representation for $\overline{\phi}_2$. But now we just divide the integral into two parts, I being on $(-1, t - K^{2\beta}\lambda(t)^2)$ and II on $(t - K^{2\beta}\lambda(t)^2, t)$.

For I, similar to (14.20), we have

$$
\begin{aligned}
\text{I} &\lesssim |x|^{-\gamma}\int_{-1}^{t-K^{2\beta}\lambda(t)^2}(t-s)^{-\frac{n+2}{4}+\frac{\gamma}{2}}\mathcal{I}(s)ds \\
&\lesssim K^{-\left(\frac{n-2}{2}-\gamma\right)\beta}\lambda(t)^{-\frac{n-2}{2}+\gamma}|x|^{-\gamma}\|\mathcal{I}\|_{L^{\infty}(-1,t-K^{2\beta}\lambda(t)^2)} \\
&\lesssim K^{-\left(\frac{n-2}{2}-\gamma\right)\beta}\varepsilon^{-\frac{n-2}{2}+\gamma}|x|^{-\gamma}\|\mathcal{I}\|_{L^{\infty}(-1,t-K^{2\beta}\lambda(t)^2)}
\end{aligned}
$$

The estimate for II is almost the same as the one for III in the proof of Lemma 14.2, that is,

$$
\begin{aligned}
\text{II} &\lesssim \int_{t-K^{2\beta}\lambda(t)^2}^{t}\frac{\mathcal{I}(s)}{L^{\alpha}\lambda(s)^{\alpha}}\left(\frac{|\lambda'(s)|}{\lambda(s)}+\frac{|\xi'(s)|}{L\lambda(s)}+\frac{1}{L^2\lambda(s)^2}\right)\left(1+\frac{\sqrt{t-s}}{|x|}\right)^{\gamma} \\
&\quad \times \left[\int_{B_{2L\lambda(s)}\setminus B_{L\lambda(s)}}(t-s)^{-\frac{n}{2}}e^{-c\frac{|x-y|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{|y|}\right)^{\gamma}dy\right] \\
&\lesssim \|\mathcal{I}\|_{L^{\infty}(t-K^{2\beta}\lambda(t)^2,t)}L^{\frac{n-2}{2}}\lambda(t)^{\frac{n-2}{2}}\int_{t-K^{2\beta}\lambda(t)^2}^{t}(t-s)^{-\frac{n}{2}}e^{-c\frac{|x|^2}{t-s}}\left(1+\frac{\sqrt{t-s}}{L\lambda(t)}\right)^{\gamma} \\
&\lesssim \|\mathcal{I}\|_{L^{\infty}(t-K^{2\beta}\lambda(t)^2,t)}L^{\frac{n-2}{2}}\lambda(t)^{\frac{n-2}{2}}\left(1+\frac{K^{\beta}}{L}\right)^{\gamma}\int_{t-K^{2\beta}\lambda(t)^2}^{t}(t-s)^{-\frac{n}{2}}e^{-c\frac{|x|^2}{t-s}} \\
&\lesssim \|\mathcal{I}\|_{L^{\infty}(t-K^{2\beta}\lambda(t)^2,t)}L^{\frac{n-2}{2}}\lambda(t)^{\frac{n-2}{2}}\left(1+\frac{K^{\beta}}{L}\right)^{\gamma}|x|^{2-n} \\
&\lesssim L^{\frac{n-2}{2}-\gamma}K^{\beta\gamma}\varepsilon^{\frac{n-2}{2}}|x|^{2-n}\|\mathcal{I}\|_{L^{\infty}(t-K^2\lambda(t)^2,t)}.
\end{aligned}
$$

Putting these two estimates together, we obtain (16.4). $\qquad\square$

Corollary 16.4. *If* $|x| \geq K\lambda(t)/2$, *then*

$$\overline{\phi}_2(x,t) \lesssim \left[K^{-\frac{n-2}{2}(\beta-1)+\beta\gamma} + \frac{L^{\frac{n-2}{2}-\gamma}}{K^{\frac{n-2}{2}-\beta\gamma}} \right] (K\varepsilon)^{-\frac{n-2}{2}}\|\mathcal{I}\|_{L^{\infty}(-1,t)}. \tag{16.4}$$

Lemma 16.5 (Estimate on $\overline{\phi}_3$). *If* $|x| \geq K\lambda(t)/2$, *then*

$$\overline{\phi}_3(x,t) \lesssim \varepsilon^{\frac{n-6}{2}}|x|^{4-n}\|\lambda\lambda'\|_{L^{\infty}(-1,t)}. \tag{16.5}$$

Proof. As in the proof of Lemma 14.3, we still have the heat kernel representation

$$
\begin{aligned}
\overline{\phi}_3(x,t) \;\lesssim\; & \int_{-81/100}^{t} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{\gamma} \lambda(s)^{\frac{n-4}{2}} |\lambda'(s)| \\
& \times \left[\int_{B_1} (t-s)^{-\frac{n}{2}} \left(1 + \frac{\sqrt{t-s}}{|y|}\right)^{\gamma} e^{-c\frac{|x-y|^2}{t-s}} |y|^{2-n} dy\right] ds \\
\lesssim\; & \varepsilon^{\frac{n-6}{2}} \|\lambda\lambda'\|_{L^\infty(-1,t)} \\
& \times \int_{-81/100}^{t} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{2\gamma} \left(\sqrt{t-s} + |x|\right)^{2-n} ds \\
\lesssim\; & \varepsilon^{\frac{n-6}{2}} \|\lambda\lambda'\|_{L^\infty(-1,t)} |x|^{4-n}.
\end{aligned}
$$

Here we have used Lemma B.1 to deduce the second inequality. $\qquad\square$

Corollary 16.6. *If* $|x| \geq K\lambda(t)/2$, *then*

$$
\overline{\phi}_3(x,t) \lesssim K^{-\frac{n-6}{2}} (K\varepsilon)^{-\frac{n-2}{2}} \|\lambda\lambda'\|_{L^\infty(-1,t)}. \tag{16.6}
$$

Lemma 16.7 (Estimate on $\overline{\phi}_4$)**.** *If* $|x| \geq K\lambda(t)/2$, *then*

$$
\overline{\phi}_4(x,t) \lesssim \varepsilon^{\frac{n-4}{2}} |x|^{3-n} \|\lambda\xi'\|_{L^\infty(-1,t)}. \tag{16.7}
$$

Proof. As in the proof of Lemma 14.4, we still have the heat kernel representation

$$
\begin{aligned}
\overline{\phi}_4(x,t) \;\lesssim\; & \int_{-81/100}^{t} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{\gamma} \lambda(s)^{\frac{n-2}{2}} |\xi'(s)| \\
& \times \left[\int_{B_1} (t-s)^{-\frac{n}{2}} \left(1 + \frac{\sqrt{t-s}}{|y|}\right)^{\gamma} e^{-c\frac{|x-y|^2}{t-s}} |y|^{1-n} dy\right] ds \\
\lesssim\; & \varepsilon^{\frac{n-4}{2}} \|\lambda\xi'\|_{L^\infty(-1,t)} \\
& \times \int_{-81/100}^{t} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{2\gamma} \left(\sqrt{t-s} + |x|\right)^{1-n} ds \\
\lesssim\; & \varepsilon^{\frac{n-4}{2}} \|\lambda\xi'\|_{L^\infty(-1,t)} |x|^{3-n}.
\end{aligned}
$$

Here we have used Lemma B.1 to deduce the second inequality. $\qquad\square$

Corollary 16.8. *If* $|x| \geq K\lambda(t)/2$, *then*

$$
\overline{\phi}_4(x,t) \lesssim K^{-\frac{n-4}{2}} (K\varepsilon)^{-\frac{n-2}{2}} \|\lambda\xi'\|_{L^\infty(-1,t)}. \tag{16.8}
$$

Lemma 16.9 (Estimate on $\overline{\phi}_5$)**.** *If* $|x| \geq K\lambda(t)/2$, *then*

$$
\overline{\phi}_5(x,t) \lesssim \varepsilon^{\frac{n-4}{2}} |x|^{3-n} \left\|\left(\frac{a}{\lambda}, \lambda a' - \mu_0\frac{a}{\lambda}, \lambda\lambda', \lambda\xi'\right)\right\|_{L^\infty(-1,t)}. \tag{16.9}
$$

Proof. For $|y| \geq L\lambda(s)$,

$$
\lambda(s)^{-\frac{n+2}{2}} e^{-c|y|/\lambda(s)} \lesssim \lambda(s)^{\frac{n-4}{2}} |y|^{1-n}.
$$

Hence

$$\partial_t \overline{\phi}_5 - \Delta \overline{\phi}_5 - \left[C + \frac{\delta}{|x|^2}\right] \overline{\phi}_5 + \xi'(t) \cdot \nabla \overline{\phi}_5$$

$$\lesssim \left[\frac{|a(t)|}{\lambda(t)} + \left|\lambda(t)a'(t) - \mu_0 \frac{a(t)}{\lambda(t)}\right| + |\lambda(t)\lambda'(t)| + |\lambda(t)\xi'(t)|\right] \lambda(t)^{\frac{n-4}{2}} |x|^{1-n} \chi_{B_{L\lambda(t)}^c}.$$

Then we can proceed as in the previous lemma to conclude. $\square$

Corollary 16.10. *If* $|x| \geq K\lambda(t)/2$, *then*

$$\overline{\phi}_5(x,t) \lesssim K^{-\frac{n-4}{2}} (K\varepsilon)^{-\frac{n-2}{2}} \left\|\left(\frac{a}{\lambda}, \lambda a' - \mu_0 \frac{a}{\lambda}, \lambda\lambda', \lambda\xi'\right)\right\|_{L^\infty(-1,t)}. \tag{16.10}$$

Combining Corollary 16.4-Corollary 16.10 with Lemma 14.1, proceeding as in the proof of Proposition 14.6 (to estimate $|\nabla\phi_1|$-$|\nabla\phi_5|$), we find two constants $C(K,L)$ and $\sigma(K,L) \ll 1$ such that, for any $t \in (-81/100, 81/100)$,

$$\mathcal{O}(t) \leq C(K,L)\varepsilon^{\frac{n-2}{2}-\gamma} + \sigma(K,L) \sup_{-1<s<t} \mathcal{D}(s). \tag{16.11}$$

Combining this inequality with Proposition 13.4, we obtain (compare this with (15.4))

$$\mathcal{D}(t) \leq \frac{1}{2} \sup_{[t-C\varepsilon^2, t+C\varepsilon^2]} \mathcal{D}(s) + C\varepsilon^{\frac{n-2}{2}-\gamma}. \tag{16.12}$$

Here we have used the fact that, by (15.9) and the definition of inner time variable τ in Section 13, we have

$$\frac{d\tau}{dt} \sim \varepsilon^{-2}. \tag{16.13}$$

Similar to Section 15, an iteration of (16.12) from $r = 1$ to $r = 8/9$ gives (16.1). This finishes the proof of Proposition 16.1. Plugging (16.1) into Lemma 16.3-Lemma 16.9, we get Proposition 16.2.

17. Improved estimates on ϕ

From now on we need to write down the dependence on i explicitly. For u_i, the decomposition in Section 12 reads as

$$\phi_i(x,t) := u_i(x,t) - W_{\xi_i(t),\lambda_i(t)}(x) - a_i(t)Z_{0,\xi_i(t),\lambda_i(t)}(x). \tag{17.1}$$

The parameters satisfy the assumptions (15.8)-(15.10). In particular,

$$\varepsilon_i := \lambda_i(0), \quad \xi_i(0) = 0.$$

In this section we prove some uniform (in ε_i) estimates on ϕ_i.

For simplicity of notation, we will denote

$$W_i(x,t) := W_{\xi_i(t),\lambda_i(t)}(x), \quad Z_{*,i}(x,t) := \left(Z_{j,\xi_i(t),\lambda_i(t)}(x)\right)_{j=0,\cdots,n+1}.$$

17.1. **Uniform L^∞ bound.** The first one is a uniform L^∞ bound on ϕ_i.

Proposition 17.1. *There exists a universal constant C (independent of i) such that*

$$|\phi_i(x,t)| \leq C \quad in \quad Q_{6/7}. \tag{17.2}$$

First by (16.2), we get

$$u_i(x,t) \lesssim \left(\frac{\varepsilon_i}{\varepsilon_i^2 + |x - \xi_i(t)|^2} \right)^{\frac{n-2}{2}} + |x - \xi_i(t)|^{-\gamma}. \tag{17.3}$$

Hence $\overline{\phi}_{i,1}$ satisfies, for some constant $C_0 > 0$,

$$\partial_t \overline{\phi}_{i,1} - \Delta \overline{\phi}_{i,1} + \xi_i'(t) \cdot \nabla \overline{\phi}_{i,1} \leq C_0 \left[\varepsilon_i^2 |x|^{-4} \chi_{B_{L\varepsilon_i}^c} + |x|^{-(p-1)\gamma} \right] \overline{\phi}_{i,1}. \tag{17.4}$$

As in the proof of Lemma 14.1, for each $\varepsilon > 0$, consider the problem

$$\begin{cases} -\Delta w_\varepsilon = C_0 \varepsilon^2 |x|^{-4} \chi_{B_{L\varepsilon}^c} w_\varepsilon & in \ B_1, \\ w_\varepsilon = 1 & on \ \partial B_1. \end{cases} \tag{17.5}$$

Lemma 17.2. *There exists a radially symmetric solution w_ε of (17.5). Moreover, if L is universally large, for any $\varepsilon > 0$,*

$$1 \leq w_\varepsilon \leq 2 \quad in \ B_1.$$

Proof. Consider the initial value problem

$$\begin{cases} w''(s) + (n-2)w'(s) + C_0 e^{-2s} w(s) = 0 & in \ [\log L, +\infty), \\ w(\log L) = 1, \quad w'(\log L) = 0. \end{cases}$$

Global existence and uniqueness of the solution follows from standard ordinary differential equation theory.

If $w > 0$ in $[\log L, s_0]$ for some $s_0 > \log L$, then $w' < 0$, and consequently, $0 < w < 1$ in $(\log L, s_0)$. Integrating the equation of w we get

$$\begin{aligned} w'(s) &= -C_0 e^{-(n-2)s} \int_{\log L}^s e^{(n-4)\tau} w(\tau) d\tau \tag{17.6} \\ &\geq -\frac{C_0}{n-4} e^{-2s}. \end{aligned}$$

Hence, if L is large enough,

$$w(s_0) \geq 1 - \frac{C_0}{n-4} L^{-2} \geq \frac{1}{2}.$$

This holds for any s_0 provided that $w > 0$ in $[\log L, s_0]$, so $w(s) \geq 1/2$ for any $s \geq \log L$.

For each $\varepsilon > 0$, the function

$$w_\varepsilon(x) := \frac{w\left(\log(\varepsilon^{-1}|x|)\right)}{w\left(\log \varepsilon^{-1}\right)}$$

satisfies all of the requirements. $\qquad \square$

Corollary 17.3. *There exists a constant C independent of ε such that*

$$|\nabla w_\varepsilon| \le C \quad in \quad B_1.$$

Proof. By definition, we have

$$|\nabla w_\varepsilon(x)| = \frac{w'\left(\log(\varepsilon^{-1}|x|)\right)}{w\left(\log \varepsilon^{-1}\right)} \frac{\varepsilon}{|x|}.$$

If $|x| \le L\varepsilon$, $w' = 0$ and there is nothing to prove. For $|x| \ge L\varepsilon$, by (17.6),

$$\left|w'\left(\log(\varepsilon^{-1}|x|)\right)\right| \le \frac{C_0}{n-4}.$$

Because $w\left(\log \varepsilon^{-1}\right) \ge 1/2$, we get

$$|\nabla w_\varepsilon(x)| \le \frac{2C_0}{n-4}. \qquad \square$$

The function $\varphi_i := \overline{\phi}_{i,1}/w_{\varepsilon_i}$ satisfies

$$\partial_t \varphi_i - w_{\varepsilon_i}^{-2}\mathrm{div}\left(w_{\varepsilon_i}^2 \nabla \varphi_i\right) + \xi_i' \cdot \nabla \varphi_i \le \left[C_0|x|^{-(p-1)\gamma} - \xi_i' \cdot \nabla w_{\varepsilon_i}\right]\varphi_i.$$

By Lemma 17.2, this is a uniformly parabolic equation. Because γ is very small, $|x|^{-(p-1)\gamma} \in L^{2n}(B_1)$. By (10.1) and Corollary 17.3, $\xi_i' \cdot \nabla w_{\varepsilon_i}$ are uniformly bounded in $L^\infty(Q_{6/7})$. Then standard Moser iteration gives

Lemma 17.4. *There exists a universal constant C independent of i such that*

$$\sup_{Q_{6/7}} \varphi_i \lesssim \int_{Q_1} \varphi_i. \tag{17.7}$$

Combining this lemma with Lemma 17.2 and the definition of $\overline{\phi}_{i,1}$ (see (14.13)), we get

$$\|\phi_{i,1}\|_{L^\infty(Q_{6/7})} \le C.$$

Starting from this estimate, following the iteration argument in Section 16, we deduce that

$$g(5/6) \lesssim \varepsilon_i^{\frac{n-2}{2}}. \tag{17.8}$$

Substituting this estimate into Lemma 16.3-Lemma 16.9, and noting that now the heat kernel for $\partial_t - \Delta - \xi_i' \cdot \nabla - V_i$ enjoys the standard Gaussian bound

$$G_i(x,t;y,s) \lesssim (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}},$$

i.e. we can take $\gamma = 0$, we obtain

Lemma 17.5. *For any $x \in B_{K\varepsilon}^c$ and $t \in (-5/6, 5/6)$, we have*

$$|\phi_{i,2}(x,t)| + |\phi_{i,3}(x,t)| + |\phi_{i,4}(x,t)| + |\phi_{i,5}(x,t)| \lesssim_K \frac{\varepsilon^{n-4}}{|x|^{n-4}}.$$

Combining (17.7), (17.8) with this lemma, we finish the proof of Proposition 17.1.

Once we know that ϕ_i are uniformly bounded in Q_1 (after a scaling of the domain), the above estimates are improved to

Proposition 17.6.

$$\|\phi_{i,1}\|_{L^\infty(Q_{5/6})} \lesssim \|\phi_i\|_{L^\infty(Q_1)},$$

$$g(5/6) \lesssim \varepsilon_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{-c\varepsilon_i^{-2}} \right),$$

and for any $(x,t) \in Q_{5/6}$,

$$|\phi_{i,2}(x,t)| + |\phi_{i,3}(x,t)| + |\phi_{i,4}(x,t)| + |\phi_{i,5}(x,t)| \lesssim_K \frac{\varepsilon^{n-4}}{|x|^{n-4}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{-c\varepsilon_i^{-2}} \right).$$

Here the term $e^{-c\varepsilon_i^{-2}}$ appears because when we apply Lemma 13.6 to estimate a/λ, there is a boundary term

$$\frac{a_i(1)}{\lambda_i(1)} e^{-c[\tau(1)-\tau(t)]},$$

where, because $d\tau/dt \sim \varepsilon_i^{-2}$,

$$\tau(1) - \tau(t) \geq c\varepsilon_i^{-2}, \quad \text{for any } t < 25/36.$$

Because $\xi_i(0) = 0$, by the estimate of $g(5/6)$ in Proposition 17.6, now (15.11) is upgraded to

$$|\xi_i(t)| \lesssim \varepsilon_i^{\frac{n-4}{2}}, \quad \text{for any } t \in [-25/36, 25/36]. \tag{17.9}$$

17.2. **Gradient estimates.** In this subsection we establish a uniform $C^{1+\theta,(1+\theta)/2}$ estimate on ϕ_i.

With the estimates in Proposition 17.6, now we have

Lemma 17.7. *In* $Q_{5/6}$,

$$|\partial_t \phi_i - \Delta \phi_i| \lesssim \left(\frac{\varepsilon_i^2}{\varepsilon_i^4 + |x|^4} + 1 \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \tag{17.10}$$

Proof. By (17.8), now (17.9) can be upgraded to

$$|\xi_i(t)| \lesssim \varepsilon_i^{\frac{n-4}{2}}, \quad \text{for any } t \in [-25/36, 25/36]. \tag{17.11}$$

Then by the definition of ϕ_i, we have

$$u_i(x,t) \lesssim W_i(x,t) + |a_i(t)|Z_{0,i}(x,t) + |\phi_i(x,t)|$$

$$\lesssim \left(\frac{\varepsilon_i}{\varepsilon_i^2 + |x|^2} \right)^{\frac{n-2}{2}} + e^{-c|x|/\varepsilon_i} + 1.$$

Therefore

$$|u_i^p - W_i^p|$$
$$\lesssim \left(u_i^{p-1} + W_i^{p-1} \right) (|\phi_i| + |a_i(t)|Z_{0,i})$$
$$\lesssim \left[\left(\frac{\varepsilon_i}{\varepsilon_i^2 + |x|^2} \right)^{\frac{n-2}{2}} + e^{-c|x|/\varepsilon_i} + 1 \right]^{\frac{4}{n-2}} \left(1 + e^{-c|x|/\varepsilon_i} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)$$

$$\lesssim \left(\frac{\varepsilon_i^2}{\varepsilon_i^4 + |x|^4} + 1 \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

By (17.8), we have

$$\begin{aligned}
|\lambda_i'(t) Z_{n+1,i}(x,t)| &\lesssim \varepsilon_i^{-2} \left(1 + \frac{|x - \xi_i(t)|}{\varepsilon_i} \right)^{2-n} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \\
&\lesssim \frac{\varepsilon_i^{n-4}}{(\varepsilon_i + |x|)^{n-2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \\
&\lesssim \frac{\varepsilon_i^2}{\varepsilon_i^4 + |x|^4} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).
\end{aligned}$$

Similar estimates hold for $(-a_i' + \mu_0 a_i \lambda_i^{-2}) Z_{0i}$, $\xi_{ji}' \cdot Z_{ji}$, $j = 1, \cdots, n$ and $a_i \partial_t Z_{0i}$. □

Lemma 17.8. *For any* $(x, t) \in Q_{4/5}$,

$$|\nabla \phi_i(x,t)| \lesssim \left(\frac{\varepsilon_i^2}{\varepsilon_i^3 + |x|^3} + 1 \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \tag{17.12}$$

Proof. We consider three cases separately.

Case 1. If $1/4 \leq |x| \leq 4/5$, this estimate (in fact, a bound on the Lipschitz seminorm) follows from standard interior gradient estimates.

Case 2. If $|x| \leq \varepsilon_i$, by looking at the inner equation (13.1) and using Proposition 17.6, we obtain

$$|\nabla \varphi_i(y, \tau)| \lesssim \varepsilon_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Scaling back to ϕ_i, this is

$$|\nabla \phi_i(x,t)| \lesssim \varepsilon_i^{-1} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

In fact, this gives a full Lipschitz estimate in space-time.

Case 3. Finally, we consider the remaining case where $\varepsilon_i \leq |x| \leq 1/4$. Take the decomposition $\phi_i = \phi_{h,i} + \phi_{n,i}$, where

$$\begin{cases} \partial_t \phi_{h,i} - \Delta \phi_{h,i} = 0, & \text{in } Q_{3/4}, \\ \phi_{h,i} = \phi_i, & \text{on } \partial^p Q_{3/4}. \end{cases}$$

By standard interior gradient estimates for heat equation, we have

$$\|\nabla \phi_{h,i}\|_{L^\infty(Q_{2/3})} \lesssim \|\phi_i\|_{L^\infty(\partial^p Q_{3/4})}. \tag{17.13}$$

Next, for the non-homogeneous part $\phi_{n,i}$, by the heat kernel representation, we have

$$\begin{aligned}
|\nabla \phi_{n,i}(x,t)| &= \left| \int_{-\frac{81}{100}}^t \int_{B_{\frac{3}{4}}} \nabla_x G(x,t;y,s) \left(\partial_t - \Delta \right) \phi_i(y,s) dy ds \right| \\
&\lesssim \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \int_{-\frac{9}{16}}^t \int_{B_{\frac{3}{4}}} \frac{|x-y|}{(t-s)^{\frac{n+2}{2}}} e^{-c\frac{|x-y|^2}{t-s}} \left(\frac{\varepsilon_i^2}{\varepsilon_i^4 + |y|^4} + 1 \right)
\end{aligned}$$

$$\lesssim \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \int_{-\frac{9}{16}}^{t} \int_{B_{\frac{3}{4}}} (t-s)^{-\frac{n+1}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(\frac{\varepsilon_i^2}{|y|^4} + 1 \right)$$

$$\lesssim \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \int_{-\frac{9}{16}}^{t} (t-s)^{-\frac{1}{2}} \left[\frac{\varepsilon_i^2}{\left(\sqrt{t-s} + |x| \right)^4} + 1 \right]$$

$$\lesssim \left(\frac{\varepsilon_i^2}{|x|^3} + 1 \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Combining this estimate with (17.13), we get (17.12) in this case. $\qquad \square$

Next we extend this Lipschitz estimate in spatial variables to the full Lipschitz estimate in space-time variables. For this we need a technical lemma transforming the spatial Lipschitz estimate into a full Lipschitz estimate in space-time.

Lemma 17.9. *Given a constant $\sigma > 0$, suppose $\psi \in C^\infty(Q_1^-)$ satisfies*

    (1) $|\nabla \psi| \leq \sigma$ *in Q_1^-;*
    (2) $|\partial_t \psi - \Delta \psi| \leq \sigma$ *in Q_1^-.*

Then the Lipschitz seminorm (with respect to the parabolic distance) of ψ in $Q_{1/2}^-$ satisfies

$$|\psi|_{Lip(Q_{1/2}^-)} \lesssim \sigma.$$

Proof. We need only to prove that, for any $x \in B_{1/2}$ and $-1/4 < t_1 < t_2 < 0$,

$$|\psi(x, t_1) - \psi(x, t_2)| \lesssim \sigma \sqrt{t_2 - t_1}. \tag{17.14}$$

Denote $h := \sqrt{t_2 - t_1}$ and define

$$\widetilde{\psi}(y, s) := \frac{1}{h} \left[\psi(x + hy, t_2 + h^2 s) - \psi(x, t_2) \right].$$

It satisfies

    (1) $|\nabla \widetilde{\psi}| \leq \sigma$ *in Q_2^-;*
    (2) for any $(y, s) \in Q_2^-$,

$$|\partial_t \widetilde{\psi}(y, s) - \Delta \widetilde{\psi}(y, s)| \leq h\sigma; \tag{17.15}$$

    (3) because $\widetilde{\psi}(0, 0) = 0$, by integrating (1) we obtain

$$\sup_{y \in B_2} |\widetilde{\psi}(y, 0)| \leq 2\sigma. \tag{17.16}$$

Fix a function $\eta \in C_0^\infty(B_2)$. Multiplying (17.15) and integrating by parts, we obtain

$$\left| \frac{d}{ds} \int_{B_2} \widetilde{\psi} \eta \right| \leq h\sigma \int_{B_2} |\eta| + \int_{B_2} |\nabla \widetilde{\psi}||\nabla \eta|$$

$$\lesssim \sigma.$$

Combining this inequality with (17.16), we obtain

$$\left| \int_{B_2} \widetilde{\psi}(y, s)\eta(y) \right| \lesssim \sigma \quad \text{for any } s \in (-4, 0).$$

With the Lipschitz estimate in (1), this implies that

$$\sup_{y \in B_2} |\widetilde{\psi}(y, s)| \lesssim \sigma \quad \text{for any } s \in (-4, 0). \tag{17.17}$$

In particular,

$$|\widetilde{\psi}(0, -1)| \lesssim \sigma.$$

Scaling back to ψ, this is (17.14). $\qquad \square$

Proposition 17.10. *For any* $(x, t) \in Q_{3/4}$,

$$|\phi_i|_{Lip(Q_{|x|/4}(x,t))} \lesssim \left(\frac{\varepsilon_i^2}{\varepsilon_i^3 + |x|^3} + 1 \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \tag{17.18}$$

Proof. When $|x| < \varepsilon_i$ and $1/4 < |x| < 3/4$, this was already covered in the proof of Lemma 17.8, see Case 1 and Case 2 therein.

If $\varepsilon_1 < |x| < 1/4$, denote $r := |x|/2$ and let

$$\widetilde{\phi}_i(y, s) := \phi_i \left(x + ry, t + r^2 s \right).$$

By Lemma 17.7 and Lemma 17.8, it satisfies the assumptions in Lemma 17.9 with the constant

$$\sigma \lesssim \left(\frac{\varepsilon_i^2}{r^2} + r \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Hence this lemma implies that

$$|\widetilde{\phi}_i|_{Lip(Q_{1/2}^-)} \lesssim \left(\frac{\varepsilon_i^2}{r^2} + r \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Scaling back to ϕ_i, this is (17.18) in this case. $\qquad \square$

17.3. **Estimate of time derivative.** In this subsection we establish the following estimate on $\partial_t \phi_i$. We can also obtain some estimates on $\nabla^2 \phi_i$, but we do not need it.

Proposition 17.11. *For any* $(x, t) \in Q_{2/3}$,

$$|\partial_t \phi_i(x, t)| \lesssim \left(\frac{\varepsilon_i^2}{\varepsilon_i^4 + |x|^4} + \frac{1}{\varepsilon_i^2 + |x|} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \tag{17.19}$$

Proof. Take an arbitrary sequence of points $(x_i, t_i) \in Q_{2/3}$. Denote $r_i := |x_i|/2$.

Case 1. First assume

$$\limsup_{i \to +\infty} \frac{r_i}{\varepsilon_i} < +\infty.$$

In this case we work in the inner coordinates introduced in Section 13. By (13.2) and (13.3), we have the following representation formulas:

$$\begin{cases} \dfrac{\dot{a}_i(\tau) - \mu_0 a_i(\tau)}{\lambda_i(\tau)} = \displaystyle\int_{\mathbb{R}^n} Z_0(y) E_{K,i}(y, \tau) dy, \\[3mm] \dfrac{\dot{\xi}_i(\tau)}{\lambda_i(\tau)} = -\displaystyle\int_{\mathbb{R}^n} Z_j(y) E_{K,i}(y, \tau) dy, \quad j = 1, \cdots, n, \\[3mm] \dfrac{\dot{\lambda}_i(\tau)}{\lambda_i(\tau)} = -\displaystyle\int_{\mathbb{R}^n} Z_{n+1}(y) E_{K,i}(y, \tau) dy. \end{cases}$$

By the form of E_K in Section 13, and the smallness of

$$\left|\frac{\dot{a}_i - \mu_0 a_i}{\lambda_i}\right| + \left|\frac{\dot{\xi}_i}{\lambda_i}\right| + \left|\frac{\dot{\lambda}_i}{\lambda_i}\right|,$$

the above three equations are solved as

$$\left(\frac{\dot{a}_i(\tau) - \mu_0 a_i(\tau)}{\lambda_i(\tau)}, \frac{\dot{\xi}_i(\tau)}{\lambda_i(\tau)}, \frac{\dot{\lambda}_i(\tau)}{\lambda_i(\tau)}\right) = \mathcal{J}\left(\nabla\varphi_i(\tau), \varphi_i(\tau), \frac{a_i(\tau)}{\lambda_i(\tau)}\right), \tag{17.20}$$

where $\mathcal{J} = (\mathcal{J}_0, \cdots, \mathcal{J}_{n+1})$ is a vector valued, nonlinear (but smooth) integral operator. Plugging this back into (13.1), we get

$$\begin{aligned}
\partial_\tau \varphi_i - \Delta_y \varphi_i &= (W + \varphi + a_i Z_0)^p - W^p + \mathcal{J} \cdot Z \\
&+ (\mathcal{J}_1, \cdots, \mathcal{J}_n) \cdot \nabla\varphi + \mathcal{J}_{n+1}\left(y \cdot \nabla\varphi_i + \frac{n-2}{2}\varphi_i\right) \\
&+ \frac{a_i}{\lambda_i}\left[(\mathcal{J}_1, \cdots, \mathcal{J}_n) \cdot \nabla Z_0 + \mathcal{J}_{n+1}\left(y \cdot \nabla Z_0 + \frac{n}{2}Z_0\right)\right].
\end{aligned}$$

Starting from the L^∞ bound of φ_i (by rescaling the estimates in Proposition 17.6) and bootstrapping parabolic estimates, we get

$$\|\varphi_i\|_{C^{2+\theta,1+\theta/2}(B_{2K}\times[\tau-1,\tau])} \lesssim \varepsilon_i^{\frac{n-2}{2}}, \quad \text{for any } \tau. \tag{17.21}$$

A byproduct of this estimate is the validity of (17.19) in this case.

Case 2. Substituting (17.21) into (17.20) gives, for any $\tau_1 < \tau_2$,

$$\left\|\left(\frac{\dot{a}_i - \mu_0 a_i}{\lambda_i}, \frac{\dot{\xi}_i}{\lambda_i}, \frac{\dot{\lambda}_i}{\lambda_i}\right)\right\|_{C^{\theta/2}([\tau_1,\tau_2])} \lesssim \varepsilon_i^{\frac{n-2}{2}}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right). \tag{17.22}$$

After scaling back to the original coordinates, by noting that $d\tau/dt \sim \varepsilon_i^{-2}$, this estimate is transformed into

$$\left\|\left(a_i' - \mu_0\frac{a_i}{\lambda_i^2}, \xi_i', \lambda_i'\right)\right\|_{C^{\theta/2}([-3/4,3/4])} \lesssim \varepsilon_i^{\frac{n-4}{2}-\theta}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right). \tag{17.23}$$

By Proposition 17.6, (17.23) and Proposition 17.10, we deduce that the terms in the right hand side of (12.8) are uniformly bounded in $C^{\theta,\theta/2}\left((B_1 \setminus B_{1/8}) \times (-1,1)\right)$. Hence standard Schauder estimates for heat equation gives

$$\sup_{\left((B_{2/3}\setminus B_{1/4})\times(-4/9,4/9)\right)} |\partial_t \phi_i| \lesssim \|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}.$$

Case 3. In this case we assume

$$\lim_{i\to+\infty}\left(r_i + \frac{\varepsilon_i}{r_i}\right) = 0.$$

Define

$$\widetilde{u}_i(x,t) := r_i^{\frac{n-2}{2}} u_i\left(r_i x, t_i + r_i^2 t\right).$$

The decomposition in Section 12 can be transferred to $\widetilde{u}_i$ as

$$\widetilde{u}_i = W_{\widetilde{\xi}_i, \widetilde{\lambda}_i} + \widetilde{a}_i Z_{0, \widetilde{\xi}_i, \widetilde{\lambda}_i} + \widetilde{\phi}_i, \tag{17.24}$$

where

$$\begin{cases} \widetilde{\xi}_i(t) := r_i^{-1}\xi_i(t_i + r_i^2 t), \\ \widetilde{\lambda}_i(t) := r_i^{-1}\lambda_i(t_i + r_i^2 t), \\ \widetilde{a}_i(t) := r_i^{-1}a_i(t_i + r_i^2 t), \\ \widetilde{\phi}_i(x,t) := r_i^{(n-2)/2}\phi_i(r_i x, t_i + r_i^2 t). \end{cases}$$

Because for any $t \in [-3/4, 3/4]$,

$$r_i > 2K\varepsilon_i > K\lambda_i(t),$$

the orthogonal relation in Proposition 12.1 still holds between $\widetilde{\phi}_i$ and $Z_{j,\widetilde{\xi}_i,\widetilde{\lambda}_i}$, $j = 0, \cdots, n+1$.

Furthermore, the error equation (12.8) now reads as

$$\begin{aligned} \partial_t \widetilde{\phi}_i - \Delta\widetilde{\phi}_i &= \left(W_{\widetilde{\xi}_i,\widetilde{\lambda}_i} + \widetilde{a}_i Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} + \widetilde{\phi}_i \right)^p - W^p_{\widetilde{\xi}_i,\widetilde{\lambda}_i} - p W^{p-1}_{\widetilde{\xi}_i,\widetilde{\lambda}_i} \widetilde{a}_i Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} \\ &\quad + \left(-\widetilde{a}_i' + \mu_0 \frac{\widetilde{a}_i}{\widetilde{\lambda}_i^2}, \widetilde{\xi}_i', \widetilde{\lambda}_i' \right) \cdot Z_{\widetilde{\xi}_i,\widetilde{\lambda}_i} - \widetilde{a}_i \partial_t Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i}. \end{aligned} \tag{17.25}$$

By (17.11) and (17.23), we get the following estimates for those terms in the right hand side of (17.25):

$$\begin{cases} \left\| W_{\widetilde{\xi}_i,\widetilde{\lambda}_i} \right\|_{L^\infty((B_2 \backslash B_{1/4}) \times (-4,4))} \lesssim \left(\dfrac{\varepsilon_i}{r_i} \right)^{\frac{n-2}{2}}; \\[3mm] \left\| \widetilde{a}_i Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} \right\|_{L^\infty((B_2 \backslash B_{1/4}) \times (-4,4))} \lesssim \varepsilon_i^{\frac{n-2}{2}} e^{-c\frac{r_i}{\varepsilon_i}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right); \\[3mm] \left\| \widetilde{\phi}_i \right\|_{L^\infty((B_1 \backslash B_{1/4}) \times (-4,4))} \lesssim r_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right); \\[3mm] \left\| \left(\widetilde{a}_i' - \mu_0 \widetilde{a}_i \widetilde{\lambda}_i^{-2}, \widetilde{\xi}_i', \widetilde{\lambda}_i' \right) \cdot Z_{\widetilde{\xi}_i,\widetilde{\lambda}_i} \right\|_{L^\infty((B_2 \backslash B_{1/4}) \times (-4,4))} \lesssim \dfrac{\varepsilon_i^{n-4}}{r_i^{\frac{n}{2}-3}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right); \\[3mm] \left\| \widetilde{a}_i \partial_t Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} \right\|_{C^{\theta,\theta/2}((B_2 \backslash B_{1/4}) \times (-4,4))} \lesssim \varepsilon_i^{n-2} e^{-c\frac{r_i}{\varepsilon_i}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \end{cases} \tag{17.26}$$

These estimates imply that in $(B_2 \setminus B_{1/4}) \times (-4,4)$,

$$\widetilde{u}_i \lesssim \left(\frac{\varepsilon_i}{r_i} \right)^{\frac{n-2}{2}} + r_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \tag{17.27}$$

and

$$\begin{aligned} &\left| \left(W_{\widetilde{\xi}_i,\widetilde{\lambda}_i} + \widetilde{a}_i Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} + \widetilde{\phi}_i \right)^p - W^p_{\widetilde{\xi}_i,\widetilde{\lambda}_i} - p W^{p-1}_{\widetilde{\xi}_i,\widetilde{\lambda}_i} \widetilde{a}_i Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} \right| \\ &\lesssim \left[W^{p-1}_{\widetilde{\xi}_i,\widetilde{\lambda}_i} + \widetilde{u}_i^{p-1} \right] \left(\left| \widetilde{\phi}_i \right| + \left| \widetilde{a}_i Z_{0,\widetilde{\xi}_i,\widetilde{\lambda}_i} \right| \right) \\ &\lesssim \left(\frac{\varepsilon_i^2}{r_i^2} + r_i^2 \right) r_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \end{aligned}$$

By Proposition 17.10 and a rescaling, we know that the Lipschitz seminorm of $\widetilde{\phi}_i$ t in $\left(B_2 \setminus B_{1/4}\right) \times (-4, 4)$ is bounded by

$$r_i^{\frac{n}{2}} \left(\frac{\varepsilon_i^2}{r_i^3} + 1\right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right).$$

By these estimates and standard $W^{2,p}$ estimates for heat equation, we obtain

$$\int_{-3}^{3} \int_{B_{9/5} \setminus B_{1/3}} |\partial_t \widetilde{\phi}_i| \lesssim \left[r_i^{\frac{n}{2}} \left(\frac{\varepsilon_i^2}{r_i^3} + 1\right) + \varepsilon_i^{-\frac{n-2}{2}} \left(\frac{\varepsilon_i}{r_i}\right)^{\frac{n-3}{2}} + r_i^{-\frac{n-2}{2}} \left(\frac{\varepsilon_i^2}{r_i^2} + r_i^2\right)\right]$$
$$\times \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right). \tag{17.28}$$

Because

$$\left(\partial_t - \Delta\right) \partial_t \widetilde{u}_i = p\widetilde{u}_i^{p-1} \partial_t \widetilde{u}_i,$$

and $\widetilde{u}_i$ are uniformly bounded in $\left(B_{9/5} \setminus B_{1/3}\right) \times (-3, 3)$, we get

$$\|\partial_t \widetilde{u}_i\|_{L^\infty((B_{3/2} \setminus B_{1/2}) \times (-1,1))}$$
$$\lesssim \int_{-3}^{3} \int_{B_{9/5} \setminus B_{1/3}} |\partial_t \widetilde{u}_i| \tag{17.29}$$
$$\lesssim \left(\int_{-3}^{3} \int_{B_{9/5} \setminus B_{1/3}} |\partial_t \widetilde{\phi}_i|\right) + \|\partial_t W_{\widetilde{\xi}_i, \widetilde{\lambda}_i}\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))}$$
$$+ \left\|\partial_t \left(\widetilde{a}_i Z_{0, \widetilde{\xi}_i, \widetilde{\lambda}_i}\right)\right\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))},$$

where we have used the expansion from (17.24),

$$\partial_t \widetilde{u}_i = \partial_t W_{\widetilde{\xi}_i, \widetilde{\lambda}_i} + \partial_t \left(\widetilde{a}_i Z_{0, \widetilde{\xi}_i, \widetilde{\lambda}_i}\right) + \partial_t \widetilde{\phi}_i.$$

Similar to (17.26), we also have

$$\begin{cases} \|\partial_t W_{\widetilde{\xi}_i, \widetilde{\lambda}_i}\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))} \lesssim \varepsilon_i^{n-4} r_i^{-\frac{n}{2}-3} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right); \\ \left\|\partial_t \left(\widetilde{a}_i Z_{0, \widetilde{\xi}_i, \widetilde{\lambda}_i}\right)\right\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))} \lesssim \varepsilon_i^{\frac{n-2}{2}} e^{-c\frac{r_i}{\varepsilon_i}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right). \end{cases}$$

Substituting these estimates and (17.28) into (17.29), we obtain

$$\left\|\partial_t \widetilde{\phi}_i\right\|_{L^\infty((B_{3/2} \setminus B_{1/2}) \times (-1,1))}$$
$$\lesssim \|\partial_t \widetilde{u}_i\|_{L^\infty((B_{3/2} \setminus B_{1/2}) \times (-1,1))}$$
$$+ \|\partial_t W_{\widetilde{\xi}_i, \widetilde{\lambda}_i}\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))} + \left\|\partial_t \left(\widetilde{a}_i Z_{0, \widetilde{\xi}_i, \widetilde{\lambda}_i}\right)\right\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))}$$
$$\lesssim \int_{-3}^{3} \int_{B_{9/5} \setminus B_{1/3}} |\partial_t \widetilde{\phi}_i|$$
$$+ \|\partial_t W_{\widetilde{\xi}_i, \widetilde{\lambda}_i}\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))} + \left\|\partial_t \left(\widetilde{a}_i Z_{0, \widetilde{\xi}_i, \widetilde{\lambda}_i}\right)\right\|_{L^\infty((B_2 \setminus B_{1/4}) \times (-4,4))}$$

$$\lesssim \quad \left[r_i^{\frac{n}{2}} \left(\frac{\varepsilon_i^2}{r_i^3} + 1 \right) + \varepsilon_i^{\frac{n-2}{2}} \left(\frac{\varepsilon_i}{r_i} \right)^{\frac{n-3}{2}} + r_i^{\frac{n-2}{2}} \left(\frac{\varepsilon_i^2}{r_i^2} + r_i^2 \right) \right] \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)$$

$$+ \quad \left[\frac{\varepsilon_i^{n-4}}{r_i^{\frac{n}{2}-3}} + \varepsilon_i^{\frac{n-2}{2}} e^{-c\frac{r_i}{\varepsilon_i}} \right] \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Scaling back to ϕ_i, we get (17.19). □

18. Linearization of Pohozaev identity

Given a smooth function v and a sphere ∂B_r, define the Pohozaev invariant as

$$\mathcal{P}_v(r) := \int_{\partial B_r} \left[\frac{|\nabla v|^2}{2} - (\partial_r v)^2 - \frac{n-2}{2r} v \partial_r v \right].$$

Choose of a sequence of r_i satisfying

$$\lim_{i \to +\infty} \frac{r_i}{\varepsilon_i} = +\infty.$$

Multiplying (1.1) by $x \cdot \nabla u_i(x, 0)$ and integrating in B_{r_i}, after integrating by parts we obtain a Pohozaev identity

$$\mathcal{P}_{u_i(0)}(r_i) - \int_{\partial B_{r_i}} \frac{u_i(0)^{p+1}}{p+1} = -\frac{1}{r_i} \int_{B_{r_i}} \partial_t u_i(0) \left[x \cdot \nabla u_i(0) + \frac{n-2}{2} u_i(0) \right]. \quad (18.1)$$

Substitute the decomposition (17.1) into this identity, and take a expansion. Let us estimate each term in this expansion. We will see that the zeroth order terms cancel with each other, because these terms form exactly the Pohozaev invariant of the bubble. The next order term then gives us some information.

18.1. The left hand side.
For the left hand side, we have the expansion

$$\mathcal{P}_{u_i(0)}(r_i) - \int_{\partial B_{r_i}} \frac{u_i(0)^{p+1}}{p+1}$$

$$= \quad \mathcal{P}_{W_i(0)}(r_i) - \int_{\partial B_{r_i}} \frac{W_i(0)^{p+1}}{p+1} + \mathcal{P}_{\phi_i(0)}(r_i) + \mathcal{P}_{a_i(0)Z_{0,i}(0)}(r_i)$$

$$+ \quad \underbrace{\int_{\partial B_{r_i}} \nabla W_i(0) \cdot \nabla \phi_i(0) - 2\partial_r W_i(0) \partial_r \phi_i(0) - \frac{n-2}{2r_i} \left[W_i(0) \partial_r \phi_i(0) + \phi_i(0) \partial_r W_i(0) \right]}_{\text{Main order term}}$$

$$- \quad \frac{1}{p+1} \int_{\partial B_{r_i}} \left[u_i(0)^{p+1} - W_i(0)^{p+1} \right] + \text{cross terms involving } a_i(0)Z_{0,i}(0).$$

Let us estimate each term in this expansion.

(1) Because $W_i(0)$ is a smooth solution of (1.9), it satisfies the standard Pohozaev identity

$$\mathcal{P}_{W_i(0)}(r_i) - \int_{\partial B_{r_i}} \frac{W_i(0)^{p+1}}{p+1} = 0.$$

(2) By Proposition 17.1 and Proposition 17.10, we get

$$\mathcal{P}_{\phi_i(0)}(r_i) = O\left(r_i^{n-3}\right)\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right)^2.$$

Here the factor $\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right)$ is kept for later use.

(3) By estimates on a_i and λ_i in Proposition 17.6, on ∂B_{r_i} we have

$$\begin{cases} |a_i(t)Z_{0,i}(t)| \lesssim e^{-cr_i/\varepsilon_i}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right), \\ |\nabla\left(a_i(t)Z_{0,i}(t)\right)| \lesssim \varepsilon_i^{-1} e^{-cr_i/\varepsilon_i}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right). \end{cases} \tag{18.2}$$

By these estimates we get

$$\left|\mathcal{P}_{a_i(0)Z_{0,i}(0)}(r)\right| \lesssim \varepsilon_i^{-2} r_i^{n-1} e^{-cr_i/\varepsilon_i}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right)^2.$$

(4) First by integrating (17.12), for any $x \in \partial B_{r_i}$, we have

$$\phi_i(x,0) = \frac{1}{|\partial B_\rho|}\int_{\partial B_\rho}\phi_i(0) + O\left(\frac{\varepsilon_i^2}{r_i^2} + \rho\right)\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right).$$

Next, because $\xi_i(0) = 0$, on ∂B_{r_i}, by Proposition 17.6 and (17.11), we obtain

$$\begin{cases} W_i(0) \lesssim \varepsilon_i^{\frac{n-2}{2}} r_i^{2-n}, \\ \partial_r W_i(0) = -c(n)\varepsilon_i^{\frac{n-2}{2}} r_i^{1-n} + O\left(\varepsilon_i^{\frac{n+2}{2}} r_i^{-1-n}\right), \\ \nabla W_i(0) = \partial_r W_i(0)\frac{x}{|x|}, \\ |\nabla\phi_i(t)| \lesssim \left(\varepsilon_i^2 r_i^{-3} + 1\right)\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right). \end{cases}$$

Therefore the main order term equals

$$-c(n)\left[\frac{1}{|\partial B_\rho|}\int_{\partial B_\rho}\phi_i + O\left(\frac{\varepsilon_i^2}{r_i^2} + \rho\right)\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right)\right]\frac{\varepsilon_i^{\frac{n-2}{2}}}{r_i}$$

$$+ \ O\left(\varepsilon_i^{\frac{n-2}{2}}\right)\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right).$$

(5) On ∂B_{r_i}, we have

$$W_i(0) \lesssim \varepsilon_i^{\frac{n-2}{2}} r_i^{2-n}, \quad |a_i Z_{0,i}(0)| \lesssim e^{-cr_i/\varepsilon_i}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right).$$

Therefore

$$\left|\int_{\partial B_{r_i}}\left[u_i(t)^{p+1} - W_i(t)^{p+1}\right]\right| \lesssim \left(\varepsilon_i^{\frac{n+2}{2}} r_i^{-3} + r_i^{n-1}\right)\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right).$$

(6) Finally, by (18.2) and estimates of $W_i(0)$ and $\phi_i(0)$, the cross terms involving $a_i Z_{0,i}(0)$ are of the order

$$O\left(\varepsilon_i^{\frac{n-4}{2}} + r_i^{n-3} + \varepsilon_i^{-1} r_i^{n-2}\right) e^{-cr_i/\varepsilon_i}\left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}}\right).$$

Putting these estimates together, we see that the left hand side of (18.1) equals

$$-c(n) \left[\frac{1}{|\partial B_\rho|} \int_{\partial B_\rho} \phi_i + O\left(\frac{\varepsilon_i^2}{r_i^2} + \rho\right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \right] \frac{\varepsilon_i^{\frac{n-2}{2}}}{r_i}$$

$$+ \quad O\left(r_i^{n-3} + \varepsilon_i^{-2} r_i^{n-1} e^{-cr_i/\varepsilon_i} + \varepsilon_i^{\frac{n-2}{2}} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right) \qquad (18.3)$$

$$+ \quad \left(\varepsilon_i^{\frac{n+2}{2}} r_i^{-3} + r_i^{n-1} + \varepsilon_i^{\frac{n-4}{2}} e^{-cr_i/\varepsilon_i} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

18.2. **The right hand side.** By (17.1), we get

$$\partial_t u_i = (a_i', \xi_i', \lambda_i') \cdot Z_{*,i} + \partial_t \phi_i + a_i \partial_t Z_{0,i} \qquad (18.4)$$

and

$$x \cdot \nabla u_i + \frac{n-2}{2} u_i = \lambda_i Z_{n+1,i} + \left(x \cdot \nabla \phi_i + \frac{n-2}{2} \phi_i \right) \qquad (18.5)$$

$$+ \quad a_i \left(x \cdot Z_{0,i} + \frac{n-2}{2} Z_{0,i} \right).$$

Therefore for any r, we have

$$\int_{B_r} \partial_t u_i(x,0) \left[x \cdot \nabla u_i(x,0) + \frac{n-2}{2} u_i(x,0) \right] dx$$

$$= \quad \lambda_i(0) \, (a_i'(0), \xi_i'(0), \lambda_i'(0)) \cdot \int_{B_r} Z_{*,i}(x,0) Z_{n+1,i}(x,0) dx$$

$$+ \quad (a_i'(0), \xi_i'(0), \lambda_i'(0)) \cdot \int_{B_r} Z_{*,i}(x,0) \left[x \cdot \nabla \phi_i(x,0) + \frac{n-2}{2} \phi_i(x,0) \right] dx$$

$$+ \quad a_i(0) \, (a_i'(0), \xi_i'(0), \lambda_i'(0)) \cdot \int_{B_r} Z_{*,i}(x,0) \left[x \cdot Z_{0,i}(x,0) + \frac{n-2}{2} Z_{0,i}(x,0) \right] dx$$

$$+ \quad \lambda_i(0) \int_{B_r} Z_{n+1,i}(x,0) \partial_t \phi_i(x,0) dx$$

$$+ \quad \int_{B_r} \partial_t \phi_i(x,0) \left[x \cdot \nabla \phi_i(x,0) + \frac{n-2}{2} \phi_i(x,0) \right] dx$$

$$+ \quad a_i(0) \int_{B_r} \partial_t \phi_i(x,0) \left[x \cdot Z_{0,i}(x,0) + \frac{n-2}{2} Z_{0,i}(x,0) \right] dx$$

$$+ \quad a_i(0) \lambda_i(0) \int_{B_r} Z_{n+1,i}(x,0) \partial_t Z_{0,i}(x,0) dx$$

$$+ \quad a_i(0) \int_{B_r} \partial_t Z_{0,i}(x,0) \left[x \cdot \nabla \phi_i(x,0) + \frac{n-2}{2} \phi_i(x,0) \right] dx$$

$$+ \quad a_i(0)^2 \int_{B_r} \partial_t Z_{0,i}(x,0) \left[x \cdot Z_{0,i}(x,0) + \frac{n-2}{2} Z_{0,i}(x,0) \right] dx$$

$$=: \quad \text{I} + \text{II} + \text{III} + \text{IV} + \text{V} + \text{VI} + \text{VII} + \text{VIII} + \text{IX}.$$

Let us take $r = r_i$ and estimate each term one by one.

(1) By (15.9), Proposition 17.6 and the orthogonal relation between Z_0 and Z_{n+1},

$$\left| \lambda_i(0) a_i'(0) \int_{B_{r_i}} Z_{0,i}(x,t) Z_{n+1,i}(x,0) dx \right| \lesssim \varepsilon_i^{\frac{n-2}{2}} e^{-cr_i/\varepsilon_i} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Similarly, for each $j = 1, \cdots, n$,

$$\left| \lambda_i(0) \xi_{j,i}'(0) \int_{B_{r_i}} Z_{j,i}(x,t) Z_{n+1,i}(x,0) dx \right| \lesssim \varepsilon_i^{\frac{3n}{2}-4} r_i^{3-n} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

By Lemma 13.5, (15.9) and Proposition 17.6,

$$\left| \lambda_i(0) \lambda_i(0)' \int_{B_{r_i}} Z_{i,n+1}(x,0)^2 dx \right| \lesssim K^{-\frac{n-2}{2}} \varepsilon_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Putting these three estimates together we obtain

$$|\mathrm{I}| \lesssim \left[e^{-cr_i/\varepsilon_i} + \left(\frac{\varepsilon_i}{r_i} \right)^{n-3} + K^{-\frac{n-2}{2}} \right] \varepsilon_i^{\frac{n-2}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

(2) By Proposition 17.6, Proposition 17.1 and Lemma 17.8, we get

$$|\mathrm{II}| \lesssim \varepsilon_i^{n-4} r_i^2 \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

(3) By (15.9) and (17.8),

$$|\mathrm{III}| \lesssim \varepsilon_i^{n-2} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

(4) To estimate IV, first note that because $r_i \gg \varepsilon_i$, the orthogonal condition (12.2) for ϕ_i reads as

$$\int_{B_{r_i}} \phi_i(x,t) \eta_{in}(x,t) Z_{n+1,i}(x,t) dx = 0.$$

Differentiating this equation in t gives

$$\left| \int_{B_{r_i}} Z_{n+1,i}(x,0) \partial_t \phi_i(x,0) \eta_{in}(x,0) dx \right|$$

$$\leq \left| \int_{B_{r_i}} \phi_i(x,0) \partial_t \eta_{in}(x,0) Z_{n+1,i}(x,0) dx \right| + \left| \int_{B_{r_i}} \phi_i(x,0) \eta_{in}(x,0) \partial_t Z_{n+1,i}(x,0) dx \right|$$

$$\lesssim K^2 \varepsilon_i^{n-3} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

On the other hand, a direct calculation using Proposition 17.11 gives

$$\left| \int_{B_{r_i}} Z_{n+1,i}(x,0) \partial_t \phi_i(x,0) \left[1 - \eta_{in}(x,0) \right] dx \right|$$

$$\lesssim \left(K^{-2} + r_i \right) \varepsilon_i^{\frac{n-4}{2}} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

Putting these two estimates together we get

$$|\text{IV}| \lesssim \left[\left(K^{-2} + r_i \right) \varepsilon_i^{\frac{n-2}{2}} + K^2 \varepsilon_i^{n-2} \right] \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

(5) By Proposition 17.1, Lemma 17.8 and Proposition 17.11,

$$|\text{V}| \lesssim \left(\varepsilon_i^2 r_i^{n-4} + r_i^{n-1} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

(6) By Proposition 17.6 and Proposition 17.11,

$$|\text{VI}| \lesssim \varepsilon_i^{n-2} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

(7) By Proposition 17.6,

$$|\text{VII}| \lesssim \varepsilon_i^{n-2} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

(8) By Proposition 17.6, Proposition 17.1 and Lemma 17.8,

$$|\text{VIII}| \lesssim \varepsilon_i^{\frac{3n}{2}-3} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

(9) By Proposition 17.6,

$$|\text{IX}| \lesssim \varepsilon_i^{2n-3} \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right)^2.$$

Putting these estimates together, we get

$$\text{RHS} \lesssim r_i^{-1} \left[\left(K^{-2} + r_i \right) \varepsilon_i^{\frac{n-2}{2}} + \varepsilon_i^2 r_i^{n-4} + r_i^{n-1} \right] \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \tag{18.6}$$

19. A weak form of Schoen's Harnack inequality

In this section, we establish a weak form of Schoen's Harnack inequality, which then finishes the proof of Theorem 10.1. In the Yamabe problem, this Harnack inequality was first introduced in Schoen [77], see also Li [47] and Li-Zhang [49] for a proof using the method of moving plane and moving sphere. The following proof instead is mainly a consequence of the Pohozaev identity calculation in Section 18.

Combining (18.3) and (18.6) in the previous section, we get

$$\frac{1}{|\partial B_\rho|} \int_{\partial B_\rho} \phi_i(0)$$

$$\lesssim \left(\rho + \frac{\varepsilon_i^2}{r_i^2} + \frac{r_i^{n-2}}{\varepsilon_i^{(n-2)/2}} + \frac{\varepsilon_i^2}{r_i^2} + K^{-2} + r_i + \frac{r_i^{n-3}}{\varepsilon_i^{\frac{n}{2}-3}} + e^{-c\frac{r_i}{\varepsilon_i}} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right).$$

By choosing $r_i = \varepsilon_i^{(n-1)/n}$, we obtain a constant $\delta(n) > 0$ such that

$$\frac{1}{|\partial B_\rho|} \int_{\partial B_\rho} \phi_i(0) \lesssim \left(\rho + K^{-2} + \varepsilon_i^{\delta(n)} \right) \left(\|\phi_i\|_{L^\infty(Q_1)} + e^{c\varepsilon_i^{-2}} \right). \tag{19.1}$$

On ∂B_ρ, $u_i(0) \to u_\infty(0)$ uniformly. In view of the estimates (15.8)-(15.10), we deduce that $\phi_i(0) \to u_\infty(0)$ uniformly on ∂B_ρ, too. Passing to the limit in (19.1), we obtain

$$\frac{1}{|\partial B_\rho|} \int_{\partial B_\rho} u_\infty(x,0) \lesssim \left(\rho + K^{-2}\right) \|u_\infty\|_{L^\infty(Q_1)} \lesssim \rho + K^{-2}. \qquad (19.2)$$

Since u_∞ is a smooth, nonnegative solution of (1.1), if u_∞ is not identically zero, by the standard Harnack ineqaulity,

$$\inf_{Q_{1/4}} u_\infty > 0. \qquad (19.3)$$

On the other hand, if we have chosen K large enough at the beginning (in Proposition 12.1) and taken ρ arbitrarily small, (19.2) contradicts (19.3). [1]

Hence in the setting of Section 10, we must have $u_\infty = 0$. This finishes the proof of Theorem 10.1.

For the study of bubble clusterings (see Section 20 below), we need a quantitative version of the above qualitative description. If the solution exists for a sufficiently large time, an iteration of the following proposition backwardly in time will lead to an quantitative upper bound on the error function, see Section 20 for details. In particular, if the solution exists globally in time (e.g. a solution independent of time), we can recover Schoen's Harnack inequality.

Proposition 19.1. *There exist two universal constants ε_0 and M_0 so that the following holds. Suppose u_ε is a positive solution of (1.1) in Q_1, satisfying* **(II.a')-(I.d')** *and the decomposition (17.1) with parameters $(a_\varepsilon(t), \xi_\varepsilon(t), \lambda_\varepsilon(t))$, where*

$$\lambda_\varepsilon(0) = \varepsilon, \quad \xi_\varepsilon(0) = 0.$$

If $\varepsilon \leq \varepsilon_0$ and

$$\|\phi_\varepsilon\|_{L^\infty(Q_1)} \geq M_0 \varepsilon^{\frac{n-2}{2}}, \qquad (19.4)$$

then

$$\|\phi_\varepsilon\|_{L^\infty(B_1 \times (-1/4, 1/8))} \leq \frac{1}{2} \|\phi_\varepsilon\|_{L^\infty(Q_1)}.$$

Proof. We prove this proposition by contradiction. Assume

- u_i is a sequence of solutions satisfying **(II.a)-(II.c)** and (10.1);
- the decomposition given in (17.1) holds, where

$$\varepsilon_i := \lambda_i(0) \to 0, \quad \xi_i(0) = 0;$$

- ϕ_i satisfies

$$\lim_{i \to +\infty} \frac{\|\phi_i\|_{L^\infty(Q_1)}}{\varepsilon_i^{\frac{n-2}{2}}} = +\infty, \qquad (19.5)$$

but

$$\|\phi_i\|_{L^\infty(B_1 \times (-1/4, 1/8))} > \frac{1}{2} \|\phi_i\|_{L^\infty(Q_1)}. \qquad (19.6)$$

---

[1] We can also avoid the use of Harnack inequality, and use only the strong maximum principle. For this approach, we need to take a sequence of $K \to +\infty$ in Proposition 12.1. This changes the error function ϕ_i, but it does not affect the argument. This is because both (19.2) and (19.3) involve only u_∞, which is the weak limit of u_i and does not depend on the construction of ϕ_i.

We show that this leads to a contradiction.

Step 1. Denote $\delta_i := \|\phi_i\|_{L^\infty(Q_1)}$. Let $\widetilde{\phi}_i := \phi_i/\delta_i$. In this step we prove

$$\widetilde{\phi}_i \to 0 \quad \text{uniformly in any compact set of } Q_1. \tag{19.7}$$

First by Proposition 17.10, for any $\rho > 0$, $\widetilde{\phi}_i$ are uniformly Lipschitz in $(B_{2/3} \backslash B_\rho) \times (-4/9, 4/9)$. Hence after passing to a subsequence, we may assume $\widetilde{\phi}_i$ converges to a limit $\widetilde{\phi}$, uniformly in any compact set of $(B_{2/3} \setminus \{0\}) \times (-4/9, 4/9)$.

Because

$$u_i = \phi_i + O\left(\varepsilon_i^{\frac{n-2}{2}} |x|^{2-n}\right),$$

in view of (19.5), u_i/δ_i also converges uniformly to $\widetilde{\phi}$ in any compact set of $(B_{2/3} \setminus \{0\}) \times (-4/9, 4/9)$. As a consequence,

$$\widetilde{\phi} \geq 0 \quad \text{in } \left(B_{2/3} \setminus \{0\}\right) \times (-4/9, 4/9). \tag{19.8}$$

Dividing (1.1) by δ_i, and then letting $i \to +\infty$, by the above uniform convergence of u_i/δ_i, we get

$$\partial_t \widetilde{\phi} - \Delta \widetilde{\phi} = 0 \quad \text{in } \left(B_{2/3} \setminus \{0\}\right) \times (-4/9, 4/9).$$

Since $\left|\widetilde{\phi}\right| \leq 1$ in Q_1, removable singularity theorem for heat equation implies that $\widetilde{\phi}$ is smooth and satisfies the heat equation in $Q_{2/3}$.

By (19.8) and the strong maximum principle, either $\widetilde{\phi} \equiv 0$ or $\widetilde{\phi} > 0$ everywhere in $Q_{2/3}$. We claim that the first case must happen.

Indeed, dividing both sides of (19.1) by δ_i and letting $i \to +\infty$, by (19.5) and the above uniform convergence of $\widetilde{\phi}_i$, we obtain

$$\frac{1}{|\partial B_\rho|} \int_{\partial B_\rho} \widetilde{\phi}(x, 0) \lesssim \rho + K^{-2}.$$

Sending $\rho \to 0$ and $K \to +\infty$ as before, we get $\widetilde{\phi}(0, 0) = 0$. By the strong maximum principle (or Harnack inequality), $\widetilde{\phi} \equiv 0$ in $Q_{2/3}$.

Step 2. Choose a small radius ρ. By results obtained in Step 1, for all i large,

$$|\phi_i| \ll \delta_i \quad \text{on } \partial^p Q_{2/3} \setminus (B_\rho \times \{-2/3\}). \tag{19.9}$$

On $B_\rho \times \{-2/3\}$, we have the trivial bound

$$|\phi_i| \leq \delta_i. \tag{19.10}$$

As in Subsection 17.1, we obtain

Claim. If ρ is sufficiently small, there exists a constant $\sigma(\rho) \ll 1$ (independent of ε_i) such that

$$|\phi_{i,1}| \leq \sigma(\rho)\delta_i \quad \text{in } Q_{7/12}. \tag{19.11}$$

Here $\phi_{i,1}$ denotes the decomposition of the outer component taken in Section 14.

By this claim, repeating the estimates of other terms in the decomposition of the outer component, $\phi_{i,2}$-$\phi_{i,5}$, and the corresponding estimate for the inner component $\phi_{in,i}$ in Section 16, we obtain

$$|\phi_i| \lesssim \sigma(\rho)\delta_i \quad \text{in } Q_{1/2}. \tag{19.12}$$

Step 3. By (19.12), we get

$$u_i \lesssim \varepsilon_i^{\frac{n-2}{2}} + \sigma(\rho)\delta_i \quad \text{in } \left(B_{1/2} \setminus B_{1/4}\right) \times (-1/4, 1/4).$$

Because u_i satisfies the standard parabolic Harnack inequality in $\left(B_1 \setminus B_{1/8}\right) \times (-1, 1)$, we get

$$u_i \lesssim \varepsilon_i^{\frac{n-2}{2}} + \sigma(\rho)\delta_i \quad \text{in } \left(B_1 \setminus B_{1/2}\right) \times (-1/4, 1/8).$$

By the definition of ϕ_i, we get

$$-\varepsilon_i^{\frac{n-2}{2}} \lesssim \phi_i \lesssim \varepsilon_i^{\frac{n-2}{2}} + \sigma(\rho)\delta_i \quad \text{in } \left(B_1 \setminus B_{1/2}\right) \times (-1/4, 1/8).$$

Combining these two estimates with (19.12), for all i sufficiently large, we have

$$\left|\widetilde{\phi_i}\right| \le 1/4 \quad \text{in } B_1 \times (-1/4, 1/8).$$

This is a contradiction with (19.6). $\qquad\square$

20. A CONDITIONAL EXCLUSION OF BUBBLE CLUSTERING

In this section, we work in the following settings. (This will appear in a suitable rescalings of bubble clustering from Theorem 21.1 in Part 3.)

(1) There exist two sequences R_i and T_i, both diverging to $+\infty$ as $i \to +\infty$;
(2) u_i is a sequence of positive, smooth solution of (1.1) in $\Omega_i := B_{R_i} \times (-T_i, T_i)$;
(3) as $i \to +\infty$,

$$\int_{-T_i}^{T_i} \int_{B_{R_i}} \partial_t u_i(x, t)^2 dxdt \to 0; \tag{20.1}$$

(4) there exists an $N \in \mathbb{N}$ such that for each i and $t \in (-T_i, T_i)$, in any compact set of $\mathbb{R}^n$, we have the bubble decomposition

$$u_i(x, t) = \sum_{j=1}^{N} W_{\xi_{ij}^*(t), \lambda_{ij}^*(t)}(x) + o_i(1), \tag{20.2}$$

where $o_i(1)$ are measured in $H_{loc}^1(\mathbb{R}^n)$;
(5) as $i \to +\infty$,

$$\max_{j=1,\cdots,N} \sup_{-T_i < t < T_i} \lambda_{ij}^*(t) \to 0; \tag{20.3}$$

(6) for any $t \in (-T_i, T_i)$,

$$\min_{1 \le j \ne k \le N} |\xi_{ij}^*(t) - \xi_{ik}^*(t)| \ge 2, \tag{20.4}$$

and at $t = 0$,

$$\xi_{i1}^*(0) = 0, \quad |\xi_{i2}^*(0)| = 1; \tag{20.5}$$

(7) after relabelling indices, assume for some $N' \le N$ and any $j = 1, \cdots, N'$,

$$\xi_{ij}^*(0) \to P_j, \quad \text{as } i \to +\infty,$$

while for any $j = N' + 1, \cdots, N$,

$$|\xi_{ij}^*(0)| \to +\infty, \quad \text{as } i \to +\infty,$$

By (20.5), $P_1 = 0$ and $P_2 \in \partial B_1$, so $N' \geq 2$.

The main result of this section is

Proposition 20.1. *Under the above assumptions, we have*

$$T_i \leq 100 \max_{1 \leq j \leq N'} |\log \lambda_{ij}^*(0)|. \tag{20.6}$$

We prove this proposition by contradiction, so assume (20.6) does not hold, that is,

$$T_i > 100 \max_{1 \leq j \leq N'} |\log \lambda_{ij}^*(0)|. \tag{20.7}$$

With this bound in hand, we can iterate Proposition 19.1 backwardly in time, leading to an optimal upper bound on the error function as Schoen's Harnack inequality in Yamabe problem. This bound allows us to define a Green function from u_i. This will be done in Subsection 20.1. Then in Subsection 20.2, still similar to the treatment in Yamabe problem, we employ Pohozaev identity again to give a sign restriction on the next order term in the expansion of this Green function at a pole. This then leads to a contradiction with the assumption that there is one bubble located at P_1 and P_2 respectively.

20.1. **Construction of Green function.** Under the above assumptions, for any i, j and $t \in (-T_i + 1, T_i - 1)$, Proposition 12.1 can be applied to u_i in $Q_1(\xi_{ij}^*(t), t)$. We denote the corresponding parameters by $\xi_{ij}(t)$, $\lambda_{ij}(t)$ and $a_{ij}(t)$, and the error function

$$\phi_{ij}(x, t) := u_i(x, t) - W_{\xi_{ij}(t), \lambda_{ij}(t)}(x) - a_{ij}(t) Z_{0, \xi_{ij}(t), \lambda_{ij}(t)}(x). \tag{20.8}$$

Denote $\varepsilon_{ij} := \lambda_{ij}(0)$, $j = 1, \cdots, N'$ and $\varepsilon_i = \max_{1 \leq j \leq N'} \varepsilon_{ij}$. By Proposition 12.1 and (20.7), we get

$$T_i > 50 \max_{1 \leq j \leq N'} |\log \varepsilon_{ij}|. \tag{20.9}$$

Lemma 20.2. *For each $j = 1, \cdots, N'$ and any $t \in (-50|\log \varepsilon_{ij}|, 50|\log \varepsilon_{ij}|)$,*

$$\frac{1}{2} \varepsilon_{ij} \leq \lambda_{ij}(t) \leq 2 \varepsilon_{ij}. \tag{20.10}$$

Proof. Applying Proposition 17.6 to each ϕ_{ij} in $Q_1(\xi_{ij}(t), t)$, we obtain

$$\left| \lambda'_{ij}(t) \right| \lesssim \lambda_{ij}(t)^{\frac{n-4}{2}}. \tag{20.11}$$

Define

$$\widetilde{T}_{ij} := \sup \left\{ T : T \leq 50|\log \varepsilon_{ij}|, \ \lambda_{ij}(t) \leq 2\varepsilon_{ij} \quad \text{in} \ [0, T] \right\}.$$

Integrating (20.11) on $[0, \widetilde{T}_{ij}]$ leads to

$$\sup_{t \in [0, \widetilde{T}_{ij}]} \lambda_{ij}(t) \leq \varepsilon_{ij} e^{C \varepsilon_{ij}^{(n-6)/2} |\log \varepsilon_{ij}|} \leq \varepsilon_{ij} e^{C \varepsilon_{ij}^{(n-6)/4}} < \frac{3}{2} \varepsilon_{ij}. \tag{20.12}$$

Hence we must have $\widetilde{T}_{ij} = 50|\log \varepsilon_{ij}|$. This gives the upper bound of (20.10) in the positive side. The lower bound and the negative side follow in the same way. $\qquad \square$

For each $t \in (-50|\log \varepsilon_{ij}|, 50|\log \varepsilon_{ij}|)$, applying Proposition 19.1 to $Q_1(\xi_{ij}(t), t)$, we obtain

$$\sup_{Q_1(\xi_{ij}(t),t)} |\phi_{ij}| \leq M_0 \varepsilon_i^{\frac{n-2}{2}} + \frac{1}{2} \sup_{Q_1(\xi_{ij}(t+1),t+1)} |\phi_{ij}|. \tag{20.13}$$

For any $R > 1$ fixed, an iteration of this estimate from $t = 50|\log \varepsilon_{ij}|$ to any $t \in [-R, R]$ leads to

$$\sup_{B_1(\xi_{ij}(0)) \times (-R,R)} |\phi_{ij}| \lesssim \varepsilon_i^{\frac{n-2}{2}}. \tag{20.14}$$

This bound depends only on the constant M_0 in Proposition 19.1, and it is independent of R.

Substituting (20.14) and estimates of parameters in Proposition 17.6 into (20.8), we get

$$\sup_{B_1(\xi_{ij}(0)) \times (-R,R)} u_i \lesssim \varepsilon_i^{\frac{n-2}{2}}.$$

In $\left(B_{7/4}(\xi_{ij}(0)) \setminus B_{3/4}(\xi_{ij}(0))\right) \times (-R, R)$, by applying standard Harnack inequality to u_i, (20.14) is extended to

$$\sup_{B_{3/2}(\xi_{ij}(0)) \times (-R,R)} |\phi_{ij}| \lesssim \varepsilon_i^{\frac{n-2}{2}}. \tag{20.15}$$

This estimate gives us the expansion

$$u_i(x,t) = \varepsilon_{ij}^{-\frac{n-2}{2}} W\left(\frac{x - \xi_{ij}(t)}{\varepsilon_{ij}}\right) + O\left(\varepsilon_i^{\frac{n-2}{2}}\right) \quad \text{in } B_{3/2}(\xi_{ij}(0)) \times (-R, R). \tag{20.16}$$

Here we note that, by Proposition 17.6,

$$|\xi'_{ij}(t)| \lesssim \lambda_{ij}(t)^{\frac{n-4}{2}} \lesssim \varepsilon_i^{\frac{n-4}{2}}.$$

Hence

$$|\xi_{ij}(t) - \xi_{ij}(0)| \lesssim_R \varepsilon_i^{\frac{n-4}{2}}, \quad \text{in } [-R, R]. \tag{20.17}$$

Assume

$$\lim_{i \to +\infty} \frac{\varepsilon_{ij}}{\varepsilon_i} = m_j \in [0, 1] \quad \text{for any } j = 1, \cdots, N'.$$

By the definition of ε_i, there exists at least one j satisfying $m_j = 1$.

It is directly verified that for some dimensional constant $c(n) > 0$,

$$\varepsilon_i^{-\frac{n-2}{2}} u_i^p dx dt \rightharpoonup c(n) m_j^{\frac{n-2}{2}} \delta_{P_j} \otimes dt \tag{20.18}$$

weakly as Radon measures in $B_{3/2}(P_j) \times (-R, R)$.

Set

$$\widehat{u}_i := \varepsilon_i^{-\frac{n-2}{2}} u_i.$$

It satisfies

$$\partial_t \widehat{u}_i - \Delta \widehat{u}_i = \varepsilon_i^{-\frac{n-2}{2}} u_i^p = \varepsilon_i^2 \widehat{u}_i^p. \tag{20.19}$$

By (20.15), in any compact set of $\left(B_{3/2}(P_j) \setminus \{P_j\}\right) \times \mathbb{R}$, $\widehat{u}_i$ are uniformly bounded. Because u_i satisfies the standard parabolic Harnack inequality in any compact set

of $\left(\mathbb{R}^n \setminus \cup_{j=1}^{N'}\{P_j\}\right) \times \mathbb{R}$, we deduce that $\widehat{u}_i$ are also uniformly bounded in any compact set of $\left(\mathbb{R}^n \setminus \cup_{j=1}^{N'}\{P_j\}\right) \times \mathbb{R}$. Then by standard parabolic regularity theory, $\widehat{u}_i$ converges to $\widehat{u}_\infty$ smoothly in any compact set of $\left(\mathbb{R}^n \setminus \cup_{j=1}^{N'}\{P_j\}\right) \times \mathbb{R}$.

By (20.18), $\widehat{u}_\infty$ satisfies

$$\partial_t \widehat{u}_\infty - \Delta \widehat{u}_\infty = c(n) \sum_{j=1}^{N'} m_j^{\frac{n-2}{2}} \delta_{P_j} \otimes dt \quad \text{in } \mathbb{R}^n \times \mathbb{R}. \tag{20.20}$$

Lemma 20.3. *For each* $j = 1 \cdots, N'$,

$$m_j > 0.$$

Proof. Because there is one m_j satisfying $m_j = 1$, $\widehat{u}_\infty \neq 0$. By the strong maximum principle,

$$\widehat{u}_\infty > 0 \quad \text{in } \left(\mathbb{R}^n \setminus \cup_{j=1}^{N'}\{P_j\}\right) \times \mathbb{R}. \tag{20.21}$$

By (20.16), for each $j = 1, \cdots, N'$,

$$\widehat{u}_i \lesssim \left(\frac{\varepsilon_{ij}}{\varepsilon_i}\right)^{\frac{n-2}{2}} \quad \text{in } \left(B_{3/2}(P_j) \setminus B_{1/2}(P_j)\right) \times (-1, 1).$$

If $m_j = 0$, we would have

$$\widehat{u}_\infty = \lim_{i \to +\infty} \widehat{u}_i = 0 \quad \text{in } \left(B_{3/2}(P_j) \setminus B_{1/2}(P_j)\right) \times (-1, 1).$$

This is a contradiction with (20.21). $\qquad\square$

Lemma 20.4. *For any* $(x, t) \in \left(\mathbb{R}^n \setminus \cup_{j=1}^{N'}\{P_j\}\right) \times \mathbb{R}$,

$$\widehat{u}_\infty(x, t) \equiv c(n) \sum_{j=1}^{N'} m_j^{\frac{n-2}{2}} |x - P_j|^{2-n}. \tag{20.22}$$

Proof. Take two arbitrary $s < t$. Take a sequence of cut-off functions $\zeta_R \in C_0^\infty(\mathbb{R}^n)$ such that

$$\begin{cases} 0 \leq \zeta_R \leq 1, \\ \zeta_R \equiv 1 \quad \text{in } B_R \setminus \cup_{j=1}^{N'} B_{1/R}(P_j), \\ \zeta_R \equiv 0 \quad \text{in } B_{2R}^c \cup \cup_{j=1}^{N'} B_{1/2R}(P_j). \end{cases}$$

By the comparison principle, for any $x \in \mathbb{R}^n \setminus \{P_1, \cdots, P_{N'}\}$,

$$\begin{aligned} \widehat{u}_\infty(x, t) \geq\ & \int_{\mathbb{R}^n} \widehat{u}_\infty(y, s) \zeta_R(y) G(x - y, t - s) dy \\ &+ c(n) \sum_{j=1}^{N'} m_j^{\frac{n-2}{2}} \int_s^t G(x - P_j, t - s) ds. \end{aligned}$$

Here G denotes the standard heat kernel on $\mathbb{R}^n$.

Letting $R \to +\infty$, by the monotone convergence theorem we obtain

$$\widehat{u}_\infty(x, t) \geq \int_{\mathbb{R}^n} \widehat{u}_\infty(y, s) G(x - y, t - s) dy + c(n) \sum_{j=1}^{N'} m_j^{\frac{n-2}{2}} \int_s^t G(x - P_j, t - s) ds.$$

Letting $s \to -\infty$, we get

$$\widehat{u}_\infty(x,t) \geq c(n) \sum_{j=1}^{N'} m_j^{\frac{n-2}{2}} |x - P_j|^{2-n}.$$

Thus their difference

$$\widehat{v}_\infty(x,t) := \widehat{u}_\infty(x,t) - c(n) \sum_{j=1}^{N'} m_j^{\frac{n-2}{2}} |x - P_j|^{2-n}$$

is a positive caloric function on $\mathbb{R}^n \times \mathbb{R}$. By [51] or [92], there exists a nonnegative Radon measure ν on $\{(\xi, \lambda) : \lambda = |\xi|^2\} \subset \mathbb{R}^n \times \mathbb{R}$ such that

$$\widehat{v}_\infty(x,t) = \int_{\{\lambda = |\xi|^2\}} e^{\lambda t + \xi \cdot x} d\nu(\xi, \lambda).$$

Because (20.16) holds for any R, $\widehat{u}_\infty$ is bounded in $\left(B_1(P_1) \setminus B_{1/2}(P_1)\right) \times \mathbb{R}$, so $\widehat{v}_\infty$ is also bounded $\left(B_1(P_1) \setminus B_{1/2}(P_1)\right) \times \mathbb{R}$. This is possible only if $\nu = 0$, and (20.22) follows. $\qquad\square$

By this lemam, near $P_1 = 0$, there exists a smooth harmonic function h such that

$$\widehat{u}_\infty(x) = c(n)m_1|x|^{2-n} + h(x).$$

Because $P_2 \in \partial B_1$, we deduce that

$$h(0) > 0. \tag{20.23}$$

20.2. **Local Pohozaev invariants.** By the expansion of $\widehat{u}_\infty(x)$ near 0, in particular, (20.23), and a direct calculation, we find a dimensional constant $c(n) > 0$ such that for all r small,

$$\mathcal{P}_{\widehat{u}_\infty}(r) = \frac{c(n)m_1 h(0)}{r} + O(1). \tag{20.24}$$

On the other hand, because $\widehat{u}_\infty$ comes from u_i, we claim that

Lemma 20.5. *For all r small,*

$$\mathcal{P}_{\widehat{u}_\infty}(r) \lesssim K^{-2} + r. \tag{20.25}$$

Proof. In (18.1), choose r_i to be a fixed, small $r > 0$. Multiply both sides of (18.1) by ε_i^{2-n}. Let us analyse the convergence of both sides.

Step 1. The left hand side of (18.1).

By the convergence of $\widehat{u}_i$ in $C_{loc}^\infty\left((\mathbb{R}^n \setminus \cup_{j=1}^{N'}\{P_j\}) \times \mathbb{R}\right)$, we get

$$\varepsilon_i^{2-n}\mathcal{P}_{u_i(0)}(r) = \mathcal{P}_{\widehat{u}_i(0)}(r) \to \mathcal{P}_{\widehat{u}_\infty(0)}(r), \quad \text{as } i \to +\infty \tag{20.26}$$

and

$$\varepsilon_i^{2-n} \int_{\partial B_r} u_i^{p+1} \lesssim \varepsilon_i^2 \to 0, \quad \text{as } i \to +\infty. \tag{20.27}$$

Hence the left hand side converges to $\mathcal{P}_{\widehat{u}_\infty}(r)$.

Step 2. The right hand side of (18.1).

For simplicity of notations, denote the parameters $(a_{i1}, \xi_{i1}, \lambda_{i1})$ just by (a_i, ξ_i, λ_i), and ϕ_{i1}, the error function in Q_1, just by ϕ_i. Plugging the L^∞ estimate (20.14) into Proposition 17.6, we get the following estimates:

$$\sup_{t \in (-1,1)} |a_i(t)| \lesssim \varepsilon_i^{n-1}, \tag{20.28}$$

$$\sup_{t \in (-1,1)} \left(\left| a_i'(t) - \mu_0 \frac{a_i(t)}{\lambda_i(t)^2} \right| + |\xi_i'(t)| + |\lambda_i'(t)| \right) \lesssim K^{-\frac{n-2}{2}} \varepsilon_i^{n-3}, \tag{20.29}$$

$$|\nabla \phi_i(x,t)| \lesssim \frac{\varepsilon_i^{(n+2)/2}}{\varepsilon_i^3 + |x|^3} + \varepsilon_i^{\frac{n-2}{2}} \quad \text{for any } (x,t) \in Q_1, \tag{20.30}$$

and

$$|\partial_t \phi_i(x,t)| \lesssim \frac{\varepsilon_i^{(n+2)/2}}{\varepsilon_i^4 + |x|^4} + \frac{\varepsilon_i^{(n-2)/2}}{\varepsilon_i + |x|} \quad \text{for any } (x,t) \in Q_1. \tag{20.31}$$

The factor $K^{-(n-2)/2}$ in (20.29) comes from Lemma 13.5.

As in Subsection 18.2, we expand the right hand side of (18.1) into nine terms, I-IX. Let us estimate them one by one.

(1) By (20.29) and the orthogonal relation between Z_0 and Z_{n+1},

$$\left| \lambda_i(0) a_i'(0) \int_{B_r} Z_{0,i}(x,0) Z_{n+1,i}(x,0) dx \right| \lesssim \varepsilon_i^{n-2} e^{-cr/\varepsilon_i}.$$

Similarly, for each $j = 1, \cdots, n$,

$$\left| \lambda_i(0) \xi_i'(0) \int_{B_r} Z_{j,i}(x,0) Z_{n+1,i}(x,0) dx \right| \lesssim \varepsilon_i^{2n-5} r^{3-n},$$

and

$$\left| \lambda_i(0) \lambda_i'(0) \int_{B_r} Z_{n+1,i}(0)^2 \right| \lesssim K^{-\frac{n-2}{2}} \varepsilon_i^{n-2}.$$

Combining these three estimates, we obtain

$$|\mathrm{I}| \lesssim \left[K^{-\frac{n-2}{2}} + \left(\frac{\varepsilon_i}{r} \right)^{n-3} + e^{-cr/\varepsilon_i} \right] \varepsilon_i^{n-2}.$$

(2) By (20.14), (20.29) and (20.30),

$$|\mathrm{II}| \lesssim \varepsilon_i^{2n-6} r^2.$$

(3) By (20.28) and (20.29),

$$|\mathrm{III}| \lesssim \varepsilon_i^{2n-4}.$$

(4) As in the treatment of IV in Subsection 18.2, we get

$$|\mathrm{IV}| \lesssim K^2 \varepsilon_i^{\frac{3n}{2}-3} + \left(K^{-2} + r \right) \varepsilon_i^{n-2}.$$

(5) By (20.14), (20.28) and (20.29),

$$|\mathrm{V}| \lesssim \varepsilon_i^n r^{n-4} + \varepsilon_i^{n-2} r^{n-1}.$$

(6) By (20.28) and (20.31),

$$|\mathrm{VI}| \lesssim \varepsilon_i^{2n-4}.$$

(7) By (20.28) and (20.29),
$$|\text{VII}| \lesssim \varepsilon_i^{2n-4}.$$

(8) By (20.28), (20.29) and (20.30),
$$|\text{VIII}| \lesssim \varepsilon_i^{3n-6}.$$

(9) By (20.28) and (20.29)
$$|\text{IX}| \lesssim \varepsilon_i^{3n-6}.$$

Putting these estimates together and using (20.26), we obtain (20.25). $\qquad\square$

Combining this lemma with (20.24), after letting $r \to 0$, we deduce that $h(0) = 0$. This is a contradiction with (20.23), so (20.7) cannot be true. The proof of Proposition 20.1 is thus complete.

A. Linearization estimates around bubbles

In this appendix, we collect some estimates about the linearized equation around the standard bubble. This is mainly used to study the inner problem in Section 13.

Recall that the standard positive bubble is

$$W(x) := \left(1 + \frac{|x|^2}{n(n-2)}\right)^{-\frac{n-2}{2}}.$$

Concerning eigenvalues and eigenfunctions for the linearized operator $-\Delta - pW^{p-1}$, we have (see for example [11, Proposition 2.2])

Theorem A.1.   (i) *There exists one and only one negative eigenvalue for $-\Delta - pW^{p-1}$, denoted by $-\mu_0$, for which there exists a unique (up to a constant), positive, radially symmetric and exponentially decaying eigenfunction Z_0.*

(ii) *There exist exactly $(n+1)$-eigenfunctions Z_i in $L^\infty(\mathbb{R}^n)$ corresponding to eigenvalue 0, given by*

$$\begin{cases} Z_i = \dfrac{\partial W}{\partial x_i}, & i = 1, \cdots, n, \\[2mm] Z_{n+1} = \dfrac{n-2}{2} W + x \cdot \nabla W. \end{cases}$$

Remark A.2. *The following decay as $|x| \to \infty$ holds for these eigenfunctions:*

$$\begin{cases} Z_0(x) \lesssim e^{-c|x|}, \\ |Z_i(x)| \lesssim |x|^{1-n}, & i = 1, \cdots, n, \\ |Z_{n+1}(x)| \lesssim |x|^{2-n}. \end{cases}$$

Throughout the paper Z_0 is normalized so that

$$\int_{\mathbb{R}^n} Z_0^2 = 1.$$

For any $\xi \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^+$, in accordance with the scalings for W, define

$$Z_{i,\xi,\lambda}(x) := \lambda^{-\frac{n}{2}} Z_i\left(\frac{x-\xi}{\lambda}\right).$$

This scaling preserves the $L^2(\mathbb{R}^n)$ norm of Z_i.

Next we study the linearized parabolic equation around W. The first one is a nondegeneracy result.

Proposition A.3 (Nondegeneracy). *Suppose $\alpha > 2$ and $\varphi \in L^\infty(-\infty, 0; \mathcal{X}_\alpha)$ is a solution of*

$$\partial_t \varphi - \Delta \varphi = pW^{p-1}\varphi, \quad in \ \mathbb{R}^n \times \mathbb{R}^-. \tag{A.1}$$

Then there exist $(N + 2)$ constants $c_0, \cdots, c_{n+1}$ such that

$$\varphi(x, t) \equiv c_0 e^{\mu_0 t} Z_0(x) + \sum_{i=1}^{n+1} c_i Z_i(x).$$

Proof. By standard parabolic regularity theory and the $O(|x|^{-4})$ decay of W^{p-1} at infinity, we deduce that $\varphi \in L^\infty(-\infty, 0; \mathcal{X}_\alpha^{2+\theta})$ and $\partial_t \varphi \in L^\infty(-\infty, 0; \mathcal{X}_{2+\alpha})$. Because $\alpha > 2$, we may assume

$$\int_{\mathbb{R}^n} \varphi(x, t) Z_i(x) dx \equiv 0, \quad \forall t < 0, \quad i = 0, \cdots, n+1.$$

Then $\partial_t \varphi$ satisfy these orthogonal conditions, too. Since it is also a solution of (A.1), by the method in [11, Section 2] (in particular, the coercivity estimate in [11, Lemma 2.3]), we deduce that $\partial_t \varphi \equiv 0$. Therefore φ is a stationary solution of (A.1). Since $\varphi \in \mathcal{X}_\alpha$ and it is orthogonal to Z_i, by Theorem A.1, $\varphi \equiv 0$. $\square$

Now we state two a priori estimates for this linearized equation.

Lemma A.4. *Suppose $\alpha' > \max\{2, n/2\}$, $\varphi_0 \in \mathcal{X}_{\alpha'}$. If φ_0 satisfies the orthogonal condition*

$$\int_{\mathbb{R}^n} \varphi_0(x) Z_i(x) dx = 0, \quad \forall i = 0, \cdots, n+1, \tag{A.2}$$

then there exists a unique $\varphi \in L^\infty_{loc}(\mathbb{R}^+; \mathcal{X}_{\alpha'})$ solving the equation

$$\begin{cases} \partial_t \varphi - \Delta \varphi = pW^{p-1}\varphi & in \ \mathbb{R}^n \times (0, 2T_D), \\ \varphi(x, 0) = \varphi_0(x) & on \ \mathbb{R}^n, \end{cases} \tag{A.3}$$

Moreover, φ satisfies the orthogonal condition

$$\int_{\mathbb{R}^n} \varphi(x, t) Z_i(x) dx = 0, \quad \forall t > 0, \quad \forall i = 0, \cdots, n+1. \tag{A.4}$$

and the decay property

$$\varphi(\cdot, \cdot + t) \to 0 \quad in \ C_{loc}^{(1+\theta)/2}(\mathbb{R}; C_{loc}^{1,\theta}(\mathbb{R}^n)) \ as \ t \to +\infty. \tag{A.5}$$

Proof. Existence of a global solution φ to the problem (A.3) follows from standard parabolic theory, if we note that the heat semigroup $e^{t\Delta}$ is uniformly bounded as an operator from $\mathcal{X}_{\alpha'}$ into itself.

By standard parabolic regularity theory, $\partial_t \varphi$, $\nabla \varphi$ and $\Delta \varphi$ belong to $L^\infty_{loc}(\mathbb{R}^+; \mathcal{X}_{\alpha'})$. The orthogonal condition (A.4) then follows by testing (A.3) with Z_i.

Testing (A.3) with φ and then applying Theorem A.1, we obtain

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^n}\varphi(x,t)^2 dx = -\int_{\mathbb{R}^n}\left[|\nabla\varphi(x,t)|^2 - pW(x)^{p-1}\varphi(x,t)^2\right]dx \le 0. \qquad \text{(A.6)}$$

Testing (A.3) with $\partial_t\varphi$, we also obtain

$$\frac{1}{2}\frac{d}{dt}\int_{\mathbb{R}^n}\left[|\nabla\varphi(x,t)|^2 - pW(x)^{p-1}\varphi(x,t)^2\right]dx = -\int_{\mathbb{R}^n}\partial_t\varphi(x,t)^2 dx \le 0. \qquad \text{(A.7)}$$

Combining these two identities with Proposition A.3, we deduce that

$$\lim_{t\to+\infty}\|\varphi(t)\|_{H^1(\mathbb{R}^n)} = 0.$$

Then (A.5) follows by an application of standard parabolic regularity theory. $\qquad\square$

The second one is inspired by [73, Lemma 5.1].

Lemma A.5. *If $\alpha > 2$, there exists a positive constant $C(\alpha)$ so that the following holds. Given $T > 1$, assume $\varphi \in L^\infty(0,T;\mathcal{X}_\alpha)$, $a_i \in L^\infty([0,T])$ and $E \in L^\infty(0,T;\mathcal{X}_{2+\alpha})$ solve the problem*

$$\begin{cases} \partial_t\varphi - \Delta\varphi = pW^{p-1}\varphi + \displaystyle\sum_{i=0}^{n+1}a_i Z_i + E & \text{in } \mathbb{R}^n \times (0,T), \\ \varphi(x,0) = 0, \end{cases} \qquad \text{(A.8)}$$

and φ satisfies the orthogonal condition

$$\int_{\mathbb{R}^n}\varphi(x,t)Z_i(x)dx = 0, \quad \forall t \in [0,T], \quad \forall i = 0,\cdots,n+1. \qquad \text{(A.9)}$$

Then

$$\begin{cases} \displaystyle\sum_{i=0}^{n+1}|a_i(t)| \le C\|E(t)\|_{2+\alpha}, \quad \forall t \in [0,T], \\ \|\varphi\|_{C^{(1+\theta)/2}\left(0,T;\mathcal{X}_\alpha^{1+\theta}\right)} \le C(\alpha)\|E\|_{L^\infty(0,T;\mathcal{X}_{2+\alpha})}. \end{cases}$$

Proof. First, for each i, multiplying (A.8) by Z_i and integrating on $\mathbb{R}^n$, we obtain

$$a_i(t) = -\frac{\int_{\mathbb{R}^n}E(x,t)Z_i(x)}{\int_{\mathbb{R}^n}Z_i(x)^2}. \qquad \text{(A.10)}$$

Because $\alpha > 2$, this gives the estimate on $a_i(t)$.

For the second estimate, by standard parabolic regularity theory, it suffices to prove

$$\|\varphi\|_{L^\infty(0,T;\mathcal{X}_\alpha)} \le C(\alpha)\|E\|_{L^\infty(0,T;\mathcal{X}_{2+\alpha})}. \qquad \text{(A.11)}$$

We will argue by contradiction. Assume there exists a sequence of $T_k > 1$, a sequence of $\varphi_k \in L^\infty(0,T_k;\mathcal{X}_\alpha)$, $a_{k,i} \in L^\infty([0,T_k])$ and $E_k \in L^\infty(0,2T_k;\mathcal{X}_{2+\alpha})$, satisfying (A.8) and (A.9), with

$$\|\varphi_k\|_{L^\infty(0,T_k;\mathcal{X}_\alpha)} = 1 \qquad \text{(A.12)}$$

but

$$\|E_k\|_{L^\infty(0,2T_k;\mathcal{X}_{2+\alpha})} \le \frac{1}{k}. \qquad \text{(A.13)}$$

Take a $t_k \in (0, T_k)$ with $\|\varphi_k(t_k)\|_\alpha \geq 1/2$ and a point $x_k \in \mathbb{R}^n$ such that

$$(1 + |x_k|)^\alpha |\varphi_k(x_k, t_k)| \geq 1/4. \tag{A.14}$$

We claim that

Sub-Lemma. $\limsup_{k \to \infty} |x_k| < +\infty$.

Proof. By the heat kernel representation,

$$\varphi(x, t_k) \;=\; \int_0^{t_k} \int_{\mathbb{R}^n} [4\pi(t_k - t)]^{-\frac{n}{2}} e^{-\frac{|x-y|^2}{4(t_k - t)}}$$

$$\times \left[pW(y)^{p-1}\varphi_k(y, t) + \sum_{i=0}^{n+1} a_{k,i}(t) Z_i(y) + E_k(y, t) \right] dy dt.$$

Following the calculation in [89, Appendix B], we get

$$\left| \int_0^{t_k} [4\pi(t_k - t)]^{-\frac{n}{2}} e^{-\frac{|x-y|^2}{4(t_k - t)}} pW(y)^{p-1}\varphi_k(y, t) dy dt \right|$$

$$\lesssim \int_{\mathbb{R}^n} |x - y|^{2-n} (1 + |y|)^{-4-\alpha}$$

$$\lesssim (1 + |x|)^{-2-\alpha}$$

and

$$\left| \int_0^{t_k} [4\pi(t_k - t)]^{-\frac{n}{2}} e^{-\frac{|x-y|^2}{4(t_k - t)}} \left[\sum_{i=0}^{n+1} a_{k,i}(t) Z_i(y) + E_k(y, t) \right] dy dt \right|$$

$$\lesssim \frac{1}{k} \int_{\mathbb{R}^n} |x - y|^{2-n} (1 + |y|)^{-2-\alpha}$$

$$\lesssim \frac{1}{k} (1 + |x|)^{-\alpha}.$$

Putting these two estimates together we get

$$|\varphi_k(x_k, t_k)| \leq C (1 + |x_k|)^{-2-\alpha} + \frac{1}{k} (1 + |x_k|)^{-\alpha}.$$

Combining this inequality with (A.14) we finish the proof of this sub-lemma. $\qquad\square$

Next we divide the proof into two cases.
Case 1. $t_k \to +\infty$.
Let

$$\widetilde{\varphi}_k(x, t) := \varphi_k(x, t_k + t).$$

An application of standard parabolic regularity theory shows that $\widetilde{\varphi}_k$ are uniformly bounded in $C_{loc}^{1+\theta,(1+\theta)/2}(\mathbb{R}^n \times (-\infty, 0])$. After passing to a subsequence, $\widetilde{\varphi}_k$ converges to a limit $\widetilde{\varphi}_\infty$, which satisfies the following conditions.

- Passing to the limit in (A.12) gives

$$\|\widetilde{\varphi}_\infty\|_{L^\infty(-\infty,0;\mathcal{X}_\alpha)} \leq 1. \tag{A.15}$$

- Since $\alpha > 2$, we can pass to the limit in (A.9) for φ_k, obtaining

$$\int_{\mathbb{R}^n} \widetilde{\varphi}_\infty(x,t) Z_i(x) dx = 0, \qquad \text{for any } i = 0, \cdots, n+1, \quad t \le 0. \tag{A.16}$$

- Passing to the limit in (A.8) for $\widetilde{\varphi}_k$ and noting (A.13) as well as the estimate on $a_{k,i}$, we see $\widetilde{\varphi}_\infty$ is a solution of (A.1).

By Proposition A.3, these three conditions imply that $\widetilde{\varphi}_\infty \equiv 0$.

On the other hand, passing to the limit in (A.14) leads to

$$(1 + |x_\infty|)^\alpha \, |\widetilde{\varphi}_\infty(x_\infty, 0)| \ge 1/4.$$

This is a contradiction with the fact that $\widetilde{\varphi}_\infty \equiv 0$.

Case 2. $t_k \to t_\infty$.

As in the previous case, now φ_k itself converges to a limit φ_∞, which solves the equation

$$\begin{cases} \partial_t \varphi_\infty - \Delta \varphi_\infty = pW^{p-1}\varphi_\infty & \text{in } \mathbb{R}^n \times (0, t_\infty), \\ \varphi_\infty(x, 0) = 0. \end{cases}$$

Passing to the limit in (A.12) still gives

$$\|\varphi_\infty\|_{L^\infty(0,t_\infty;\mathcal{X}_\alpha)} \le 1. \tag{A.17}$$

By standard parabolic theory, we also get $\varphi_\infty \equiv 0$. Then we get the same contradiction as in Case 1. $\qquad\square$

B. Estimates on some integrals

In this appendix we give some technical integral estimates involving the heat kernel associated to the outer equation in Section 14.

Lemma B.1. *Assume* $0 < \nu < n$, $0 \le \gamma < n - \nu$, $t > s$. *Then*

$$\int_{\mathbb{R}^n} (t-s)^{-\frac{n}{2}} e^{-c\frac{|x-y|^2}{t-s}} \left(1 + \frac{\sqrt{t-s}}{|y|}\right)^\gamma |y|^{-\nu} dy \lesssim \left(|x| + \sqrt{t-s}\right)^{-\nu} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^\gamma.$$

Proof. After a change of variables $\hat{x} := x/\sqrt{t-s}$, $\hat{y} := y/\sqrt{t-s}$, this integral is transformed into

$$(t-s)^{-\nu/2} \int_{\mathbb{R}^n} e^{-c|\hat{x}-\hat{y}|^2} \left(1 + \frac{1}{|\hat{y}|}\right)^\gamma |\hat{y}|^{-\nu} d\hat{y}.$$

To estimate the integral in this formula, we consider two cases separately.

Case 1. If $|\hat{x}| \le 1$, this integral is bounded by a universal constant.

Case 2. If $|\hat{x}| \ge 1$, we divide this integral into two parts, I outside $B_{|\hat{x}|/3}(0)$, and II in $B_{|\hat{x}|/3}(0)$.

Outside $B_{|\hat{x}|/3}(0)$,

$$\left(1 + \frac{1}{|\hat{y}|}\right)^\gamma |\hat{y}|^{-\nu} \lesssim \left(1 + \frac{1}{|\hat{x}|}\right)^\gamma |\hat{x}|^{-\nu}.$$

Hence

$$
\begin{aligned}
\text{I} &\lesssim \left(1 + \frac{1}{|\hat{x}|}\right)^{\gamma} |\hat{x}|^{-\nu} \int_{B_{|\hat{x}|/3}(0)^c} e^{-c|\hat{x}-\hat{y}|^2} d\hat{y} \\
&\lesssim (t-s)^{\nu/2} |x|^{-\nu} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{\gamma}.
\end{aligned}
$$

For II, by noting that for $\hat{y} \in B_{|\hat{x}|/3}(0)$,

$$
e^{-c|\hat{x}-\hat{y}|^2} \lesssim e^{-c|\hat{x}|^2/9}.
$$

we obtain

$$
\begin{aligned}
\text{II} &\lesssim e^{-c|\hat{x}|^2/9} \int_{B_{|\hat{x}|/3}(0)} |\hat{y}|^{-\nu-\gamma} d\hat{y} \\
&\lesssim e^{-c\frac{|x|^2}{t-s}} \\
&\lesssim (t-s)^{\nu/2} |x|^{-\nu}.
\end{aligned}
$$

Adding I and II together, we get

$$
\int_{\mathbb{R}^n} e^{-c|\hat{x}-\hat{y}|^2} \left(1 + \frac{1}{|\hat{y}|}\right)^{\gamma} |\hat{y}|^{-\nu} d\hat{y} \lesssim (t-s)^{\nu/2} |x|^{-\nu} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{\gamma}. \qquad \square
$$

Lemma B.2. *Assume $t > s$, c_1, c_2, L and λ are four positive constants. Then*

$$
\int_{B_{L\lambda}^c} (t-s)^{-\frac{n}{2}} e^{-c_1 \frac{|x-y|^2}{t-s}} \left(1 + \frac{\sqrt{t-s}}{|y|}\right)^{\gamma} e^{-c_2 \frac{|y|}{\lambda}} dy
$$

$$
\lesssim e^{-\frac{c_2}{3} \frac{|x|}{\lambda}} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{\gamma} + e^{-\frac{c_2 L}{2} - \frac{c_1}{9} \frac{|x|^2}{t-s}} \left(\frac{\lambda}{\sqrt{t-s}}\right)^{n-\gamma} \left(1 + \frac{\lambda}{\sqrt{t-s}}\right)^{\gamma}.
$$

Proof. After a change of variables $\hat{x} := x/\sqrt{t-s}$, $\hat{y} := y/\sqrt{t-s}$, this integral is transformed into

$$
\int_{B_{L\lambda/\sqrt{t-s}}^c} e^{-c_1|\hat{x}-\hat{y}|^2} \left(1 + \frac{1}{|\hat{y}|}\right)^{\gamma} e^{-c_2 \frac{\sqrt{t-s}}{\lambda}|\hat{y}|} d\hat{y}.
$$

We divide this integral into two parts, I outside $B_{|\hat{x}|/3}(0)$, and II in $B_{|\hat{x}|/3}(0)$.

Outside $B_{|\hat{x}|/3}(0)$,

$$
\begin{cases}
e^{-c_2 \frac{\sqrt{t-s}}{\lambda}|\hat{y}|} \leq e^{-\frac{c_2}{3} \frac{\sqrt{t-s}}{\lambda}|\hat{x}|}, \\
\left(1 + \frac{1}{|\hat{y}|}\right)^{\gamma} \lesssim \left(1 + \frac{1}{|\hat{x}|}\right)^{\gamma}.
\end{cases}
$$

Therefore

$$
\begin{aligned}
\text{I} &\lesssim e^{-\frac{c_2}{3} \frac{\sqrt{t-s}}{\lambda}|\hat{x}|} \left(1 + \frac{1}{|\hat{x}|}\right)^{\gamma} \int_{B_{|\hat{x}|/3}(0)^c} e^{-c_1|\hat{x}-\hat{y}|^2} d\hat{y} \\
&\lesssim e^{-\frac{c_2}{3} \frac{|x|}{\lambda}} \left(1 + \frac{\sqrt{t-s}}{|x|}\right)^{\gamma}.
\end{aligned}
$$

In $B_{|\hat{x}|/3}(0)$,

$$e^{-c_1|\hat{x}-\hat{y}|^2} \leq e^{-c_1|\hat{x}|^2/9}.$$

Hence

$$
\begin{aligned}
\mathrm{II} \;&\lesssim\; e^{-c_1|\hat{x}|^2/9} \int_{B_{|\hat{x}|/3}\setminus B_{L\lambda/\sqrt{t-s}}} \left(1 + \frac{1}{|\hat{y}|}\right)^\gamma e^{-c_2 \frac{\sqrt{t-s}}{\lambda}|\hat{y}|} d\hat{y} \\
&\lesssim\; e^{-\frac{c_2 L}{2} - \frac{c_1}{9}\frac{|x|^2}{t-s}} \left(\frac{\lambda}{\sqrt{t-s}}\right)^n \left(1 + \frac{\sqrt{t-s}}{\lambda}\right)^\gamma \\
&\lesssim\; e^{-\frac{c_2 L}{2} - \frac{c_1}{9}\frac{|x|^2}{t-s}} \left(\frac{\lambda}{\sqrt{t-s}}\right)^n \left(1 + \frac{\lambda}{\sqrt{t-s}}\right)^\gamma. \quad \square
\end{aligned}
$$

Part 3. **Energy concentration in the general case**

21. SETTING

In this part, we still consider a sequence of *smooth, positive solutions* u_i to the nonlinear heat equation (1.1) (with $p = (n+2)/(n-2)$) in Q_1, but now satisfying the following three assumptions.

(IIII.a) Weak limit: u_i converges weakly to u_∞ in $L^{p+1}(Q_1)$, and ∇u_i converges weakly to ∇u_∞ in $L^2(Q_1)$. Here u_∞ is a *smooth solution* of (1.1) in Q_1.

(IIII.b) Energy concentration behavior: there exists an $N \geq 2$ such that

$$\begin{cases} |\nabla u_i|^2 dx dt \rightharpoonup |\nabla u_\infty|^2 dx dt + N\Lambda \delta_0 \otimes dt, \\ u_i^{p+1} dx dt \rightharpoonup u_\infty^{p+1} dx dt + N\Lambda \delta_0 \otimes dt, \end{cases}$$

weakly as Radon measures.

(IIII.c) Convergence of time derivatives: as $i \to \infty$, $\partial_t u_i$ converges to $\partial_t u_\infty$ strongly in $L^2(Q_1)$.

The only difference with the assumptions in Part 2 is **(IIII.b)**, where now we do not assume there is only one bubble, in other words, now we are in *the higher multiplicity case* instead of *the multiplicity one case*.

The main result in this part is

Theorem 21.1. *After passing to a subsequence, the followings hold for u_i.*

(1) *For all i and any $t \in [-9/16, 9/16]$, there exist exactly N local maximal point of $u_i(\cdot, t)$ in the interior of $B_1(0)$.*

Denote these points by $\xi_{ij}^*(t)$ ($j = 1, \cdots, N$) and let $\lambda_{ij}^*(t) := u_i(\xi_{ij}^*(t), t)^{-\frac{2}{n-2}}$.

(2) *Both λ_{ij}^* and ξ_{ij}^* are continuous functions on $[-9/16, 9/16]$.*

(3) *There exists a constant C_2 such that, for all i and any $x \in B_1$,*

$$\min_{1 \leq j \leq N} |x - \xi_{ij}^*(t)|^{\frac{n-2}{2}} u_i(x, t) + \min_{1 \leq j \leq N} |x - \xi_{ij}^*(t)|^{\frac{n}{2}} |\nabla u_i(x, t)| \qquad (21.1)$$

$$+ \min_{1 \leq j \leq N} |x - \xi_{ij}^*(t)|^{\frac{n+2}{2}} \left(|\nabla^2 u_i(x, t)| + |\partial_t u_i(x, t)| \right) \leq C_2.$$

(4) *For each $j = 1, \cdots, N$, as $i \to \infty$,*

$$\lambda_{ij}^*(t) \to 0, \quad \xi_{ij}^*(t) \to 0, \quad \text{uniformly on } [-9/16, 9/16],$$

and the function

$$u_{ij}^t(y, s) := \lambda_{ij}^*(t)^{\frac{n-2}{2}} u\left(\xi_{ij}^*(t) + \lambda_{ij}^*(t)y, t + \lambda_{ij}^*(t)^2 s \right),$$

converges to $W(y)$ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.

(5) *For any $1 \leq j \neq k \leq N$ and $t \in [-9/16, 9/16]$,*

$$\lim_{i \to +\infty} \frac{|\xi_{ij}^*(t) - \xi_{ik}^*(t)|}{\max\left\{\lambda_{ij}^*(t), \lambda_{ik}^*(t)\right\}} = +\infty. \qquad (21.2)$$

(6) $u_\infty \equiv 0.$

Remark 21.2. *Item (5) says there is no bubble towering. Item (6) will be used (in an inductive way) to prove this property.*

The proof of this theorem uses an induction argument on N. This part is organized in the following way.

(1) In Section 22, we establish several preliminary results.
(2) In Section 23, we find all bubbles under the inductive assumption that Theorem 21.1 holds up to $N - 1$.
(3) In Section 24, we prove the Lipschitz hypothesis (10.1) in Part 2, under the assumptions **(II.a)-(II.c)**. This also provides the first step of our inductive argument.
(4) In Section 20, we finish the proof of Theorem 21.1 by employing Proposition 19.1 in Part 2.

22. Preliminaries

Before starting the bubble tree construction, we recall several technical results.

The first one is a uniform Morrey space estimate on u_i, which is just a special case of Corollary 3.5 in Part 1

Lemma 22.1. *There exists a constant $M > 0$ such that for each i and $Q_r(x,t) \subset Q_{3/4}$,*

$$\int_{Q_r(x,t)} \left(|\nabla u_i|^2 + u_i^{p+1} \right) \leq M r^2.$$

The second one is the non-concentration estimate of $\partial_t u_i$ in Lemma 11.3 from Part 2, which still holds in the current setting (thanks to **(III.c)**).

The following lemma is about the convergence of rescalings of u_i.

Lemma 22.2. *Given a sequence $(x_i, t_i) \in Q_{1/2}$, $r_i \to 0$, define*

$$u_i^{r_i}(x,t) := r_i^{\frac{n-2}{2}} u_i \left(x_i + r_i x, t_i + r_i^2 t \right). \tag{22.1}$$

Then we have

(1) *After passing to a subsequence, $u_i^{r_i}$ converges weakly to 0 or $W_{\xi,\lambda}$ for some $(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^+$.*
(2) *If the defect measure associated to $u_i^{r_i}$ is nontrivial, it must be of the form*

$$\sum_j k_j \Lambda \delta_{P_j} \otimes dt,$$

where P_j are distinct points in $\mathbb{R}^n$ and

$$\sum_j k_j \leq M/\Lambda, \quad k_j \in \mathbb{N}.$$

(3) *If there is no defect measure, then $u_i^{r_i}$ converges in $C_{loc}^{\infty}(\mathbb{R}^n \times \mathbb{R})$.*

Proof. As in Part 1, $u_i^{r_i}$ converges weakly to a nonnegative solution of (1.1) in $L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R})$. By Lemma 11.3 (non-concentration of time derivatives), $\partial_t u_i^{r_i}$ converges to 0 in $L_{loc}^2(\mathbb{R}^n \times \mathbb{R})$. Hence this weak limit is independent of time. By

Caffarelli-Gidas-Spruck [8], it must be 0 or $W_{\xi,\lambda}$ for some $(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^+$. Furthermore, the defect measure is also independent of time, that is, there exists a Radon measure μ_0 on $\mathbb{R}^n$ such that the defect measure equals $\mu_0 \otimes dt$. By Lemma 22.1, for any $R > 0$,

$$\int_{-R^2}^{R^2} \int_{B_R} d\mu_0 dt \le MR^2.$$

Hence

$$\mu_0(\mathbb{R}^n) \le M. \tag{22.2}$$

By Lemma 9.2, there exist finitely many distinct points $P_j \in \mathbb{R}^n$ and constants $k_j \in \mathbb{N}$ such that

$$\mu_0 = \sum_j k_j \Lambda \delta_{P_j}.$$

Finally, if there is no defect measure, the smooth convergence of $u_i^{r_i}$ follows from a direct application of the ε-regularity theorem, Theorem 3.9, by noting the smoothness of its weak limit. $\qquad\square$

We also need a technical result about the blow up of each time slice of u_i. This will be used to find the first bubble in the next section.

Lemma 22.3. *For each $t \in [-9/16, 9/16]$,*

$$\lim_{i \to +\infty} \max_{B_1} u_i(x, t) = +\infty.$$

Proof. For each u_i and $(x, t) \in Q_{3/4}$, define $\rho_i(x, t)$ to be the unique (if it exists) solution to the equation

$$\Theta_{r^2}(x, t; u_i) = \varepsilon_*.$$

Here ε_* is the small constant in the ε-regularity theorem, Theorem 3.7.

By **(III.b)**, if $x \ne 0$, there is a uniform, positive lower bound for $\rho_i(x, t)$. By **(III.a)** and **(III.b)**, for any $s > 0$ fixed,

$$\lim_{i \to +\infty} \Theta_s(0, t; u_i) = (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} N + o_s(1).$$

Therefore

$$\lim_{i \to +\infty} \rho_i(0, t) = 0.$$

As a consequence, $\rho_i(\cdot, t)$ has an interior minimal point in B_1, say $x_i(t)$. Denote $r_i := \rho_i(x_i(t), t)$. Define $u_i^{r_i}$ as in (22.1), with base point at $(x_i(t), t)$.

Claim. For some $(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^+$, $u_i^{r_i}$ converges to $W_{\xi,\lambda}$ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.

By this claim we get

$$\lim_{i \to +\infty} \rho_i(x_i(t), t)^{\frac{n-2}{2}} u_i(x_i(t), t) = W_{\xi,\lambda}(0) > 0,$$

and the proof of this lemma is complete.

Proof of the Claim. By the definition of r_i and the scaling invariance of $\Theta_s(\cdot)$, for any $y \in B_{r_i^{-1}}$,

$$\Theta_1(y, 0; u_i^{r_i}) \le \varepsilon_*.$$

Moreover, the equality is attained at $y = 0$, that is,

$$\varepsilon_* = \int_{\mathbb{R}^n} \left[\frac{|\nabla u_i^{r_i}(y, -1)|^2}{2} - \frac{u_i^{r_i}(y, -1)^{p+1}}{p + 1} \right] \psi \left(\frac{y - x_i(t)}{r_i} \right)^2 G(y, 1) dy$$

$$+ \frac{1}{2(p-1)} \int_{\mathbb{R}^n} u_i^{r_i}(y, -1)^2 \psi \left(\frac{x - x_i(t)}{r_i} \right)^2 G(y, 1) dy + C e^{-c r_i^{-2}}. \quad (22.3)$$

By Theorem 3.9, $u_i^{r_i}$ are uniformly bounded in $C^2(B_{r_i^{-1}-1} \times (-\delta_*, \delta_*))$. This then implies that there is no defect measure appearing when we apply Lemma 22.2 to $u_i^{r_i}$. In fact, if the defect measure is nontrivial, by Lemma 22.2, it has the form

$$\sum_j k_j \Lambda \delta_{P_j} \otimes dt, \quad P_j \in \mathbb{R}^n, \ k_j \in \mathbb{N}.$$

Then by Lemma 9.3, for a.e. $t \in (-\delta_*, \delta_*)$, $u_i^{r_i}(\cdot, t)$ should develop nontrivial Dirac measures at P_j. This is a contradiction with the uniform $C^2(B_{r_i^{-1}-1} \times (-\delta_*, \delta_*))$ regularity of $u_i^{r_i}$.

Since there is no defect measure, Lemma 22.2 implies that $u_i^{r_i}$ converges to u_∞ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$. In view of the Liouville theorem in Caffarelli-Gidas-Spruck [8], it suffices to show that $u_\infty \neq 0$. (By the strong maximum principle, either $u_\infty \equiv 0$ or $u_\infty > 0$ everywhere.)

First by the monotonicity formula (Proposition 3.2), we have

$$\varepsilon_* \leq \int_{-2}^{-1} (-s)^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla u_i^{r_i}(y, s)|^2}{2} - \frac{u_i^{r_i}(y, s)^{p+1}}{p + 1} \right] \psi \left(\frac{y - x_i(t)}{r_i} \right)^2 G(y, -s)$$

$$+ \frac{1}{2(p-1)} \int_{-2}^{-1} (-s)^{\frac{2}{p-1}} \int_{\mathbb{R}^n} u_i^{r_i}(y, s)^2 \psi \left(\frac{x - x_i(t)}{r_i} \right)^2 G(y, -s) + C e^{-c r_i^{-2}/4}.$$

By the $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$ convergence of $u_i^{r_i}$, the Morrey space bound in Lemma 22.1 for $u_i^{r_i}$ and the exponential decay of $G(y, -s)$ as $|y| \to +\infty$, we can let $i \to +\infty$ in the above inequality to get

$$\varepsilon_* \leq \int_{-2}^{-1} (-s)^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla u_\infty(y, s)|^2}{2} - \frac{u_\infty(y, s)^{p+1}}{p + 1} \right] G(y, -s)$$

$$+ \frac{1}{2(p-1)} \int_{-2}^{-1} (-s)^{\frac{2}{p-1}} \int_{\mathbb{R}^n} u_\infty(y, s)^2 G(y, -s).$$

Hence it is impossible that $u_\infty \equiv 0$. $\qquad\square$

23. Bubble tree construction

In this section, we construct bubbles by finding local maximal points of $u_i(\cdot, t)$. The construction is divided into six steps. During the course of this construction, we will also prove Theorem 21.1 except the last point (6), under the inductive assumption that this theorem holds when the multiplicity is not larger than $N - 1$ ($N \geq 2$).

Step 1. Construction of the first maximal point. By **(IIII.a)** and Lemma 22.3, for each t, $\max_{B_1} u_i(x,t)$ is attained at an interior point, say $\xi_{i1}^*(t)$. Denote

$$\lambda_{i1}^*(t) := u_i\left(\xi_{i1}^*(t), t\right)^{-\frac{2}{n-2}}.$$

By Lemma 22.3,

$$\lim_{i \to +\infty} \lambda_{i1}^*(t) = 0.$$

Let

$$u_{i1}(y, s) := \lambda_{i1}^*(t)^{\frac{n-2}{2}} u_i\left(\xi_{i1}^*(t) + \lambda_{i1}^*(t)y, t + \lambda_{i1}^*(t)^2 s\right).$$

It satisfies the following conditions.

- By Lemma 22.1, for any $R > 0$,

$$\int_{Q_R} \left[|\nabla u_{i1}|^2 + u_{i1}^{p+1}\right] \le MR^2. \tag{23.1}$$

- By Lemma 11.3, for any $R > 0$,

$$\lim_{i \to +\infty} \int_{Q_R} |\partial_t u_{i1}|^2 = 0. \tag{23.2}$$

- At $s = 0$,

$$\max_{|y| \le \lambda_{i1}^*(0)^{-1}/2} u_{i1}(y, 0) = u_{i1}(0, 0) = 1. \tag{23.3}$$

Using these conditions we show

Lemma 23.1. *As $i \to +\infty$, u_{i1} are uniformly bounded in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$, and it converges to W in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.*

Proof. We use Lemma 22.2 to study the convergence of u_{i1}. We need only to show that there is no defect measure appearing. Once this is established, u_{i1} would converge in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$. Then by (23.3), this limit must be W.

Assume by the contrary that the defect measure is nonzero. By Lemma 22.2, it has the form

$$\sum_j k_j \Lambda \delta_{P_j} \otimes dt, \quad P_j \in \mathbb{R}^n, \ k_j \in \mathbb{N}.$$

Then applying Lemma 22.3 to u_{i1} at $t = 0$, we obtain

$$\lim_{i \to +\infty} \max_{B_1(P_j)} u_{i1}(x, 0) = +\infty.$$

This is a contradiction with (23.3). $\qquad\qquad\square$

By this lemma, there exists a sequence $R_{i1} \to +\infty$ such that

$$\lim_{i \to +\infty} \sup_{y \in B_{R_{i1}}} |y|^{\frac{n-2}{2}} u_{i1}(y, 0) < +\infty. \tag{23.4}$$

Step 2. Iterative construction. Suppose $\xi_{i1}^*(t), \cdots, \xi_{i,j-1}^*(t)$ have been constructed. If

$$\max_{x \in B_1} \min_{1 \le k \le j-1} |x - \xi_{ik}^*(t)|^{\frac{n-2}{2}} u_i(x, t)$$

are uniformly bounded in B_1, we stop at this step. Otherwise,

$$\lim_{i \to \infty} \max_{x \in B_1} \min_{1 \le k \le j-1} |x - \xi_{ik}^*(t)|^{\frac{n-2}{2}} u_i(x,t) = +\infty. \tag{23.5}$$

By **(IIII.a)**, this function has an interior maximal point, say $\xi_{ij}^*(t)$. Define $\lambda_{ij}^*(t)$, u_{ij} as in Step 1. The estimates (23.1) and (23.2) still hold for u_{ij}, while (23.3) now reads as

$$\begin{cases} u_{ij}(0,0) = 1, \\ u_{ij}(y,0) \le \dfrac{\min_{1 \le k \le j-1} |\xi_{ij}^*(t) - \xi_{ik}^*(t)|^{\frac{n-2}{2}}}{\min_{1 \le k \le j-1} |\xi_{ij}^*(t) - \xi_{ik}^*(t) + \lambda_{ij}^*(t)y|^{\frac{n-2}{2}}}, \end{cases} \tag{23.6}$$

the right hand side of which converges to 1 uniformly in any compact set of $\mathbb{R}^n$, by noting that we have

$$\min_{1 \le k \le j-1} |\xi_{ij}^*(t) - \xi_{ik}^*(t)|^{\frac{n-2}{2}} u_i\left(\xi_{ij}^*(t), t\right) = +\infty.$$

(This follows by letting $x = \xi_{ij}^*(t)$ in (23.5).)

We can argue as in Step 1 to deduce that u_{ij} converges to W in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$. Moreover, there exists a sequence $R_{ij} \to +\infty$ such that

$$\lim_{i \to +\infty} \sup_{y \in B_{R_{ij}}} |y|^{\frac{n-2}{2}} u_{ij}(y,0) < +\infty. \tag{23.7}$$

For any $k < j$, combining (23.5) (evaluated at $x = \xi_{ij}^*(t)$) and (23.4) (with j replaced by k) together, we obtain (21.2).

Step 3. $\xi_{ij}^*(t)$ **are local maximal points of** $u_i(\cdot, t)$**.** Because 0 is the maximal point of W and the Hessian $\nabla^2 W(0)$ is strictly negative definite, by the above smooth convergence of u_{ij} in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$, we deduce that $\xi_{ij}^*(t)$ are local maximal points of $u_i(\cdot, t)$.

The strict concavity of u_i near $\xi_{ij}^*(t)$ (with the help of the implicit function theorem) also implies the continuous dependence of $\xi_{ij}^*(t)$ with t. The continuity of $\lambda_{ij}^*(t)$ then follows from its definition.

Step 4. Fix a large positive constant R. For each $t \in (-9/16, 9/16)$, let

$$\Omega_i(t) := \cup_j B_{R\lambda_{ij}^*(t)}\left(\xi_{ij}^*(t)\right).$$

By (21.2), for all i large enough, these balls are disjoint from each other, and all of them are contained in $B_{1/2}$. Hence the above iterative construction of $\xi_{ij}^*(t)$ must stop in finitely many steps. (At this stage, we do not claim any uniform in i bound on the number of these points.)

By our construction, there exists a constant C_2 such that for any i and $x \in B_1 \setminus \Omega_i(t)$,

$$u_i(x,t) \le C_2 \max_j |x - \xi_{ij}^*(t)|^{-\frac{n-2}{2}}. \tag{23.8}$$

Then (21.1) follows by applying standard parabolic regularity theory.

Step 5. In this step and the next one we show that there is no bubble other than those constructed in the previous steps, and the number of bubbles is exactly N.

In this step we consider the case that there are at least two bubbles. For each t, let

$$d_i(t) := \max_{j \neq k} |\xi_{ij}^*(t) - \xi_{ik}^*(t)|. \tag{23.9}$$

Without loss of generality, assume this is attained between $\xi_{i1}^*(t)$ and $\xi_{i2}^*(t)$. By (21.2),

$$\lim_{i \to +\infty} \frac{\max_j \lambda_{ij}^*(t)}{d_i(t)} = 0. \tag{23.10}$$

Lemma 23.2. *For any $\sigma \in (0, 1)$, there exists an $R(\sigma)$ such that for each i, j and t,*

$$\Theta_{R(\sigma)^2 d_i(t)^2}\left(\xi_{ij}^*(t), t; u_i\right) \geq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma).$$

Proof. Assume by the contrary, there exists a $\sigma \in (0, 1)$, a sequence of r_i satisfying

$$\lim_{i \to +\infty} \frac{d_i(t)}{r_i} = 0, \tag{23.11}$$

but

$$\limsup_{i \to +\infty} \Theta_{r_i^2}\left(\xi_{ij}^*(t), t; u_i\right) \leq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma).$$

By the weak convergence of u_i and **(III.a)** and **(III.b)**, for any $s > 0$ fixed,

$$\lim_{i \to +\infty} \Theta_s\left(\xi_{ij}^*(t), t; u_i\right) = (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} N + \Theta_s(0, t; u_\infty).$$

Because u_∞ is smooth,

$$\lim_{s \to 0} \Theta_s(0, t; u_\infty) = 0.$$

Therefore by choosing s sufficiently small, for all i large

$$(4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} N \leq \Theta_s\left(\xi_{ij}^*(t), t; u_i\right) \leq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} N + \sigma. \tag{23.12}$$

By the monotonicity and continuity of $\Theta_s(\cdot; u_i)$ in s, we can enlarge r_i so that

$$\Theta_{r_i^2}\left(\xi_{ij}^*(t), t; u_i\right) = (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma). \tag{23.13}$$

By this choice, (23.11) still holds, while in view of (23.12), we must have $r_i \to 0$.

Define $u_i^{r_i}$ as in (22.1), with respect to the base point $(\xi_{ij}^*(t), t)$ for some j. A scaling of (23.13) gives

$$
\begin{aligned}
&(4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma) \\
&= \Theta_1\left(0, 0; u_i^{r_i}\right) \\
&= \int_{\mathbb{R}^n} \left[\frac{|\nabla u_i^{r_i}(y, -1)|^2}{2} - \frac{u_i^{r_i}(y, -1)^{p+1}}{p+1}\right] \psi\left(\frac{y - \xi_{ij}^*(t)}{r_i}\right)^2 G(y, 1) dy \\
&\quad + \frac{1}{2(p-1)} \int_{\mathbb{R}^n} u_i^{r_i}(y, -1)^2 \psi\left(\frac{y - \xi_{ij}^*(t)}{r_i}\right)^2 G(y, 1) dy + C e^{-cr_i^{-2}}.
\end{aligned}
\tag{23.14}
$$

By (23.8) and (23.11), for any $\rho > 0$, if i is large enough,

$$u_i^{r_i}(y, 0) \leq C_2 |y|^{-\frac{n-2}{2}}, \quad \text{for any } y \in B_{r_i^{-1}/2} \setminus B_\rho.$$

Arguing as in the proof of Lemma 23.1, we deduce that the defect measure of $u_i^{r_i}$ is $N'\delta_0 \otimes dt$, for some $N' \in \mathbb{N}$. Letting $i \to +\infty$ in (23.14), we see $N' \leq N - 1$.

By our inductive assumption, $\nabla u_i^{r_i}$ converges to 0 weakly in $L_{loc}^2(\mathbb{R}^n \times \mathbb{R})$, $u_i^{r_i}$ converges to 0 weakly in $L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R})$ and strongly in $C_{loc}(\mathbb{R}; L_{loc}^2(\mathbb{R}^n))$. Then by Lemma 9.2, both $|\nabla u_i^{r_i}|^2 dxdt$ and $(u_i^{r_i})^{p+1} dxdt$ converges to $N'\delta_0 \otimes dt$ weakly as Radon measures. These convergence imply that the right hand side of (23.14) converges to

$$(4\pi)^{-\frac{n}{2}} N'\Lambda/n.$$

(Here we need to use the monotonicity formula, integrate in time s and then argue as in the last step of the proof of Lemma 22.3.) This is a contradiction with (23.13). $\quad\square$

For $(y, s) \in Q_{d_i(t)^{-1}/2}$, define

$$\widetilde{u}_i(y, s) := d_i(t)^{\frac{n-2}{2}} u_i \left(\xi_{i1}^*(t) + d_i(t)y, t + d_i(t)^2 s \right) \tag{23.15}$$

and

$$\widetilde{\xi}_{ij}^*(t) := \frac{\xi_{ij}^*(t) - \xi_{i1}^*(t)}{d_i(t)}$$

Then we have

- $\widetilde{\xi}_{i1}^*(t) = 0$;
- $|\widetilde{\xi}_{i2}^*(t)| = 1$;
- for each j, $|\widetilde{\xi}_{ij}^*(t)| \leq 1$.

After passing to a subsequence, we may assume each $\widetilde{\xi}_{ij}^*(t)$ converges to a point $\widetilde{\xi}_j^*(t) \in \overline{B_1}$, and $\widetilde{\xi}_{i2}^*(t)$ converges to a point $\widetilde{\xi}_2^*(t) \in \partial B_1$. (These limiting points need not to be distinct.)

By (23.10), $\widetilde{u}_i$ develops bubbles at each $\xi_j^*(t)$. By a scaling of (23.8), $\widetilde{u}_i$ does not develop any bubble outside $\cup_j \{\xi_j^*(t)\}$. Therefore the defect measure of $\widetilde{u}_i$ is

$$\Lambda \sum_j \delta_{\xi_j^*(t)} \otimes dt = \Lambda \sum_k m_k \delta_{P_k}.$$

In the above, we write these $\widetilde{\xi}_j^*(t)$ as distinct points P_k, each one with multiplicity $m_k \in \mathbb{N}$. There are at least two different points, $P_1 = 0$ and $P_2 \in \partial B_1$. Therefore for each k, $m_k \leq N - 1$. By our inductive assumption, Theorem 21.1 holds for $\widetilde{u}_i$. In particular, $\widetilde{u}_i$ converges to 0 weakly in $L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R})$, and there is no bubble towering for $\widetilde{u}_i$.

By (23.12) and Proposition 3.2, we have

$$\limsup_{i \to +\infty} \Theta_{R(\sigma)^2 d_i(t)^2} \left(\xi_{i1}^*(t), t; u_i \right) \leq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N + \sigma).$$

Combining this inequality with Lemma 23.2, we get

$$(4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma)$$

$$
\begin{aligned}
\leq \quad & R(\sigma)^{\frac{n}{2}} \int_{\mathbb{R}^n} \left[\frac{|\nabla \widetilde{u}_i(x,-1)|^2}{2} - \frac{\widetilde{u}_i(x,-1)^{p+1}}{p+1} \right] \psi \left(\frac{x - \xi_{i1}^*(t)}{r_i} \right)^2 G(x, R(\sigma)^2) dx \\
+ \quad & \frac{R(\sigma)^{\frac{n-2}{2}}}{2(p-1)} \int_{\mathbb{R}^n} \widetilde{u}_i(x,-1)^2 \psi \left(\frac{x - \xi_{i1}^*(t)}{r_i} \right)^2 G(x, R(\sigma)^2) dx + C e^{-cd_i(t)^{-2}} \quad (23.16) \\
\leq \quad & (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N + \sigma).
\end{aligned}
$$

Letting $i \to +\infty$, we deduce that the defect measure from $\widetilde{u}_i$ satisfies

$$
N - \sigma \leq \sum_k m_k e^{-\frac{|P_k|^2}{4R(\sigma)^2}} \leq N + \sigma.
$$

Letting $\sigma \to 0$, which implies that $R(\sigma) \to +\infty$, we get

$$
\sum_k m_k = N.
$$

By Theorem 21.1 and our inductive assumption, for all i large, there are exactly N bubbles located at $\widetilde{\xi}_{ij}^*(t)$ for $\widetilde{u}_i(\cdot, 0)$. Coming back to u_i, this says for each t, there are exactly N bubbles located at $\xi_{ij}^*(t)$, $j = 1, \cdots, N$.

Step 6. In this step we consider the case where there is only one $\xi_{ij}^*(t)$ at time t. We show that this is impossible, because we have assumed the multiplicity $N \geq 2$.

In this case, (23.8) reads as

$$
u_i(x,t) \leq C_2 |x - \xi_{i1}^*(t)|^{-\frac{n-2}{2}}, \quad \text{for any } x \in B_1. \quad (23.17)
$$

Under this assumption, we expect that there are N bubbles towering at $\xi_{i1}^*(t)$. We will determine the scale for the lowest bubble, and then perform a rescaling at this scale. This gives a sequence of solutions weakly converging to some nontrivial $W_{\xi,\lambda}$ and exhibiting at most $N - 1$ bubbles. By our inductive assumption, this is impossible.

More precisely, similar to Lemma 23.2, we have

Lemma 23.3. *For any $\sigma \in (0, 1)$, there exists an $R(\sigma)$ such that*

$$
\Theta_{R(\sigma)^2 \lambda_{i1}^*(t)^2} (\xi_{i1}^*(t), t; u_i) \geq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma). \quad (23.18)
$$

Proof. Assume by the contrary that (23.18) does not hold. Arguing as in the proof of Lemma 23.2, we find a sequence of $r_i \to 0$, satisfying

$$
\lim_{i \to +\infty} \frac{\lambda_{i1}^*(t)}{r_i} = 0, \quad (23.19)
$$

and

$$
\Theta_{r_i^2} (\xi_{i1}^*(t), t; u_i) = (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} (N - \sigma). \quad (23.20)
$$

Define $u_i^{r_i}$ as before, with base point at $(\xi_{i1}^*(t), t)$. By (23.17), the defect measure for $u_i^{r_i}$ is $N' \delta_0 \otimes dt$. By (23.20), $N' \leq N - 1$, while by (23.19), $N' \geq 1$. By our inductive assumption, $u_i^{r_i}$ converges weakly to 0. Then as in the proof of Lemma 23.2, we get

$$
\Theta_{r_i^2} (\xi_{i1}^*(t), t; u_i) = \Theta_1 (0, 0; u_i^{r_i}) \to (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} N', \quad \text{as } i \to +\infty.
$$

This is a contradiction with (23.20), because $N' \le N - 1$ and $\sigma \in (0, 1)$. $\qquad\square$

Remark 23.4. *The scale r_i satisfying* (23.20) *is the scale of the lowest bubble.*

Choose a $\sigma \in (0, 1)$ and set $r_i := R(\sigma)\lambda_{i1}^*(t)$ according to the previous lemma. Define $u_i^{r_i}$ as before, with respect to the base point $(\xi_{i1}^*(t), t)$. By definition,

$$u_i^{r_i}(0, 0) = \max_{y \in B_{r_i^{-1}/2}(0)} u_i^{r_i}(y, 0) = R(\sigma)^{-\frac{n-2}{2}}.$$

Similar to Lemma 23.1, this implies that $u_i^{r_i}$ converges to $W_{0, R(\sigma)}$ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$. Arguing as in the proof of Lemma 22.3, we deduce that

$$
\begin{aligned}
&\Theta_{R(\sigma)^2 \lambda_{i1}^*(t)^2}\left(\xi_{i1}^*(t), t; u_i\right) \\
={}& R(\sigma)^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla \widetilde{u}_i(x, -1)|^2}{2} - \frac{\widetilde{u}_i(x, -1)^{p+1}}{p+1}\right] \psi\left(\frac{x - \xi_{ij}^*(t)}{r_i}\right)^2 G(x, R(\sigma)^2) \\
+{}& \frac{1}{2(p-1)} R(\sigma)^{\frac{2}{p-1}} \int_{\mathbb{R}^n} \widetilde{u}_i(x, -1)^2 \psi\left(\frac{x - \xi_{ij}^*(t)}{r_i}\right)^2 G(x, R(\sigma)^2) + Ce^{-cR(\sigma)^{-2}r_i^{-2}} \\
\to{}& R(\sigma)^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla W_{0, R(\sigma)}(x)|^2}{2} - \frac{W_{0, R(\sigma)}(x)^{p+1}}{p+1}\right] G(x, R(\sigma)^2) dx \\
+{}& \frac{1}{2(p-1)} R(\sigma)^{\frac{2}{p-1}} \int_{\mathbb{R}^n} W_{0, R(\sigma)}(x)^2 G(x, R(\sigma)^2) dx \\
<{}& (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n}.
\end{aligned}
$$

This is a contradiction with Lemma 23.3.

24. A REMARK ON THE LIPSCHITZ HYPOTHESIS IN PART 2

In this section, under the hypothesis **(II.a)-(II.c)** in Part 2, we prove the Lipschitz hypothesis (10.1).

First we claim that Proposition 11.1 and Lemma 11.5 still hold, even now we do not assume (10.1). This follows from the general analysis in the previous section. Here we give more details.

Proof of the claim. The proof is divided into three steps.

Step 1. By Lemma 22.3, for any $t \in (-1, 1)$, $u_i(\cdot, t)$ blows up as $i \to +\infty$. This allows us to find the first blow up point $\xi_i^*(t)$.

Step 2. With $u_i^{r_i}$ defined as in (22.1) (with $x_i \to 0$), we claim that

- either there is no defect measure and $u_i^{r_i}$ converges in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$,
- or there exists a point $P \in \mathbb{R}^n$ such that

$$|\nabla u_i^{r_i}(x, t)|^2 dx dt \rightharpoonup \Lambda \delta_P \otimes dt.$$

First, similar to (23.12), for any $\sigma \in (0, 1)$, there exists an s such that

$$(4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} \le \lim_{i \to +\infty} \Theta_s\left(\xi_i^*(t), t; u_i\right) \le (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} + \sigma.$$

Then because $r_i \to 0$, by the monotonicity formula we get, for any $R > 0$,

$$R^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla u_i^{r_i}(x, -R^2)|^2}{2} - \frac{u_i^{r_i}(x, -R^2)^{p+1}}{p+1} \right] \psi \left(\frac{x - x_i}{r_i} \right)^2 G(x, R^2) dx$$

$$+ \frac{R^{\frac{p-1}{2}}}{2(p-1)} \int_{\mathbb{R}^n} u_i^{r_i}(x, -R^2)^2 \psi \left(\frac{x - x_i}{r_i} \right)^2 G(x, R^2) dx + Ce^{-cR^2 r_i^{-2}}$$

$$= \Theta_{R^2 r_i^2}(\xi_i^*(t), t; u_i)$$

$$\leq \Theta_s (\xi_i^*(t), t; u_i)$$

$$\leq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} + \sigma.$$

Denote the weak limit of $u_i^{r_i}$ by u_∞ (which, by Lemma 22.2, is 0 or $W_{\xi,\lambda}$ for some $(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^+$), the defect measure associated to them by $\Lambda \sum_j k_j \delta_{P_j} \otimes dt$ (with notations as in Lemma 22.2). Passing to the limit in the above inequality leads to

$$R^{-2} \int_{-2R^2}^{-R^2} \Theta_s(0, 0; u_\infty) ds + (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} R^{-2} \int_{-2R^2}^{-R^2} \sum_j k_j e^{\frac{|P_j|^2}{4s}} \leq (4\pi)^{-\frac{n}{2}} \frac{\Lambda}{n} + \sigma.$$

By the monotonicity formula and the smoothness of u_∞, we deduce that

$$\Theta_s(0, 0; u_\infty) \geq \lim_{s \to 0} \Theta_s(0, 0; u_\infty) = 0.$$

Hence we have

$$R^{-2} \int_{-2R^2}^{-R^2} \sum_j k_j e^{\frac{|P_j|^2}{4s}} \leq 1 + O(\sigma).$$

Letting $R \to +\infty$, we deduce that there exists at most one blow up point, whose multiplicity is exactly 1.

If there is no defect measure, then Lemma 22.2 implies that $u_i^{r_i}$ converges to u_∞ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.

Finally, we show that if the defect measure is nontrivial, then $u_\infty \equiv 0$. Assume this is not the case. By Lemma 22.2, $u_\infty = W_{\xi,\lambda}$ for some $(\xi, \lambda) \in \mathbb{R}^n \times \mathbb{R}^+$. By the weak convergence of $|\nabla u_i^{r_i}|^2 dxdt$, there exists a sequence R_i satisfying

$$R_i \to +\infty \quad \text{and} \quad R_i r_i \to 0,$$

such that the defect measure associated to $u_i^{R_i r_i}$ is $2\Lambda \delta_0 \otimes dt$. This is a contradiction with the claim that the multiplicity of this defect measure is 1.

Step 3. Denote $\lambda_i^*(t) := u_i(\xi_i^*(t), t)$. Similar to Lemma 23.1, we deduce that $u_i^{\lambda_i^*(t)}$ (with base point at $(\xi_i^*(t), t)$) converges to W in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$. On the other hand, if the sequence r_i satisfies

$$r_i \to 0 \quad \text{and} \quad \frac{r_i}{\lambda_i^*(t)} \to +\infty,$$

then by results in Step 2, we deduce that the sequence $u_i^{r_i}$ satisfies

$$|\nabla u_i^{r_i}(x, t)|^2 dxdt \rightharpoonup \Lambda \delta_0 \otimes dt.$$

Combining this fact with Theorem 3.7, we deduce that

$$|x - \xi_i^*(t)|^{\frac{n-2}{2}} u_i(x, t) + |x - \xi_i^*(t)|^{\frac{n}{2}} |\nabla u_i(x, t)| \tag{24.1}$$
$$+ |x - \xi_i^*(t)|^{\frac{n+2}{2}} \left(|\nabla^2 u_i(x, t)| + |\partial_t u_i(x, t)| \right) \leq C_2. \quad \square$$

Now we come to the proof of the Lipschitz hypothesis (10.1). Assume there exist a sequence of smooth solutions u_i to (1.1), satisfying **(II.a-II.c)**, but

$$\max_{|t| \leq 1/2} \int_{B_1} |\partial_t u_i(x, t)|^2 \, dx \to +\infty. \tag{24.2}$$

For each i, take a $t_i \in (-1, 1)$ attaining

$$\max_{|t| \leq 1} (1 - |t|) \int_{B_1} |\partial_t u_i(x, t)|^2 \, dx.$$

Denote

$$\lambda_i^{-2} := \int_{B_1} |\partial_t u_i(x, t_i)|^2 \, dx.$$

By (12.1), as $i \to +\infty$,

$$\frac{1 - |t_i|}{\lambda_i^2} \geq \frac{1}{2} \max_{|t| \leq 1/2} \int_{B_1} |\partial_t u_i(x, t)|^2 \, dx \to +\infty. \tag{24.3}$$

As a consequence,

$$\lim_{i \to +\infty} \lambda_i = 0.$$

If $|t - t_i| < (1 - |t_i|)/2$, by the definition of t_i, we get

$$\int_{B_1} |\partial_t u_i(x, t)|^2 \, dx \leq 2\lambda_i^{-2}. \tag{24.4}$$

At time t_i, there exists a unique maximal point of $u_i(\cdot, t_i)$ in B_1, denoted by $\xi_i(t_i)$. Define

$$\widetilde{u}_i(x, t) := \lambda_i^{\frac{n-2}{2}} u_i \left(\xi_i(t_i) + \lambda_i x, t_i + \lambda_i^2 t \right).$$

By a scaling, we have

$$\begin{cases} \displaystyle\int_{B_{\lambda_i^{-1}}} |\partial_t \widetilde{u}_i(x, 0)|^2 \, dx = 1, \\[3mm] \displaystyle\int_{B_{\lambda_i^{-1}}} |\partial_t \widetilde{u}_i(x, t)|^2 \, dx \leq 2, \quad \text{for any } |t| < \dfrac{1 - |t_i|}{2\lambda_i^2} \end{cases} \tag{24.5}$$

On the other hand, we claim that if i is large,

$$\int_{B_{\lambda_i^{-1}}} |\partial_t \widetilde{u}_i(x, 0)|^2 \, dx \leq \frac{1}{2}. \tag{24.6}$$

This contradiction implies that (24.2) cannot be true. In other words, if u_i satisfies **(II.a-II.c)**, then there must exist a constant L such that

$$\limsup_{i \to +\infty} \sup_{|t| \leq 1/2} \int_{B_1} \partial_t u_i(x, t)^2 dx \leq L,$$

that is, the Lipschitz hypothesis (10.1) holds.

To prove (24.6), first note that for u_i, by (24.1) we have

$$|\partial_t u_i(x, t_i)| \lesssim |x - \xi_i(t_i)|^{-\frac{n+2}{2}}.$$

For $\widetilde{u}_i$, this reads as

$$|\partial_t \widetilde{u}_i(x, 0)| \lesssim |x|^{-\frac{n+2}{2}}.$$

Therefore there exists a large R (independent of i) such that

$$\int_{B_{\lambda_i^{-1}} \setminus B_R} |\partial_t \widetilde{u}_i(x, 0)|^2 \, dx \leq C \int_{B_R^c} |x|^{-n-2} \leq \frac{1}{4}. \tag{24.7}$$

After establishing this inequality, (24.6) will follow from

Lemma 24.1. *For all i large,*

$$\int_{B_R} |\partial_t \widetilde{u}_i(x, 0)|^2 \, dx \leq \frac{1}{4}. \tag{24.8}$$

Proof. **Case 1.** $\widetilde{u}_i$ converges in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.

By Lemma 11.3,

$$\lim_{i \to +\infty} \int_{-1}^1 \int_{B_R} |\partial_t \widetilde{u}_i|^2 = 0.$$

Hence by the smooth convergence of $\widetilde{u}_i$, we deduce that $\partial_t \widetilde{u}_i$ converges to 0 in $C_{loc}(\mathbb{R}^n \times \mathbb{R})$ and the conclusion follows.

Case 2. $\widetilde{u}_i$ does not converge in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.

In this case, in view of the estimates in (24.5), Theorem 10.1 is applicable to $\widetilde{u}_i$, which gives

$$|\nabla \widetilde{u}_i|^2 dx dt \rightharpoonup \Lambda \delta_0 \otimes dt. \tag{24.9}$$

By (18.4), estimates on the scaling parameters etc. in Proposition 17.6 and the estimate on $\partial_t \phi_i$ in Proposition 17.11, there exists an $\rho > 0$ such that for all i large,

$$\begin{aligned}
\int_{B_\rho} |\partial_t \widetilde{u}_i(x, 0)|^2 \, dx &\leq& C \int_{B_\rho} |x|^{-4} + o(1) \\
&\leq& C\rho^{n-4} + o(1) \leq 1/8.
\end{aligned}$$

Next, in view of (24.9), we deduce that as $i \to +\infty$, $\partial_t \widetilde{u}_i(x, 0) \to 0$ uniformly in $B_R \setminus B_\rho$. Hence for all i large, we also have

$$\int_{B_R \setminus B_\rho} |\partial_t \widetilde{u}_i(x, 0)|^2 \, dx \leq \frac{1}{8}.$$

Putting these two inequalities together we get (24.8). $\qquad\square$

25. Exclusion of bubble towering

In this section, we prove Item (6) in Theorem 21.1. If $N = 1$, in view of the analysis in the previous section, this is exactly Theorem 10.1, so here we consider the case $N \geq 2$.

We will prove this by a contradiction argument. Recall that we have assumed u_∞ is smooth, so it is a classical solution of (1.1). If $u_\infty \neq 0$, by standard Harnack inequality,

$$\inf_{Q_{7/8}} u_\infty > 0. \tag{25.1}$$

First we note that

Lemma 25.1. *There exists a constant $c_\infty > 0$, depending only on $\inf_{Q_{7/8}} u_\infty$, such that*

$$u_i \geq c_\infty \quad in \quad Q_{6/7}.$$

Proof. Because

$$\partial_t u_i - \Delta u_i > 0 \quad in \quad Q_1,$$

the following weak Harnack inequality holds for u_i: for any $(x, t) \in Q_{6/7}$,

$$u_i(x, t) \geq c \int_{Q_{1/100}^-(x, t - 1/50)} u_i.$$

Because $u_i \to u_\infty$ in $L^2_{loc}(Q_1)$, by (25.1) we get

$$\lim_{i \to +\infty} \int_{Q_{1/100}^-(x, t - 1/50)} u_i = \int_{Q_{1/100}^-(x, t - 1/50)} u_\infty \geq c \inf_{Q_{7/8}} u_\infty > 0.$$

The conclusion follows by combining these two inequalities. $\square$

For each $t \in (-1, 1)$, let

$$\rho_i(t) := \min_{1 \leq j \neq k \leq N} |\xi_{ij}^*(t) - \xi_{ik}^*(t)|.$$

Set $t_0 = 0$ and $\rho_0 = \rho_i(t_0)$, which is very small if i is large enough. Fix a large index i. For each $k \in \mathbb{N}$, set

$$t_k := \sup \left\{ t : \rho_i(s) \geq \frac{1}{2} \rho_i(t_{k-1}) \quad \text{for any } s \in [t_{k-1}, t] \right\},$$

and

$$\rho_k := \rho_i(t_k) = 2^{-1} \rho_i(t_{k-1}) = \cdots = 2^{-k} \rho_0. \tag{25.2}$$

For each k, assume $\rho_i(t_k)$ is attained between $\xi_{i1}^*(t_k)$ and $\xi_{i2}^*(t_k)$. Consider

$$u_i^k(x, t) := \rho_k^{\frac{n-2}{2}} u_i \left(\xi_{i1}^*(t_k) + \rho_k x, t_k + \rho_k^2 t \right).$$

By Lemma 25.1, we have

$$u_i^k(x, 0) \geq c_\infty \rho_k^{\frac{n-2}{2}}, \quad \text{for } x \in B_1 \setminus B_{1/2}.$$

By our construction, for any $t \in [0, (t_{k+1} - t_k)/\rho_k^2]$, the distance between different bubble points of u_i^k is not smaller than $1/2$. This allows us to iteratively apply Proposition 19.1 (backwardly in time), which gives

$$u_i^k(x, 0) \lesssim e^{-c\frac{t_{k+1} - t_k}{\rho_k^2}}, \quad \text{for } x \in B_1 \setminus B_{1/2}.$$

Combining these two inequalities, we obtain

$$t_{k+1} - t_k \lesssim \rho_k^2 |\log \rho_k| \lesssim \rho_k.$$

Combining this inequality with (25.2), we get

$$t_\infty := \lim_{k \to \infty} t_k \lesssim \sum_{k=0}^{\infty} \rho_k \lesssim \rho_0 \ll 1.$$

This then implies $\rho_i(t_\infty) = 0$ (by the continuity of $\rho_i(t)$, which follows from the continuity of $\xi_{ij}^*(t)$). This is a contradiction. In other words, we must have $u_\infty \equiv 0$.

Part 4. **Analysis of first time singularity**

26. SETTING

In this part we assume $u \in C^\infty(Q_1^-)$ is a *smooth* solution of (1.1), satisfying

$$\int_{-1}^{t} \int_{B_1} \left(|\nabla u|^2 + |u|^{p+1} \right) \leq K(T-t)^{-K} \tag{26.1}$$

for some constant K. In particular, u may not be smoothly extended to $B_1 \times \{0\}$.

Define

$$\mathcal{R}(u) := \left\{ a \in B_1 : \exists\, r > 0 \text{ such that } u \in L^\infty(Q_r^-(a,T)) \right\}$$

to be the regular set, and $\mathcal{S}(u) := B_1 \setminus \mathcal{R}(u)$ to be the set of blow up points. By standard parabolic estimates, if $a \in \mathcal{R}(u)$, u can be extended smoothly up to $t = 0$ in a small backward parabolic cylinder centered at $(a,0)$.

The main result of this part is

Theorem 26.1. *If $n \geq 7$, $p = \frac{n+2}{n-2}$ and $u > 0$, then there exists a constant C such that*

$$\|u(t)\|_{L^\infty(B_{1/2})} \leq C(T-t)^{-\frac{1}{p-1}}. \tag{26.2}$$

We will also show how to get Theorem 1.1 and Theorem 1.2 from this theorem. The proof of this theorem consists of the following four steps.

(1) After some preliminary estimates about the monotonicity formula and Morrey space bounds in Section 27, we perform the tangent flow analysis in Section 28 at a possible singular point (a, T). Here we mainly use results from Part 1. This tangent flow analysis shows that Type II blow up points are isolated in $\mathcal{S}(u)$, see Section 29, which allows us to apply results in Part 2 and Part 3.

(2) In Section 30, we apply Theorem 21.1 in Part 3 and Proposition 20.1 in Section 20 to exclude the bubble clustering formation.

(3) In Section 31, we exclude the formation of only one bubble. Here we mainly use the differential inequality on the scaling parameter λ derived in Part 2.

(4) Finally, we use a stability result on Type I blow ups to derive the Type I estimate (32.2).

27. SOME INTEGRAL ESTIMATES

In this section and the next one, we assume only $p > 1$, and u needs not to be positive. This section is devoted to showing a Morrey space bound on u, which will be used in the next section to perform the tangent flow analysis. This is similar to Section 3 in Part 1.

Denote the standard heat kernel on $\mathbb{R}^n$ by G. Take a function $\psi \in C^\infty(B_{1/4})$, which satisfies $0 \leq \psi \leq 1$, $\psi \equiv 1$ in $B_{1/8}$ and $|\nabla \psi| + |\nabla^2 \psi| \leq C$. Take an arbitrary point $a \in B_{1/2}$. For any $s \in (0,1)$, set

$$\Theta_s(a) \;\; := \;\; s^{\frac{p+1}{p-1}} \int_{B_1} \left[\frac{|\nabla u(y,-s)|^2}{2} - \frac{|u(y,-s)|^{p+1}}{p+1} \right] G(y-a,s)\psi(y)^2 dy$$

$$+ \quad \frac{1}{2(p-1)} s^{\frac{2}{p-1}} \int_{B_1} u(y,-s)^2 G(y-a,s)\psi(y)^2 dy + Ce^{-cs^{-1}}.$$

The following is another localized version of the monotonicity formula, which is a little different from Proposition 3.2.

Proposition 27.1 (Localized monotonicity formula II). *For any $0 < s_1 < s_2 < 1$,*

$$\Theta_{s_2}(x) - \Theta_{s_1}(x)$$
$$\geq \int_{s_1}^{s_2} \tau^{\frac{2}{p-1}-1} \int_{B_1} \left| (-\tau)\partial_t u(y,-\tau) + \frac{u(y,-\tau)}{p-1} + \frac{y\nabla u(y,-\tau)}{2} \right|^2$$
$$\times G(y-x,-\tau)\psi(y)^2 dyd\tau.$$

This almost monotonicity formula allows us to define

$$\Theta(x) := \lim_{s\to 0} \Theta_s(x).$$

As in Part 1, we still have

Lemma 27.2. Θ *is nonnegative and it is upper semi-continuous in x.*

Proposition 27.3. *For any $x \in B_{1/2}$, $r < 1/4$ and $\theta \in (0,1/2)$, there exists a constant $C(\theta, u)$ such that*

$$r^{-2} \int_{Q_r^-(x,-\theta r^2)} \left(|\nabla u|^2 + |u|^{p+1} \right) + \int_{Q_r^-(x,-\theta r^2)} |\partial_t u|^2 \leq C(\theta, u). \qquad (27.1)$$

By the ε-regularity theorem (see Theorem 3.9), we also get

Proposition 27.4. $\mathcal{R}(u) = \{\Theta = 0\}$ *and* $\mathcal{S}(u) = \{\Theta \geq \varepsilon_*\}$.

28. Tangent flow analysis, II

In this section, we perform the tangent flow analysis for u. This is similar to the one in Section 5, with the only difference that now we work in backward parabolic cylinders, instead of the full parabolic cylinders. Some special consequences will also follow from the assumption that $p = (n+2)/(n-2)$, which will be used in the next section to classify Type I and Type II blow up points.

Take an arbitrary point $a \in B_{1/2}$. For any $\lambda > 0$ sufficiently small, define

$$u_\lambda(x,t) := \lambda^{\frac{2}{p-1}} u(a + \lambda x, T + \lambda^2 t).$$

By Proposition 27.3, for any $R > 0$ and $\theta \in (0,1/2)$, we have

$$R^{-2} \int_{Q_R^-(0,-\theta R^2)} \left[|\nabla u_\lambda|^2 + |u_\lambda|^{p+1} \right] + \int_{Q_R^-(0,-\theta R^2)} |\partial_t u_\lambda|^2 \leq C(\theta, u), \qquad (28.1)$$

where $C(\theta, u)$ is the constant in Proposition 27.3. Therefore, as in Section 4, for any $\lambda_i \to 0$, we can subtract a subsequence λ_i (not relabelling) such that

- u_{λ_i} converges to u_∞, weakly in $L_{loc}^{p+1}(\mathbb{R}^n \times \mathbb{R}^-)$ and strongly in $L_{loc}^q(\mathbb{R}^n \times \mathbb{R}^-)$ for any $q < p+1$;
- ∇u_{λ_i} converges to ∇u_∞ weakly in $L_{loc}^2(\mathbb{R}^n \times \mathbb{R}^-)$;
- $\partial_t u_{\lambda_i}$ converges to $\partial_t u_\infty$ weakly in $L_{loc}^2(\mathbb{R}^n \times \mathbb{R}^-)$;

- there exist two Radon measures μ and ν such that in any compact set of $\mathbb{R}^n \times \mathbb{R}^-$,

$$\begin{cases} |\nabla u_i|^2 dxdt \rightharpoonup |\nabla u_\infty|^2 dxdt + \mu, \\ |u_i|^{p+1} dxdt \rightharpoonup |u_\infty|^{p+1} dxdt + \mu, \\ |\partial_t u_i|^2 dxdt \rightharpoonup |\partial_t u_\infty|^2 dxdt + \nu \end{cases}$$

weakly as Radon measures.

Passing to the limit in the equation for u_{λ_i}, we deduce that u_∞ is a weak solution of (1.1) in $\mathbb{R}^n \times \mathbb{R}^-$. Passing to the limit in (28.1) gives

$$R^{-2} \int_{Q_R^-(0,-\theta R^2)} \left[|\nabla u_\infty|^2 + |u_\infty|^{p+1} \right] + \int_{Q_R^-(0,-\theta R^2)} |\partial_t u_\infty|^2 \leq C(\theta, u), \qquad (28.2)$$

where $C(\theta, u)$ is the constant in Proposition 27.3.

Similar to Lemma 5.2 in Part 1, we deduce that both u_∞ and μ are backwardly self-similar.

Results obtained so far hold for any $p > 1$. If $p = (n+2)/(n-2)$, we can say more about (u_∞, μ).

Proposition 28.1. *If $p = (n+2)/(n-2)$, then the followings hold.*

(1) *Either $u_\infty \equiv 0$ or $u_\infty \equiv \pm [-(p-1)t]^{-\frac{1}{p-1}}$.*
(2) *There exists a constant $M > 0$ such that*

$$\mu = M\delta_0 \otimes dt.$$

(3) $\nu = 0$.

Proof. A rescaling of (28.2) gives

$$\int_{B_1(x)} \left(|\nabla w_\infty|^2 + |w_\infty|^{p+1} \right) \leq C, \quad \forall x \in \mathbb{R}^n. \qquad (28.3)$$

Because $p = (n+2)/(n-2)$, by Theorem 7.2, we obtain (1).

Because μ is self-similar, by Theorem 9.1, we find a discrete set $\{\xi_j\} \subset \mathbb{R}^n$, $M_j > 0$ such that

$$\mu = \mu_t dt, \quad \mu_t = \sum_j M_j \delta_{\sqrt{-t}\xi_j}. \qquad (28.4)$$

Because u_{λ_i} are all smooth in $\mathbb{R}^n \times \mathbb{R}^-$, the energy inequalities (2.2) for them are in fact identities. Letting $\lambda_i \to 0$, we see the limiting energy inequality (4.9) for (u_∞, μ) is also an identity. Furthermore, because u_∞ is smooth in $\mathbb{R}^n \times \mathbb{R}^-$, it also satisfies the energy identity. From these facts we deduce that

$$\partial_t \mu = -n\nu \quad \text{in the distributional sense in } \mathbb{R}^n \times \mathbb{R}^-.$$

Combining this relation with (28.4), we get (2) and (3). $\qquad \square$

Furthermore, if u is positive, this proposition implies that the blowing up sequence u_λ satisfy the assumptions **(III.a-III.c)** in Part 3.

29. No coexistence of Type I and Type II blow ups

From now on it is always assumed that u is a positive solution, $p = (n+2)/(n-2)$ and $n \geq 7$. In this section we classify Type I and Type II blow up points, and give some qualitative description of the set of Type I and Type II blow up points.

Lemma 29.1. *Either $u_\infty = 0$ or $\mu = 0$.*

Proof. If $\mu \neq 0$, by the positivity, there exists an $N \in \mathbb{N}$ such that the constant in (28.4) satisfies

$$M = N\Lambda.$$

Since u_∞ is smooth in $\mathbb{R}^n \times \mathbb{R}^-$, **(III.a-III.c)** are satisfied by any blow up sequence u_{λ_i} at (a, T). Then an application of Theorem 21.1 implies that $u_\infty = 0$. Alternatively, if $u_\infty \neq 0$, then $\mu = 0$. $\square$

A direct consequence of this lemma is

Proposition 29.2. *Any point $a \in B_{1/2}$ belongs to one of the following three classes:*

Regular point: $\Theta(a) = 0$;

Type I blow up point: $\Theta(a) = n^{-1} \left(\frac{n-2}{4} \right)^{\frac{n}{2}}$;

Type II blow up point: $\Theta(a) = n^{-1} (4\pi)^{-\frac{n}{2}} N\Lambda$ *for some $N \in \mathbb{N}$.*

Since all of the possible tangent flows form a discrete set (classified according to the value of $\Theta(a)$), we obtain

Corollary 29.3. *The tangent flow of u at any point $a \in B_{1/2}$ is unique.*

Next we study the sets of Type I and Type II blow up points.

Lemma 29.4. *The set of Type I blow up points is relatively open in $\mathcal{S}(u)$.*

Proof. It is directly verified that

$$n^{-1} \left(\frac{n-2}{4} \right)^{\frac{n}{2}} < n^{-1} (4\pi)^{-\frac{n}{2}} \Lambda. \tag{29.1}$$

The conclusion then follows from the upper semi-continuity of Θ, see Lemma 27.2. $\square$

Lemma 29.5. *The set of Type II blow up points is isolated in $\mathcal{S}(u)$.*

Proof. Assume x_0 is Type II, that is, $\Theta(x_0) = n^{-1} (4\pi)^{-\frac{n}{2}} N\Lambda$ for some $N \in \mathbb{N}$. Assume by the contrary that there exists a sequence of $x_i \in \mathcal{S}(u)$ converging to x_0.

Denote

$$\lambda_i := |x_i - x_0|, \quad \widehat{x}_i := \frac{x_i - x_0}{\lambda_i}.$$

Define the blow up sequence

$$u_i(x, t) := \lambda_i^{\frac{2}{p-1}} u(x_0 + \lambda_i x, T + \lambda_i^2 t).$$

Then the tangent flow analysis together with the fact that x_0 is Type II implies that

$$|\nabla u_i|^2 dx dt \rightharpoonup N\Lambda \delta_0 \otimes dt.$$

As a consequence, there exists a fixed, small $s > 0$ such that

$$\Theta_s\left(\widehat{x}_i; u_i\right) \leq \varepsilon_*/2.$$

However, by the monotonicity formula (Proposition 27.1), the fact that $x_i \in \mathcal{S}$ and the scaling invariance of Θ_s, we always have

$$\Theta_s\left(\widehat{x}_i; u_i\right) = \Theta_{\lambda_i^2 s}(x_i; u) \geq \Theta(x_i; u_i) \geq \varepsilon_*.$$

This is a contradiction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

30. Exclusion of bubble clustering

In this section we exclude Type II blow ups with bubble clustering. From now on we assume there exists a Type II blow up point. By Lemma 29.5, this blow up point is isolated in $\mathcal{S}(u)$. Hence after a translation and a scaling, we may assume $u \in C^\infty(\overline{Q_1^-} \setminus \{(0,0)\})$, and $(0,0)$ is a Type II blow up point.

Under these assumptions, Theorem 21.1 is applicable to any rescaling of u at $(0,0)$. Combining this theorem with a continuity argument in time, we obtain

Proposition 30.1. *There exists an $N \in \mathbb{N}$ so that for all t sufficiently close to 0, the following results hold.*

(1) *There exist exactly N local maximal point of $u(\cdot, t)$ in the interior of $B_1(0)$.*
   Denote this point by $\xi_j^*(t)$ $(j = 1, \cdots, N)$ and let $\lambda_j^*(t) := u(\xi_{ij}^*(t), t)^{-\frac{2}{n-2}}$.
(2) *Both λ_j^* and ξ_j^* depend continuously on t.*
(3) *There exists a constant K such that, for all i and any $x \in B_1$,*

$$u(x, t) \leq K \max_{1 \leq j \leq N} |x - \xi_j^*(t)|^{-\frac{n-2}{2}}. \tag{30.1}$$

(4) *As $t \to 0$,*

$$\frac{\lambda_j^*(t)}{\sqrt{-t}} \to 0, \quad \frac{|\xi_j^*(t)|}{\sqrt{-t}} \to 0,$$

   and the function

$$u_j^t(y, s) := \lambda_j^*(t)^{\frac{n-2}{2}} u\left(\xi_j^*(t) + \lambda_j^*(t)y, t + \lambda_j^*(t)^2 s\right),$$

   converges to $W(y)$ in $C_{loc}^\infty(\mathbb{R}^n \times \mathbb{R})$.
(5) *For any $1 \leq j \neq k \leq N$ and $t \in [-9/16, 0)$,*

$$\lim_{t \to 0} \frac{|\xi_j^*(t) - \xi_k^*(t)|}{\max\left\{\lambda_j^*(t), \lambda_k^*(t)\right\}} = +\infty. \tag{30.2}$$

The main result of this section is

Proposition 30.2. *The multiplicity N in Proposition 30.1 must be 1.*

The proof is by a contradiction argument, so assume $N \geq 2$. For each $t \in (-1, 0)$, denote

$$\rho(t) := \min_{1 \leq k \neq \ell \leq N} |\xi_k^*(t) - \xi_\ell^*(t)|.$$

By Point (4) in Proposition 30.1, for any $t < 0$, $\rho(t) > 0$, and

$$\lim_{t \to 0} \rho(t) = 0. \tag{30.3}$$

Then by the continuity of $\rho(t)$ (thanks to the continuity of $\xi_j^*(t)$), there exists a large positive integer i_0 so that the following is well-defined. For each $i \geq i_0$, define t_i to be the first time satisfying $\rho(t) = 2^{-i}$. In other words,

$$\rho(t_i) = 2^{-i} \quad \text{and} \quad \rho(t) > 2^{-i} \text{ in } [-1, t_i). \tag{30.4}$$

On the other hand, we claim that

Lemma 30.3. *For all i large,*

$$\rho(t_{i+1}) \geq \frac{2}{3}\rho(t_i). \tag{30.5}$$

This is clearly a contradiction with the definition of $\rho(t_i)$. The proof of Proposition 30.2 is thus complete. (This contradiction in fact shows that if $N \geq 2$, $\xi_i^*(t)$ cannot converges to the same point as $t \to 0^-$, that is, (30.3) does not hold.)

The proof of Lemma 30.3 uses two facts: the first one is the reduction equation for scaling parameters, which will be applied in a forward in time manner; the second one is the weak form of Schoen's Harnack inequality, Proposition 19.1, which will applied in a backward in time manner.

Proof of Lemma 30.3. Without loss of generality, assume

$$\rho(t_{i+1}) = |\xi_1^*(t_{i+1}) - \xi_2^*(t_{i+1})| = 2^{-i-1}.$$

Choose $\widetilde{t}_i$ to be the first time in $[t_i, t_{i+1}]$ satisfying

$$|\xi_1^*(t) - \xi_2^*(t)| = 2^{-i},$$

that is,

$$|\xi_1^*(\widetilde{t}_i) - \xi_2^*(\widetilde{t}_i)| = 2^{-i}, \quad \text{and} \quad |\xi_1^*(t) - \xi_2^*(t)| > 2^{-i} \text{ for any } t \in [t_i, \widetilde{t}_i).$$

This is well defined, if we note the continuous dependence of $|\xi_1^*(t) - \xi_2^*(t)|$ with respect to t, and the fact that

$$|\xi_1^*(t_i) - \xi_2^*(t_i)| \geq \rho(t_i) = 2^{-i}.$$

Let

$$u_i(x, t) := 2^{-\frac{n-2}{2}i} u\left(\xi_1^*(\widetilde{t}_i) + 2^{-i}x, \widetilde{t}_i + 4^{-i}t\right).$$

It is well defined for $x \in B_{2^{i-1}}$ and $t \in (-4^{i-1}, -4^i\widetilde{t}_i)$.

For each $t \in (-4^{i-1}, -4^i\widetilde{t}_i)$, u_i has N bubbles, located at

$$\xi_{ij}^*(t) := 2^i \left[\xi_j^*\left(\widetilde{t}_i + 4^i t\right) - \xi_1^*\left(\widetilde{t}_i\right)\right],$$

with bubble scale

$$\lambda_{ij}^*(t) := 2^i \lambda_j^*\left(\widetilde{t}_i + 4^i t\right).$$

Denote

$$T_i := 4^i\left(t_{i+1} - \widetilde{t}_i\right).$$

By the definition of $\widetilde{t}_i$ and t_{i+1}, for any $1 \le j \ne k \le N$, we have

$$|\xi_{ij}^*(t) - \xi_{ik}^*(t)| \ge \frac{1}{2}, \quad \forall t \in \left[-4^{i-1}, T_i\right]. \tag{30.6}$$

We also have the normalization condition

$$|\xi_{i1}^*(0) - \xi_{i2}^*(0)| = 1. \tag{30.7}$$

Moreover, by Point (5) in Proposition 30.1, as $i \to +\infty$, if $|\xi_{ij}^*(t)|$ does not escape to infinity, then

$$\lambda_{ij}^*(t) \to 0 \quad \text{uniformly in any compact set of } \mathbb{R}.$$

Then Proposition 20.1 is applicable if i is large enough, which gives

$$T_i \le 100 \min \left\{|\log \lambda_{i1}^*(0)|, |\log \lambda_{i2}^*(0)|\right\}. \tag{30.8}$$

For any $t \in [0, T_i]$, Proposition 12.1 can be applied to u_i in $Q_{1/2}(\xi_{i1}^*(t), t)$ and $Q_{1/2}(\xi_{i2}^*(t), t)$. Denote the scaling parameter by $\lambda_{i1}(t)$ and $\lambda_{i2}(t)$. By Proposition 12.1,

$$\lambda_{i1}(t) = [1 + o(1)]\lambda_{i1}^*(t), \quad |\xi_{i1}^*(t) - \xi_{i1}(t)| = o(\lambda_{i1}^*(t)). \tag{30.9}$$

The reduction equation for $\lambda_{i1}(t)$ (see Proposition 17.6) gives

$$|\lambda_{i1}'(t)| \lesssim \lambda_{i1}(t)^{\frac{n-4}{2}}. \tag{30.10}$$

For any $t \in (0, T_i]$, integrating (30.10) on $[0, t]$ leads to

$$\lambda_{i1}(t)^{-\frac{n-6}{2}} \ge \lambda_{i1}(0)^{-\frac{n-6}{2}} - CT_i.$$

Plugging (30.8) and (30.9) into this inequality, after a simplification we get

$$\begin{aligned}
\lambda_{i1}(t) &\le \lambda_{i1}(0)\left[1 + C\lambda_{i1}(0)^{\frac{n-6}{2}}|\log \lambda_{i1}(0)|\right] \\
&\le 2\lambda_{i1}(0),
\end{aligned} \tag{30.11}$$

where we have used the fact that $\lambda_{i1}(0) \ll 1$ in the last step.

The reduction equation for $\xi_{i1}(t)$ is

$$|\xi_{i1}'(t)| \lesssim \lambda_{i1}(t)^{\frac{n-4}{2}}.$$

Integrating this on $[0, T_i]$ and using (30.11) and (30.8) again, we get

$$|\xi_{i1}(T_i) - \xi_{i1}(0)| \lesssim \lambda_{i1}(0)^{\frac{n-4}{2}}|\log \lambda_{i1}(0)| \ll 1.$$

The same estimates hold for ξ_{i2}. Combining these two estimates with the second equality in (30.9) (and a similar estimate for ξ_{i2}), we obtain

$$|\xi_{i1}^*(T_i) - \xi_{i2}^*(T_i)| \ge |\xi_{i1}^*(0) - \xi_{i2}^*(0)| - \frac{1}{3} = \frac{2}{3}|\xi_{i1}^*(0) - \xi_{i2}^*(0)|.$$

Scaling back to u, this is

$$\rho(t_{i+1}) = |\xi_1(t_{i+1}) - \xi_2(t_{i+1})| \ge \frac{2}{3}|\xi_1(\widetilde{t}_i) - \xi_2(\widetilde{t}_i)| = \frac{2}{3}\rho(t_i). \qquad \square$$

31. Exclusion of blow up with only one bubble

In this section, we prove

Proposition 31.1. *Under the setting of Theorem 26.1, any point $a \in \mathcal{S}(u)$ must be Type I, in the sense that*

$$\Theta(a) = n^{-1} \left(\frac{n-2}{4} \right)^{\frac{n}{2}}.$$

In view of Proposition 30.2, we need to consider the remaining case in Proposition 30.1, that is, when there is only one bubble, $N = 1$. Under this assumption, Proposition 30.1 now reads as

Proposition 31.2. *For any t sufficiently close to 0, there exists a unique maxima point of $u(\cdot, t)$ in $B_1(0)$. Denote this point by $\xi^*(t)$ and let $\lambda^*(t) := u(\xi^*(t), t)^{-\frac{2}{n-2}}$.*
As $t \to 0$,

$$\lambda^*(t) \to 0, \quad |\xi^*(t)| \to 0,$$

and the function

$$u^t(y, s) := \lambda^*(t)^{\frac{n-2}{2}} u\left(\xi^*(t) + \lambda^*(t)y, t + \lambda^*(t)^2 s\right),$$

converges to $W(y)$ in $C^\infty_{loc}(\mathbb{R}^n \times \mathbb{R})$.

Next, applying Proposition 12.1 (with the notation of η_K etc. as used in Part 2) to a suitable rescaling of u at $(\xi^*(t), t)$ gives

Proposition 31.3 (Orthogonal condition)**.** *For any t sufficiently close 0, there exists a unique $(\xi(t), \lambda(t), a(t)) \in \mathbb{R}^n \times \mathbb{R}_+ \times \mathbb{R}$ with*

$$\frac{|\xi(t) - \xi^*(t)|}{\lambda(t)} + \left| \frac{\lambda(t)}{\lambda^*(t)} - 1 \right| + \left| \frac{a(t)}{\lambda(t)} \right| = o(1), \tag{31.1}$$

such that for any $i = 0, \cdots, n+1$,

$$\int_{B_1} \left[u(x, t) - W_{\xi(t), \lambda(t)}(x) - a(t) Z_{0, \xi(t), \lambda(t)}(x) \right] \eta_K \left(\frac{x - \xi(t)}{\lambda(t)} \right) Z_{i, \xi(t), \lambda(t)}(x) dx = 0. \tag{31.2}$$

Starting with this decomposition, the analysis in Part 2 is applicable to

$$\widetilde{u}^t(y, s) := (-t)^{\frac{n-2}{4}} u\left(\sqrt{-t}y, -ts\right), \quad \forall t \in (-1/4, 0).$$

This gives an ordinary differential inequality for $\lambda(t)$:

$$|\lambda'(t)| \lesssim (-t)^{\frac{n-10}{4}} \lambda(t)^{\frac{n-4}{2}}. \tag{31.3}$$

This inequality can be rewritten as

$$\left| \frac{d}{dt} \lambda(t)^{-\frac{n-6}{2}} \right| \lesssim (-t)^{\frac{n-10}{4}} \tag{31.4}$$

Because $n \geq 7$,

$$\frac{n-10}{4} > -1.$$

Integrating (31.4) leads to

$$\lim_{t \to 0} \lambda(t)^{-\frac{n-6}{2}} < +\infty.$$

Thus as $t \to 0^-$, $\lambda(t)$ does not converges to 0. By (31.1) and Proposition 31.2,

$$\limsup_{t \to 0^-} \sup_{B_1} u(x,t) < +\infty.$$

Hence u does not blow up at time $t = 0$. This is a contradiction.

32. Proof of main results

In this section we first finish the proof of Theorem 26.1, and then show how Theorem 1.1, Theorem 1.2 and Corollary 1.3 follow from this theorem.

32.1. **Proof of Theorem 26.1.** Combining Proposition 31.1 with the tangent flow analysis in Section 28, we deduce that for any $a \in \mathcal{S}(u)$,

$$\limsup_{t \to T}(T-t)^{\frac{1}{p-1}} u\left(a + \sqrt{T-t}y, t + (T-t)s\right) = (p-1)^{-\frac{1}{p-1}}(-s)^{\frac{1}{p-1}} \qquad (32.1)$$

uniformly in any compact set of $\mathbb{R}^n \times \mathbb{R}^-$.

However, this is only a qualitative description and we cannot obtain (26.2) from it. Now we show how to obtain a quantitative estimate. We will mainly use the stability of Type I blow ups, as characterized in (29.1). (For other notions of stability of Type I blow ups, see Merle-Zaag [66], Kammerer -Merle-Zaag [29], Collot-Merle–Raphaël [12], Collot-Raphaël-Szeftel [14].)

Lemma 32.1. *For each $a \in B_{1/2}$, there exist two constants $C(a) > 1$ and $\rho(a) < 1/8$ such that for any $0 < s < \rho(a)^2$,*

$$\sup_{x \in B_{\rho(a)}(a)} u(x, T-s) \leq C(a)s^{-\frac{1}{p-1}}. \qquad (32.2)$$

Proof. If $a \in \mathcal{R}(u)$, by definition, there exists an $\rho(a) > 0$ such that $u \in L^\infty(Q_{\rho(a)}^-(a,T))$. Then (32.2) holds trivially by choosing a suitable constant $C(a)$.

Next, assume $a \in \mathcal{S}(u)$. By (29.1), we can take a small $\varepsilon > 0$ so that

$$n^{-1}\left(\frac{n-2}{4}\right)^{\frac{n}{2}} + 4\varepsilon < n^{-1}(4\pi)^{-\frac{n}{2}}\Lambda. \qquad (32.3)$$

By (32.1), there exists an $r(a) > 0$ such that

$$\Theta_{r(a)^2}(a) < n^{-1}\left(\frac{n-2}{4}\right)^{\frac{n}{2}} + \varepsilon. \qquad (32.4)$$

Next, combining (32.1) with Proposition 27.3 and the monotonicity formula (Proposition 27.1), we find another constant $\rho(a) < r(a)$ such that

$$\Theta_{r(a)^2}(x) \leq \Theta_{r(a)^2}(a) + \varepsilon, \quad \forall x \in B_{2\rho(a)}(a).$$

Substituting this inequality into (32.4) and noting (32.3), we get

$$\Theta_{r(a)^2}(x) < n^{-1}(4\pi)^{-\frac{n}{2}}\Lambda - 2\varepsilon, \quad \forall x \in B_{2\rho(a)}(a). \qquad (32.5)$$

In the following, we argue by contradiction. Assume with $\rho(a)$ defined as above, there does not exist such a constant $C(a)$ so that (32.2) can hold, that is, there exists a sequence of $x_k \in B_{\rho(a)}(a)$, $s_k \to 0$ with

$$u(x_k, T - s_k) \geq k s_k^{-\frac{1}{p-1}}. \tag{32.6}$$

Define the blow up sequence

$$u_k(x, t) := s_k^{\frac{1}{p-1}} u(x_k + \sqrt{s_k} y, T + s_k t).$$

As in Section 28, by applying results in Part 1, we may assume u_k converges weakly to u_∞, and

$$|\nabla u_k|^2 dx dt \rightharpoonup |\nabla u_\infty|^2 dx dt + \mu$$

weakly as Radon measures in $\mathbb{R}^n \times \mathbb{R}^-$. By (32.6),

$$u_k(0, -1) \geq k.$$

Hence u_k cannot converge smoothly to u_∞. On the other hand, we claim that u_k does converge smoothly to u_∞. This contradiction then finishes the proof of this lemma.

Proof of the claim. We first prove $\mu = 0$. By scaling (32.5) and passing to the limit, we get

$$\Theta_s(x, 0; u_\infty, \mu) \leq n^{-1} (4\pi)^{-\frac{n}{2}} \Lambda - 2\varepsilon, \quad \forall x \in \mathbb{R}^n, \quad s > 0. \tag{32.7}$$

By Corollary 9.7, for μ-a.e. (x, t),

$$\Theta(x, t; u_\infty, \mu) \geq n^{-1} (4\pi)^{-\frac{n}{2}} \Lambda. \tag{32.8}$$

On the other hand, if s is large enough, by the Morrey space bound as in (5.1), we get

$$\Theta_s(x, t; u_\infty, \mu) \leq \Theta_s(x, 0; u_\infty, \mu) + \varepsilon. \tag{32.9}$$

By the monotonicity of $\Theta_s(\cdot)$, these three inequalities lead to a contradiction unless $\mu = 0$.

Next we show that $u_\infty \in C^\infty(\mathbb{R}^n \times \mathbb{R}^-)$. First, because $\mu = 0$, as in Theorem 6.1, we deduce that u_∞ is a suitable weak solution of (1.1) in $\mathbb{R}^n \times \mathbb{R}^-$. If there is a singular point of u_∞, say (x, t), by the analysis in Subsection 9.2 and Theorem 9.1, any tangent flow of u_∞ at (x, t) would be of the form

$$N_1 \Lambda \delta_0 \otimes dt \lfloor_{\mathbb{R}^-} + N_2 \Lambda \delta_0 \otimes dt \lfloor_{\mathbb{R}^+},$$

where $N_1 \geq 1$, $N_2 \geq 0$. Then by Corollary 9.7, we get

$$\Theta(x, t; u_\infty) \geq n^{-1} (4\pi)^{-\frac{n}{2}} \Lambda.$$

Using this inequality to replace (32.8) and aruging as above, we get a contradiction with (32.7). In conclusion, the singular set of u_∞ is empty.

Now $\mu = 0$ implies that u_k converges to u_∞ strongly. The fact that u_∞ is smooth, implies that, for any $(x, t) \in \mathbb{R}^n \times \mathbb{R}^-$, we can apply the ε-regularity theorem, Theorem 3.9, to u_k in a sufficiently small cylinder $Q_r^-(x, t)$. This implies that u_k converges to u_∞ in $C^\infty(\mathbb{R}^n \times \mathbb{R}^-)$. $\qquad \square$

These balls $\{B_{\rho(a)}(a)\}$ form a covering of $B_{1/2}$. Take a finite sub-cover of $B_{1/2}$ from $\{B_{\rho(a)}(a)\}$, $\{B_{\rho_i}(a_i)\}$. For

$$t \geq T - \min_i \rho_i^2,$$

(32.2) is a direct consequence of (26.2), if we choose the constant in (32.2) to be

$$C := \max_i C(a_i).$$

Since (26.2) holds trivially for $t \leq T - \min_i \rho_i^2$, we finish the proof of Theorem 26.1.

32.2. Cauchy-Dirichlet problems. Define

$$H(t) := \int_\Omega u(x,t)^2 dx, \quad E(t) := \int_\Omega \left(\frac{|\nabla u(x,t)|^2}{2} - \frac{|u(x,t)|^{p+1}}{p+1} \right) dx.$$

The following estimate is Blatt-Struwe's [4, Lemma 6.3].

Lemma 32.2. *There exists a constant C such that for any $t \in (T/2, T)$,*

$$\begin{cases} H(t) \leq C(T-t)^{-\frac{2}{p-1}}, \\ \int_{T/2}^t \int_\Omega \left[|\nabla u(y,s)|^2 + |u(y,s)|^{p+1} \right] dy ds \leq C(T-t)^{-\frac{2}{p-1}}. \end{cases}$$

With this estimate in hand, we know that u locally satisfies the assumptions in Section 26. Theorem 1.2 then follows from Theorem 26.1 and a covering argument.

32.3. Cauchy problem. Now we come to the proof of Theorem 1.1. As in [39], by the regularizing effect of parabolic equations, we may assume $u_0 \in C^2(\mathbb{R}^n)$.

For each $a \in \mathbb{R}^n$ and $s \in [0, T)$, define

$$\begin{aligned} \Theta_s(a) \quad := \quad & s^{\frac{p+1}{p-1}} \int_{\mathbb{R}^n} \left[\frac{|\nabla u(y, T-s)|^2}{2} - \frac{|u(y, T-s)|^{p+1}}{p+1} \right] G(y-x, s) dy \\ + \quad & \frac{1}{2(p-1)} s^{\frac{2}{p-1}} \int_{\mathbb{R}^n} u(y, T-s)^2 G(y-x, s) dy. \end{aligned}$$

Then we have (see the summarization in [39, Section 3.2])

(1) $\Theta_s(a)$ is non-decreasing in s;
(2) for any $a \in \mathbb{R}^n$, $r \in (0, \sqrt{T}/2)$ and $\theta \in (0, 1/2)$,

$$\begin{cases} \int_{Q_r(a, T-\theta r^2)} |\nabla u|^2 + |u|^{p+1} \leq C(\theta) \max \left\{ \Theta_T(a), \Theta_T(a)^{\frac{2}{p+1}} \right\} r^2, \\ \int_{Q_r(a, T-\theta r^2)} |\partial_t u|^2 \leq C(\theta) \max \left\{ \Theta_T(a), \Theta_T(a)^{\frac{2}{p+1}} \right\}. \end{cases}$$

With these estimates in hand, we can repeat the proof of Theorem 26.1 to get Theorem 1.1.

32.4. **Energy collapsing and the blow up of L^{p+1} norm.** In this subsection, we prove Corollary 1.3.

Recall that we have assumed that there exists a blow up point $a \in \Omega$. We define the blow up sequence u_λ at (a, T) as in Section 28. Combining Proposition 28.1 with Proposition 31.1, we see

$$u_\lambda(x, t) \to u_\infty(x, t) = (p-1)^{-\frac{1}{p-1}}(-t)^{-\frac{1}{p-1}} \quad \text{in } C^\infty_{loc}(\mathbb{R}^n \times \mathbb{R}^-).$$

Hence there exists a universal constant $c > 0$ such that

$$\lim_{\lambda \to 0} \int_{Q^-_\lambda(a, T-\lambda^2)} |\partial_t u|^2 = \lim_{\lambda \to 0} \int_{Q^-_1(0,-1)} |\partial_t u_\lambda|^2 \geq \int_{Q_1(0,-1)} |\partial_t u_\infty|^2 \geq c.$$

Taking a subsequence λ_i so that $Q^-_{\lambda_i}(a, T - \lambda_i^2)$ are disjoint from each other. Then

$$\lim_{t \to T} \int_0^t \int_\Omega |\partial_t u|^2 \geq \sum_{i=1}^{+\infty} \int_{Q^-_{\lambda_i}(a, T-\lambda_i^2)} |\partial_t u|^2 = +\infty.$$

Then by the energy identity for u, we get (1.6).

In the same way, there exists a universal constant $c > 0$ such that, for any $R > 1$,

$$\lim_{\lambda \to 0} \int_{B_{\lambda R}(a, T-\lambda^2)} u(x,t)^{p+1} dx = \lim_{\lambda \to 0} \int_{B_R} u_\lambda(x,t)^{p+1} dx \geq cR^n.$$

Therefore

$$\lim_{t \to T} \int_\Omega u(x,t)^{p+1} dx \geq \lim_{t \to T} \int_{B_{R\sqrt{T-t}}(a,T)} u(x,t)^{p+1} dx \geq cR^n.$$

Because R can be arbitrarily large, (1.7) follows.

REFERENCES

[1] Luigi Ambrosio and Halil Mete Soner. A measure-theoretic approach to higher codimension mean curvature flows. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 25(1-2):27–49 (1998), 1997. Dedicated to Ennio De Giorgi.

[2] Adolfo Arroyo-Rabasa, Guido De Philippis, Jonas Hirsch, and Filip Rindler. Dimensional estimates and rectifiability for measures satisfying linear PDE constraints. *Geom. Funct. Anal.*, 29(3):639–658, 2019.

[3] P. Baras and L. Cohen. Complete blow-up after $T_{\max}$ for the solution of a semilinear heat equation. *J. Funct. Anal.*, 71(1):142–174, 1987.

[4] Simon Blatt and Michael Struwe. An analytic framework for the supercritical Lane-Emden equation and its gradient flow. *Int. Math. Res. Not. IMRN*, 2015(9):2342–2385, 2015.

[5] Simon Blatt and Michael Struwe. Well-posedness of the supercritical Lane-Emden heat flow in Morrey spaces. *ESAIM Control Optim. Calc. Var.*, 22(4):1370–1381, 2016.

[6] Kenneth A. Brakke. *The motion of a surface by its mean curvature*, volume 20 of *Mathematical Notes*. Princeton University Press, Princeton, N.J., 1978.

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

[8] 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.

[9] Kai-Seng Chou, Shi-Zhong Du, and Gao-Feng Zheng. On partial regularity of the borderline solution of semilinear parabolic problems. *Calc. Var. Partial Differential Equations*, 30(2):251–275, 2007.

[10] Charles Collot. Nonradial type II blow up for the energy-supercritical semilinear heat equation. *Anal. PDE*, 10(1):127–252, 2017.

[11] 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.

[12] Charles Collot, Frank Merle, and Pierre Raphaël. Stability of ODE blow-up for the energy critical semilinear heat equation. *C. R. Math. Acad. Sci. Paris*, 355(1):65–79, 2017.

[13] Charles Collot, Frank Merle, and Pierre Raphaël. Strongly anisotropic type II blow up at an isolated point. *J. Amer. Math. Soc.*, 33(2):527–607, 2020.

[14] Charles Collot, Pierre Raphaël, and Jeremie Szeftel. On the stability of type I blow up for the energy super critical heat equation. *Mem. Amer. Math. Soc.*, 260(1255):v+97, 2019.

[15] Carmen Cortázar, Manuel del Pino, and Monica Musso. Green's function and infinite-time bubbling in the critical nonlinear heat equation. *J. Eur. Math. Soc. (JEMS)*, 22(1):283–344, 2020.

[16] Juan Dávila, Manuel del Pino, and Juncheng Wei. Singularity formation for the two-dimensional harmonic map flow into $\mathbb{S}^2$. *Invent. Math.*, 219(2):345–466, 2020.

[17] Manuel del Pino. Bubbling blow-up in critical parabolic problems. In *Nonlocal and nonlinear diffusions and interactions: new methods and directions*, volume 2186 of *Lecture Notes in Math.*, pages 73–116. Springer, Cham, 2017.

[18] Manuel del Pino, Chen-Chih Lai, Monica Musso, Juncheng Wei, and Yifu Zhou. New type II finite time blow-up for the energy supercritical heat equation. *arXiv: Analysis of PDEs*, 2020.

[19] Manuel del Pino, Monica Musso, and Jun Cheng Wei. Type II blow-up in the 5-dimensional energy critical heat equation. *Acta Math. Sin. (Engl. Ser.)*, 35(6):1027–1042, 2019.

[20] Manuel del Pino, Monica Musso, and Juncheng Wei. Infinite-time blow-up for the 3-dimensional energy-critical heat equation. *Anal. PDE*, 13(1):215–274, 2020.

[21] Manuel del Pino, Monica Musso, and Juncheng Wei. Geometry driven type II higher dimensional blow-up for the critical heat equation. *J. Funct. Anal.*, 280(1):108788, 49, 2021.

[22] Manuel del Pino, Monica Musso, Juncheng Wei, Qidi Zhang, and Yifu Zhang. Type II finite time blow-up for the three dimensional energy critical heat equation. *arXiv: Analysis of PDEs*, 2020.

[23] Manuel del Pino, Monica Musso, Juncheng Wei, and Yifu Zhou. Type II finite time blow-up for the energy critical heat equation in $\mathbb{R}^4$. *Discrete Contin. Dyn. Syst.*, 40(6):3327–3355, 2020.

[24] Olivier Druet, Emmanuel Hebey, and Frédéric Robert. *Blow-up theory for elliptic PDEs in Riemannian geometry*, volume 45 of *Mathematical Notes*. Princeton University Press, Princeton, NJ, 2004.

[25] 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.

[26] Shi-Zhong Du. The singular set of stationary solution for semilinear elliptic equations with supercritical growth. *J. Differential Equations*, 256(7):2392–2405, 2014.

[27] Shi-Zhong Du. Energy non-collapsing and refined blowup for a semilinear heat equation. *J. Differential Equations*, 266(9):5942–5970, 2019.

[28] Herbert Federer and William P. Ziemer. The Lebesgue set of a function whose distribution derivatives are p-th power summable. *Indiana Univ. Math. J.*, 22:139–158, 1972/73.

[29] C. Fermanian Kammerer, F. Merle, and H. Zaag. Stability of the blow-up profile of non-linear heat equations from the dynamical system point of view. *Math. Ann.*, 317(2):347–387, 2000.

[30] Stathis Filippas, Miguel A. Herrero, and Juan J. L. Velázquez. Fast blow-up mechanisms for sign-changing solutions of a semilinear parabolic equation with critical nonlinearity. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 456(2004):2957–2982, 2000.

[31] Stathis Filippas and Robert V. Kohn. Refined asymptotics for the blowup of $u_t - \Delta u = u^p$. *Comm. Pure Appl. Math.*, 45(7):821–869, 1992.

[32] Stathis Filippas and Wen Xiong Liu. On the blowup of multidimensional semilinear heat equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 10(3):313–344, 1993.

[33] Hiroshi Fujita. On the blowing up of solutions of the Cauchy problem for $u_t = \Delta u + u^{1+\alpha}$. *J. Fac. Sci. Univ. Tokyo Sect. I*, 13:109–124 (1966), 1966.

[34] Victor A. Galaktionov and Juan L. Vazquez. Continuation of blowup solutions of nonlinear heat equations in several space dimensions. *Comm. Pure Appl. Math.*, 50(1):1–67, 1997.

[35] Yoshikazu Giga. A bound for global solutions of semilinear heat equations. *Comm. Math. Phys.*, 103(3):415–421, 1986.

[36] Yoshikazu Giga and Robert V. Kohn. Asymptotically self-similar blow-up of semilinear heat equations. *Comm. Pure Appl. Math.*, 38(3):297–319, 1985.

[37] Yoshikazu Giga and Robert V. Kohn. Characterizing blowup using similarity variables. *Indiana Univ. Math. J.*, 36(1):1–40, 1987.

[38] Yoshikazu Giga and Robert V. Kohn. Nondegeneracy of blowup for semilinear heat equations. *Comm. Pure Appl. Math.*, 42(6):845–884, 1989.

[39] Yoshikazu Giga, Shin'ya Matsui, and Satoshi Sasayama. Blow up rate for semilinear heat equations with subcritical nonlinearity. *Indiana Univ. Math. J.*, 53(2):483–514, 2004.

[40] Yoshikazu Giga, Shin'ya Matsui, and Satoshi Sasayama. On blow-up rate for sign-changing solutions in a convex domain. *Math. Methods Appl. Sci.*, 27(15):1771–1782, 2004.

[41] Junichi Harada. A higher speed type II blowup for the five dimensional energy critical heat equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 37(2):309–341, 2020.

[42] Junichi Harada. A type II blowup for the six dimensional energy critical heat equation. *Ann. PDE*, 6(2):Paper No. 13, 63, 2020.

[43] Miguel A. Herrero and Juan J. L. Velázquez. Explosion de solutions d'équations paraboliques semilinéaires supercritiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 319(2):141–145, 1994.

[44] John E. Hutchinson. Second fundamental form for varifolds and the existence of surfaces minimising curvature. *Indiana Univ. Math. J.*, 35(1):45–71, 1986.

[45] Kota Kasai and Yoshihiro Tonegawa. A general regularity theory for weak mean curvature flow. *Calc. Var. Partial Differential Equations*, 50(1-2):1–68, 2014.

[46] M. A. Khuri, F. C. Marques, and R. M. Schoen. A compactness theorem for the Yamabe problem. *J. Differential Geom.*, 81(1):143–196, 2009.

[47] Yan Yan Li. Harnack type inequality: the method of moving planes. *Comm. Math. Phys.*, 200(2):421–444, 1999.

[48] Yan Yan Li and Lei Zhang. Compactness of solutions to the Yamabe problem. II. *Calc. Var. Partial Differential Equations*, 24(2):185–237, 2005.

[49] YanYan Li and Lei Zhang. Liouville-type theorems and Harnack-type inequalities for semilinear elliptic equations. *J. Anal. Math.*, 90:27–87, 2003.

[50] Yanyan Li and Meijun Zhu. Yamabe type equations on three-dimensional Riemannian manifolds. *Commun. Contemp. Math.*, 1(1):1–50, 1999.

[51] F.-H. Lin and Q. S. Zhang. On ancient solutions of the heat equation. *Comm. Pure Appl. Math.*, 72(9):2006–2028, 2019.

[52] Fang-Hua Lin. Gradient estimates and blow-up analysis for stationary harmonic maps. *Ann. of Math. (2)*, 149(3):785–829, 1999.

[53] Fang-Hua Lin and Tristan Rivière. Energy quantization for harmonic maps. *Duke Math. J.*, 111(1):177–193, 2002.

[54] Fanghua Lin. Mapping problems, fundamental groups and defect measures. *Acta Math. Sin. (Engl. Ser.)*, 15(1):25–52, 1999.

[55] FangHua Lin and ChangYou Wang. Harmonic and quasi-harmonic spheres. *Comm. Anal. Geom.*, 7(2):397–429, 1999.

[56] FangHua Lin and ChangYou Wang. Harmonic and quasi-harmonic spheres. II. *Comm. Anal. Geom.*, 10(2):341–375, 2002.

[57] FangHua Lin and ChangYou Wang. Harmonic and quasi-harmonic spheres. III. Rectifiability of the parabolic defect measure and generalized varifold flows. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 19(2):209–259, 2002.

[58] Fanghua Lin and Changyou Wang. *The analysis of harmonic maps and their heat flows*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008.

[59] Fanghua Lin and Xiaoping Yang. *Geometric measure theory—an introduction*, volume 1 of *Advanced Mathematics (Beijing/Boston)*. Science Press Beijing, Beijing; International Press, Boston, MA, 2002.

[60] Wen Xiong Liu. Blow-up behavior for semilinear heat equations: multi-dimensional case. *Rocky Mountain J. Math.*, 23(4):1287–1319, 1993.

[61] John M. Marstrand. The (φ, s) regular subsets of n-space. *Trans. Amer. Math. Soc.*, 113:369–392, 1964.

[62] Yvan Martel. Complete blow up and global behaviour of solutions of $u_t - \Delta u = g(u)$. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 15(6):687–723, 1998.

[63] Hiroshi Matano and Frank Merle. On nonexistence of type II blowup for a supercritical nonlinear heat equation. *Comm. Pure Appl. Math.*, 57(11):1494–1541, 2004.

[64] Hiroshi Matano and Frank Merle. Classification of type I and type II behaviors for a supercritical nonlinear heat equation. *J. Funct. Anal.*, 256(4):992–1064, 2009.

[65] F. Merle and H. Zaag. Refined uniform estimates at blow-up and applications for nonlinear heat equations. *Geom. Funct. Anal.*, 8(6):1043–1085, 1998.

[66] Frank Merle and Hatem Zaag. Stability of the blow-up profile for equations of the type $u_t = \Delta u + |u|^{p-1}u$. *Duke Math. J.*, 86(1):143–195, 1997.

[67] Frank Merle and Hatem Zaag. Optimal estimates for blowup rate and behavior for nonlinear heat equations. *Comm. Pure Appl. Math.*, 51(2):139–196, 1998.

[68] Frank Merle and Hatem Zaag. A Liouville theorem for vector-valued nonlinear heat equations and applications. *Math. Ann.*, 316(1):103–137, 2000.

[69] Noriko Mizoguchi. Boundedness of global solutions for a supercritical semilinear heat equation and its application. *Indiana Univ. Math. J.*, 54(4):1047–1059, 2005.

[70] Noriko Mizoguchi and Philippe Souplet. Optimal condition for blow-up of the critical L^q norm for the semilinear heat equation. *Adv. Math.*, 355:106763, 24, 2019.

[71] Luisa Moschini and Alberto Tesei. Harnack inequality and heat kernel estimates for the Schrödinger operator with Hardy potential. *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.*, 16(3):171–180 (2006), 2005.

[72] Roger Moser. Stationary measures and rectifiability. *Calc. Var. Partial Differential Equations*, 17(4):357–368, 2003.

[73] Monica Musso, Manuel Del Pino, Juncheng Wei, and Youquan Zheng. Sign-changing blowing-up solutions for the critical nonlinear heat equation. *Annali Della Scuola Normale Superiore Di Pisa-classe Di Scienze*, page 1, 2019.

[74] Aaron Naber and Daniele Valtorta. Energy identity for stationary Yang Mills. *Invent. Math.*, 216(3):847–925, 2019.

[75] Pavol Quittner and Philippe Souplet. *Superlinear parabolic problems*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser/Springer, Cham, 2019. Blow-up, global existence and steady states, Second edition of [MR2346798].

[76] Laurent Saloff-Coste. *Aspects of Sobolev-type inequalities*, volume 289 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2002.

[77] Richard M. Schoen. Courses at Stanford University. 1988.

[78] Richard M. Schoen. Variational theory for the total scalar curvature functional for Riemannian metrics and related topics. In *Topics in calculus of variations (Montecatini Terme, 1987)*, volume 1365 of *Lecture Notes in Math.*, pages 120–154. Springer, Berlin, 1989.

[79] Rémi Schweyer. Type II blow-up for the four dimensional energy critical semi linear heat equation. *J. Funct. Anal.*, 263(12):3922–3983, 2012.

[80] Yukihiro Seki. Type II blow-up mechanisms in a semilinear heat equation with critical Joseph-Lundgren exponent. *J. Funct. Anal.*, 275(12):3380–3456, 2018.

[81] Philippe Souplet. Morrey spaces and classification of global solutions for a supercritical semilinear heat equation in $\mathbb{R}^n$. *J. Funct. Anal.*, 272(5):2005–2037, 2017.

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

[83] J. J. L. Velázquez. Higher-dimensional blow up for semilinear parabolic equations. *Comm. Partial Differential Equations*, 17(9-10):1567–1596, 1992.

[84] J. J. L. Velázquez. Classification of singularities for blowing up solutions in higher dimensions. *Trans. Amer. Math. Soc.*, 338(1):441–464, 1993.

[85] J. J. L. Velázquez. Estimates on the $(n-1)$-dimensional Hausdorff measure of the blow-up set for a semilinear heat equation. *Indiana Univ. Math. J.*, 42(2):445–476, 1993.

[86] Kelei Wang and Juncheng Wei. Analysis of blow-up locus and existence of weak solutions for nonlinear supercritical problems. *Int. Math. Res. Not. IMRN*, (10):2634–2670, 2015.

[87] Kelei Wang and Juncheng Wei. Finite Morse index implies finite ends. *Comm. Pure Appl. Math.*, 72(5):1044–1119, 2019.

[88] Kelei Wang and Juncheng Wei. Second order estimate on transition layers. *Adv. Math.*, 358:106856, 85, 2019.

[89] Juncheng Wei and Shusen Yan. Infinitely many solutions for the prescribed scalar curvature problem on $\mathbb{S}^N$. *J. Funct. Anal.*, 258(9):3048–3081, 2010.

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

[91] Brian White. Stratification of minimal surfaces, mean curvature flows, and harmonic maps. *J. Reine Angew. Math.*, 488:1–35, 1997.

[92] D. V. Widder. The role of the Appell transformation in the theory of heat conduction. *Trans. Amer. Math. Soc.*, 109:121–134, 1963.

[93] Hatem Zaag. A remark on the energy blow-up behavior for nonlinear heat equations. *Duke Math. J.*, 103(3):545–556, 2000.

[94] Hatem Zaag. On the regularity of the blow-up set for semilinear heat equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 19(5):505–542, 2002.

[95] Hatem Zaag. One-dimensional behavior of singular N-dimensional solutions of semilinear heat equations. *Comm. Math. Phys.*, 225(3):523–549, 2002.

[96] Hatem Zaag. Determination of the curvature of the blow-up set and refined singular behavior for a semilinear heat equation. *Duke Math. J.*, 133(3):499–525, 2006.

†School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China
E-mail address: wangkelei@whu.edu.cn

∗Department of Mathematics, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z2.
E-mail address: jcwei@math.ubc.ca