# STRUCTURE THEORY OF PARABOLIC NODAL AND SINGULAR SETS

MAX HALLGREN, ROBERT KOIRALA, AND ZILU MA

Abstract. We establish new estimates for the size and structure of the nodal set $\{u = 0\}$ and the singular set $\{u = |\nabla u| = 0\}$ of solutions $u$ to parabolic inequalities with parabolic Lipschitz coefficients. In particular, we show that almost all of the nodal and singular sets are covered by *regular* parabolic Lipschitz graphs with estimates, and that both sets satisfy parabolic Minkowski estimates depending only on a doubling quantity at a point. Many of our results are new even for the heat equation on $\mathbb{R}^n \times \mathbb{R}$.

Contents



*Date*: November 21, 2025.

## 1. Introduction

Let $u$ be a non-trivial solution to a parabolic inequality with parabolic Lipschitz coefficients defined on an open subset of $\mathbb{R}^n \times \mathbb{R}$:

$$
\begin{gathered}
|\partial_t u - \operatorname{div}(a \cdot \nabla u)| \le M(|u| + |\nabla u|), \\
(1+M)^{-1}\delta^{ij} \le a^{ij} \le (1+M)\delta^{ij}, \\
|a^{ij}(x,t) - a^{ij}(y,s)| \le M \max\{|x-y|, \sqrt{|t-s|}\},
\end{gathered} \tag{1.1}
$$

where $a^{ij} = a^{ji}$ and $M \ge 0$.

In this paper, we study the geometric and measure-theoretic structure of the nodal and singular sets of $u$,

$$
Z(u) \coloneqq \{u = 0\}, \qquad S(u) \coloneqq \{u = |\nabla u| = 0\}. \tag{1.2}
$$

Our analysis assumes one of several doubling conditions on $u$, which excludes pathological cases such as solutions that vanish identically on a time slab (see, for instance, [DSP05]). Under these assumptions, we show that both $Z(u)$ and $S(u)$ can be covered, up to negligible subsets, by regular parabolic Lipschitz graphs over affine subspaces of dimensions $(n+1)$ and $n$, respectively. Furthermore, we establish sharp $n+1$ and $n$-dimensional parabolic Minkowski content estimates for $Z(u)$ and $S(u)$, respectively, depending on the doubling condition. As a consequence, we obtain new $(n-2)$-dimensional bounds for $S(u)$ on almost every time slice. Precise statements of our results are presented in Section 1.2, while Section 1.3 provides an outline of the proofs and highlights the main new ideas.

### 1.1. Background.

To place our setting in context, consider first the case $M = 0$, when $u$ satisfies the heat equation

$$
\partial_t u = \Delta u. \tag{1.3}
$$

Such functions $u$ are called *caloric.* If $u$ is globally defined and satisfies appropriate growth bounds, then $u$ is real analytic in both space and time variables. Consequently, both $Z(u)$ and $S(u)$ are real analytic subsets of $\mathbb{R}^n \times \mathbb{R}$ of dimension $n$ and $n-1$, respectively. In particular (cf. [Fed14, §3.4.10]), each time slice $Z(u) \cap (\mathbb{R}^n \times \{t\})$ has locally finite $(n-1)$-dimensional Hausdorff measure for any $t \in \mathbb{R}$, and $Z(u)$ has locally finite $n$-dimensional Hausdorff measure. Similarly, $S(u)$ has locally finite $(n-1)$-dimensional Hausdorff measure.

Quantitative bounds for these measures, however, are substantially more difficult to obtain. For caloric polynomials of degree $m$, the Hausdorff measure of their nodal and singular sets grows polynomially on $m$ (see, for example, [Won93]). For general solutions $u$, one introduces quantities such as the *parabolic frequency* and the *doubling index* (see Definition 2.4), which play roles analogous to the degree of a caloric polynomial. Han and Lin in [HL94a] obtained $n$-dimensional Hausdorff estimates, depending on the doubling index, for nodal sets of solutions to more general parabolic equations. In [ABG25, Kuk95, Lin91], the authors established estimates for the time slices $Z(u) \cap (\mathbb{R}^n \times \{t\})$ of the nodal sets of solutions to parabolic equations with real-analytic coefficients. See [AB23, Ang91, CM11, CMI20, CNV15, DF88, DF90, Don92, HHL98, HHOHON99, HJ23, HL94a, HL94b, HS89, HS12, KZZ22, Lin91, LL16, Log18a, Log18b, LM18, LM20, LMNN21, LS19, LTY24, NV17b, SZ12, Zel13, Zel15] and the references therein for a non-exhaustive list of related works. More recently, Huang and Jiang in [HJ24] extended these results to obtain Hausdorff measure estimates for the time slices of the nodal sets of general solutions to (1.1), following the elliptic framework developed by Naber and Valtorta in [NV17b].

Further extensions of the work of Naber and Valtorta have remained open even in the caloric case. In the setting of (1.1), it is natural to formulate estimates using the parabolic metric $d_{\mathcal{P}}$ on $\mathbb{R}^n \times \mathbb{R}$ induced by the "norm"

$$|(x,t)|_{\mathcal{P}}^2 := \max\{|x|^2, |t|\},$$

which better reflects the underlying space-time geometry. With respect to $d_{\mathcal{P}}$, $\mathbb{R}^n \times \mathbb{R}$ has parabolic dimension $n+2$. Until now, parabolic Minkowski estimates for the nodal and singular sets have been unavailable; in particular, even Hausdorff estimates for the singular sets were unknown.

A related question concerns parabolic rectifiability. Following [Mat22], one may ask whether $Z(u)$ and $S(u)$ are covered, up to negligible sets, by a union of parabolic Lipschitz graphs. In the PDE literature, a stronger notion of Lipschitz property called the *regular* parabolic Lipschitz graphs exists, which requires estimates on the half-time derivative. Such graphs play for parabolic problems the same role as Euclidean Lipschitz graphs play for elliptic equations. In fact, it was shown in [BHMN25] that if $\Omega$ has boundary locally given by a Lipschitz graph, then the $L^p$ Dirichlet boundary problem is well-posed if and only if the Lipschitz graphs are regular.

In this paper, we prove that $Z(u)$ and $S(u)$ satisfy the stronger notion of rectifiability as well as the aforementioned parabolic Minkowski estimates.

## 1.2. **Main results.**

1.2.1. *General setting.* We now state our main results, first assuming that $u$ solves (1.1) on $P(\mathbf{0}, 5)$, where

$$P(\mathbf{x}, r) := B(x,r) \times (t - r^2, t + r^2) = \{\mathbf{y} \in \mathbb{R}^{n+1} : |\mathbf{y} - \mathbf{x}|_{\mathcal{P}} < r\}$$

denotes the parabolic ball centered at $\mathbf{x} = (x,t)$. For $A \subseteq \mathbb{R}^{n+1}$, we write $P(A, r)$ for the $r$-neighborhood of $A$ with respect to $d_{\mathcal{P}}$. We also assume that $u$ satisfies the *doubling condition*:

$$\Theta := \sup_{t \in [-2,2]} \frac{\int_{B(0,5)\times[-25,t]} u^2 \, d\mathcal{H}_{\mathcal{P}}^{n+2}}{\int_{B(0,\frac{1}{2})} u^2(x,t)\, dx} < \infty. \tag{1.4}$$

Asking that $\Theta$ is finite is equivalent to asking that $u$ does not vanish identically on any time slice (see [AV04, Fer03]).

For convenience, we set

$$\mathscr{C}^k(u) := \begin{cases} Z(u) & \text{when } k = n+1, \\ S(u) & \text{when } k = n. \end{cases}$$

We first establish that the nodal and singular sets are almost everywhere covered by regular parabolic Lipschitz graphs, with quantitative control on their content and on the mean oscillation of the half-time derivative. (See Definitions 3.1 and 4.13 for the precise notions.)

**Theorem 1.1.** *Suppose $u$ is a solution to (1.1) on $P(\mathbf{0}, 5)$ satisfying (1.4). For any $k \in \{n, n+1\}$ and any $\epsilon \in (0,1]$, there exist points $\mathbf{x}_i \in P(\mathbf{0}, \frac{1}{2})$, vertical planes $V_i \in \mathrm{Aff}_{\mathcal{P}}(k)$, radii $r_i \in (0,1]$, and $\epsilon$-regular parabolic Lipschitz functions $F_i \colon V_i \to V_i^{\perp}$ such that*

$$\sum_i r_i^k \le C(\epsilon, M, \Theta),$$

*and the graphs $G_i := \{\mathbf{x} + F_i(\mathbf{x}) \colon \mathbf{x} \in V_i \cap P(\mathbf{x}_i, 10r_i)\}$ satisfy*

$$\mathcal{H}^k_{\mathcal{P}}\left( (P(\mathbf{0}, \tfrac{1}{4}) \cap \mathscr{C}^k(u)) \setminus \bigcup_i G_i \right) = 0.$$

*Here $\mathcal{H}^k_{\mathcal{P}}$ denotes the $k$-dimensional parabolic Hausdorff measure induced by $d_{\mathcal{P}}$.*

*Remark* 1.2.
(1) Hence $\mathscr{C}^k(u)$ is vertically $k$-rectifiable (see Definition 6.3) in the sense of [Mat22, Definition 3.7].
(2) Regular parabolic Lipschitz graphs form a strictly smaller class than parabolic rectifiable or parabolic Lipschitz graphs. In fact, [Mat22, Example 8.2] constructs $\mathcal{H}^{n+1}$-measurable parabolic rectifiable subsets $E \subseteq \mathbb{R}^{n+1}$ with $\mathcal{H}^{n+1}_{\mathcal{P}}(E) \in (0, \infty)$ yet $\mathcal{H}^{n+1}_{\mathcal{P}}(E \cap G) = 0$ for every regular parabolic Lipschitz graph $G$.
(3) There is no analog of regular Lipschitz graphs in the Euclidean case: Euclidean Lipschitz functions are differentiable almost everywhere with uniformly bounded gradients, while parabolic Lipschitz functions need not admit half-time derivatives even almost everywhere.

As a consequence of the proof of Theorem 1.1, we get a qualitative regularity of the nodal and singular sets of solutions to (1.1).

**Corollary 1.3.** *Let $u$ be a solution to (1.1) on an open subset $\Omega \subset \mathbb{R}^n \times \mathbb{R}$ such that $u(\cdot, t)$ does not vanish identically on a non-empty connected component of $\Omega \cap (\mathbb{R}^n \times \{t\})$. Then there exists a countable collection of regular parabolic Lipschitz graphs $G_i$ such that*

$$\mathcal{H}^k_{\mathcal{P}}\left( \mathscr{C}^k(u) \setminus \bigcup_i G_i \right) = 0.$$

*In particular, $\mathscr{C}^k(u)$ is vertically parabolic $k$-rectifiable with locally finite $\mathcal{H}^k_{\mathcal{P}}$ measure.*

We next quantify the size of $\mathscr{C}^k$.

**Theorem 1.4.** *Let $u$ satisfy (1.1) and (1.4). Then, for any $0 < r \leq \overline{r}(M, \Theta)$,*

$$\mathcal{H}^{n+2}_{\mathcal{P}}(P(\mathscr{C}^k(u), r) \cap P(\mathbf{0}, \tfrac{1}{4})) \leq C(M, \Theta) r^{n+2-k}.$$

*Hence the parabolic Minkowski dimension of $\mathcal{C}^k$ is at most $k$.*

This theorem follows from stronger volume estimates for quantitative strata (cf. Theorem 7.45), which are themselves parabolically rectifiable (see Definition 6.3).

As a consequence of the proof of Theorem 1.4, we also have estimates for almost every time slice.

**Corollary 1.5.** *Let $k \in \{n, n+1\}$.*

*(1) For $\mathcal{H}^{\frac{1}{2}}$-almost every $t \in (-1, 1)$, $\mathscr{C}^k_t(u) := \mathscr{C}^k(u) \cap (B(0^n, 1) \times \{t\})$ satisfies*

$$\mathcal{H}^{k-1}(\mathscr{C}^k_t(u)) = 0.$$

*(2) For $\mathcal{H}^1$-almost every $t \in (-1, 1)$, $\mathscr{C}^k_t(u)$ is $(k-2)$-rectifiable, and*

$$\mathcal{H}^{k-2}(\mathscr{C}^k_t(u)) \leq C(M, \Theta).$$

*Remark* 1.6. The stated estimates for $S_t(u)$ cannot hold for every time slice. For instance, $u(x, t) = x^2 + 2t$ satisfies $S_0(u) = Z_0(u) = \{(0, 0)\}$.

1.2.2. *Caloric setting.* We next state our results in the caloric setting. Here the estimates are sharper and depend only on the parabolic frequency at a point, defined using a backward heat kernel measure as follows. For a base point $\mathbf{x}_0 := (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and $t < t_0$, set

$$d\nu_{\mathbf{x}_0;t}(x) := \frac{1}{(4\pi(t_0 - t))^{\frac{n}{2}}} \exp\left(-\frac{|x - x_0|^2}{4(t_0 - t)}\right) dx.$$

When $\mathbf{x}_0 = \mathbf{0} := (0^n, 0)$, we write $d\nu_t$ for short. For any $\tau := t_0 - t > 0$, define the *frequency* (introduced in [Poo96]) by

$$N_{\mathbf{x}_0}(\tau) := \frac{2\tau \int_{\mathbb{R}^n} |\nabla u|^2 \, d\nu_{\mathbf{x}_0;t}}{\int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_0;t}},$$

and write $N(\tau) := N_{\mathbf{x}_0}(\tau)$ when $\mathbf{x}_0 = \mathbf{0} = (0^n, 0)$.

We assume that $u := \mathbb{R}^n \times (t_0, t_1) \to \mathbb{R}$ is a solution to (1.3) satisfying the following *mild growth condition*

$$\sup_{s \in [t_0 + \epsilon, t_1 - \epsilon]} \int_{\mathbb{R}^n} u^2 \, d\nu_{0, t_1; s} < \infty, \tag{1.5}$$

for all $\epsilon > 0$.

In this setting, we get estimates for the following effective nodal and singular sets.

**Definition 1.7.** For $r > 0$, define

$$Z_r(u) := \left\{ \mathbf{x} \in P(\mathbf{0}, 1) \colon \inf_{P(\mathbf{x}, \frac{s}{16})} |u|^2 \leq \frac{1}{8} \int_{\mathbb{R}^n} |u|^2 \, d\nu_{\mathbf{x}; t - s^2} \text{ for all } s \in [r, 1] \right\},$$

$$S_r(u) := \left\{ \mathbf{x} \in P(\mathbf{0}, 1) \colon \inf_{P(\mathbf{x}, \frac{s}{16})} \left(|u|^2 + 2s^2 |\nabla u|^2\right) \leq \frac{1}{16} \int_{\mathbb{R}^n} |u|^2 \, d\nu_{\mathbf{x}; t - s^2} \text{ for all } s \in [r, 1] \right\}.$$

Observe that $Z(u) \subseteq Z_r(u)$ and $S(u) \subseteq S_r(u)$ for all $r > 0$. We define $\mathscr{C}_r^k$ analogously.

We now state the quantitative volume estimate of effective nodal and singular sets.

**Theorem 1.8.** *Let $k \in \{n, n+1\}$ and let $u$ be a caloric function with $N(10^5) \leq \Lambda$. Then, for any $r \in (0, 1]$,*

$$\mathcal{H}_{\mathcal{P}}^{n+2}(P(\mathscr{C}^k(u), r) \cap P(\mathbf{0}, 1)) \leq \mathcal{H}_{\mathcal{P}}^{n+2}(P(\mathscr{C}_r^k(u), r) \cap P(\mathbf{0}, 1)) \leq C^{\Lambda^4} r^{n+2-k}.$$

*Remark* 1.9. We do not expect that the stated dependence of the right hand side on $\Lambda$ is sharp. For example, in the elliptic setting, it is known that the dependence on $\Lambda$ is polynomial (see [LM20, LMNN21, Log18a] and the references therein).

We also get estimates at almost every time.

**Corollary 1.10.** *Let $k \in \{n, n+1\}$ and let $u$ be a caloric function with $N(10^5) \leq \Lambda$.*

*(1) For $\mathcal{H}^{\frac{1}{2}}$-almost every $t \in (-1, 1)$, $\mathscr{C}_t^k(u) := \mathscr{C}^k(u) \cap (B(0^n, 1) \times \{t\})$ satisfies*

$$\mathcal{H}^{k-1}(\mathscr{C}_t^k(u)) = 0.$$

*(2) For $\mathcal{H}^1$-almost every $t \in (-1, 1)$, $\mathscr{C}_t^k(u)$ is $(k-2)$-rectifiable, and*

$$\mathcal{H}^{k-2}(\mathscr{C}_t^k(u)) \leq C^{\Lambda^4}.$$

1.3. **Outline of the proof.** For simplicity, we describe the proof in the caloric setting. Our approach follows the general strategy of *dimension reduction*, first introduced by Federer in [Fed70], later quantified using *quantitative stratification* by Cheeger and Naber in [CN13], and further refined through the concept of *neck regions* introduced in [JN21, NV19]. Similar techniques have been used for other problems in both the elliptic [CJN21, HJ23, Nab18, NV24] and parabolic [FL25a, FL25b, FWWZ25, Gia25, GL25, HJ24, HJ25] settings. However, our regularity result Theorem 1.1 is new in the parabolic setting, which necessitates new techniques, as explained below.

An essential ingredient in our analysis is a monotone quantity: the frequency $N_{\mathbf{x}}(r^2)$. For each base point $\mathbf{x}$, the frequency is an increasing function of the scale $r$. The monotonicity and rigidity of the frequency allows us to obtain quantitative uniqueness of tangent flows (Lemma 2.5 and Theorem 2.13) and quantitative cone splitting (Theorem 3.6) at multiple scales and locations.

Using these tools, we obtain a *finite-resolution neck decomposition* (Theorem 5.2) of $P(\mathbf{0},1)$:

$$P(\mathbf{0},1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_b P(\mathbf{x}_b, r_b) \cup \bigcup_f P(\mathbf{x}_f, r_f),$$

into $k$-*neck regions* $\mathcal{N}^a \subseteq P(\mathbf{x}_a, r_a)$ (cf. Definition 4.1), almost $(k+1)$-*symmetric regions* $P(\mathbf{x}_b, r_b)$, and residual region $P(\mathbf{x}_r, r_f)$ with a $k$-content estimate:

$$\sum_a r_a^k + \sum_b r_b^k + \sum_f r_f^k \le C,$$

where $r \le r_a, r_b$ and $r_f \approx r$, see Theorem 5.2. In previous settings, the residual region could be ignored by approximating solutions by solutions with empty singular sets. However, in our setting, this is impossible, so the finite-resolution neck decomposition becomes crucial for proving Minkowski estimates. To prove the content estimate, we first obtain a quantitative dimension reduction (Proposition 6.9) for which we have to argue at multiple scales depending on the proximity to the time slab of the symmetry plane. Using this, $\epsilon$-regularity (Proposition 6.5), and the neck structure theorem (Theorem 4.6), we reduce the volume bound for $\mathscr{C}_r^k$ to the above content estimate, thereby proving the parabolic Minkowski estimates in the caloric case (Theorem 1.8). Once the parabolic Minkowski estimates are established, we use a disintegration of $\mathcal{H}_{\mathcal{P}}^{n+2}$ (Lemma 6.12) and the neck structure theorem to obtain Hausdorff estimates for almost every time slice (Theorem 1.10). An important distinction from [HJ24] is that our neck regions are constructed in space–time rather than on individual time slices. This is essential for obtaining estimates on $S(u)$, since such estimates fail on individual time slices.

Although our Minkowski estimates follow the overall framework of [CJN21, JN21], the proof of Theorem 1.1 requires a fundamentally new argument. The lack of any control on the half-time derivative under a parabolic Lipschitz assumption creates a genuine parabolic obstruction with no analogue in the elliptic theory, and the existing strategies cannot be fully adapted to prove the strong neck structure theorem (Theorem 4.14). This theorem will apply to necks in a refined decomposition (Theorem 5.3):

$$P(\mathbf{0},1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_g P(\mathbf{x}_g, r_g) \cup \left( \widetilde{\mathcal{C}} \cup \bigcup_a \mathcal{C}_{a,0} \right).$$

The packing measures of these neck regions now satisfy an Ahlfors regularity condition (Theorem 4.6 and Proposition 4.9), which guarantees that the average frequency drop with respect to the packing measures satisfies a Carleson-type estimate (Lemma 4.20). To apply this, we prove a new sharp cone-splitting inequality (Theorem 3.11) based on commutator identities (3.8a) and (3.8b) for space-time vector fields, rather than linear combinations of radial fields as in the elliptic case ([NV24, Theorem 2.20]). Using the sharp cone splitting, we show that the parabolic Jones $\beta$-numbers at points in the center set of $\mathcal{N}^a$ are controlled by the average frequency drop (Lemma 3.14). Combining this with the Carleson–type estimate for the average frequency drop, we conclude that the $\beta$-numbers also satisfy a Carleson-type estimate (Lemma 4.20). We then use the Carleson-type estimate for $\beta$-numbers to prove the strong neck structure Theorem (4.14). To achieve this, we use a Whitney-type covering (Lemma 4.8) to construct a parabolic Lipschitz graph whose neighborhood contains the center set of $\mathcal{N}^a$ (Lemma 4.15 and Lemma 4.18). A corresponding Carleson-type estimate for a variant of $\beta$-number called $\kappa$-number provides bounds on the mean oscillation of the half-time derivative of the graph (Proposition B.1). We show that such a Carleson-type estimate for the $\kappa$-numbers of our extended graph follows from the Carleson estimate for $\beta$-numbers of the packing measure (Lemma 4.19). Combining these with the containment of the nodal and singular sets in the center set, we show that the nodal and singular sets are contained in a countable union of regular parabolic Lipschitz graphs up to negligible subsets.

Our estimates in the more general setting of parabolic inequalities are analogous to the caloric setting, but contain scale-dependent error terms. A significant new technical challenge is to ensure that these error terms are summably small across scales. While the methods of [HJ24] suffice for volume estimates of nodal sets, they do not guarantee the summability needed for the strong neck structure theorem, and overcoming this requires substantially new arguments. Along the way, we also establish sharp quantitative uniqueness (Theorem 7.13) via a linear stability analysis of the operator $\Delta_f + \frac{m}{2} = \Delta - \frac{x}{2\tau} \cdot \nabla + \frac{m}{2}$, following ideas of Simon [Sim85, §II.3]. The operator $\Delta_f + \frac{m}{2}$ appears when $u$ is rescaled dynamically. See [HJ24, Theorem 4.1] for a related quantitative uniqueness result, proved by a different method and in a non-sharp form. With these tools, we can extend the results from the caloric case to general solutions to (1.1), proving Theorem 1.1, Corollary 1.3, Theorem 1.4, and Corollary 1.5.

**Notation.** Unless otherwise specified, we omit the dependence of constants on the dimension $n$. We write $\mathbb{N}_0 = \{0, 1, 2 \ldots\}$. Usually, $C < \infty$ denotes a large constant greater than 1, while $c$ denotes a small positive constant less than 1, and we do not rename constants from line to line. When the context is clear we write $|\cdot| := |\cdot|_{\mathcal{P}}$. A statement "if $\delta \leq \bar{\delta}(\Lambda, \epsilon)$, then ..." means "for any $\Lambda > 0$ and any $\epsilon > 0$, there exists $\bar{\delta} > 0$ depending on $\Lambda$ and $\epsilon$ such that if $\delta \leq \bar{\delta}$, then ...". A statement "Apply Lemma L with $\alpha \leftarrow A$, $\beta \leftarrow B$ ..." means "Apply Lemma L with the parameters $\alpha$ replaced by $A$, $\beta$ replaced by $B$ ...".

**Acknowledgments.** The authors would like to thank Bennett Chow, Dennis Kriventsov, and Alec Payne for helpful comments and suggestions. They are especially grateful to Aaron Naber for numerous insightful discussions and recommendations. MH was supported in part by the National Science Foundation under Grant No. DMS-2202980. ZM was supported by an AMS–Simons Travel Grant.

## 2. Frequency

In the following five sections, we prove Theorem 1.8 and Corollary 1.10, thereby estimating the parabolic Minkowski content of the effective nodal and singular sets of a *caloric function* $u$, i.e., a solution to

$$\partial_t u - \Delta u = 0.$$

We treat caloric functions separately for multiple reasons. First, the results and techniques are more effective and cleaner than in the general setting. Second, many estimates in Section 7 are proved using a caloric approximation and appealing to the corresponding results in the caloric framework.

### 2.1. Conjugate heat kernel.

The *conjugate heat kernel measure* based at $\mathbf{x}_0 := (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ is

$$d\nu_{\mathbf{x}_0;t}(x) := \frac{1}{(4\pi(t_0 - t))^{\frac{n}{2}}} \exp\left(-\frac{|x - x_0|^2}{4(t_0 - t)}\right) dx.$$

We set
$$f_{\mathbf{x}_0}(x,t) := \frac{|x-x_0|^2}{4(t_0-t)}.$$
When $\mathbf{x}_0 = \mathbf{0} := (0^n, 0)$, we write $\nu_t := \nu_{\mathbf{x}_0;t}$ and $f := f_{\mathbf{x}_0}$.
**Convention:** We assume that a caloric function $u \in C^\infty(\mathbb{R}^n \times (t_0, t_1))$ has *mild growth*, i.e.,

$$\sup_{s\in[t_0+\epsilon, t_1-\epsilon]} \int_{\mathbb{R}^n} u^2 \, d\nu_{0,t_1;s} < \infty. \tag{2.1}$$

We impose such a bound for the following three reasons. First, to ensure uniqueness of solutions by ruling out Tychonoff's example [Eva10, §2.3 Theorem 7]. Second, to freely integrate by parts (see Lemma A.1). Finally, such a bound allows us to have a spectral decomposition (2.7) of $u$ into an infinite sum of eigenfunctions of the drift Laplacian where the sum converges locally smoothly and in weighted $L^2$, see [MPP02, Proposition 3.1].

For any function $\phi \in C^1(\mathbb{R}^{n+1})$, we define the *drift Laplacian* $\Delta_\phi$ as
$$\Delta_\phi u := \Delta u - \langle \nabla\phi, \nabla u\rangle.$$
**Convention:** In the remainder of the paper, we will use the following convention for ease of notation. Given any function $v : \mathbb{R}^n \times I \to \mathbb{R}$, where $I$ is an interval containing $t \in \mathbb{R}$, and $v(\cdot, t) \in L^1(\mathbb{R}^n, \nu_{\mathbf{x}_0;t})$, we write
$$\int_{\mathbb{R}^n} v \, d\nu_{\mathbf{x}_0;t} := \int_{\mathbb{R}^n} v(x,t)\, d\nu_{\mathbf{x}_0;t}(x).$$
We will need the following comparison of conjugate heat kernels based at nearby points.

**Lemma 2.1.** *The following holds for any $\sigma \in (0,1]$ and $\alpha_0, \alpha$ such that $\alpha - 1 < \alpha_0 < \alpha < 1$. If $\mathbf{x}_0 = (x_0, t_0), \mathbf{x}_1 = (x_1, t_1)$ are points in $\mathbb{R}^n \times \mathbb{R}$ satisfying $|x_0 - x_1| < 1$ and $|t_0 - t_1| \le \sigma$, then, for $\tau \ge \frac{2\sigma}{\alpha-\alpha_0}$,*
$$e^{\alpha_0 f_{\mathbf{x}_0}} d\nu_{\mathbf{x}_0;t_1-\tau} \le C^{\frac{1}{\sigma}} e^{\alpha f} \, d\nu_{\mathbf{x}_1;t_1-\tau}.$$
*Proof.* By spacetime translation, we may assume that $\mathbf{x}_1 = \mathbf{0}$. Set $\tau := -t$ and $\tau_0 := t_0 - t$. If $2\sigma \le (\alpha - \alpha_0)\tau$, then

$$1 - \frac{\alpha-\alpha_0}{2} \le \frac{\tau_0}{\tau} \le 1 + \frac{\alpha - \alpha_0}{2}. \tag{2.2}$$

Since $\alpha_0 < \alpha < 1$, we have
$$\tau\left(1 - \frac{\tau_0(1-\alpha)}{\tau(1-\alpha_0)}\right) \ge \frac{\tau(\alpha-\alpha_0)}{2} \ge \sigma.$$
Therefore,
$$\begin{aligned} &((1-\alpha_0)f_{\mathbf{x}_0} - (1-\alpha)f)(x,t) \\ &= \left(\frac{1-\alpha_0}{4\tau_0} - \frac{1-\alpha}{4\tau}\right)\left|x - x_0\frac{1-\alpha_0}{4\tau_0\left(\frac{1-\alpha_0}{4\tau_0} - \frac{1-\alpha}{4\tau}\right)}\right|^2 - \frac{1-\alpha}{4\tau\left(1 - \frac{\tau_0(1-\alpha)}{\tau(1-\alpha_0)}\right)}|x_0|^2 \ge -\frac{1}{4\sigma}. \end{aligned}$$
This together with (2.2) gives
$$(4\pi\tau_0)^{-\frac{n}{2}} e^{\alpha_0 f_{\mathbf{x}_0}} e^{-f_{\mathbf{x}_0}} \le (4\pi(1 - \tfrac{\alpha-\alpha_0}{2})\tau)^{-\frac{n}{2}} e^{\frac{1}{4\sigma} + \alpha f} e^{-f}.$$
□

We recall the following hypercontractivity estimate, which is a well-known consequence of the Gaussian log Sobolev inequality.

**Proposition 2.2.** *Suppose $u$ is a caloric function. For any $0 < \tau_1 < \tau_2$ and $1 < q \le p < \infty$ with $\frac{\tau_1}{\tau_2} \ge \frac{p-1}{q-1}$, the following holds:*

$$\left( \int_{\mathbb{R}^n} |u|^p \, d\nu_{\mathbf{x}_0; t_0 - \tau_1} \right)^{\frac{1}{p}} \le \left( \int_{\mathbb{R}^n} |u|^q \, d\nu_{\mathbf{x}_0; t_0 - \tau_2} \right)^{\frac{1}{q}}.$$

*Proof.* See [Gro75]. □

As an application, we can compare the $L^2$ norm of caloric functions based at nearby points.

**Lemma 2.3.** *Suppose $u$ is a caloric function. For any $\theta \in (0, \frac{1}{4}]$, $\sigma \in (0, 1]$, $r > 0$, if $|\mathbf{x}_1 - \mathbf{x}_0| < r$ and $|t_1 - t_0| \le \sigma r^2$, then, for any $\tau \ge \frac{6\sigma}{\theta} r^2$*

$$\int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_0; t_0 - \tau} \le C^{\frac{1}{\sigma}} \int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_1; t_1 - (1+\theta)\tau}.$$

*The same estimate holds for $\partial_t u$ and $\nabla u \cdot z$ for any $z \in \mathbb{R}^n$.*

*Proof.* After a space-time translation and parabolic rescaling, we assume that $\mathbf{x}_0 = \mathbf{0}$, and $r = 1$. Take $\alpha = \frac{\theta}{3}$ and $p = \frac{1 + \frac{7}{12}\theta}{1 + \frac{1}{6}\theta}$. If $\tau \ge \frac{6\sigma}{\theta}$, then $|t_1| \le \sigma \le \frac{\theta}{6}\tau$, and

$$\frac{(1+\theta)\tau}{t_1 + \tau} \ge \frac{(1+\theta)\tau}{\frac{\theta}{6}\tau + \tau} = 2p - 1, \qquad \frac{\alpha p}{p - 1} = \frac{\theta}{3} \frac{1 + \frac{7}{12}\theta}{\frac{5}{12}\theta} < .92 < 1.$$

Apply Lemma 2.1 with $\mathbf{x}_0 \leftarrow \mathbf{x}_1$, $t \leftarrow -\tau$ and $\alpha_0 \leftarrow 0$, Hölder's inequality, and then Proposition 2.2 with $q \leftarrow 2$ and $p \leftarrow 2p$ to obtain

$$\begin{aligned} \int_{\mathbb{R}^n} u^2 \, d\nu_{-\tau} &\le C^{\frac{1}{\sigma}} \int_{\mathbb{R}^n} u^2 e^{\alpha f_{\mathbf{x}_1}} \, d\nu_{\mathbf{x}_1; -\tau} \le C^{\frac{1}{\sigma}} \left( \int_{\mathbb{R}^n} u^{2p} \, d\nu_{\mathbf{x}_1; t_1 - (t_1 + \tau)} \right)^{\frac{1}{p}} \left( \int_{\mathbb{R}^n} e^{\frac{\alpha p}{p-1} f_{\mathbf{x}_1}} \, d\nu_{\mathbf{x}_1; -\tau} \right)^{\frac{1}{2}} \\ &\le C^{\frac{1}{\sigma}} \int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_1; t_1 - (1+\theta)\tau}. \end{aligned}$$

The second claim follows because $\partial_t u$ and $\nabla u \cdot z$ are caloric. □

## 2.2. **Energy functionals.**

We record the definitions of pointed energy functionals (cf. Definition 2.4) and their associated derivative formulae (cf. Lemma 2.5 and Lemma 2.6) for functions on $\mathbb{R}^n \times \mathbb{R}$.

**Definition 2.4.** Given a function $w \in C^\infty(\mathbb{R}^n \times I)$, $\tau > 0$, and a base point $\mathbf{x}_0 = (x_0, t_0)$, we define

$$\begin{aligned} H^w_{\mathbf{x}_0}(\tau) &\coloneqq \int_{\mathbb{R}^n} w^2 \, d\nu_{\mathbf{x}_0; t}, & E^w_{\mathbf{x}_0}(\tau) &\coloneqq 2\tau \int_{\mathbb{R}^n} |\nabla w|^2 \, d\nu_{\mathbf{x}_0; t}, \\ N^w_{\mathbf{x}_0}(\tau) &\coloneqq \frac{E^w_{\mathbf{x}_0}(\tau)}{H^w_{\mathbf{x}_0}(\tau)}, & D^w_{\mathbf{x}_0}(\tau) &\coloneqq \log_2 \frac{H^w_{\mathbf{x}_0}(2\tau)}{H^w_{\mathbf{x}_0}(\tau)}. \end{aligned}$$

In the literature, $D^w_{\mathbf{x}_0}(\tau)$ and $N^w_{\mathbf{x}_0}(\tau)$ are called the *doubling index* and the *(parabolic) frequency* based at $\mathbf{x}_0$ at scale $\sqrt{\tau}$, respectively. When the context is clear, we omit $w$ in the superscript and write $E_{\mathbf{x}_0}(\tau) \coloneqq E^w_{\mathbf{x}_0}(\tau)$ and so on. Furthermore, we write $E(\tau) = E_{\mathbf{x}_0}(\tau)$ and so on when $\mathbf{x}_0 = \mathbf{0} \coloneqq (0^n, 0)$.

For simplicity, we will state our results at $\mathbf{0}$ but they hold for any other base point $\mathbf{x}_0$ after a space-time translation. We now recall the monotonicity formulae. We write

$$\square \coloneqq \partial_t - \Delta, \qquad \square^* = -\partial_t - \Delta.$$

**Lemma 2.5.** *Suppose $w \in C^\infty(\mathbb{R}^n \times \mathbb{R})$ and its derivatives have mild growth as in (2.1) and let $\tau > 0$. Then*

$$
\begin{aligned}
D'(\tau) &= \frac{N(2\tau) - N(\tau)}{\tau \log 2} + \frac{2}{\log 2}\left( \int_{\mathbb{R}^n} w \square w \, d\nu_{-\tau} - 2 \int_{\mathbb{R}^n} w \square w \, d\nu_{-2\tau} \right),\\
E'(\tau) &= 4\tau \int_{\mathbb{R}^n} (\Delta_f w)^2 + (\Delta_f w)(\square w)\, d\nu_{-\tau},\\
H'(\tau) &= \frac{E(\tau)}{\tau} - 2 \int_{\mathbb{R}^n} w \square w \, d\nu_{-\tau},\\
N'(\tau) &= \frac{4\tau}{H(\tau)} \int_{\mathbb{R}^n} \left( \left( \Delta_f w + \frac{N(\tau)}{2\tau} w \right)^2 + \left( \Delta_f w + \frac{N(\tau)}{2\tau} w \right) \square w \right) d\nu_{-\tau}.
\end{aligned}
$$

*Proof.* The derivative formulae for $H$ and $E$ are standard (see [Poo96]), and follow from the evolution equations

$$
\square w^2 = 2 w \square w - 2|\nabla w|^2, \quad \square |\nabla w|^2 = 2 \nabla w \cdot \nabla \square w - 2|\nabla^2 w|^2,
$$

and integrating by parts the $f$-Bochner formula

$$
\Delta_f |\nabla w|^2 = 2|\nabla^2 w|^2 + 2\langle \nabla w, \nabla \Delta_f w \rangle + \frac{1}{\tau}|\nabla w|^2. \tag{2.5}
$$

To compute $N'$, we use $N(\tau) = \frac{E(\tau)}{H(\tau)}$ to get

$$
\begin{aligned}
H^2 N' &= 4\tau \int_{\mathbb{R}^n} w^2 \, d\nu_{-\tau} \int_{\mathbb{R}^n} (\Delta_f w)^2 \, d\nu_{-\tau} - 4\tau \left( \int_{\mathbb{R}^n} w \Delta_f w \, d\nu_{-\tau} \right)^2\\
&\quad + 4\tau H \int_{\mathbb{R}^n} (\Delta_f w)(\square w)\, d\nu_{-\tau} + 2E \int_{\mathbb{R}^n} w \square w \, d\nu_{-\tau}\\
&= 4\tau H \int_{\mathbb{R}^n} \left( \Delta_f w + \frac{N(\tau)}{2\tau} w \right)^2 d\nu_{-\tau} + 4\tau H \int_{\mathbb{R}^n} \left( \Delta_f w + \frac{N(\tau)}{2\tau} w \right) \square w \, d\nu_{-\tau}.
\end{aligned}
$$

Finally, $N(\tau) = \frac{E(\tau)}{H(\tau)} = \frac{\tau(H'(\tau) + 2\int_{\mathbb{R}^n} w \square w \, d\nu_{-\tau})}{H(\tau)}$ implies that

$$
\log \frac{H(2\tau)}{H(\tau)} = \int_\tau^{2\tau} \frac{N(\overline{\tau})}{\overline{\tau}} \, d\overline{\tau} - 2 \int_\tau^{2\tau} \int_{\mathbb{R}^n} w \square w \, d\nu_{-\overline{\tau}} d\overline{\tau}. \tag{2.6}
$$

In particular,

$$
D'(\tau) = \frac{N(2\tau) - N(\tau)}{\tau \log 2} + \frac{2}{\log 2}\left( \int_{\mathbb{R}^n} w \square w \, d\nu_{-\tau} - 2 \int_{\mathbb{R}^n} w \square w \, d\nu_{-2\tau} \right).
$$

□

To compare the frequency at a fixed scale at nearby points, for $\mathbf{x} = (x,t)$, $s > 0$, we write $H_s(\mathbf{x}) = H_{\mathbf{x}}(s)$ and so on. We also define the restricted functional with respect to a $k$-plane $L^k \subset \mathbb{R}^n$ and with respect to the temporal direction as:

$$
N_{s;L}(\mathbf{x}) := \frac{2s \int_{\mathbb{R}^n} |\pi_L \nabla w|^2 \, d\nu_{\mathbf{x};t-s}}{H_s(\mathbf{x})}, \qquad T_s(\mathbf{x}) := \frac{2s^2 \int_{\mathbb{R}^n} |\partial_t w|^2 \, d\nu_{\mathbf{x};t-s}}{H_s(\mathbf{x})}.
$$

**Lemma 2.6.** *Suppose $w \in C^\infty(\mathbb{R}^n \times I)$ and $L \subset \mathbb{R}^n$ is a $k$-plane. Then*

$$
s|\pi_L \nabla D_s|^2 \le C(N_{2s;L} + N_{s;L}), \qquad s^2 |\partial_t D_s|^2 \le C(T_s + T_{2s}).
$$

*Proof.* After a rotation, we may assume $L \ni v = e_1$, so $\nabla w \cdot v = \partial_{x_1} w$. At $\mathbf{x} = (x,t)$, we compute

$$\partial_{x_1} H_s(\mathbf{x}) = (4\pi s)^{-\frac{n}{2}} \int_{\mathbb{R}^n} w^2(y,t-s)\partial_{x_1} e^{-\frac{|x-y|^2}{4s}}\, dy = -(4\pi s)^{-\frac{n}{2}} \int_{\mathbb{R}^n} w^2(y,t-s)\partial_{y_1} e^{-\frac{|x-y|^2}{4s}}\, dy$$

$$= \int_{\mathbb{R}^n} 2w(y)\partial_{y_1} w(y)\, d\nu_{\mathbf{x};t-s}(y) \le 2\left(\int_{\mathbb{R}^n} w^2\, d\nu_{\mathbf{x};t-s}\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^n} |\pi_L \nabla w|^2\, d\nu_{\mathbf{x};t-s}\right)^{\frac{1}{2}}.$$

Then,

$$s(\log 2)^2 |\partial_{x_1} D_s|^2 \le 2s\left(|\partial_{x_1} \log H_{2s}|^2 + |\partial_{x_1} \log H_s|^2\right) \le 4(N_{2s;L} + N_{s;L}).$$

The second inequality is proved similarly. ∎

We often use the following $L^2(\nu_{-\tau})$-orthogonal decomposition of a function $w \in L^2(\nu_{-\tau})$

$$w = \sum_{j=0}^{\infty} p_j, \tag{2.7}$$

where $p_j$ is the projection of $w$ onto the $L^2(\nu_t)$-eigenspace of $2\tau\Delta_f$ with eigenvalue $j \in \mathbb{N}_0$, i.e., $2\tau\Delta_f p_j + jp_j = 0$ and is a *homogeneous caloric polynomial* (at $\mathbf{0}$) of degree $j$ in the following sense. When $u$ is caloric, the same decomposition (2.7) holds for all time. The existence of such decomposition follows from a spectral analysis of $2\tau\Delta_f$, see [MPP02, Proposition 3.1].

**Definition 2.7.** A caloric polynomial $p$ is *homogeneous* at $\mathbf{x}_0$ of degree $m$ if for any $\lambda > 0$ we have

$$p(x_0 + \lambda(x - x_0), t_0 + \lambda^2(t - t_0)) = \lambda^m p(x,t).$$

Denote $\mathcal{P}_m$ to be the space of homogeneous caloric polynomials of degree $m$.

In the following lemma, we prove that if $w(\cdot,\tau)$ is almost an eigenfunction of $2\tau\Delta_f$ with eigenvalue $N(\tau)$, then $w(\cdot,\tau)$ can be approximated by a caloric homogeneous polynomial and $N(\tau)$ is almost an integer.

**Lemma 2.8.** *For any $w \in C^\infty(\mathbb{R}^n \times I)$ such that $w(\cdot,-\tau) \in L^2(\nu_{-\tau})$, $\tau > 0$, suppose $m \in \arg\min_{k\in\mathbb{N}_0} |N(\tau) - k|$ and $N(\tau) \in [j, j+1)$ for some $j \in \mathbb{N}_0$. Let $p_m \in \mathcal{P}_m$ be the $L^2(\mathbb{R}^n, \nu_t)$-orthogonal projection onto the $\frac{m}{2}$-eigenspace of $-\Delta_{f_t}$. Then*

$$(N(\tau) - j)(j + 1 - N(\tau))H(\tau) + \int_{\mathbb{R}^n} (w - p_m)^2\, d\nu_{-\tau} \le 20\tau^2 \int_{\mathbb{R}^n} \left(\Delta_f w + \frac{N(\tau)}{2\tau} w\right)^2 d\nu_{-\tau}.$$

*Proof.* Note that

$$\max_{a\in\mathbb{R}} \min_{k\in\mathbb{N}_0} \{(N(\tau) - k)(a - k)\} = (N(\tau) - j)(j + 1 - N(\tau)),$$

which is achieved when $a = 2j + 1 - N(\tau)$. Using the spectral decomposition $w = \sum_{j=0}^{\infty} p_j$, and the $L^2(\nu_{-\tau})$-orthogonality of $p_j$, we get

$$\begin{aligned}
\int_{\mathbb{R}^n} (2\tau\Delta_f w + N(\tau)w)^2\, d\nu_{-\tau} &= \int_{\mathbb{R}^n} (2\tau\Delta_f w + N(\tau)w)(2\tau\Delta_f w + aw)\, d\nu_{-\tau} \\
&= \sum_j (N(\tau) - j)(a - j) \int_{\mathbb{R}^n} p_j^2\, d\nu_{-\tau} \\
&\ge (N(\tau) - j)(j + 1 - N(\tau))H(\tau),
\end{aligned}$$

where we used $\int_{\mathbb{R}^n} (2\tau\Delta_f u + N(\tau)u)u\, d\nu_{-\tau} = 0$ for the first equality. Using this again, the conclusion follows by combining the following:

$$\int_{\mathbb{R}^n} (w - p_m)^2\, d\nu_{-\tau} = \sum_{k\ne m} \int_{\mathbb{R}^n} p_k^2\, d\nu_{-\tau} \le 4 \sum_k (N(\tau) - k)(N(\tau) - k) \int_{\mathbb{R}^n} p_k^2\, d\nu_{-\tau}.$$

∎

2.3. **Monotonicity and quantitative uniqueness in caloric setting.** For the rest of the section, we will restrict our discussion to the case of a caloric function $u$. Since $\square u = 0$, the functionals $D$, $E$, $H$, $N$ are monotone by Lemma 2.5. We list some consequences of the monotonicity.

First, we get the following version of interior estimates on two-sided parabolic balls that we need later. Note that the standard interior estimates hold on backward parabolic balls; see, for example, [Eva10, §2.3].

**Lemma 2.9.** *If $u$ is a caloric function, then*

$$\sup_{\mathbf{x}\in P(\mathbf{0},\frac{r}{8})} r^{2(2l+j)}|\partial_t^l \nabla^j u|^2(\mathbf{x}) \le C_{jl} H(4r^2).$$

*Proof.* After parabolic scaling, we may assume $r = 1$. Since $\partial_t^l \nabla^j u = \nabla^j \Delta^l u$ is caloric, using the monotonicity formula in Lemma 2.5 and the comparison of energy in Lemma 2.3, we deduce that for all $\mathbf{x} \in P(\mathbf{0}, \frac{1}{8})$

$$|\partial_t^l \nabla^j u|^2(\mathbf{x}) \le \int_{\mathbb{R}^n} |\partial_t^l \nabla^j u|^2 d\nu_{\mathbf{x};t-\frac{1}{2^{2l+j}}} \le C_{jl} \int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{x};t-1} \le C_{jl} H_{\mathbf{0}}(3).$$

□

We now prove that the frequency at nearby points at smaller scale is bounded above by the frequency at a reference point at a larger scale (cf. Theorem 2.2.8 in [HL94a], Lemma 2.7 in [CNV15], or Theorem 3.14 in [NV17b]). This allows us to assume a bound on the frequency just at one point.

**Lemma 2.10.** *Suppose $u$ is caloric. Then*

$$\sup_{\mathbf{x}\in P(\mathbf{0},\frac{1}{8})} N_{\mathbf{x}}(1) \le C(N(4)+1).$$

*Proof.* Applying Lemma 2.3 with $\theta = \frac{1}{4}$, for $\mathbf{x} \in P(\mathbf{0}, \frac{1}{8})$, we get

$$H_{\mathbf{0}}\left(\tfrac{1}{1+\theta}\right) \le C H_{\mathbf{x}}(1), \qquad H_{\mathbf{x}}(2) \le C H_{\mathbf{0}}(2+2\theta).$$

Therefore,

$$N_{\mathbf{x}}(1) \le D_{\mathbf{x}}(1) = \log_2 \frac{H_{\mathbf{x}}(2)}{H_{\mathbf{x}}(1)} \le C + \frac{1}{\log 2}\int_{\frac{1}{1+\theta}}^{2+2\theta} \frac{N(\tau)}{\tau}\, d\tau \le C + (1+2\theta) N_{\mathbf{0}}^u(4).$$

□

So far, we have used the monotonicity of the frequency. However, by Lemma 2.5 and Lemma 2.8, we get refined monotonicity

$$\tau N'(\tau) \ge (N(\tau)-m)(m+1-N(\tau)), \tag{2.8}$$

when $N(\tau) \in [m, m+1)$ for some $m \in \mathbb{N}_0$. In the lemma below we use (2.8) to prove a definite drop in the frequency and pinching of the frequency near integers.

**Lemma 2.11.** *The following hold:*

1. *If $N(\tau) \le m+1-\epsilon$ for some $\epsilon \in (0, \frac{1}{4})$, then*
$$N\left(\left(\frac{\epsilon}{1-\epsilon}\right)^2 \tau\right) \le m+\epsilon.$$
2. *If $5\epsilon = \min_{k\in\mathbb{N}_0} |N(\tau)-k|$, then $N(\tau) - N\left(\frac{\tau}{2}\right) \ge \epsilon$.*
3. *Further, if $N(\tau) - N\left(\frac{\tau}{2}\right) \le \delta$, then there exists $m \in \mathbb{N}_0$ such that for $s \in \left[\frac{\tau}{2}, \tau\right]$,*
$$|N(s) - m| \le 6\delta.$$

*Proof.* (1) We use (2.8) to argue as in [NV17b, Lemma 3.16]. We may assume that $\tau = 1$ and $N(1) \in [m, m+1-\epsilon]$. Then $x(\tau) = N(\tau) - m$ satisfies $\tau x' \geq x(1-x)$ as long as $x > 0$. Let $z(\tau) = \frac{a\tau}{1+a\tau}$, where $a = \frac{1-\delta}{\delta}$ for some $\delta < \epsilon$ so that $z(1) = 1-\delta$. Then $\tau z' = z(1-z)$ and $\tau \frac{d}{d\tau} \log\left(\frac{x}{z}\right) \geq z - x$. By an ODE argument, $x(\tau) \leq z(\tau)$ as long as $x(\tau) > 0$ holds. The conclusion follows by taking $\delta \to \epsilon$.

(2) We may assume $\epsilon > 0$. Let $m = [N(\tau)]$ be the largest integer not greater than $N(\tau)$. Then $m + 5\epsilon \leq N(\tau) \leq m + 1 - 5\epsilon$. For $s \leq \tau$, as long as $N(s) \in [m+3\epsilon, m+1-5\epsilon]$, by Lemma 2.11,

$$sN'(s) \geq (N-m)(m+1-N) \geq 3\epsilon(1-3\epsilon) \geq 2\epsilon.$$

If $N\left(\frac{\tau}{2}\right) < m + 3\epsilon$, then we are done. Otherwise, integrating the differential inequality yields

$$N(\tau) - N(\tfrac{\tau}{2}) \geq 2\epsilon \log 2 > \epsilon.$$

(3) Let $5\epsilon = \min_k |N(\tau) - k| = |N(\tau) - m|$. By (2), $\epsilon \leq N(\tau) - N(\frac{\tau}{2}) \leq \delta$. The conclusion follows by the monotonicity of the frequency. □

We prove that the frequency is almost constant at smaller scales using the monotonicity and a bound on the frequency at a larger scale.

**Lemma 2.12.** *For any $\epsilon \in (0, \frac{1}{10})$, if*

$$N(r_2^2) - N(r_1^2) \leq \Lambda, \qquad 0 < r_1 \leq \epsilon^{4\Lambda+10} r_2,$$

*then there is $s \in (r_1, r_2)$ such that*

$$N(\epsilon^{-2}s^2) - N\left(\epsilon^2 s^2\right) < \epsilon.$$

*Proof.* By scaling, we may assume $r_2 = 1$. Define $\tau_k = \epsilon^{4k}$. Let $m$ be the smallest integer such that $N(\tau_0) < m + \frac{1}{2}\epsilon$. If $N(\tau_1) \geq m - \frac{1}{2}\epsilon$, then we can take $s = \epsilon^2$. Otherwise, $N(\tau_1) < m - \frac{1}{2}\epsilon$, and by Lemma 2.11, $N(\tau_2) \leq N(\frac{1}{4}\epsilon^2\tau_1) < m - 1 + \frac{1}{2}\epsilon$. If $N(\tau_3) \geq m - 1 + \frac{1}{2}\epsilon$, then we can take $s = \epsilon^5$. Otherwise, we can proceed as above. If we fail to find such $s$ in $k = [\Lambda] + 2$ steps, then

$$N(r_1^2) \leq N(\tau_{2k}) < m - k + \tfrac{1}{2}\epsilon \leq N(1) + 1 - k \leq N(r_1^2) + \Lambda - [\Lambda] - 1,$$

a contradiction. □

The following theorem shows that when the frequency is pinched at $m$, $u$ is close to $p_m$ where $u = \sum_{j=0}^{\infty} p_j$ is the $L^2(\nu_{-\tau})$-orthogonal decomposition in terms of homogeneous caloric polynomials.

**Theorem 2.13.** *Fix $0 \leq \tau_1 < \frac{\tau_2}{2}$. Assume that $N(\tau_2) - N(\tau_1) \leq \delta < \frac{1}{10}$, and let $m \in \mathbb{N}_0$ given by Lemma 2.11(3) with $\tau \leftarrow \tau_2$. Then, for any $l, j \in \mathbb{N}_0$ and $0 < \tau \leq \frac{\tau_1}{2}$, we have*

$$\tau^{2l+j} \int_{\mathbb{R}^n} |\partial_t^l \nabla^j (u - p_m)|^2 \, d\nu_{-\tau} \leq C_{j,l} \delta H(\tau_2),$$

*where $p_m$ is the projection of $u$ onto the eigenspace of $2\tau\Delta_f$ with eigenvalue $m$.*

*Proof.* First assume that $j = l = 0$. By Lemma 2.5 and Lemma 2.8, we know that

$$\int_{\tau_1}^{\tau_2} \frac{1}{\tau H(\tau)} \int_{\mathbb{R}^n} |u - p_m|^2 \, d\nu_{-\tau} \, d\tau = \ 4\int_{\tau_1}^{\tau_2} N'(\tau) \, d\tau = 4(N(\tau_2) - N(\tau_1)) \leq 4\delta.$$

Since $u - p_m$ is caloric, by the monotonicity of $H$,

$$\int_{\mathbb{R}^n} |u - p_m|^2 \, d\nu_{-\tau_1} \leq \frac{4\tau_2}{\tau_2 - \tau_1} \delta H(\tau_2).$$

To prove the general case, we can use the monotonicity formula to get

$$\frac{d}{dt} \int_{\mathbb{R}^n} |u - p_m|^2 \, d\nu_{-\tau} = -2 \int_{\mathbb{R}^n} |\nabla(u - p_m)|^2 \, d\nu_{-\tau},$$

where $\tau = -t$. By the monotonicity of $\tau \mapsto \int_{\mathbb{R}^n} |\nabla(u - p_m)|^2 d\nu_{-\tau}$,

$$\frac{\tau}{2}\int_{\mathbb{R}^n} |\nabla(u-p_m)|^2\, d\nu_{-\frac{\tau}{2}} \le \int_{\frac{\tau}{2}}^{\tau}\int_{\mathbb{R}^n} |\nabla(u-p_m)|^2\, d\nu_{-s} ds \le \frac{1}{2}\int_{\mathbb{R}^n} (u-p_m)^2\, d\nu_{-\tau}.$$

Furthermore, since $\partial_t^l \nabla^j (u - p_m) = \nabla^j \Delta^l (u - p_m)$ is caloric, we can prove the general case by induction. □

As a consequence of Theorem 2.13, we show that if the frequencies based at two nearby points are almost constant at a given scale, then they should be pinched at the same integer.

**Corollary 2.14.** *Suppose $N_{\mathbf{x}_1}(8) \le \Lambda$ and $\mathbf{x}_2 \in P(\mathbf{x}_1, \frac{1}{8})$ such that for both $i = 1, 2$, we have $N_{\mathbf{x}_i}(8) - N_{\mathbf{x}_i}(\frac{1}{8}) < \delta$. If $\delta \le C^{-\Lambda}$, then there exists $m \in \mathbb{N}_0$ such that*

$$|N_{\mathbf{x}_i}(\tau) - m| < C\delta$$

*for all $\tau \in \left[\frac{1}{8}, 8\right]$ and both $i = 1, 2$.*

*Proof.* By Lemma 2.11(3), there exist $m_1, m_2 \in \mathbb{N}_0$ such that $|N_{\mathbf{x}_i}(\tau) - m_i| < C\delta$ for both $i = 1, 2$ and all $\tau \in \left[\frac{1}{8}, 8\right]$, so it suffices to show $m_1 = m_2$. Suppose by way of contradiction that $m_1 < m_2$.

Let $p_{m_i}(\cdot, t)$ be the projection of $u(\cdot, t)$ onto the $L^2(\nu_{\mathbf{x}_i;t})$-eigenspace of $2\tau_i \Delta_{f_{\mathbf{x}_i}}$ with eigenvalue $m_i$, which is a polynomial at $x_i$ of degree $m_i$. Using $[\Delta_f, \nabla] = \frac{1}{2(t_i - t)}\nabla$ and integrating $\frac{1}{2}\Delta_{f_i}|\nabla^k p_{m_i}|^2 = |\nabla^{k+1} p_{m_i}|^2 + \langle \nabla^k p_{m_i}, \Delta_{f_i} \nabla^k p_{m_i}\rangle$ yields

$$\int_{\mathbb{R}^n} |\nabla^{m_2} p_{m_2}|^2 d\nu_{\mathbf{x}_2;t_2-\tau} = \frac{m_2!}{(2\tau)^{m_2}} \int_{\mathbb{R}^n} p_{m_2}^2 d\nu_{\mathbf{x}_2;t_2-\tau}$$

for any $\tau > 0$. Thus Lemma 2.3 and Theorem 2.13 give the following, where $C_0 \coloneqq m_2! 4^{m_2} \ge C^{\Lambda^2}$:

$$\begin{aligned}
C_0(1 - C^{\Lambda}\delta) H_{\mathbf{x}_2}(1/8) &\le C_0 \int_{\mathbb{R}^n} |p_{m_2}|^2\, d\nu_{\mathbf{x}_2;t_2-\frac{1}{8}} = \int_{\mathbb{R}^n} |\nabla^{m_2} p_{m_2}|^2\, d\nu_{\mathbf{x}_2;t_2-\frac{1}{8}} \\
&= \int_{\mathbb{R}^n} |\nabla^{m_2} p_{m_1} - \nabla^{m_2} p_{m_2}|^2\, d\nu_{\mathbf{x}_2;t_2-\frac{1}{8}} \\
&\le 2\int_{\mathbb{R}^n} |\nabla^{m_2} p_{m_1} - \nabla^{m_2} u|^2\, d\nu_{\mathbf{x}_2;t_2-\frac{1}{8}} \\
&\quad + 2\int_{\mathbb{R}^n} |\nabla^{m_2} u - \nabla^{m_2} p_{m_2}|^2\, d\nu_{\mathbf{x}_2;t_2-\frac{1}{8}} \\
&\le C^{\Lambda}\delta H_{\mathbf{x}_2}(1/8),
\end{aligned}$$

which is a contradiction if $\delta \le C^{-\Lambda}$. □

## 3. Symmetry and Cone Splitting

In this section, we define $k$-pinching and quantitative $k$-symmetry. In Theorem 3.6, we show that almost $k$-symmetry is equivalent to smallness of the $k$-pinching by establishing a cone splitting inequality in Theorem 3.11. In Lemma 3.12, we prove the existence of a uniform (across scales and locations) plane of symmetry. Finally, in Lemma 3.14, we prove that Jones $\beta$-numbers are controlled by the average frequency drop.

### 3.1. Quantitative symmetry.

Before we define symmetry, we define a quantitative notion of linear independence of space-time points, following [NV17a, Definition 4.5] from the elliptic setting.

**Definition 3.1.** Let $\mathrm{Gr}_{\mathcal{P}}(k)$ be the set of linear subspaces of $\mathbb{R}^{n+1}$ of one of the following types:

(i) $L \times \{0\}$, where $L \subseteq \mathbb{R}^n$ is a $k$-dimensional subspace,
(ii) $L \times \mathbb{R}$, where $L \subseteq \mathbb{R}^n$ is a $(k-2)$-dimensional subspace.

Note that $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ are the only linear subspaces that are invariant under parabolic scaling. If $V = L \times \{0\} \in \mathrm{Gr}_{\mathcal{P}}(k)$, define $V^{\perp} = L^{\perp} \times \mathbb{R}$. If $V = L \times \mathbb{R} \in \mathrm{Gr}_{\mathcal{P}}(k)$, define $V^{\perp} = L^{\perp} \times \{0\}$. Let $\mathrm{Aff}_{\mathcal{P}}(k)$ be the set of affine subspaces of the form $W = \mathbf{x} + V$, where $\mathbf{x} \in \mathbb{R}^{n+1}$ and $V \in \mathrm{Gr}_{\mathcal{P}}(k)$. For $W = \mathbf{x} + V \in \mathrm{Aff}_{\mathcal{P}}(k)$, we write $\hat{W} = V$ and $W^{\perp} = V^{\perp}$.

**Definition 3.2** (Independence)**.** We say $S \subseteq \mathbb{R}^{n+1}$ is $(k,\alpha)$-*independent* if there is no $W \in \mathrm{Aff}_{\mathcal{P}}(k-1)$ such that $S \subseteq P(W,\alpha)$. Given a $(k,\alpha)$-independent $S$, we say $S$ is $(k,\alpha)$-*spatially independent* if $|t_1 - t_2| < \alpha^2$ for any $\mathbf{x}_i = (x_i, t_i) \in S$, $i = 1, 2$; otherwise we say $S$ is $(k,\alpha)$-*temporally independent.*

As a result of Gram-Schmidt orthogonalization, we can prove that independent points give rise to a basis.

**Lemma 3.3.** *Let $S = \{\mathbf{x}_j\}_{j=0}^{K}$ be $(k,\alpha r)$-independent. Suppose either $K = k$ and $S$ is spatial, or $K = k-2$ and $S$ is temporal. Denote*

$$L \coloneqq \mathrm{Span}\{x_1 - x_0, \cdots, x_K - x_0\}.$$

*Let $V \coloneqq L \times \{0\}$ if $S$ is spatial, and $V \coloneqq L \times \mathbb{R}$ if $S$ is temporal. Then, for all $\mathbf{y} \in (\mathbf{x}_0 + V) \cap P(\mathbf{x}_0, 2r)$, there exists a unique set $\{q_j\}_{j=1}^{K} \subseteq \mathbb{R}$ such that*

$$y = x_0 + \sum_{j=1}^{K} q_j (x_j - x_0), \qquad\qquad |q_i| \le C\alpha^{-n} \frac{|y - x_0|}{r}.$$

*Proof.* Apply the argument of [NV17a, Lemma 4.6] to the set $\{x_j\}_{j=0}^{K}$. □

**Definition 3.4** (Frequency pinching)**.** Given a caloric function $u$, define

$$\mathcal{E}_r(\mathbf{x}) \coloneqq \mathcal{E}_r(\mathbf{x}; u) \coloneqq N_{\mathbf{x}}(8r^2) - N_{\mathbf{x}}\left(\tfrac{1}{8} r^2\right), \qquad\qquad \mathcal{E}_r(\{\mathbf{x}_i\}_{i=0}^{k}) \coloneqq \max_i \mathcal{E}_r(\mathbf{x}_i).$$

Define the $(k,\alpha r)$-*pinching* at $\mathbf{x}_0$

$$\mathcal{E}_r^{k,\alpha}(\mathbf{x}_0) \coloneqq \mathcal{E}_r^{k,\alpha}(\mathbf{x}_0; u) \coloneqq \inf \mathcal{E}_r(\{\mathbf{x}_i\}_{i=0}^{K}),$$

where the infimum is over $(k, \frac{1}{20}\alpha r)$-independent subsets $\{\mathbf{x}_i\}_{i=0}^{K} \subseteq P(\mathbf{x}_0, \frac{1}{10} r)$.

**Definition 3.5.** A caloric function $u$ is $(k,\epsilon,r)$-*symmetric* at $\mathbf{x} = (x,t)$ (with respect to $V$), if one of the following holds:

(1) There is a $k$-plane $L \subseteq \mathbb{R}^n$ such that $V = L \times \{0\}$ satisfies

$$r^2 \int_{\mathbb{R}^n} |\pi_L \nabla u|^2 \, d\nu_{\mathbf{x};t-r^2} \le \epsilon \int_{\mathbb{R}^n} |u|^2 \, d\nu_{\mathbf{x};t-r^2}.$$

(2) There is a $(k-2)$-plane $L \subseteq \mathbb{R}^n$ such that $V = L \times \mathbb{R}$ satisfies

$$r^2 \int_{\mathbb{R}^n} |\pi_L \nabla u|^2 \, d\nu_{\mathbf{x};t-r^2} + r^4 \int_{\mathbb{R}^n} |\partial_t u|^2 \, d\nu_{\mathbf{x};t-r^2} \le \epsilon \int_{\mathbb{R}^n} |u|^2 \, d\nu_{\mathbf{x};t-r^2}.$$

We say the symmetry is *spatial* if (1) holds and *temporal* otherwise.

In the following, we roughly show that $\mathcal{E}_r^{k,\alpha}(\mathbf{x})$ is small if and only if $\mathbf{x}$ is frequency-pinched and almost $k$-symmetric at scale $r$.

**Theorem 3.6.** *Let $u$ be a caloric function with $N_{\mathbf{x}}(10^5 r^2) \le \Lambda$.*

*(1) $u$ is $(k, C^{\Lambda}\alpha^{-n}\mathcal{E}_r^{k,\alpha}(\mathbf{x}), r)$-symmetric at $\mathbf{x}$.*

*(2) Suppose $u$ is $(k,\epsilon,10^2 r)$-symmetric at $\mathbf{x}$ with respect to $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ and*

$$|N_{\mathbf{x}}(10^{-2}\kappa^2 r^2) - N_{\mathbf{x}}(10^2 r^2)| \le \epsilon.$$

*Then*

$$\left|N_{\mathbf{v}}(50r^2) - N_{\mathbf{v}}\left(50^{-1}\kappa^2 r^2\right)\right| \le C(\kappa)^{\Lambda}\sqrt{\epsilon} \tag{3.1}$$

*for any* $\mathbf{v} \in (\mathbf{x}+V) \cap P(\mathbf{x}, 10r)$. *As a consequence,* $\mathcal{E}_s^{k,1}(\mathbf{v}) \leq C(\kappa)^\Lambda\sqrt{\epsilon}$, *and* $u$ *is* $(k, C(\kappa)^\Lambda\sqrt{\epsilon}, s)$*-symmetric at* $\mathbf{v}$ *with respect to* $V$ *for all* $s \in [\kappa r, r]$.

We first prove (2) and defer the proof of (1) to the Section 3.3.

*Proof of Theorem 3.6(2).* By parabolic rescaling and translation, we may assume $\mathbf{x} = \mathbf{0}$ and $r = 1$. First, suppose the symmetry is spatial: $V = L^k \times \{0\}$ and fix an arbitrary $\mathbf{v} = (v, 0) \in V \cap P(\mathbf{0}, 10)$. By the monotonicity and Lemma 2.3 with $\theta \leftarrow 10^{-1}, \sigma \leftarrow 10^{-5}$, $t_0 = t_1 \leftarrow 0$, $\tau \leftarrow 90$, $r \leftarrow 10$, if $w$ is caloric, then

$$\int_{\mathbb{R}^n} w^2 \, d\nu_{\mathbf{v};-s^2} \leq \int_{\mathbb{R}^n} w^2 \, d\nu_{\mathbf{v};-90} \leq C \int_{\mathbb{R}^n} w^2 \, d\nu_{-10^4},$$

for all $s \in [\kappa, 9]$. Applying this to $w = \nabla u \cdot v_i$ for an orthonormal basis $\{v_i\}_{i=1}^k$ of $L$, and using the assumption that $u$ is $(k, \epsilon, 10^2)$-symmetric at $\mathbf{0}$ with respect to $V$ yields

$$\begin{aligned}\int_{\mathbb{R}^n} |\pi_L \nabla u|^2 d\nu_{\mathbf{v};-s^2} &\leq C \int_{\mathbb{R}^n} |\pi_L \nabla u|^2 d\nu_{-10^4} \leq C\epsilon \int_{\mathbb{R}^n} u^2 d\nu_{-10^4} \\ &\leq C\epsilon \int_{\mathbb{R}^n} u^2 d\nu_{\mathbf{v};-2\cdot 10^4} \leq C\kappa^{-2\Lambda}\epsilon \int_{\mathbb{R}^n} u^2 d\nu_{\mathbf{v};-s^2}\end{aligned}$$

for all $s \in [\kappa, 9]$, where for the last two inequalities, we applied Lemma 2.3 and the assumption $N_{\mathbf{0}}(10^5) \leq \Lambda$. In other words, $N_{s^2;L}(\mathbf{w}) \leq C(\kappa)^\Lambda \epsilon$ for all $\mathbf{w} \in P(\mathbf{0}, 10) \cap V$ and $s \in [\kappa, 9]$. By Lemma 2.6, we have

$$\left| \frac{d}{d\sigma} D_{s^2}(\sigma v, 0) \right|^2 \leq C s^{-2} (N_{2s^2;L} + N_{s^2;L})(\sigma v, 0) \leq C(\kappa)^\Lambda \epsilon$$

for all $s \in [\kappa, 4]$ and $\sigma \in [0, 1]$. Thus

$$N_{\mathbf{v}}(s^2) \leq D_{\mathbf{v}}(s^2) \leq D_{\mathbf{0}}(s^2) + C(\kappa)^\Lambda \sqrt{\epsilon} \leq N_{\mathbf{0}}(2s^2) + C(\kappa)^\Lambda \sqrt{\epsilon}, \tag{3.2}$$

and similarly $N_{\mathbf{v}}(s^2) \geq N_{\mathbf{x}}(\frac{1}{2}s^2) - C(\kappa)^\Lambda\sqrt{\epsilon}$ for all $s \in [\kappa, 4]$. In the above computation, we used

$$N(\tau) \leq D(\tau) \leq N(2\tau), \tag{3.3}$$

which follows from the monotonicity of the frequency and (2.6).

Since $|N_{\mathbf{0}}(10^{-2}\kappa^2) - N_{\mathbf{0}}(10^2)| \leq \epsilon$, (3.1) holds for any $\mathbf{v} \in V \cap P(\mathbf{0}, 10)$ because of (3.2), (3.3) and the monotonicity of the frequency. As a consequence, $\mathcal{E}_s^{k,1}(\mathbf{v}) \leq C(\kappa)^\Lambda\sqrt{\epsilon}$, and $u$ is $(k, C(\kappa)^\Lambda\sqrt{\epsilon}, s)$-symmetric at $\mathbf{v}$ with respect to $V$ by the proof of (1).

If instead $V = L^k \times \mathbb{R}$, Lemma 2.3 with $r \leftarrow 10$, $\sigma = 1$, $\tau \leftarrow 9 \cdot 10^3$, and $\theta \leftarrow 10^{-1}$, we have

$$\int_{\mathbb{R}^n} |\partial_t u|^2 d\nu_{0,t;t-s^2} \leq \int_{\mathbb{R}^n} |\partial_t u|^2 d\nu_{0,t;t-9\cdot 10^3} \leq C \int_{\mathbb{R}^n} |\partial_t u|^2 d\nu_{\mathbf{0};-10^4}$$

for all $s \in [\kappa, 9]$ and $t \in [-100, 100]$. Given this, the proof of (3.1) proceeds as the case $V = L^k \times \{0\}$, using Lemma 2.6. □

3.2. **Propagation of symmetry.** In the following lemma, we prove that almost symmetry propagates to a smaller scale, with some loss in the estimates. Furthermore, if we assume that the frequency is already pinched, then almost symmetry propagates to a larger scale as well.

**Lemma 3.7.** *For any* $\alpha > 0$, $0 < \beta \leq 1$, *and* $\delta > 0$, *the following statements hold. Suppose* $u$ *is a caloric function and* $r > 0$ *is such that* $N(10^5 r^2) \leq \Lambda$. *Fix* $\mathbf{x} \in P(\mathbf{0}, r)$.

*(1) If* $u$ *is* $(k, \delta, r)$*-symmetric with respect to* $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ *at* $\mathbf{x}$, *then* $u$ *is* $(k, \beta^{-2\Lambda}\delta, s)$*-symmetric with respect to* $V$ *at* $\mathbf{x}$ *for all* $s \in [\beta r, r]$.

*(2) If* $0 < r_1 \leq r$, $|N_{\mathbf{x}}(r_1^2) - N_{\mathbf{x}}(4r^2)| < \delta < C^{-\Lambda}$, *and for some* $r_2 \in [r_1, r]$, $u$ *is* $(k, \delta, r_2)$*-symmetric at* $\mathbf{x}$ *with respect to* $V \in \mathrm{Gr}_{\mathcal{P}}(k)$, *then, for all* $s \in [r_1, r]$, $u$ *is* $(k, C^\Lambda\delta, s)$*-symmetric at* $\mathbf{x}$ *with respect to* $V$.

*Proof.* (1) We may assume $\mathbf{x} = \mathbf{0}$ and $r = 1$. Because $\partial_t u$ and $\nabla u$ are caloric, if $u$ is $(k, \delta, 1)$-symmetric with respect to $L \times \mathbb{R}$, then, for any $s \in [\beta, 1]$, we have

$$\int_{\mathbb{R}^n} (|\pi_L \nabla u|^2 + |\partial_t u|^2)\, d\nu_{-s^2} \le \delta \int_{\mathbb{R}^n} |u|^2\, d\nu_{-1} \le \delta\beta^{-2\Lambda} \int_{\mathbb{R}^n} |u|^2\, d\nu_{-s^2}.$$

The spatially symmetric case is identical.

(2) We may assume $\mathbf{x} = \mathbf{0}$ and $r_2 = 1$. We may also assume $u$ is $(k, \delta, 1)$-symmetric with respect to $V = W \times \{0\}$ for some $k$-plane $W$, as the other case can be dealt with similarly. Then

$$\int_{\mathbb{R}^n} |\pi_W \nabla u|^2\, d\nu_{-1} \le \delta \int_{\mathbb{R}^n} u^2\, d\nu_{-1}. \tag{3.4}$$

Theorem 2.13 implies that

$$s^2 \int_{\mathbb{R}^n} |\nabla(u - p_m)|^2\, d\nu_{-s^2} + s^4 \int_{\mathbb{R}^n} |\partial_t(u - p_m)|^2 d\nu_{-s^2} \le C^\Lambda \delta \int_{\mathbb{R}^n} u^2\, d\nu_{-s^2}, \tag{3.5}$$

for any $s \in [r_1, r]$, where $p_m$ is a caloric homogeneous polynomial of degree $m$, which is independent of $s$ and satisfies

$$\frac{1}{2} \int_{\mathbb{R}^n} u^2\, d\nu_{-s^2} \le \int_{\mathbb{R}^n} p_m^2\, d\nu_{-s^2} \le \int_{\mathbb{R}^n} u^2\, d\nu_{-s^2} \tag{3.6}$$

for any $s \in [r_1, r]$. Combining (3.4), (3.5), and (3.6), we obtain

$$\int_{\mathbb{R}^n} |\pi_W \nabla p_m|^2\, d\nu_{-1} \le C^\Lambda \delta \int_{\mathbb{R}^n} p_m^2\, d\nu_{-1}.$$

Because $p_m$ is homogeneous, it follows that

$$s^2 \int_{\mathbb{R}^n} |\pi_W \nabla p_m|^2\, d\nu_{-s^2} \le C^\Lambda \delta \int_{\mathbb{R}^n} p_m^2\, d\nu_{-s^2} \tag{3.7}$$

for any $s \in [r_1, r]$. By (3.5) and (3.7), for any $s \in [r_1, r]$

$$s^2 \int_{\mathbb{R}^n} |\pi_W \nabla u|^2\, d\nu_{-s^2} \le C^\Lambda \delta \int_{\mathbb{R}^n} u^2\, d\nu_{-s^2}.$$

□

3.3. **Sharp cone splitting inequality.** In this subsection, we prove a sharp *cone splitting inequality* (cf. Theorem 3.11) which implies Theorem 3.6(1). To motivate the ideas, we first consider the exact case. Suppose the frequency is constant at two distinct points $\mathbf{x}_1 \neq \mathbf{x}_2$. By Lemma 2.5 and Corollary 2.14, $u$ is an eigenfunction of $2\tau_i \Delta_{f_i}$ with the same eigenvalue $m \in \mathbb{N}_0$, i.e.,

$$2\tau_i \Delta_{f_i} u + mu = 0,$$

where $f_i = f_{\mathbf{x}_i}$, $\tau_i = t_i - t$. The following computation shows that $u$ is spatially invariant in the direction of $(x_2 - x_1)$ if $x_2 \neq x_1$; similarly, $u$ is static if $t_1 \neq t_2$ :

$$\nabla u \cdot (x_2 - x_1) = 4\tau_2 \Delta_{f_2} u - 4\tau_1 \Delta_{f_1} u - [2\tau_1 \Delta_{f_1}, 2\tau_2 \Delta_{f_2}] u = 0, \tag{3.8a}$$

$$2(t_2 - t_1)\partial_t u = [2\tau_1 \Delta_{f_1}, 2\tau_2 \Delta_{f_2}] u + 2\tau_1 \Delta_{f_1} u - 2\tau_2 \Delta_{f_2} u = 0. \tag{3.8b}$$

We shall make the arguments above effective. We first prove that homogeneity at a point arises from the smallness of the frequency pinching at that point.

**Lemma 3.8.** *Let $u$ be a caloric function with $N(\tau_2) - N(\tau_1) \le \delta$ for some $0 < \tau_1 < \frac{\tau_2}{2}$, and let $m \in \mathbb{N}_0$ be as in Lemma 2.11(3). Then, for any $\tau \in (\tau_1, \tau_2)$,*

$$\int_{\mathbb{R}^n} v^2\, d\nu_{-\tau} \le \frac{C\tau_2}{\tau_2 - \tau} \delta \int_{\mathbb{R}^n} u^2\, d\nu_{-\tau_2},$$

*where $v := 2\tau \Delta_f u + mu$.*

*Proof.* By parabolic rescaling, we may assume $\tau_2 = 1$. Since $v$ is caloric, we may apply the monotonicity formula in Lemma 2.5 so that for any $s \in (-1, -\tau_1)$, we obtain

$$\int_{\mathbb{R}^n} v^2\, d\nu_s \le \frac{1}{s+1}\int_{-1}^{s}\int_{\mathbb{R}^n} v^2\, d\nu_t dt \le \frac{2}{s+1}\int_{-1}^{s}\int_{\mathbb{R}^n} |2\tau\Delta_f u + N(\tau)u|^2 + (N(\tau)-m)^2 u^2\, d\nu_t dt$$
$$\le \frac{2}{s+1} H(1)\int_{-1}^{s} \tau N'(\tau)\, dt + C\delta^2 H(1) \le \frac{C}{s+1}\delta H(1).$$

□

The following lemma allows us to get rid of one of the $2\tau_i\Delta_{f_i}$ in $2\tau_1\Delta_{f_1}(2\tau_2\Delta_{f_2})$ in the commutator identities (3.8a) and (3.8b).

**Lemma 3.9.** *Suppose $u$ is a caloric function. Then, for any $0 < \tau < \tau_1$, we have*

$$\int_{\mathbb{R}^n} (2\tau\Delta_f u)^2\, d\nu_{-\tau} \le \frac{4\tau\tau_1}{(\tau_1-\tau)^2}\int_{\mathbb{R}^n} u^2\, d\nu_{-\tau_1}.$$

*Proof.* By parabolic rescaling, we may assume $\tau_1 = 1$. Pick $-1 < s < t < 0$. Since $2\tau\Delta_f u$ is caloric, by the monotonicity formula in Lemma 2.5,

$$\int_{\mathbb{R}^n} (2\tau\Delta_f u)^2\, d\nu_{-\tau} \le \frac{1}{t-s}\int_{s}^{t}\int_{\mathbb{R}^n} (2\tau\Delta_f u)^2\, d\nu_r dr \le \frac{1}{t-s}\int_{s}^{t} |r|E'(|r|)\, dr$$
$$\le \frac{-s}{t-s}E(-s) = \frac{2s^2}{t-s}\int_{\mathbb{R}^n} |\nabla u|^2\, d\nu_s \le \frac{2s^2}{(t-s)(s+1)}\int_{-1}^{s}\int_{\mathbb{R}^n} |\nabla u|^2\, d\nu_t dt$$
$$\le \frac{s^2}{(t-s)(s+1)}\int_{\mathbb{R}^n} u^2\, d\nu_{-1}.$$

Noting that $s = \frac{2t}{1-t}$ minimizes the coefficient above, we get the desired inequality. □

In the following proposition, we prove a cone splitting inequality for two points.

**Proposition 3.10.** *Let $u$ be a caloric function. Fix $\mathbf{x}_2 \in P(\mathbf{x}_1, \frac{1}{10}r)$. If $N_{\mathbf{x}_j}(r^2) \le \Lambda$, and $\delta := \mathcal{E}_r(\{\mathbf{x}_1, \mathbf{x}_2\}) \le \bar\delta(\Lambda) = C^{-\Lambda}$, then, for $j = 1, 2$, we have*

$$\int_{\mathbb{R}^n} |(x_2 - x_1)\cdot\nabla u|^2 + |(t_2-t_1)\cdot\partial_t u|^2\, d\nu_{\mathbf{x}_j; t_j - r^2} \le C(1+\Lambda^2)\delta\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{x}_j; t_j - 2r^2}.$$

*Proof.* Without loss of generality, we assume that $r = 1$. We only prove the first inequality for $j = 1$. By Corollary 2.14, if $\delta \le C^{-\Lambda}$, there is a common $m \in \mathbb{N}_0$ such that

$$\sup_{\tau\in[\frac{1}{8},8]} |N_{\mathbf{x}_j}(\tau) - m| < C\delta.$$

For $i = 1, 2$, denote $f_i = f_{\mathbf{x}_i}, \tau_i = t_i - t$, $v_i = 2(t_i - t)\Delta_{f_i}u + mu$. Recall that

$$\begin{aligned}\nabla u\cdot(x_2 - x_1) &= 4\tau_2\Delta_{f_2}u - 4\tau_1\Delta_{f_1}u - [2\tau_1\Delta_{f_1}, 2\tau_2\Delta_{f_2}]u \\ &= -2\tau_1\Delta_{f_1}v_2 + 2\tau_2\Delta_{f_2}v_1 - (2-m)v_1 + (2-m)v_2.\end{aligned} \tag{3.9}$$

Let $\eta = \frac{9}{8}$. Since $v_i$ is caloric, by using Lemma 3.9, changing the base points (Lemma 2.3 with $\theta = \frac{1}{8}$), and using Lemma 3.8, we get

$$\int_{\mathbb{R}^n} |2\tau_1\Delta_{f_1}v_2|^2\, d\nu_{\mathbf{x}_1; t_1 - 1} \le C\int_{\mathbb{R}^n} v_2^2\, d\nu_{\mathbf{x}_1; t_1-\eta} \le C\int_{\mathbb{R}^n} v_2^2\, d\nu_{\mathbf{x}_2; t_2 - \eta^2} \le C\delta\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{x}_2; t_2-\eta^3}$$
$$\le C\delta\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{x}_1; t_1 - 2}.$$

Using similar techniques, we can estimate

$$\int_{\mathbb{R}^n} |2\tau_2 \Delta_{f_2} v_1|^2 \, d\nu_{\mathbf{x}_1;t_1-1} \leq C\delta \int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_1;t_1-2};$$

$$\int_{\mathbb{R}^n} (v_1^2 + v_2^2) \, d\nu_{\mathbf{x}_1;t_1-1} \leq C\delta \int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_1;t_1-2}.$$

Combining the estimates above and using (3.9), we get the desired estimates. □

Applying the above Proposition multiple times, we get the following cone splitting inequality.

**Theorem 3.11.** *Let $u$ be a caloric function with $N_{\mathbf{x}_0}(10^5 r^2) \leq \Lambda$. If $\{\mathbf{x}_i\}_{i=0}^K \subset P(\mathbf{x}_0, \frac{1}{10} r)$ is $(k, \alpha r)$-independent and $K = k$, then*

$$r^2 \int_{\mathbb{R}^n} |\pi_L \nabla u|^2 \, d\nu_{\mathbf{x}_0;t_0-r^2} \leq C^\Lambda \alpha^{-n} \mathcal{E}_r(\{\mathbf{x}_i\}) \int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_0;t_0-r^2},$$

*where $L = \operatorname{Span}\{x_1 - x_0, \cdots, x_K - x_0\}$. If $\{\mathbf{x}_i\}_{i=0}^K$ is $(k, \alpha r)$-temporally independent and $K = k-2$, then*

$$\int_{\mathbb{R}^n} r^2 |\pi_L \nabla u|^2 + r^4 |\partial_t u|^2 \, d\nu_{\mathbf{x}_0;t_0-r^2} \leq C^\Lambda \alpha^{-n} \mathcal{E}_r(\{\mathbf{x}_i\}) \int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_0;t_0-r^2}.$$

*Proof.* The inequalities hold by applying Proposition 3.10 for $\mathbf{x}_1 \leftarrow \mathbf{x}_0$, $\mathbf{x}_2 \leftarrow \mathbf{x}_j$ for $j = 1, \cdots, K$, and by Lemma 3.3 and (2.6). □

*Proof of Theorem 3.6(1).* Apply Theorem 3.11. □

3.4. **Existence of a plane of symmetry.** In the lemma below, we prove that there exists a plane of symmetry that is uniform across scales and locations. This will be the main tool for proving the neck decomposition theorem (cf. Theorem 5.5).

**Lemma 3.12.** *For any $\delta, \eta, \kappa \in (0, 1]$ and $\Lambda \in [1, \infty)$, the following holds for any $r_0 > 0$ and any caloric function $u$ satisfying $N(10^5 \kappa^{-2} r_0^2) \leq \Lambda$. Suppose $\mathcal{C} \subseteq P(\mathbf{0}, r_0)$ and $\mathbf{r}_\bullet : \mathcal{C} \to [0, r_0]$ is a function satisfying the following for some $m \in \mathbb{N}_0$:*

*(a) there exists $\mathbf{x}_0 \in \mathcal{C}$ and $s_0 \in [\mathbf{r}_{\mathbf{x}_0}, r_0]$ such that $u$ is $(k, \delta, s_0)$-symmetric with respect to $V$;*
*(b) $\sup_{\mathbf{x} \in \mathcal{C}} \sup_{s \in [10^{-3}\kappa \mathbf{r}_{\mathbf{x}}, 10^3 \kappa^{-1} r_0]} |N_{\mathbf{x}}(s^2) - m| < \delta$.*

*Then the following statements hold for any $\mathbf{x} \in \mathcal{C}$:*

*(1) for any $s \in [\kappa \mathbf{r}_{\mathbf{x}}, \kappa^{-1} r_0]$, $u$ is $(k, C^\Lambda \delta, s)$-symmetric with respect to $V$ at $\mathbf{x}$;*
*(2) for any $s \in [\mathbf{r}_{\mathbf{x}}, r_0]$,*

$$(\mathbf{x} + V) \cap P(\mathbf{x}, 10s) \subset \{\mathbf{y} \in P(\mathbf{x}, 10s) \colon |N_{\mathbf{y}}(\kappa^2 s^2) - m| < C(\kappa)^\Lambda \sqrt{\delta}\};$$

*(3) if in addition we assume that $u$ is not $(k+1, \eta, s_0)$-symmetric at $\mathbf{x}_0$, then, for any $s \in [\mathbf{r}_{\mathbf{x}}, 10^{-2} r_0]$ and $\zeta \geq C^\Lambda \sqrt{\delta}$,*

$$\{\mathbf{y} \in P(\mathbf{x}, 10s) : \mathcal{E}_{100s}(\mathbf{y}) < \zeta\} \subseteq P(\mathbf{x} + V, C^\Lambda \eta^{-\frac{1}{n}} \zeta^{\frac{1}{n}} s).$$

*Proof.* By parabolic rescaling, we may assume $r_0 = 1$.

(1) By Lemma 3.7(2) with $r \leftarrow 100$, $r_1 \leftarrow \mathbf{r}_{\mathbf{x}_0}$, $r_2 \leftarrow s_0$, $u$ is $(k, C^\Lambda \delta, 100)$-symmetric at $\mathbf{x}_0$ with respect to $V$. By Lemma 2.3, $u$ is $(k, C^\Lambda \delta, 1)$-symmetric with respect to $V$ at any $\mathbf{x} \in P(\mathbf{0}, 1)$. The claim then follows by applying Lemma 3.7(2) to each point $\mathbf{x} \in \mathcal{C}$.

(2) Apply (1) and Theorem 3.6(2).

(3) Suppose $\mathbf{x} \in \mathcal{C}$, $s \in [\mathbf{r}_{\mathbf{x}}, 10^{-2}]$, $\mathbf{y} \in P(\mathbf{x}, 10s)$, and $\mathcal{E}_{100s}(\mathbf{y}) < \zeta$ for some $\zeta \geq C^\Lambda \sqrt{\delta}$. Set $\epsilon := C^\Lambda \eta^{-\frac{1}{n}} \zeta^{\frac{1}{n}}$, and assume by way of contradiction that $\mathbf{y} \notin P(\mathbf{x} + V, \epsilon s)$. By applying (2) twice with $s \leftarrow s$ and $s \leftarrow 100s$, there exists a $(k, 100s)$-independent set $\{\mathbf{v}_i\}_{i \in I} \subseteq (\mathbf{x} + V) \cap P(\mathbf{x}, 10s)$ containing $\mathbf{x}$ and satisfying $\mathcal{E}_{100s}(\mathbf{v}_i) < C^\Lambda \sqrt{\delta} \leq \zeta$ for all $i \in I$. Thus $\mathcal{E}_{100s}^{k+1,\epsilon}(\mathbf{x}) < \zeta$, so by Theorem 3.6(1) and Lemma 3.7(2), $u$ is $(k+1, C^{-\Lambda}\eta, s)$-symmetric at $\mathbf{x}$. By Lemma 3.7(2), it follows that

$u$ is $(k+1, C^{-\Lambda}\eta, 100)$-symmetric at $\mathbf{x}$, hence by Lemma 2.3, $u$ is also $(k+1, C^{-\Lambda}\eta, 1)$-symmetric at $\mathbf{x}_0$. Another application of Lemma 3.7(2) finally yields that $u$ is $(k+1, \eta, s_0)$-symmetric at $\mathbf{x}_0$, a contradiction. $\square$

3.5. **$L^2$-subspace approximation.** In this subsection, we make precise use of Theorem 3.11 to prove Lemma 3.14. The lemma states that the average frequency pinching bounds from above the Jones beta number (defined below), which quantifies, in the $L^2$ sense, how well the support of a measure $\mu$ can be approximated by an affine subspace. The strategy of the proof is from [NV17a, Section 7].

**Definition 3.13.** For any measure $\mu$ on $\mathbb{R}^{n+1}$, an integer $0 \leq k \leq n+2$, and any $\mathbf{x} \in \mathbb{R}^{n+1}$, we define the *parabolic Jones $\beta$-number* by

$$\beta^2_{\mathcal{P},k}(\mathbf{x}, r) := \beta^2_{\mathcal{P},k}(\mathbf{x}, r; \mu) := \inf_{V \in \mathrm{Aff}_{\mathcal{P}}(k)} \frac{1}{r^k} \int_{P(\mathbf{x},r)} \left( \frac{d_{\mathcal{P}}(\mathbf{y}, V)}{r} \right)^2 d\mu(\mathbf{y}).$$

**Lemma 3.14.** *Let $k \in \{n, n+1\}$. For any finite measure $\mu$ supported in $P(\mathbf{0}, 1)$ on $\mathbb{R}^{n+1}$ and any caloric function $u$ satisfying $N_{\mathbf{0}}(10^5) \leq \Lambda$, if $u$ is $(k, \epsilon, 2r)$-symmetric but not $(k+1, \eta, r)$-symmetric at $\mathbf{0}$, then, for any $P(\mathbf{x}, r) \subseteq P(\mathbf{0}, 2)$, if $\epsilon \leq C^{-\Lambda}\eta$, we have*

$$\beta^2_{\mathcal{P},k}(\mathbf{x}, r; \mu) \leq \frac{C^\Lambda}{r^k \eta} \left( \int_{P(\mathbf{x},r)} \mathcal{E}_{20r}(\mathbf{y}) d\mu(\mathbf{y}) \right).$$

We prove the lemma first by characterizing $\beta$ in terms of the eigenvalues of the *covariance matrix* $Q$ of $\mu$ defined below and then estimating the eigenvalues.

**Definition 3.15.** For a probability measure $\mu$ supported on $P(\mathbf{0}, 1)$, we define the corresponding center of mass and covariance matrix by

$$x_{\mathrm{cm}} := \int_{P(\mathbf{0},1)} x \, d\mu(x, t), \quad Q := \int_{P(\mathbf{0},1)} (x - x_{\mathrm{cm}})(x - x_{\mathrm{cm}})^T \, d\mu(x, t), \tag{3.10}$$

respectively. Let $\lambda_1 \leq \cdots \leq \lambda_n$ be the eigenvalues of $Q$, with corresponding unit eigenvectors $v_1, \ldots, v_n$.

**Lemma 3.16.** *Let $\mu$ be a probability measure supported in $P(\mathbf{0}, 1)$. Then*

$$\beta^2_{\mathcal{P},k}(\mathbf{0}, 1; \mu) \leq \lambda_1 + \cdots + \lambda_{n+2-k}. \tag{3.11}$$

*Proof.* For any $L \in \mathrm{Gr}(k-2)$, we have $d_{\mathcal{P}}(y, x_{\mathrm{cm}} + V) = |\pi_V^\perp(\mathbf{y} - \mathbf{x}_{\mathrm{cm}})|$, so that

$$\begin{aligned}\beta^2_{\mathcal{P},k}(\mathbf{0}, 1) &\leq \inf_{L \in \mathrm{Gr}(k-2)} \int_{P(\mathbf{0},1)} |\pi_L^\perp(y - x_{\mathrm{cm}})|^2 \, d\mu(\mathbf{y}) = \inf_{L \in \mathrm{Gr}(n+2-k)} \int_{P(\mathbf{0},1)} |\pi_L(y - x_{\mathrm{cm}})|^2 \, d\mu(\mathbf{y}) \\ &\leq \lambda_1 + \cdots + \lambda_{n+2-k}.\end{aligned}$$

$\square$

In the lemma below, we estimate the eigenvalues of $Q$ in terms of the average frequency pinching.

**Lemma 3.17.** *Let $v_j$ be the eigenvector of the covariance matrix $Q$ with eigenvalue $\lambda_j$. Then*

$$\lambda_j \int_{\mathbb{R}^n} (v_j \cdot \nabla u)^2 \, d\nu_{-400} \leq C^\Lambda \int_{P(\mathbf{0},1)} \mathcal{E}_{20}(\mathbf{x}) \, d\mu(\mathbf{x}) \cdot \int_{\mathbb{R}^n} u^2 \, d\nu_{-400}.$$

*Proof.* For any $z \in \mathbb{R}^n, t_* < 0$,

$$\lambda_j v_j \cdot \nabla u(z, t_*) = Q v_j \cdot \nabla u(z, t_*) = \int_{P(\mathbf{0},1)} \left( (x - x_{\mathrm{cm}}) \cdot v_j \right) \nabla u(z, t_*) \cdot (x - x_{\mathrm{cm}}) \, d\mu(x, t)$$

$$= \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} ((x - x_{\mathrm{cm}}) \cdot v_j) \, \nabla u(z, t_*) \cdot (x - y) \, d\mu(x,t) d\mu(y,s).$$

Integration against $\nu_{t_*}$ and Cauchy–Schwarz yield

(3.12)

$$\begin{aligned}
&\lambda_j^2 \int_{\mathbb{R}^n} (v_j \cdot \nabla u(z,t_*))^2 \, d\nu_{t_*}(z) \\
&= \int_{\mathbb{R}^n} \left( \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} ((x - x_{\mathrm{cm}}) \cdot v_j) \, \nabla u(z,t_*) \cdot (x-y) \, d\mu(\mathbf{x}) d\mu(\mathbf{y}) \right)^2 d\nu_{t_*}(z) \\
&\leq \left( \int_{P(\mathbf{0},1)} ((x - x_{\mathrm{cm}}) \cdot v_j)^2 d\mu(\mathbf{x}) \right) \int_{\mathbb{R}^n} \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} \left( \nabla u(z,t_*) \cdot (x-y) \right)^2 \, d\mu(\mathbf{x}) d\mu(\mathbf{y}) d\nu_{t_*}(z) \\
&= \lambda_j \int_{\mathbb{R}^n} \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} \left( \nabla u(z,t_*) \cdot (x-y) \right)^2 \, d\mu(\mathbf{x}) d\mu(\mathbf{y}) d\nu_{t_*}(z).
\end{aligned}$$

On the other hand, for any fixed $\mathbf{x}, \mathbf{y} \in P(\mathbf{0}, 1)$, Lemma 2.3 and Proposition 3.10 with $r \leftarrow 20$ yield

$$\begin{aligned}
\int_{\mathbb{R}^n} (\nabla u(z, -400) \cdot (x-y))^2 \, d\nu_{-400}(z) &\leq \int_{\mathbb{R}^n} (\nabla u(z,-400) \cdot (x-y))^2 \, d\nu_{\mathbf{x};t-440}(z) \\
&\leq C^\Lambda \mathcal{E}_{20}(\{\mathbf{x}, \mathbf{y}\}) \int_{\mathbb{R}^n} u^2 \, d\nu_{-400}.
\end{aligned}$$

Integrating over $\mathbf{x}, \mathbf{y} \in P(\mathbf{0}, 1)$ and combining the result with (3.12) yields

$$\begin{aligned}
\lambda_j \int_{\mathbb{R}^n} (v_j \cdot \nabla u)^2 \, d\nu_{-400} &\leq \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} \int_{\mathbb{R}^n} ((x-y) \cdot \nabla u(z,-400))^2 d\nu_{-400}(z) \, d\mu(\mathbf{x}) d\mu(\mathbf{y}) \\
&\leq C^\Lambda \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} (\mathcal{E}_{20}(\mathbf{x}) + \mathcal{E}_{20}(\mathbf{y})) \, d\mu(\mathbf{x}) d\mu(\mathbf{y}) \cdot \int_{\mathbb{R}^n} u^2 \, d\nu_{-400} \\
&\leq C^\Lambda \int_{P(\mathbf{0},1)} \mathcal{E}_{20}(\mathbf{x}) \, d\mu(\mathbf{x}) \cdot \int_{\mathbb{R}^n} u^2 \, d\nu_{-400}.
\end{aligned}$$

□

*Proof of Lemma 3.14.* By translation and rescaling, we may assume $\mathbf{x} = \mathbf{0}$ and $r = 1$. By rescaling $\mu$, we may also assume that $\mu$ is a probability measure. By shifting $\mu$ we may assume $\mathbf{x}_{\mathrm{cm}} = \mathbf{0}$.

Recall that $u$ is $(k, \epsilon, 2)$-symmetric. We first consider the case where the symmetry is temporal. Because $u$ is not $(k+1, \eta, 1)$-symmetric at $\mathbf{0}$,

$$\int_{\mathbb{R}^n} |\pi_L \nabla u|^2 \, d\nu_{-1} \geq (\eta - \epsilon) \int_{\mathbb{R}^n} u^2 \, d\nu_{-1},$$

where $L \coloneqq \mathrm{span}(v_{n+2-k}, \ldots, v_n)$. Thus there exists $j \in \{n+2-k, \ldots, n\}$ such that

$$\int_{\mathbb{R}^n} |v_j \cdot \nabla u|^2 \, d\nu_{-1} \geq C^{-\Lambda} (\eta - \epsilon) \int_{\mathbb{R}^n} u^2 \, d\nu_{-1}.$$

Combining this with Lemma 3.16 and 3.17, we get

$$\beta^2_{\mathcal{P},k}(\mathbf{0}, 1) \leq \lambda_{n+2-k} \leq C^\Lambda (\eta - \epsilon)^{-1} \int_{P(\mathbf{0},1)} \mathcal{E}_{20}(\mathbf{x}) \, d\mu(\mathbf{x}).$$

Suppose instead that the symmetry is spatial (so that $k = n$). Then

$$\int_{\mathbb{R}^n} |\partial_t u|^2 \, d\nu_{-1} = \int_{\mathbb{R}^n} |\Delta u|^2 \, d\nu_{-1} \leq C \int_{\mathbb{R}^n} |\nabla^2 u|^2 \, d\nu_{-1} \leq \int_{\mathbb{R}^n} |\nabla u|^2 \, d\nu_{-4} \leq \epsilon C^\Lambda \int_{\mathbb{R}^n} u^2 \, d\nu_{-1},$$

which contradicts our assumption that $u$ is not $(n+2, \eta, 1)$-symmetric at $\mathbf{0}$. □

## 4. Neck Regions

In this section, we define neck regions and study their structure. The neck regions (see Definition 4.1) serve as transition zones between quantitative singular regions and regular (i.e., highly symmetric) regions. They capture the presence of *multi-scale symmetry* and *homogeneity.* In the neck structure theorem (Theorem 4.6), we show that the singular sets associated with these neck regions align into $k$-dimensional structures across scales and locations, a property that is essential for establishing the content estimates.

Throughout this section, fix $\gamma \coloneqq 10^{-10n}$. Define $\overline{P}(\mathcal{C}, \mathbf{r}_\bullet) \coloneqq \cup_{\mathbf{x}\in\mathcal{C}} \overline{P}(\mathbf{x}, \mathbf{r}_\mathbf{x})$ when $\mathbf{r}_\bullet : \mathcal{C} \to [0,1]$.

Our definitions of neck region and strong neck region are adaptations of [JN21, Definition 7.10] and [CJN21, Definition 2.4] to the parabolic setting. However, in our setting we can use quantitative uniqueness (through Lemma 3.12) to show that the approximating planes of symmetry can be taken independent of both scales and locations. This greatly simplifies our proof of the weak neck structure theorem (Theorem 4.6).

**Definition 4.1** (Neck Region)**.** Let $\mathcal{C} \subseteq P(\mathbf{x}_0, 2r)$ be a closed subset such that $\mathbf{x}_0 \in \mathcal{C}$, $\mathbf{r}_\bullet : \mathcal{C} \to [0, \gamma r]$ be a continuous function and $k, m \in \mathbb{N}_0$, $\delta, \eta > 0$. We say that the set

$$\mathcal{N} \coloneqq P(\mathbf{x}_0, 2r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$$

is an $(m, k, \delta, \eta)$*-neck region (of scale $r$)* modeled on $V \in \mathrm{Gr}_\mathcal{P}(k)$ if the following hold:

- (n1) $\{P(\mathbf{x}, \gamma^2\mathbf{r}_\mathbf{x})\}_{\mathbf{x}\in\mathcal{C}}$ is pairwise disjoint;
- (n2) $\sup_{s\in[\mathbf{r}_\mathbf{x},\gamma^{-3}r]} |N_\mathbf{x}(s^2) - m| < \delta$ for all $\mathbf{x} \in \mathcal{C}$;
- (n3) For all $s \in [\mathbf{r}_\mathbf{x}, \gamma^{-3}r]$, $u$ is $(k, \delta, s)$-symmetric at $\mathbf{x}$ with respect to $V$ but not $(k+1, \eta, s)$-symmetric;
- (n4) For all $\mathbf{x} \in \mathcal{C}$ and $s \in [\mathbf{r}_\mathbf{x}, \gamma^{-3}r]$ such that $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_0, 2r)$, we have

$$\mathcal{C} \cap P(\mathbf{x}, s) \subseteq P(\mathbf{x} + V, \delta s), \tag{n4.a}$$

$$(\mathbf{x} + V) \cap P(\mathbf{x}, s) \subseteq \bigcup_{\mathbf{y}\in\mathcal{C}} P(\mathbf{y}, 10\gamma(s + \mathbf{r}_\mathbf{y})). \tag{n4.b}$$

We say $\mathcal{N}$ is a *strong* $(m, k, \delta, \eta)$*-neck region* if (n1)-(n3) and (n4.a) hold, and if in addition we have the following:

- (n4') For all $\mathbf{x} \in \mathcal{C}$ and $s \in [\mathbf{r}_\mathbf{x}, \gamma^{-3}r]$ such that $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_0, 2r)$, we have

$$(\mathbf{x} + V) \cap P(\mathbf{x}, s) \subseteq P(\mathcal{C}, \gamma s); \tag{n4.b'}$$

- (n5) $\mathbf{r}_\bullet : \mathcal{C} \to [0, \gamma r]$ is parabolically $\delta$-Lipschitz.

**Example 4.2.** If $p$ is a homogeneous caloric polynomial of degree $m$ that is invariant with respect to $V \in \mathrm{Gr}_\mathcal{P}(k)$, then defining $\mathcal{C} = \mathcal{C}_0 \coloneqq V$ and $\mathbf{r} \equiv 0$ yields an $(m, k, 0, \eta)$-neck region modeled on $V$.

We refer to $\mathcal{C}$ as the *center set* associated to $\mathcal{N}$, and define the subsets

$$\mathcal{C}_+ \coloneqq \{\mathbf{x} \in \mathcal{C} \colon \mathbf{r}_\mathbf{x} > 0\}, \qquad \mathcal{C}_0 \coloneqq \{\mathbf{x} \in \mathcal{C} \colon \mathbf{r}_\mathbf{x} = 0\}.$$

We also define the *packing measure* associated to the center set $\mathcal{C}$.

$$\mu \coloneqq \mu_0 + \mu_+ \coloneqq \mathcal{H}^{n+1}_\mathcal{P}|_{\mathcal{C}_0} + \sum_{\mathbf{x}\in\mathcal{C}_+} \mathbf{r}_\mathbf{x}^{n+1}\delta_\mathbf{x}. \tag{4.2}$$

In the following lemma, we prove that restriction of a strong neck region still defines a strong neck region.

**Lemma 4.3.** *Suppose* $\mathcal{N} \coloneqq P(\mathbf{x}_0, 2r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$ *is a strong* $(m, k, \delta, \eta)$*-neck region with* $\mathbf{r}_\bullet \le \frac{1}{4}\gamma r$*, and with* (n4.b') *replaced by*

$$(\mathbf{x} + V) \cap P(\mathbf{x}, s) \subseteq P(\mathcal{C}, \tfrac{\gamma}{10}s). \tag{n4.b''}$$

*Then* $\mathcal{N}' := P(\mathbf{x}_0, \frac{r}{2}) \setminus \bigcup_{\mathbf{y}\in\mathcal{C}\cap P(\mathbf{x}_0,\frac{r}{2})} \overline{P}(\mathbf{y}, \mathbf{r}_\mathbf{y})$ *is a strong* $(m, k, \delta, \eta)$*-neck region of scale* $\frac{r}{4}$.

*Proof.* We may assume $\mathbf{x}_0 = \mathbf{0}$ and $r = 1$. It suffices to justify (n4.b') since the other statements hold because $\mathcal{N}$ is a neck region. Fix $\mathbf{x} \in \widetilde{\mathcal{C}} := \mathcal{C} \cap P(\mathbf{0}, \frac{1}{2})$. Suppose $s \in [\mathbf{r}_\mathbf{x}, \frac{1}{4}\gamma^{-3}]$ such that $P(\mathbf{x}, s) \subset P(\mathbf{0}, \frac{1}{2})$. Fix $\mathbf{z} \in (\mathbf{x} + V) \cap P(\mathbf{x}, s)$. Let $\mathbf{z}' \in (\mathbf{x} + V) \cap P(\mathbf{x}, (1 - \frac{\gamma}{2})s)$ such that $|\mathbf{z} - \mathbf{z}'| = \frac{\gamma}{2}s$. By (n4.b”), there exists $\mathbf{y} \in \mathcal{C}$ such that $|\mathbf{y} - \mathbf{z}'| < \frac{\gamma}{10}s$. Moreover, $|\mathbf{y} - \mathbf{z}| \leq \gamma s$ and

$$|\mathbf{y}| \leq |\mathbf{y} - \mathbf{z}'| + |\mathbf{z}' - \mathbf{x}| + |\mathbf{x}| \leq \frac{\gamma}{10}s + \left(1 - \frac{\gamma}{2}\right)s + (\tfrac{1}{2} - s) < \frac{1}{2}.$$

Thus $\mathbf{y}' \in \widetilde{\mathcal{C}}$, and (n4.b') follows. □

4.1. **Bi-Lipschitz property and Ahlfors regularity.** In this subsection, we prove that the center set $\mathcal{C}$ of a neck region is bi-Lipschitz equivalent to the plane $V$ with respect to which the neck region is defined. As a consequence, the $k$-dimensional content of $\mathcal{C}$ is bounded above. Moreover, we prove Ahlfors regularity for the packing measure $\mu$ associated with the center set. The proof of the lower Ahlfors regularity uses the strategy of [JN21, Theorem 3.24].

We record the definition of parabolically bi-Lipschitz maps and some properties of Lipschitz graphs.

**Definition 4.4** (Projection)**.** Given $V \in \mathrm{Gr}_\mathcal{P}(k)$, we define the *orthogonal projection* $\pi_V$ onto $V$ as follows.

- If $V = L \times \{0\}$ for some $k$-plane $L \subseteq \mathbb{R}^n$, then
$$\pi_V(x, t) = (\pi_L(x), 0),$$
where $\pi_L : \mathbb{R}^n \to L$ is the standard orthogonal projection.
- If instead $V = L \times \mathbb{R}$ for some $(k-2)$-plane $L \subseteq \mathbb{R}^n$, then
$$\pi_V(x, t) = (\pi_L(x), t).$$

In both cases, we define the *orthogonal complement projection*

$$\pi_V^\perp := \mathrm{id}_{\mathbb{R}^n\times\mathbb{R}} - \pi_V,$$

so that for any $\mathbf{y} \in \mathbb{R}^n \times \mathbb{R}$, the parabolic distance to $V$ satisfies

$$d_\mathcal{P}(\mathbf{y}, V) = |\pi_V^\perp(\mathbf{y})|.$$

Given $V \in \mathrm{Gr}(k)$, for any $\mathbf{y} \in V$ and $s \geq 0$, we define the restricted ball as

$$P^V(\mathbf{y}, s) := P(\mathbf{y}, s) \cap V. \tag{4.3}$$

**Definition 4.5** ([Mat22, Definition 3.1])**.** We say that $G \subseteq \mathbb{R}^{n+1}$ is a $(k, \ell)$*-Lipschitz graph* if there exist an affine $k$-plane $V \in \mathrm{Aff}_\mathcal{P}(k)$, a subset $A \subseteq V$, and an $\ell$-Lipschitz function (with respect to $d_\mathcal{P}$) $\mathbf{f}\colon A \to V^\perp$ such that

$$G = \{\mathbf{v} + \mathbf{f}(\mathbf{v})\colon \mathbf{v} \in A\}.$$

We now prove the following structural result for $\mathcal{C}$.

**Theorem 4.6** (Neck Structure Theorem)**.** *Let* $\mathcal{N} = P(\mathbf{x}_0, 2r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$ *be a (strong)* $(m, k, \delta, \eta)$*-neck region modeled on* $V \in \mathrm{Gr}_\mathcal{P}(k)$*. Then the following statements hold if* $\delta \leq \overline{\delta}$*:*

*(i) For all* $\mathbf{x}, \mathbf{y} \in \mathcal{C}$*, we have*

$$(1 - C\delta)|\mathbf{x} - \mathbf{y}| \leq |\pi_V(\mathbf{x}) - \pi_V(\mathbf{y})| \leq |\mathbf{x} - \mathbf{y}|. \tag{4.4}$$

*(ii)* $\mathcal{C}$ *is a* $(k, C\delta)$*-graph over* $V$*, and satisfies the Ahlfors regularity upper bound*

$$\mu(P(\mathbf{x}, s)) \leq Cs^k$$

*for all* $\mathbf{x} \in \mathcal{C}$ *and* $s \in [\mathbf{r}_\mathbf{x}, r]$*.*

*Proof.* By rescaling and translation, we can assume $r = 1$ and $\mathbf{x}_0 = \mathbf{0}$.

(i) The upper bound in (4.4) holds trivially. To prove the lower bound, choose any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$. By (n1), $s \coloneqq \gamma^{-2}|\mathbf{x}_1 - \mathbf{x}_2| \geq \mathbf{r}_{\mathbf{x}_1} + \mathbf{r}_{\mathbf{x}_2}$. By (n4.a), $\mathbf{x}_2 \in \mathcal{C} \cap P(\mathbf{x}_1, s) \subseteq P(\mathbf{x}_1 + V, \delta s)$. It follows that

$$|\pi_V^\perp(\mathbf{x}_1 - \mathbf{x}_2)| = d(\mathbf{x}_2, \mathbf{x}_1 + V) < \delta s = \delta\gamma^{-2}|\mathbf{x}_1 - \mathbf{x}_2|,$$

which implies (4.4).

(ii) Set $A \coloneqq \pi_V(\mathcal{C})$. By (i), we can define $\mathbf{f} : A \to V^\perp$ so that for any $\mathbf{x} \in \mathcal{C}$, $\mathbf{v} + \mathbf{f}(\mathbf{v}) = \mathbf{x}$ if $\mathbf{v} = \pi_V(\mathbf{x})$. For any $\mathbf{x}_i \in \mathcal{C}$, $\mathbf{v}_i = \pi_V(\mathbf{x}_i)$, $i = 1, 2$,

$$|\mathbf{f}(\mathbf{v}_1) - \mathbf{f}(\mathbf{v}_2)| = |(\mathbf{x}_1 - \mathbf{v}_1) - (\mathbf{x}_2 - \mathbf{v}_2)| = |\pi_V^\perp(\mathbf{x}_1 - \mathbf{x}_2)| \leq C\delta|\mathbf{x}_1 - \mathbf{x}_2|.$$

Thus, $\mathcal{C} = \mathbf{f}(A)$ is a $(k, C\delta)$-graph.

Fix $\mathbf{x} \in \mathcal{C}$ and $s \in [\mathbf{r}_\mathbf{x}, 1]$. By (4.4) and recalling (4.3), we have

$$\mu_0(P(\mathbf{x}, s)) \leq \mathcal{H}^k_{\mathcal{P}}\left((\pi_V|_{\mathcal{C}})^{-1}(P^V(\pi_V(\mathbf{x}), 2s))\right) \leq C\mathcal{H}^k_{\mathcal{P}}(P^V(\pi_V(\mathbf{x}), 2s)) \leq Cs^k.$$

By (n1) and (i), $\{P^V(\pi_V(\mathbf{y}), \frac{1}{2}\gamma^2\mathbf{r}_\mathbf{y})\}_{\mathbf{y} \in \mathcal{C}_+ \cap P(\mathbf{x}, s)}$ is pairwise disjoint in $P^V(\mathbf{x}, 2s)$ so that

$$\mu_+(P(\mathbf{x}, s)) = \sum_{\mathbf{x} \in \mathcal{C}_+ \cap P(\mathbf{x}, s)} \mathbf{r}_\mathbf{x}^k \leq C\mathcal{H}^k_{\mathcal{P}}\left(P^V(\pi_V(\mathbf{x}), \frac{1}{2}\gamma^2\mathbf{r}_\mathbf{x})\right) \leq C\mathcal{H}^k_{\mathcal{P}}(P^V(\pi_V(\mathbf{x}), 2s)) \leq Cs^k.$$

Combining these estimates, we obtain $\mu(P(\mathbf{x}, s)) \leq Cs^k$. □

*Remark* 4.7. Note that for any $\mathbf{x} \in \mathcal{C}$ and $r \in [\mathbf{r}_\mathbf{x}, \gamma^{-1}]$ satisfying $P(\mathbf{x}, 2r) \subseteq P(\mathbf{0}, 2)$, we use (n4.a) and Theorem 4.6(ii) to get

$$\beta^2_{\mathcal{P},k}(\mathbf{x}, r) \leq \frac{\delta^2}{r^k}\mu(P(\mathbf{x}, r)) \leq C\delta^2. \tag{4.5}$$

We now turn to proving Ahlfors regularity estimates for packing measure of a strong neck region. As a first step, using the bi-Lipschitz property, (4.4), of the orthogonal projection

$$\pi \coloneqq \pi_V|_{\mathcal{C}} \colon \mathcal{C} \to V,$$

we will prove in Lemma 4.8 that the restriction of $\mathcal{C} \cap P(\mathbf{z}, \frac{9}{5}s)$ at $\mathbf{z} \in \mathcal{C}$ discretely looks like a ball in $V$ centered at $\pi(\mathbf{z})$.

**Lemma 4.8.** *Let $\mathcal{N} = P(\mathbf{x}_0, 2r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$ be a strong $(m, k, \delta, \eta)$-neck region modeled on $V \in \mathrm{Gr}_{\mathcal{P}}(k)$, where $\delta \leq \bar{\delta}$. For all $\mathbf{x} \in \mathcal{C}$ and $s \in [\gamma^{-1}\mathbf{r}_\mathbf{x}, r]$ such that $P(\mathbf{x}, \frac{3}{2}s) \subseteq P(\mathbf{x}_0, 2r)$,*

$$P^V(\pi(\mathbf{x}), \tfrac{7}{4}s) \subseteq \bigcup_{\mathbf{z} \in \mathcal{C} \cap P(\mathbf{x}, \frac{9}{5}s)} \overline{P}^V(\pi(\mathbf{z}), \mathbf{r}_\mathbf{z}).$$

*Proof.* By translation and rescaling, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $r = 1$. Fix $\mathbf{x} \in \mathcal{C}$ and $s \in (0, 1]$ such that $P(\mathbf{x}, 2s) \subseteq P(\mathbf{0}, 2)$.

First, we prove that (n4.b’) is preserved under the projection map.

**Claim 4.8.1.** *If $\delta \leq \bar{\delta}$, the following holds. For any $\mathbf{z} \in \mathcal{C} \cap P(\mathbf{x}, \frac{19}{10}s)$, $s' \in [\mathbf{r}_\mathbf{z}, \gamma^{-1}]$ such that $P(\mathbf{z}, s') \subseteq P(\mathbf{x}, \frac{19}{10}s)$, we have*

$$P^V(\pi(\mathbf{z}), s') \subseteq P^V(\pi(\mathcal{C} \cap P(\mathbf{z}, s')), 7\gamma s'). \tag{4.6}$$

*Proof.* Note that for any $\mathbf{y} \in \mathcal{C}$, if $P(\mathbf{y}, \gamma s') \cap P(\mathbf{z}, (1 - 2\gamma)s') \neq \emptyset$, then $|\mathbf{y} - \mathbf{z}| \leq (1 - \gamma)s'$. It follows that

$$P(\mathcal{C}, \gamma s') \cap P(\mathbf{z}, (1 - 2\gamma)s') \subset P(\mathcal{C} \cap P(\mathbf{z}, s'), \gamma s'). \tag{4.7}$$

By (n4.b') with $\mathbf{x} \leftarrow \mathbf{z}$ and $s \leftarrow (1-2\gamma)s'$ we get $(\mathbf{z}+V) \cap P(\mathbf{z},(1-2\gamma)s') \subseteq P(\mathcal{C},\gamma s') \cap P(\mathbf{z},(1-2\gamma)s')$, which combined with (4.7) yields

$$(\mathbf{z}+V) \cap P(\mathbf{z},(1-2\gamma)s') \subset P(\mathcal{C} \cap P(\mathbf{z},s'),\gamma s').$$

Applying $\pi$ to both sides, using (4.4) with $\delta \leq \overline{\delta}$ as well as

$$P^V(\pi(\mathbf{z}),s') \subseteq P^V\left(P^V(\pi(\mathbf{z}),(1-3\gamma)s'),3\gamma s'\right)$$

yield

$$\begin{aligned} P^V(\pi(\mathbf{z}),s') &\subseteq P^V\left(P^V(\pi(\mathbf{z}),(1-3\gamma)s'),3\gamma s'\right) \subseteq P^V\left(P^V(\pi(\mathcal{C} \cap P(\mathbf{z},s')),2\gamma s'),5\gamma s'\right) \\ &= P^V(\pi(\mathcal{C} \cap P(\mathbf{z},s')),7\gamma s'). \end{aligned}$$

□

Suppose by way of contradiction that there exists

$$\overline{\mathbf{y}} \in P^V(\pi(\mathbf{x}),\tfrac{7}{4}s) \setminus \bigcup_{\mathbf{z} \in \mathcal{C} \cap P(\mathbf{x},\frac{9}{5}s)} \overline{P}^V(\pi(\mathbf{z}),\mathbf{r}_{\mathbf{z}}).$$

Since $\mathbf{r}_{\mathbf{x}} \leq \gamma s < \frac{7}{4}s < \frac{19}{10}s$, we can apply Claim 4.8.1 with $\mathbf{z} \leftarrow \mathbf{x}$ and $s' \leftarrow \frac{7}{4}s$ to get

$$P^V(\pi(\mathbf{x}),\tfrac{7}{4}s) \subseteq P^V(\pi(\mathcal{C} \cap \overline{P}(\mathbf{x},\tfrac{9}{5}s)),20\gamma s), \tag{4.8}$$

Choose $\mathbf{y} \in \arg\min d_{\mathcal{P}}(\overline{\mathbf{y}},\pi(\mathcal{C} \cap \overline{P}(\mathbf{x},\frac{9}{5}s)))$. Since $\overline{\mathbf{y}} \in P^V(\pi(\mathbf{x}),\frac{7}{4}s)$ we can use (4.8) to obtain $|\mathbf{y}-\overline{\mathbf{y}}| < 20\gamma s$. Therefore, from (4.4), it follows that $P(\pi^{-1}(\mathbf{y}),2|\mathbf{y}-\overline{\mathbf{y}}|) \subseteq P(\mathbf{x},\frac{19}{10}s)$. By the definition of $\overline{\mathbf{y}}$, we know that $\mathbf{r}_{\pi^{-1}(\mathbf{y})} < |\mathbf{y}-\overline{\mathbf{y}}|$. Hence we can apply Claim 4.8.1 with $\mathbf{z} \leftarrow \pi^{-1}(\mathbf{y})$ and $s' \leftarrow 2|\mathbf{y}-\overline{\mathbf{y}}|$ to obtain

$$P^V(\mathbf{y},2|\mathbf{y}-\overline{\mathbf{y}}|) \subseteq P^V\big(\pi\big(\mathcal{C} \cap P\big(\pi^{-1}(\mathbf{y}),2|\mathbf{y}-\overline{\mathbf{y}}|\big)\big),14\gamma|\mathbf{y}-\overline{\mathbf{y}}|\big).$$

In particular, there exists $\mathbf{y}' \in \pi\big(\mathcal{C} \cap P\big(\pi^{-1}(\mathbf{y}),2|\mathbf{y}-\overline{\mathbf{y}}|\big)\big)$ such that

$$|\mathbf{y}'-\overline{\mathbf{y}}| < 14\gamma|\mathbf{y}-\overline{\mathbf{y}}| < |\mathbf{y}-\overline{\mathbf{y}}|. \tag{4.9}$$

On the other hand, by (4.4) if $\delta \leq \overline{\delta}$, we have

$$|\mathbf{x}-\pi^{-1}(\mathbf{y}')| \leq (1+\gamma)(|\pi(\mathbf{x})-\overline{\mathbf{y}}|+|\overline{\mathbf{y}}-\mathbf{y}'|) \leq (1+\gamma)(\tfrac{7}{4}+280\gamma^2)s \leq \tfrac{9}{5}s.$$

Hence $\pi^{-1}(\mathbf{y}') \in \mathcal{C} \cap \overline{P}(\mathbf{x},\frac{9}{5}s)$, which together with (4.9) contradicts the choice of $\mathbf{y}$. □

Using Lemma 4.8, we prove a lower bound $\mu(P(\mathbf{z},s)) \geq C^{-1}s^k$, proving Ahlfors regularity.

**Proposition 4.9.** *Let $\mathcal{N} = P(\mathbf{x}_0,2r) \setminus \overline{P}(\mathcal{C},\mathbf{r}_\bullet)$ be a strong $(m,k,\delta,\eta)$-neck region modeled on $V \in \mathrm{Gr}_{\mathcal{P}}(k)$, where $\delta \leq \overline{\delta}$. Then the packing measure $\mu$ of $\mathcal{C}$ defined in (4.2) satisfies the Ahlfors regularity condition*

$$C^{-1}s^k \leq \mu(P(\mathbf{x},s)), \tag{4.10}$$

*for all $\mathbf{x} \in \mathcal{C}$ and $s \in (0,r]$ such that $P(\mathbf{x},\frac{3}{2}s) \subseteq P(\mathbf{x}_0,2r)$, where $C$ is universal.*

*Proof.* By translation and rescaling, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $r=1$. Fix $\mathbf{x} \in \mathcal{C}$ and $s \in (0,1]$ such that $P(\mathbf{x},2s) \subseteq P(\mathbf{0},2)$. If $s < \gamma^{-1}\mathbf{r}_{\mathbf{x}}$, then $\mathbf{x} \in \mathcal{C}_+$. Thus, (4.10) follows because

$$\mu(P(\mathbf{x},s)) \geq \mathbf{r}_{\mathbf{x}}^k \geq C^{-1}s^k.$$

Suppose instead $s \in [\gamma^{-1}\mathbf{r}_{\mathbf{x}},1]$. By Lemma 4.8, we have

$$C^{-1}s^k \leq \sum_{\mathbf{y} \in \mathcal{C} \cap P(\mathbf{x},\frac{9}{5}s)} \mathbf{r}_{\mathbf{y}}^k \leq \mu(P(\mathbf{x},\tfrac{9}{5}s)).$$

Thus (4.10) follows by replacing $s$ in the above computation with $\frac{5}{9}s$. □

4.2. **Best subspaces on a temporal strong neck region.** Suppose $\mathcal{N} = P(\mathbf{0},2)\setminus\bigcup_{\mathbf{x}\in\mathcal{C}}\overline{P}(\mathbf{x},\mathbf{r_x})$ is a temporal strong $(m,k,\delta,\eta)$-neck region modeled on $V\in\mathrm{Gr}_{\mathcal{P}}(k)$. In general, $V$ is not the best plane that approximates $\mathcal{C}$, and we often rely on the best plane. In this subsection, we provide tilting estimates for the best planes across different locations and scales. The proofs in this section are modifications of [Nab18, Section 11.1] from the elliptic setting.

For each $\mathbf{x}\in\mathcal{C}$ with $P(\mathbf{x},r)\subseteq P(\mathbf{0},2)$, we define a best vertical plane at $\mathbf{x}$ at scale $r$ by

$$V_{\mathbf{x},r}\in\arg\min_{W\in\mathrm{Aff}_{\mathcal{P},\mathrm{v}}(k)}\frac{1}{r^k}\int_{P(\mathbf{x},r)}\left(\frac{d_{\mathcal{P}}(\mathbf{y},W)}{r}\right)^2 d\mu(\mathbf{y}), \tag{4.11}$$

where $\mathrm{Aff}_{\mathcal{P},\mathrm{v}}(k)$ is the set of all vertical planes in $\mathrm{Aff}_{\mathcal{P}}(k)$. Let $\widetilde{V}_{\mathbf{x},r}\in\mathrm{Gr}_{\mathcal{P}}(k)$ be the corresponding linear subspace, and write $\pi_{\mathbf{x},r}\coloneqq\pi_{\widetilde{V}_{\mathbf{x},r}}$, $\pi^{\perp}_{\mathbf{x},r}\coloneqq\pi^{\perp}_{\widetilde{V}_{\mathbf{x},r}}$.

In the lemma below, we estimate the distance between $V_{\mathbf{x},r}$ and the points in $r$-neighborhood of $\mathbf{x}$.

**Lemma 4.10.** *For any $\mathbf{x}\in\mathcal{C}$ and $r>0$ such that $P(\mathbf{x},r)\subseteq P(\mathbf{0},2)$, if $\delta\le\overline{\delta}$, we have*

$$\sup_{\mathbf{w}\in P(\mathbf{x},\frac{9}{10}r)} d_{\mathcal{P}}(\mathbf{w},V_{\mathbf{x},r})\le C\delta^{\frac{2}{k+2}}r.$$

*Proof.* Fix $\mathbf{w}\in P(\mathbf{x},\frac{9}{10}r)$ and set $\rho\coloneqq d_{\mathcal{P}}(\mathbf{w},V_{\mathbf{x},r})$, $s=\min\{r,\rho\}$.

$$C\delta^2\ge\frac{1}{r^{k+2}}\int_{P(\mathbf{w},\frac{1}{10}s)} d^2_{\mathcal{P}}(\mathbf{y},V_{\mathbf{x},r})d\mu(\mathbf{y})\ge\frac{s^k}{Cr^{k+2}}\rho^2, \tag{4.12}$$

where we used (4.5) in the first inequality and Ahlfors regularity (4.10) of $\mu$ in the second inequality. Now suppose $\min\{r,\rho\}=r$. In this case,

$$C\delta^2\ge\frac{\rho^2}{Cr^2}\ge\frac{1}{C}.$$

If $\delta\le\overline{\delta}$, then this case cannot occur. We may thus assume $\min\{r,\rho\}=\rho$, so that (4.12) implies $\rho\le C\delta^{\frac{2}{k+2}}r$. Therefore, the claim follows. □

For $\mathbf{x}\in\mathcal{C}$ and $r>0$, we define the *parabolic mean:*

$$\mathbf{z}_{\mathbf{x},r}\coloneqq\frac{1}{\mu(P(\mathbf{x},r))}\int_{P(\mathbf{x},r)}\mathbf{z}\,d\mu(\mathbf{z}).$$

In the lemma below, we estimate the distance between $V_{\mathbf{x},r}$ and a parabolic mean at different locations and scales.

**Lemma 4.11.** *If $\mathbf{y}\in\mathcal{C}$ and $P(\mathbf{y},s)\subseteq P(\mathbf{x},r)$ and $P(\mathbf{x},2r)\subseteq P(\mathbf{0},2)$, then*

$$d^2_{\mathcal{P}}(\mathbf{z}_{\mathbf{y},s},V_{\mathbf{x},r})\le\frac{r^{k+2}}{\mu(P(\mathbf{y},s))}\beta^2_{\mathcal{P},k}(\mathbf{x},r).$$

*Proof.* Recall that for any choice of $\mathbf{x}_0\in V_{\mathbf{x},r}$, we have $d_{\mathcal{P}}(\mathbf{z},V_{\mathbf{x},r})=|\pi^{\perp}_{\mathbf{x},r}(\mathbf{z}-\mathbf{x}_0)|$ for all $\mathbf{z}\in\mathbb{R}^{n+1}$. Thus Jensen's inequality gives

$$\begin{aligned}d^2_{\mathcal{P}}(\mathbf{z}_{\mathbf{y},s},V_{\mathbf{x},r})&=\left|\frac{1}{\mu(P(\mathbf{y},s))}\int_{P(\mathbf{y},s)}\pi^{\perp}_{\mathbf{x},r}(\mathbf{z}-\mathbf{x}_0)\,d\mu(\mathbf{z})\right|^2\le\frac{1}{\mu(P(\mathbf{y},s))}\int_{P(\mathbf{y},s)}|\pi^{\perp}_{\mathbf{x},r}(\mathbf{z}-\mathbf{x}_0)|^2\,d\mu(\mathbf{z})\\&\le\frac{1}{\mu(P(\mathbf{y},s))}\int_{P(\mathbf{x},r)}d^2_{\mathcal{P}}(\mathbf{z},V_{\mathbf{x},r})\,d\mu(\mathbf{z})\le\frac{r^{k+2}}{\mu(P(\mathbf{y},s))}\beta^2_{\mathcal{P},k}(\mathbf{x},r).\end{aligned}$$

□

Below, we prove that the subspaces $V_{\mathbf{x},r}$ are comparable for nearby points and scales.

**Proposition 4.12.** *Suppose $\mathcal{N} = P(\mathbf{0},2) \setminus \bigcup_{\mathbf{x}\in\mathcal{C}} P(\mathbf{x},\mathbf{r_x})$ is a temporal strong $(m,k,\delta,\eta)$-neck region modeled on $V \in \mathrm{Gr}_{\mathcal{P}}(k)$. There exists $C$ such that the following holds if $\delta \leq \overline{\delta}(\eta)$.*

*(1) For any $\mathbf{x},\mathbf{y} \in \mathcal{C}$ and $s>0$ such that $P(\mathbf{x},10^{-2}r) \subseteq P(\mathbf{y},s) \subseteq P(\mathbf{x},10^2 r)$, we have*

$$d_{\mathcal{P},H}(V_{\mathbf{x},r} \cap P(\mathbf{x},10^3 r), V_{\mathbf{y},s} \cap P(\mathbf{x},10^3 r)) \leq C\beta_{\mathcal{P},k}(\mathbf{x},10^4 r). \tag{4.13}$$

*(2) Moreover, for any $\mathbf{x} \in \mathcal{C}$ and $r \in [\gamma^{-2}\mathbf{r_x},\gamma^{-1}]$ such that $P(\mathbf{x},10^3 r) \subseteq P(\mathbf{0},2)$, we have*

$$|\pi_V^{\perp} - \pi_{\mathbf{x},r}^{\perp}| \leq C\delta^{\frac{2}{k+2}}. \tag{4.14}$$

*Proof.* (1) Fix $\mathbf{x},\mathbf{y} \in \mathcal{C}$ and $r,s>0$ such that

$$P(\mathbf{x},10^{-2}r) \subseteq P(\mathbf{y},s) \subseteq P(\mathbf{x},10^2 r) \subseteq P(\mathbf{0},2), \qquad P(\mathbf{x},10^3 r) \subseteq P(\mathbf{0},2).$$

By (n4.b'), we have

$$(\mathbf{y}+V) \cap P(\mathbf{y},s) \subseteq P(\mathcal{C},\gamma s).$$

Since $V$ is vertical, we can find a spatially $(k-2,\frac{s}{2})$-independent subset $\{\mathbf{y}_i\}_{i=0}^{k-2} \subseteq \mathcal{C} \cap P(\mathbf{y},s)$ with $\mathbf{y}_0 = \mathbf{y}$. Because $P(\mathbf{y}_i,\frac{3}{2}\gamma s) \subseteq P(\mathbf{y},2s) \subseteq P(\mathbf{0},2)$, Proposition 4.9 gives

$$\mu(P(\mathbf{y}_i,\gamma s)) \geq C^{-1}s^k.$$

Fix a collection $\{\mathbf{y}_i\}_{i=0}^{k-2} \subseteq \mathcal{C} \cap P(\mathbf{y},s)$. Set $\mathbf{z}_0 := \mathbf{z}_{\mathbf{y},\gamma s}$ and $\mathbf{z}_i := \mathbf{z}_{\mathbf{y}_i,\gamma s}$. Since $P(\mathbf{y}_i,\gamma s) \subseteq P(\mathbf{y},s) \subseteq P(\mathbf{x},10^2 r)$, Lemma 4.11 with $\mathbf{y} \leftarrow \mathbf{y}_i$, $s \leftarrow \gamma s$, $\mathbf{x} \leftarrow \mathbf{x}$, and $r \leftarrow 10^2 r$ implies

$$d_{\mathcal{P}}^2(\mathbf{z}_i, V_{\mathbf{x},10^2 r}) \leq \frac{Cr^{k+2}}{\mu(P(\mathbf{y}_i,\gamma s))}\beta_{\mathcal{P},k}^2(\mathbf{x},10^2 r) \leq C\beta_{\mathcal{P},k}^2(\mathbf{x},10^2 r)r^2, \tag{4.15}$$

where we used (4.10) in the second inequality. Similarly, we apply Lemma 4.11 with $\mathbf{y} \leftarrow \mathbf{y}_i$, $s \leftarrow \gamma s$, $\mathbf{x} \leftarrow \mathbf{y}$, and $r \leftarrow s$ to obtain

$$d_{\mathcal{P}}^2(\mathbf{z}_i, V_{\mathbf{y},s}) \leq \frac{s^{k+2}}{\mu(P(\mathbf{y}_i,\gamma s))}\beta_{\mathcal{P},k}^2(\mathbf{y},s) \leq C\beta_{\mathcal{P},k}^2(\mathbf{y},s)r^2. \tag{4.16}$$

From $P(\mathbf{y},s) \subseteq P(\mathbf{x},10^2 r)$, we also have

$$\beta_{\mathcal{P},k}^2(\mathbf{y},s) \leq \frac{1}{s^{k+2}}\int_{P(\mathbf{x},10^2 r)} d_{\mathcal{P}}^2(\mathbf{z},V_{\mathbf{x},10^2 r})d\mu(\mathbf{z}) \leq C\beta_{\mathcal{P},k}^2(\mathbf{x},10^2 r). \tag{4.17}$$

Because $\mathbf{z}_i \in P(\mathbf{y}_i,\gamma s)$, we know $\{z_i\}_{i=0}^{k-2} \subseteq \mathbb{R}^n$ is $(k-2,\frac{1}{4}s)$-independent. By (4.5), (4.15), (4.16), and (4.17), we have

$$d_{\mathcal{P}}(\mathbf{z}_i, V_{\mathbf{x},10^2 r}) + d_{\mathcal{P}}(\mathbf{z}_i, V_{\mathbf{y},s}) \leq C\beta_{\mathcal{P},k}(\mathbf{x},10^2 r)r \leq C\delta r. \tag{4.18}$$

On the other hand, we apply Lemma 4.10 with $\mathbf{w} \leftarrow \mathbf{x}$ and $r \leftarrow 10^2 r$ to obtain

$$d_{\mathcal{P}}(\mathbf{x}, V_{\mathbf{x},10^2 r}) \leq C\delta^{\frac{2}{k+2}}r. \tag{4.19}$$

Following the proof of [Nab18, Lemma 11.8], we can use (4.18) and (4.19) to obtain

$$d_{\mathcal{P},H}(V_{\mathbf{x},10^2 r} \cap P(\mathbf{x},10^3 r), V_{\mathbf{y},s} \cap P(\mathbf{x},10^3 r)) \leq C\beta_{\mathcal{P},k}(\mathbf{x},10^2 r)r. \tag{4.20}$$

We can argue as above and take $\mathbf{y} \leftarrow \mathbf{x}$ and $s \leftarrow r$ in (4.20) to get

$$d_{\mathcal{P},H}(V_{\mathbf{x},10^2 r} \cap P(\mathbf{x},10^3 r), V_{\mathbf{x},r} \cap P(\mathbf{x},10^3 r)) \leq C\beta_{\mathcal{P},k}(\mathbf{x},10^2 r)r. \tag{4.21}$$

Combining (4.20) and (4.21) yields (4.13).

(2) First, we apply Lemma 4.10 with $\mathbf{w} \leftarrow \mathbf{y}_i$ and $r \leftarrow s$ to get

$$d_{\mathcal{P}}(\mathbf{y}_i, V_{\mathbf{y},s}) \leq C\delta^{\frac{2}{k+2}}s. \tag{4.22}$$

Further, (n4.a) implies $\mathbf{y}_i \in \operatorname{supp}(\mu) \cap P(\mathbf{y}, s) \subseteq P(\mathbf{y} + V, \delta s)$, hence

$$d_{\mathcal{P}}(\mathbf{y}_i, \mathbf{y} + V) \leq \delta s. \tag{4.23}$$

Because $\{y_i\}_{i=0}^{k-2} \subseteq \mathbb{R}^n$ is $(k-2, \frac{s}{4})$-independent, we can apply (4.22), (4.23), and again appeal to the proof of [Nab18, Lemma 11.8] to obtain

$$d_{\mathcal{P},H}((\mathbf{y} + V) \cap P(\mathbf{y}, 10^3 s), V_{\mathbf{y},s} \cap P(\mathbf{y}, 10^3 s)) \leq C\delta^{\frac{2}{k+2}} s.$$

This implies that the underlying linear subspace $\widetilde{V}_{\mathbf{y},s}$ of $V_{\mathbf{y},s}$ satisfies

$$d_{\mathcal{P},H}(V \cap P(\mathbf{0}, 1), \widetilde{V}_{\mathbf{y},s} \cap P(\mathbf{0}, 1)) \leq C\delta^{\frac{2}{k+2}}.$$

We may therefore follow the proof of [NV17a, Lemma 4.3 and Lemma 4.4] to obtain (4.14). $\square$

4.3. **Strong neck structure theorem.** In Theorem 4.14, we upgrade the $C\delta$-graphicality of $\mathcal{C}_0$ from Theorem 4.6 under the assumption that the neck region is strong and $k \in \{n, n+1\}$. More precisely, we prove that $\mathcal{C}$ is contained in a $C\delta\mathbf{r}$ neighborhood of a regular parabolic Lipschitz graph defined below. Some of the methods of this section are inspired by [BHH$^+$23, DS91].

For a compactly supported function $\phi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n+2-k}$, we define

$$\partial_t^{\frac{1}{2}} \phi(\mathbf{x}) \coloneqq \check{c} \int_{\mathbb{R}} \frac{\phi(x, s) - \phi(x, t)}{|s - t|^{\frac{3}{2}}} \, ds \tag{4.24}$$

for $\mathbf{x} = (x, t) \in \mathbb{R}^n \times \mathbb{R}$, where $\check{c} \in \mathbb{R}$ is a normalizing constant, chosen so that for any Schwartz function $\psi \in C^\infty(\mathbb{R})$, if $\mathcal{F}\psi$ denotes the Fourier transform of $\psi$, then

$$(\mathcal{F}\partial_t^{\frac{1}{2}} \psi)(\tau) = |\tau|^{\frac{1}{2}} (\mathcal{F}\psi)(\tau)$$

for all $\tau \in \mathbb{R}$. We also define the *parabolic* BMO *norm* on $V \in \operatorname{Gr}(k)$ as

$$\|\phi\|_{\mathrm{BMO}_{\mathcal{P}}(V)} \coloneqq \sup_{\mathbf{x} \in V} \sup_{r > 0} ⨍_{P^V(\mathbf{x}, r)} \left| \phi(\mathbf{y}) - ⨍_{P^V(\mathbf{x}, r)} \phi \, d\mathcal{H}^k_{\mathcal{P}} \right| d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}). \tag{4.25}$$

**Definition 4.13.** Given a vertical $V \in \operatorname{Gr}_{\mathcal{P}}(k)$, we say that a compactly supported function $\phi : V \to V^\perp$ is *$\delta$-regular parabolic Lipschitz* if it is $\delta$-Lipschitz with respect to $d_{\mathcal{P}}$ and $\|\partial_t^{\frac{1}{2}} \phi\|_{\mathrm{BMO}_{\mathcal{P}}(V)} < \delta$.

For any function $\phi \colon V \to V^\perp$, we let $\mathcal{G}(\phi) \coloneqq \{\mathbf{y} + \phi(\mathbf{y}) \colon \mathbf{y} \in V\}$ denote the *graph of $\phi$*.

We now state our main result on the regularity of $\mathcal{C}$.

**Theorem 4.14** (Strong Neck Structure Theorem)**.** *Let $k \in \{n, n+1\}$ and let $\mathcal{N} = P(\mathbf{x}_0, 2r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$ be an $(m, k, \delta, \eta)$-neck region modeled on $V \in \operatorname{Gr}_{\mathcal{P}}(k)$. Then there exists a compactly supported $C\delta^{\frac{1}{k+2}}$-regular parabolic Lipschitz function $G : V \to V^\perp$ such that for all $\mathbf{x} \in \mathcal{C} \cap P(\mathbf{x}_0, \frac{4}{3} r)$*

$$d_{\mathcal{P}}(\mathbf{x}, \mathcal{G}(G)) < C(\eta, \Lambda)\delta \mathbf{r}_{\mathbf{x}}, \tag{4.26}$$

*and in particular, $\mathcal{C}_0 \cap P(\mathbf{x}_0, \frac{4}{3} r) \subseteq \mathcal{G}(G)$.*

In our setting, since $\mathcal{C}$ is a bi-Lipschitz graph over $V$ by Theorem 4.6, we can substitute the submanifold approximation theorem used in neck structure theorems, for instance [Nab18], with a regular parabolic Lipschitz extension problem. The main idea of the proof is to construct a Lipschitz almost extension $F$ of $\pi^{-1}$, which is Lipschitz by Theorem 4.6. At large scales $F$ will look like $\pi^{-1}$, while at scales smaller than $\mathbf{r}_\bullet$ near $\mathcal{C}_+$, $F$ is well approximated by its tangent plane.

Let $\mu$ be the packing measure of $\mathcal{N}$ as defined in (4.2). Set

$$\widehat{\mathcal{C}} \coloneqq \pi(\mathcal{C}), \qquad \widehat{\mathcal{C}}_0 \coloneqq \pi(\mathcal{C}_0), \qquad \widehat{\mathcal{C}}_+ \coloneqq \pi(\mathcal{C}_+), \qquad \widehat{\mu} \coloneqq \pi_* \mu.$$

Observe that $\widehat{\mathcal{C}}$ and $\widehat{\mathcal{C}}_0$ are closed subsets of $V$ by (4.4). Further, (4.4) and the proof of Theorem 4.6 imply that the map

$$F^*\colon \widehat{\mathcal{C}} \to \mathbb{R}, \quad \mathbf{x} \mapsto (\pi|_{\mathcal{C}}^{-1}(\mathbf{x}) - \mathbf{x}) \cdot e_1$$

is well-defined and $C\delta$-Lipschitz. By definition, $F^*(\mathbf{x})e_1 + \mathbf{x} = \pi^{-1}(\mathbf{x}) \in \mathcal{C}$ for all $\mathbf{x} \in \widehat{\mathcal{C}}$. Define

$$\widehat{\mathbf{r}}_{\mathbf{x}} := \mathbf{r}_{\pi^{-1}(\mathbf{x})}, \qquad \widehat{\beta}_{\mathcal{P},k}(\mathbf{x}, s) := \beta_{\mathcal{P},k}(\pi^{-1}(\mathbf{x}), s), \qquad \widehat{V}_{\mathbf{x},r} := V_{\pi^{-1}(\mathbf{x}),r} \qquad \text{for } \mathbf{x} \in \widehat{\mathcal{C}}. \tag{4.27}$$

By Theorem 4.6 and Proposition 4.9, $\widehat{\mu}$ satisfies the Ahlfors regularity condition

$$C^{-1}s^k \le \widehat{\mu}(P^V(\mathbf{x}, s)) \le Cs^k \tag{4.28}$$

for all $\mathbf{x} \in \widehat{\mathcal{C}}$ and all $s \in [\widehat{\mathbf{r}}_{\mathbf{x}}, \gamma^{-1}]$ such that $P^V(\mathbf{x}, 3s) \subseteq P^V(\mathbf{0}, 2)$.

In the following lemma, we prove that $F^*$ can be well approximated at scales larger than $\mathbf{r}$ by Lipschitz affine functions with vanishing time derivative and small Lipschitz constant. Furthermore, the approximating functions vary slowly across scales.

**Lemma 4.15.** *There exist affine functions $\ell_{\mathbf{x},r} : V \to \mathbb{R}^{n+2-k}$ for any $\mathbf{x} \in \widehat{\mathcal{C}}$ and $r \in [\widehat{\mathbf{r}}_{\mathbf{x}}, \gamma^{-1}]$ with $P^V(\mathbf{x}, 10^3 r) \subseteq P^V(\mathbf{0}, 2)$, such that the following hold:*

*(1) For any $\mathbf{x}_0 \in \widehat{\mathcal{C}}$ and $r_0 \in [\widehat{\mathbf{r}}_{\mathbf{x}_0}, 1]$, $\partial_t \ell_{\mathbf{x}_0,r_0} \equiv 0$.*

*(2) We have*

$$\frac{1}{r^k}\int_{P(\mathbf{x},r)} \left(\frac{|F^*(\mathbf{y}) - \ell_{\mathbf{x},r}(\mathbf{y})|}{r}\right)^2 d\widehat{\mu}(\mathbf{y}) \le C\widehat{\beta}^2_{\mathcal{P},k}(\mathbf{x}, 2r).$$

*(3) For any $\mathbf{y} \in P^V(\mathbf{0}, 2)$ and $s \ge \widehat{\mathbf{r}}_{\mathbf{y}}$ satisfying $P^V(\mathbf{x}, 10^{-2}r) \subseteq P^V(\mathbf{y}, s) \subseteq P^V(\mathbf{x}, 10^2 r)$,*

$$\sup_{P^V(\mathbf{x},10^3 r)} |\ell_{\mathbf{x},r} - \ell_{\mathbf{y},s}| \le C\widehat{\beta}_{\mathcal{P},k}(\mathbf{x}, 10^4 r)r.$$

*(4) For each $\mathbf{x} \in \widehat{\mathcal{C}}$ and $r \in [\widehat{\mathbf{r}}_{\mathbf{x}}, \gamma^{-1}]$, the function $\ell_{\mathbf{x},r}$ is $C\delta^{\frac{2}{k+2}}$-Lipschitz.*

*Proof.* By rotation and translation, we may assume that $V = \{0\} \times \mathbb{R}^{k-2} \times \mathbb{R}$. If $k = n+1$, we define $\ell_{\mathbf{x},r} : V \to \mathbb{R}$, by

$$\ell_{\mathbf{x},r}(\mathbf{y}) := \Big((\pi_V|_{\widehat{V}_{\mathbf{x},r}})^{-1}(\mathbf{y}) - \mathbf{y}\Big) \cdot e_1,$$

so that $\mathbf{y} + \ell_{\mathbf{x},r}(\mathbf{y})e_1 \in \widehat{V}_{\mathbf{x},r}$ for $\mathbf{y} \in V$. Otherwise, we define $\ell_{\mathbf{x},r} : V \to \mathbb{R}^2$ by

$$\ell_{\mathbf{x},r}(\mathbf{y}) := \Big(\Big((\pi_V|_{\widehat{V}_{\mathbf{x},r}})^{-1}(\mathbf{y}) - \mathbf{y}\Big) \cdot e_1, \Big((\pi_V|_{\widehat{V}_{\mathbf{x},r}})^{-1}(\mathbf{y}) - \mathbf{y}\Big) \cdot e_2\Big).$$

For ease of exposition, we restrict our attention to $k = n+1$ since the other case is similar.

(1) This holds because $\widehat{V}_{\mathbf{x},r}$ are vertical.

(2) Notice that $\pi^{-1}(\mathbf{x}) = \mathbf{x} + F^*(\mathbf{x})e_1 \in \mathcal{C}$. For all $\mathbf{z} \in \mathcal{C}$, we have $\pi_V^{\perp}(\mathbf{z} - \pi(\mathbf{z})) = (F^* \circ \pi)(\mathbf{z})$ and $\pi(\mathbf{z}) + (\ell_{\mathbf{x},r} \circ \pi)(\mathbf{z}) \in \widehat{V}_{\mathbf{x},r}$. From this and $\pi_V^{\perp}(e_1) = e_1$, we have

$$\begin{aligned} d_{\mathcal{P}}(\mathbf{z}, \widehat{V}_{\mathbf{x},r}) &= \left|\pi^{\perp}_{\pi^{-1}(\mathbf{x}),r}\left(\mathbf{z} - (\pi(\mathbf{z}) + (\ell_{\mathbf{x},r} \circ \pi)(\mathbf{z}))\right)\right| \\ &\ge |\pi_V^{\perp}(\mathbf{z} - \pi(\mathbf{z}) - (\ell_{\mathbf{x},r} \circ \pi)(\mathbf{z})e_1)| - |(\pi^{\perp}_{\pi^{-1}(\mathbf{x}),r} - \pi_V^{\perp})(\mathbf{z} - \pi(\mathbf{z}) - (\ell_{\mathbf{x},r} \circ \pi)(\mathbf{z}))| \\ &\ge (1 - C\delta^{\frac{2}{n+3}})|(F^* \circ \pi)(\mathbf{z}) - (\ell_{\mathbf{x},r} \circ \pi_V)(\mathbf{z})| \end{aligned} \tag{4.29}$$

for $\mathbf{z} \in \mathcal{C}$, where we used (4.14) to obtain the last inequality.

Since $\pi(P(\pi^{-1}(\mathbf{x}), 2r)) \supseteq P^V(\mathbf{x}, r)$, we can use (4.29) to estimate

$$\frac{1}{r^{n+1}}\int_{P^V(\mathbf{x},r)} \left(\frac{|F^*(\mathbf{y}) - \ell_{\mathbf{x},r}(\mathbf{y})|}{r}\right)^2 d\widehat{\mu}(\mathbf{y})$$

$$
\begin{aligned}
&\leq \frac{C}{r^{n+1}} \int_{\pi(P(\pi^{-1}(\mathbf{x}),2r))} \left( \frac{|F^*(\mathbf{y}) - \ell_{\mathbf{x},r}(\mathbf{y})|}{r} \right)^2 d\widehat{\mu}(\mathbf{y}) \\
&= \frac{C}{r^{n+1}} \int_{P(\pi^{-1}(\mathbf{x}),2r)} \left( \frac{|(F^* \circ \pi)(\mathbf{z}) - (\ell_{\mathbf{x},r} \circ \pi)(\mathbf{z})|}{r} \right)^2 d\mu(\mathbf{z}) \\
&\leq \frac{C}{r^{n+1}} \int_{P(\pi^{-1}(\mathbf{x}),2r)} \left( \frac{d_{\mathcal{P}}(\mathbf{z}, \widehat{V}_{\mathbf{x},r})}{r} \right)^2 d\mu(\mathbf{z}) \\
&= C\widehat{\beta}^2_{\mathcal{P}}(\mathbf{x}, 2r).
\end{aligned}
$$

Thus (2) holds.

(3) Now assume $\mathbf{y} \in P^V(\mathbf{0}, 2)$ and $s \geq \widehat{\mathbf{r}}_{\mathbf{y}}$ satisfy $P^V(\mathbf{x}, 10^{-2}r) \subseteq P^V(\mathbf{y}, s) \subseteq P^V(\mathbf{x}, 10^2 r)$. For any $\mathbf{w} \in P^V(\mathbf{0}, 2)$, we have $\mathbf{w} + \ell_{\mathbf{x},r}(\mathbf{w})e_1 \in \widehat{V}_{\mathbf{x},r}$, so that from (4.14), we have

$$
\begin{aligned}
|(\ell_{\mathbf{x},r} - \ell_{\mathbf{y},s})(\mathbf{w})| &= \left| \pi_V^\perp \left( (\ell_{\mathbf{x},r}(\mathbf{w}) - \ell_{\mathbf{y},s}(\mathbf{w}))e_1 \right) \right| \\
&\leq |\pi_V^\perp - \pi^\perp_{\pi^{-1}(\mathbf{x}),r}| \cdot |(\ell_{\mathbf{x},r} - \ell_{\mathbf{y},s})(\mathbf{w})| \\
&\quad + \left| \pi^\perp_{\pi^{-1}(\mathbf{x}),r} \left( (\mathbf{w} + \ell_{\mathbf{x},r}(\mathbf{w})e_1) - (\mathbf{w} + \ell_{\mathbf{y},s}(\mathbf{w})e_1) \right) \right| \\
&\leq \frac{1}{2} |(\ell_{\mathbf{x},r} - \ell_{\mathbf{y},s})(\mathbf{w})| + d_{\mathcal{P}}(\mathbf{w} + \ell_{\mathbf{y},s}(\mathbf{w})e_1, \widehat{V}_{\mathbf{x},r})
\end{aligned}
$$

if $\delta \leq \overline{\delta}$. If in addition, $\mathbf{w} \in P(\mathbf{x}, 10^3 r)$, then (4.13) gives

$$
\begin{aligned}
d_{\mathcal{P}}(\mathbf{w} + \ell_{\mathbf{y},s}(\mathbf{w})e_1, \widehat{V}_{\mathbf{x},r}) &\leq d_{\mathcal{P},H}(\widehat{V}_{\mathbf{x},r} \cap P(\mathbf{x}, 10^3 r), \widehat{V}_{\mathbf{y},r} \cap P(\mathbf{x}, 10^3 r)) \\
&\leq C\widehat{\beta}_{\mathcal{P},k}(\mathbf{x}, 10^4 r),
\end{aligned}
$$

and the claim follows by combining expressions.

(4) For any $\mathbf{x}_1, \mathbf{x}_2 \in V$, we have $\mathbf{x}_i + \ell_{\mathbf{x},r}(\mathbf{x}_i)e_1 \in \widehat{V}_{\mathbf{x},r}$, so that

$$
\mathbf{w} \coloneqq (\mathbf{x}_1 + \ell_{\mathbf{x},r}(\mathbf{x}_1)e_1) - (\mathbf{x}_2 + \ell_{\mathbf{x},r}(\mathbf{x}_2)e_1) \in \widehat{V}_{\mathbf{x},r}.
$$

Thus $\pi^\perp_{\pi^{-1}(\mathbf{x}),r}(\mathbf{w}) = 0$, hence (4.14) yields

$$
\begin{aligned}
|\ell_{\mathbf{x},r}(\mathbf{x}_1) - \ell_{\mathbf{x},r}(\mathbf{x}_2)| =& |\pi_V^\perp(\mathbf{w})| = |(\pi_V^\perp - \pi^\perp_{\pi^{-1}(\mathbf{x}),r})(\mathbf{w})| \leq C\delta^{\frac{2}{n+3}} |\mathbf{w}| \\
\leq& C\delta^{\frac{2}{n+3}} (|\mathbf{x}_1 - \mathbf{x}_2| + |\ell_{\mathbf{x},r}(\mathbf{x}_1) - \ell_{\mathbf{x},r}(\mathbf{x}_2)|),
\end{aligned}
$$

and the claim follows assuming $\delta \leq \overline{\delta}$. □

Now we will extend $F^*$ to $P^V(\mathbf{0}, \frac{7}{4})$. We will smoothly patch together the linear approximations of $F^*$ around each point in $\widehat{\mathcal{C}}_+$ at scale comparable to $\mathbf{r}_\bullet$. To do so, we have the following Whitney-type decomposition of $P^V(\mathbf{0}, \frac{7}{4})$ as an application of Lemma 4.8 with $\mathbf{x} \leftarrow 0$ and $s \leftarrow 1$:

$$
P^V(\mathbf{0}, \tfrac{7}{4}) \setminus \widehat{\mathcal{C}}_0 \subseteq \bigcup_{\mathbf{y} \in \widehat{\mathcal{C}}_+ \cap P(\mathbf{0}, \frac{9}{5})} P(\mathbf{y}, 2\widehat{\mathbf{r}}_{\mathbf{y}}), \tag{4.30}
$$

*Remark* 4.16. (1) For any $\mathbf{z} \in \widehat{\mathcal{C}}_+$ and $\mathbf{y} \in P^V(\mathbf{z}, 10\widehat{\mathbf{r}}_{\mathbf{z}})$, we use the $C\delta$-Lipschitz property of $\widehat{\mathbf{r}}_\bullet$ to obtain

$$
\widehat{\mathbf{r}}_{\mathbf{z}} < \widehat{\mathbf{r}}_{\mathbf{y}} + C\delta|\mathbf{y} - \mathbf{z}| < \widehat{\mathbf{r}}_{\mathbf{y}} + C\delta\widehat{\mathbf{r}}_{\mathbf{z}}.
$$

In particular, $\widehat{\mathbf{r}}_{\mathbf{z}} \leq (1 + C\delta)\widehat{\mathbf{r}}_{\mathbf{y}} \leq (1 + C\delta)C\delta|\mathbf{x} - \mathbf{y}|$, so that $|\mathbf{y} - \mathbf{z}| \leq 11\widehat{\mathbf{r}}_{\mathbf{y}}$, hence we have

$$
(1 - C\delta)\widehat{\mathbf{r}}_{\mathbf{y}} \leq \widehat{\mathbf{r}}_{\mathbf{z}} \leq (1 + C\delta)\widehat{\mathbf{r}}_{\mathbf{y}}. \tag{4.31}
$$

(2) There is a dimensional constant $C$ such that the following holds when $\delta \le \overline{\delta}$. For any $\mathbf{x} \in P^V(\mathbf{0}, \frac{7}{4})$, there exist at most $C$ points $\mathbf{y} \in \widehat{\mathcal{C}}_+$ satisfying $\mathbf{x} \in P^V(\mathbf{y}, 20\widehat{\mathbf{r}}_{\mathbf{y}})$. To see this, let $\widehat{\mathcal{C}}_{\mathbf{x}}$ be the set of $\mathbf{y} \in P^V(\mathbf{0}, \frac{7}{4})$ such that $\mathbf{x} \in P^V(\mathbf{y}, 20\widehat{\mathbf{r}}_{\mathbf{y}})$. Fix $\mathbf{y}_0 \in \widehat{\mathcal{C}}_{\mathbf{x}}$. For any other $\mathbf{y} \in \widehat{\mathcal{C}}_{\mathbf{x}}$, we then have $|\mathbf{y} - \mathbf{y}_0| \le 2(\widehat{\mathbf{r}}_{\mathbf{y}} + \widehat{\mathbf{r}}_{\mathbf{y}_0})$, so by arguing as in (4.31), we have

$$(1 - C\delta)\widehat{\mathbf{r}}_{\mathbf{y}_0} \le \widehat{\mathbf{r}}_{\mathbf{y}} \le (1 + C\delta)\widehat{\mathbf{r}}_{\mathbf{y}_0}.$$

By (n1), it follows that $\{P(\mathbf{y}, \gamma^3\widehat{\mathbf{r}}_{\mathbf{y}_0})\}_{\mathbf{y}\in\widehat{\mathcal{C}}_{\mathbf{x}}}$ is a pairwise-disjoint collection whose union is contained in $P(\mathbf{x}, 100\widehat{\mathbf{r}}_{\mathbf{y}_0})$. It follows that the cardinality of $\widehat{\mathcal{C}}_{\mathbf{x}}$ is bounded above by $C$.

We also need to construct a partition of unity subordinate to the cover (4.30).

**Lemma 4.17** (Partition of unity)**.** *There exist nonnegative* $\psi_{\mathbf{z}} \in C_c^\infty(P^V(\mathbf{z}, 10\widehat{\mathbf{r}}_{\mathbf{z}}) \cap P(\mathbf{0}, \frac{7}{4}))$ *for* $\mathbf{z} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{0}, \frac{9}{5})$ *such that the following hold:*

*(1)* $\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{0},\frac{9}{5})} \psi_{\mathbf{z}} = 1$ *on* $P^V(\mathbf{0}, \frac{7}{4})$.

*(2) There exists a universal constant $C$ such that for any* $\mathbf{x} \in P^V(\mathbf{0}, \frac{7}{4})$*, there are at most $C$ distinct* $\mathbf{z} \in \widehat{\mathcal{C}}_+$ *such that* $\mathbf{x} \in \operatorname{supp}(\psi_{\mathbf{z}})$*.*

*(3) For any* $k, \ell \in \mathbb{N}_0$*, there exist universal constants* $C_{k,\ell}$ *such that*

$$\widehat{\mathbf{r}}_{\mathbf{z}}^{k+2\ell}|\nabla^k \partial_t^\ell \psi_{\mathbf{z}}| \le C_{k,\ell}.$$

*Proof.* Fix $\phi \in C_c^\infty(P^V(\mathbf{0}, 10))$ such that $0 \le \phi \le 1$ and $\phi|_{P^V(\mathbf{0},2)} \equiv 1$. For $\mathbf{z} = (z, t_z)$ and $\mathbf{x} = (x, t) \in P^V(\mathbf{0}, \frac{7}{4})$, we define

$$\phi_{\mathbf{z}}(\mathbf{x}) := \phi(\widehat{\mathbf{r}}_{\mathbf{z}}^{-1}(x - z), \widehat{\mathbf{r}}_{\mathbf{z}}^{-2}(t - t_z)), \qquad \psi_{\mathbf{z}}(\mathbf{x}) := \frac{\phi_{\mathbf{z}}(\mathbf{x})}{\sum_{\mathbf{w}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{0},\frac{9}{5})} \phi_{\mathbf{w}}(\mathbf{x})}.$$

Using the cover (4.30), we know that

$$\sum_{\mathbf{w}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{0},\frac{9}{5})} \phi_{\mathbf{w}}(\mathbf{x}) \ge 1 \tag{4.32}$$

for every $\mathbf{x} \in P^V(\mathbf{0}, \frac{7}{4})$. In particular, $\psi_{\mathbf{z}}$ is well-defined and smooth.

(1) This claim holds by construction.

(2) Fix $\mathbf{x} \in P^V(\mathbf{0}, \frac{7}{4})$. By Remark 4.16(2), there is a universal constant $C$ such that $\mathbf{x} \in \operatorname{supp}(\phi_{\mathbf{z}})$ for at most $C$ distinct $\mathbf{z} \in \widehat{\mathcal{C}}_+$. In particular, (2) holds.

(3) This follows from (4.32), Remark 4.16(1) and the proof of (2). □

We now define the candidate almost-extension $F : V \to \mathbb{R}$ of $F^*$ as

$$F(\mathbf{x}) := \begin{cases} \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} \psi_{\mathbf{z}}(\mathbf{x})\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{x}) & \text{if } \mathbf{x} \notin \widehat{\mathcal{C}}_0, \\ F^*(\mathbf{x}) & \text{if } \mathbf{x} \in \widehat{\mathcal{C}}_0. \end{cases}$$

In the lemma below, we prove that $F$ is a $C\delta^{\frac{2}{k+2}}$-Lipschitz almost extension of $F^*$.

**Lemma 4.18.** *(1) For any* $\mathbf{y} \in \widehat{\mathcal{C}} \cap P^V(\mathbf{0}, \frac{7}{4})$*,*

$$|F(\mathbf{y}) - F^*(\mathbf{y})| \le C\delta\widehat{\mathbf{r}}_{\mathbf{y}}. \tag{4.33}$$

*In particular, $F$ is continuous.*

*(2) $F$ is* $C\delta^{\frac{2}{k+2}}$*-Lipschitz function on* $P^V(\mathbf{0}, \frac{7}{4})$*.*

*Proof.* (1) Because $F|_{\widehat{\mathcal{C}}_0} = F^*|_{\widehat{\mathcal{C}}_0}$, it suffices to consider $\mathbf{y} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{0}, \frac{7}{4})$. For any $\mathbf{z} \in \widehat{\mathcal{C}}_+$ such that $\psi_{\mathbf{z}}(\mathbf{y}) > 0$, we use (4.31), Lemma 4.15(2), and (4.5) to get

$$
\begin{aligned}
C^{-1}\frac{|F^*(\mathbf{y}) - \ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{y})|^2}{\widehat{\mathbf{r}}_{\mathbf{y}}^2} &\le \frac{1}{\widehat{\mathbf{r}}_{\mathbf{z}}^{n+1}} \int_{P(\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}})} \left(\frac{|F^*(\mathbf{w}) - \ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{w})|}{\widehat{\mathbf{r}}_{\mathbf{z}}}\right)^2 d\widehat{\mu}(\mathbf{w}) \\
&\le \frac{\widehat{\mathbf{r}}_{\mathbf{y}}^{n+1}}{\widehat{\mathbf{r}}_{\mathbf{z}}^{n+1}} C\beta_{\mathcal{P}}^2(\mathbf{z}, 2\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}) \le C\delta^2.
\end{aligned}
\tag{4.34}
$$

Therefore, we can estimate

$$
|F^*(\mathbf{y}) - F(\mathbf{y})| \le \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} \psi_{\mathbf{z}}(\mathbf{y})|\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{y}) - F^*(\mathbf{y})| \le C\delta\widehat{\mathbf{r}}_{\mathbf{y}}.
$$

(2) First, suppose $\mathbf{x}, \mathbf{y} \in \widehat{\mathcal{C}}_0$. Then the claim follows from the corresponding property for $F^*$.

Second, let $\mathbf{x} \in V \setminus \widehat{\mathcal{C}}_0$ and $\mathbf{y} \in \widehat{\mathcal{C}}_0$. For any $\mathbf{z} \in \widehat{\mathcal{C}}_+$ such that $\psi_{\mathbf{z}}(\mathbf{x}) > 0$, we use (4.31), Lemma 4.15(4) and $C\delta$-Lipschitz property of $\widehat{\mathbf{r}}_{\bullet}$ to obtain

$$
|\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{x}) - \ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{z})| \le C\delta^{\frac{2}{k+2}}|\mathbf{x} - \mathbf{z}| < C\delta^{\frac{2}{k+2}}\widehat{\mathbf{r}}_{\mathbf{z}} < C\delta^{\frac{2}{k+2}}\widehat{\mathbf{r}}_{\mathbf{x}} \le C\delta^{\frac{2}{k+2}}\delta|\mathbf{x} - \mathbf{y}|.
$$

Using this, the bi-Lipschitz property of $F^*$ and $\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}$, we can estimate

$$
\begin{aligned}
|F(\mathbf{x}) - F(\mathbf{y})| = |F(\mathbf{x}) - F^*(\mathbf{y})| &\le \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} \psi_{\mathbf{z}}(\mathbf{x})|\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{x}) - F^*(\mathbf{y})| \\
&\le C\delta|\mathbf{x} - \mathbf{y}| + \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} \psi_{\mathbf{z}}(\mathbf{x})|\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{z}) - F^*(\mathbf{z})| \\
&\le C\delta|\mathbf{x} - \mathbf{y}|,
\end{aligned}
$$

where we used that $|\mathbf{x} - \mathbf{z}| \le |\mathbf{x} - \mathbf{y}| + |\mathbf{y} - \mathbf{z}| \le (1 + C\delta)|\mathbf{x} - \mathbf{y}|$ in the second line and we used $\widehat{\mathbf{r}}_{\mathbf{y}} \le \delta|\mathbf{x} - \mathbf{y}|$, (4.31) and a computation similar to (4.34) in the third line. This implies the claim when $\mathbf{x} \in V \setminus \widehat{\mathcal{C}}_0$ and $\mathbf{y} \in \widehat{\mathcal{C}}_0$.

Finally, we consider the case $\mathbf{x}_1, \mathbf{x}_2 \in V \setminus \widehat{\mathcal{C}}_0$. We will break up the proof into two cases depending on whether the points are far or near compared to their corresponding $\mathbf{r}$. In the first case, $F$ can be approximated by $F^*$ evaluated at points in $\widehat{\mathcal{C}}_+$ that are nearby $\mathbf{x}_1, \mathbf{x}_2$. Then we can use Lipschitz property of $F^*$. In the latter case, since $F$ is approximated by $\ell$, we can use its derivative estimates. More precisely, suppose $|\mathbf{x}_1 - \mathbf{x}_2| \ge 100 \max\{\widehat{\mathbf{r}}_{\mathbf{x}_1}, \widehat{\mathbf{r}}_{\mathbf{x}_2}\}$. By (4.30), there exist $\mathbf{y}_1, \mathbf{y}_2 \in \widehat{\mathcal{C}}_+$ such that $\mathbf{x}_i \in P(\mathbf{y}_i, 2\widehat{\mathbf{r}}_{\mathbf{y}_i})$. Now using the computation in the preceding paragraph, we have

$$
|F(\mathbf{x}_i) - F^*(\mathbf{y}_i)| \le C\delta|\mathbf{x}_i - \mathbf{y}_i| \le C\delta\widehat{\mathbf{r}}_{\mathbf{y}_i}.
$$

Therefore,

$$
\begin{aligned}
|F(\mathbf{x}_1) - F(\mathbf{x}_2)| &\le |F(\mathbf{x}_1) - F^*(\mathbf{y}_1)| + |F(\mathbf{x}_2) - F^*(\mathbf{y}_2)| + |F^*(\mathbf{y}_1) - F^*(\mathbf{y}_2)| \\
&\le C\delta(\widehat{\mathbf{r}}_{\mathbf{y}_1} + \widehat{\mathbf{r}}_{\mathbf{y}_2}) + C\delta(|\mathbf{y}_1 - \mathbf{x}_1| + |\mathbf{x}_1 - \mathbf{x}_2| + |\mathbf{x}_2 - \mathbf{y}_2|) \\
&\le C\delta|\mathbf{x}_1 - \mathbf{x}_2|.
\end{aligned}
$$

Now we consider the case $|\mathbf{x}_1 - \mathbf{x}_2| \le 100 \max\{\widehat{\mathbf{r}}_{\mathbf{x}_1}, \widehat{\mathbf{r}}_{\mathbf{x}_2}\}$. Without loss of generality, let $\widehat{\mathbf{r}}_{\mathbf{x}_1} = \max\{\widehat{\mathbf{r}}_{\mathbf{x}_1}, \widehat{\mathbf{r}}_{\mathbf{x}_2}\}$. Then using the $C\delta$-Lipschitz property of $\widehat{\mathbf{r}}$, we get $\widehat{\mathbf{r}}_{\mathbf{x}_1} \le \widehat{\mathbf{r}}_{\mathbf{x}_2} + C\delta|\mathbf{x}_1 - \mathbf{x}_2| \le \widehat{\mathbf{r}}_{\mathbf{x}_2} + C\delta\widehat{\mathbf{r}}_{\mathbf{x}_1}$. In particular,

$$
\widehat{\mathbf{r}}_{\mathbf{x}_2} \le \widehat{\mathbf{r}}_{\mathbf{x}_1} \le (1 + C\delta)\widehat{\mathbf{r}}_{\mathbf{x}_2}. \tag{4.35}
$$

Let $\mathbf{x}'$ denote a point in the line segment connecting $\mathbf{x}_1$ and $\mathbf{x}_2$. Since $\mathbf{x}' \in P^V(\mathbf{0}, \frac{7}{4})$, using (4.30), we can choose $\mathbf{y}' \in \widehat{\mathcal{C}}_+$ satisfying $\mathbf{x}' \in P(\mathbf{y}', 2\widehat{\mathbf{r}}_{\mathbf{y}'})$. Then Lemma 4.15(4), and Lemma 4.15((3))

with $\mathbf{x} \leftarrow \mathbf{y}'$, $\mathbf{x}' \leftarrow \mathbf{z}$ where $\mathbf{z}$ satisfies $\nabla\psi_{\mathbf{z}}(\mathbf{x}') \neq 0$, $r \leftarrow \gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{y}'}$, $r' \leftarrow \gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}$ (where the hypotheses are satisfied due to (4.31)) give

$$|\nabla F(\mathbf{x}')| \leq C\delta^{\frac{2}{k+2}} + \left|\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} \nabla\psi_{\mathbf{z}}(\mathbf{x}')(\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{x}') - \ell_{\mathbf{y}',\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{y}'}}(\mathbf{x}'))\right| \leq C\delta^{\frac{2}{k+2}},$$

$$|\partial_t F(\mathbf{x}')| \leq \left|\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} \partial_t\psi_{\mathbf{z}}(\mathbf{x}')(\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{x}') - \ell_{\mathbf{y}',\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{y}'}}(\mathbf{x}'))\right| \leq C\frac{\delta}{\widehat{\mathbf{r}}_{\mathbf{y}'}}.$$

Integrating this along a spacetime path from $\mathbf{x}_1$ to $\mathbf{x}_2$ yields the claim. In fact, the error coming from spatial direction is bounded above by $C\delta^{\frac{2}{k+2}}|\mathbf{x}_1 - \mathbf{x}_2|$. To estimate the error coming from integrating in temporal direction, note that

$$|\mathbf{y}' - \mathbf{x}_i| \leq |\mathbf{y}' - \mathbf{x}'| + |\mathbf{x}' - \mathbf{x}_i| \leq 2\widehat{\mathbf{r}}_{\mathbf{y}'} + |\mathbf{x}_1 - \mathbf{x}_2|.$$

Using this and the $C\delta$-Lipschitz property of $\widehat{\mathbf{r}}$, we get

$$\widehat{\mathbf{r}}_{\mathbf{y}'} \geq \widehat{\mathbf{r}}_{\mathbf{x}_i} - C\delta|\mathbf{y}' - \mathbf{x}_i| \geq \widehat{\mathbf{r}}_{\mathbf{x}_i} - C\delta\widehat{\mathbf{r}}_{\mathbf{x}_i} \geq \frac{1}{2}\widehat{\mathbf{r}}_{\mathbf{x}_i}$$

if $\delta \leq \overline{\delta}$, where we used (4.35) in the second inequality. Therefore, the contribution of integrating in the temporal direction is bounded above by

$$C\frac{\delta|t_1 - t_2|}{\widehat{\mathbf{r}}_{\mathbf{x}_i}} \leq C\frac{\delta|\mathbf{x}_1 - \mathbf{x}_2|^2}{\widehat{\mathbf{r}}_{\mathbf{x}_i}} \leq C\delta|\mathbf{x}_1 - \mathbf{x}_2|\frac{\max\{\widehat{\mathbf{r}}_{\mathbf{x}_1}, \widehat{\mathbf{r}}_{\mathbf{x}_2}\}}{\widehat{\mathbf{r}}_{\mathbf{x}_i}} \leq C\delta|\mathbf{x}_1 - \mathbf{x}_2|,$$

where we used (4.35) in the last inequality. □

We now work towards establishing the BMO estimates for $\partial_t^{\frac{1}{2}}F$. By Proposition B.1, this follows from a Carleson estimate for an analog of the $\beta$-number for measures $\mu$ supported on a Lipschitz graph. For $\mathbf{x} \in V$ and $r > 0$, we define

$$\kappa^2(\mathbf{x}, r) := \kappa_k^2(\mathbf{x}, r; F) := \inf_{\ell} \frac{1}{r^k}\int_{P^V(\mathbf{x},r)} \left(\frac{|F(\mathbf{y}) - \ell(\mathbf{y})|}{r}\right)^2 \mathcal{H}^k_{\mathcal{P}}(\mathbf{y}),$$

where the infimum is taken over all affine functions $\ell\colon V \to \mathbb{R}^{n+2-k}$ satisfying $\partial_t\ell \equiv 0$.

Inspired by [DS91, §19] (see also [BHH+23, Appendix]), we first prove that the Carleson norm of $\kappa^2$ is bounded above by that of $\beta^2$.

**Lemma 4.19.** *For any* $\mathbf{x} \in P(\mathbf{0}, \frac{3}{2})$ *and* $r \in (0, \frac{1}{100}]$,

$$\int_{P^V(\mathbf{x},r)}\int_0^r \kappa^2(\mathbf{y}, s)\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \leq C\delta^{\frac{2}{k+2}}r^k + \int_{P^V(\mathbf{x},10r)}\int_{\frac{1}{8}\widehat{\mathbf{r}}_{\mathbf{y}}}^r \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{y}, 10^5 s)\frac{ds}{s}d\widehat{\mu}(\mathbf{y}). \tag{4.36}$$

The idea is to break up the integral over scale $s$ into two ranges depending on whether it is smaller or larger than $\mathbf{r}_\bullet$. In the case where the scale is smaller than $\mathbf{r}$, we just use higher derivative estimates of $F$. When $s$ is large compared to $\mathbf{r}_\bullet$, we can approximate $\mathcal{H}^k_{\mathcal{P}}$ by $\widehat{\mu}$ using the cover (4.30). On such large scales, $F$ is well approximated by an affine function given by Lemma 4.15. This gives rise to two terms involving $\widehat{\beta}$ and $\widehat{\mu}$, where the second term appears because we use Lipschitz approximation so that we can apply Lemma 4.15. To estimate the term involving $\widehat{\mu}$, we use its Ahlfors regularity (4.28).

*Proof.* First, we assume that $\mathbf{x} \in P(\mathbf{0}, \frac{3}{2}) \cap \widehat{\mathcal{C}}$ and $r \in (0, \frac{1}{100}]$. Note that

$$\int_{P^V(\mathbf{x},r)}\int_0^r \kappa^2(\mathbf{y}, s)\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) = \int_{P^V(\mathbf{x},r)}\int_0^{\widehat{\mathbf{r}}_{\mathbf{y}}} \kappa^2(\mathbf{y}, s)\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) + \int_{P^V(\mathbf{x},r)}\int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \kappa^2(\mathbf{y}, s)\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})$$

$$=: I_1 + I_2,$$

so that the integrand in $I_1$ vanishes when $\mathbf{y} \in \widehat{\mathcal{C}}_0$.

We now estimate $I_2$. Fix $\mathbf{y} \in P^V(\mathbf{x}, r) \subset P^V(\mathbf{0}, \frac{7}{4})$. Using (4.30), we choose $\overline{\mathbf{y}} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{0}, \frac{9}{5})$ such that $\mathbf{y} \in P^V(\overline{\mathbf{y}}, 2\widehat{\mathbf{r}}_{\overline{\mathbf{y}}})$. Now fix $s \in [\widehat{\mathbf{r}}_{\mathbf{y}}, r]$. Let $\ell_{\overline{\mathbf{y}},s}$ be as in Lemma 4.15. By the $C\delta^{\frac{2}{k+2}}$-Lipschitz properties of $\ell$ and $F$ in Lemma 4.15 and Lemma 4.18, respectively, for any $\mathbf{z} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{0}, \frac{9}{5})$, we have

$$\begin{aligned}
&\int_{P^V(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} |F(\mathbf{w}) - \ell_{\overline{\mathbf{y}},s}(\mathbf{w})|^2 \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{w}) \\
&\le C \int_{P^V(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} \left(|F(\mathbf{w}) - F(\mathbf{z})|^2 + |\ell_{\overline{\mathbf{y}},s}(\mathbf{w}) - \ell_{\overline{\mathbf{y}},s}(\mathbf{z})|^2\right) \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{w}) \\
&\qquad + C\widehat{\mathbf{r}}_{\mathbf{z}}^k |\ell_{\overline{\mathbf{y}},s}(\mathbf{z}) - F^*(\mathbf{z})|^2 + C\widehat{\mathbf{r}}_{\mathbf{z}}^k |F^*(\mathbf{z}) - F(\mathbf{z})|^2 \\
&\le C\delta^{\frac{2}{k+2}} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + C\widehat{\mathbf{r}}_{\mathbf{z}}^k |\ell_{\overline{\mathbf{y}},s}(\mathbf{z}) - F^*(\mathbf{z})|^2 \\
&\le C\delta^{\frac{2}{k+2}} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + C \int_{P(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} |\ell_{\overline{\mathbf{y}},s}(\mathbf{z}) - F^*(\mathbf{z})|^2 \, d\widehat{\mu}(\mathbf{w}) \\
&\le C\delta^{\frac{2}{k+2}} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + C \int_{P(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} |\ell_{\overline{\mathbf{y}},s}(\mathbf{w}) - F^*(\mathbf{w})|^2 \, d\widehat{\mu}(\mathbf{w}),
\end{aligned}$$

where we used Ahlfors regularity (4.28) in the third inequality and the Lipschitz properties in the last line. Using the cover (4.30) and summing over $\mathbf{z} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{x}, s)$, we get

$$\begin{aligned}
s^{k+2}\kappa^2(\mathbf{y}, s) &\le C\delta^{\frac{2}{k+2}} \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\overline{\mathbf{y}},4s)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + C \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\overline{\mathbf{y}},4s)} \int_{P(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} |\ell_{\overline{\mathbf{y}},s}(\mathbf{w}) - F^*(\mathbf{w})|^2 d\widehat{\mu}(\mathbf{w}) \\
&\le C\delta^{\frac{2}{k+2}} \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\overline{\mathbf{y}},4s)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + C \int_{P(\overline{\mathbf{y}},8s)} |\ell_{\overline{\mathbf{y}},s}(\mathbf{w}) - F^*(\mathbf{w})|^2 \, d\widehat{\mu}(\mathbf{w}) \\
&\le C\delta^{\frac{2}{k+2}} \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\overline{\mathbf{y}},4s)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + C \int_{P(\overline{\mathbf{y}},8s)} |\ell_{\overline{\mathbf{y}},8s}(\mathbf{w}) - F^*(\mathbf{w})|^2 \, d\widehat{\mu}(\mathbf{w}) \\
&\qquad + Cs^{k+2}\widehat{\beta}^2_{\mathcal{P},k}(\overline{\mathbf{y}}, 8\cdot 10^4 s) \\
&\le C\delta^{\frac{2}{k+2}} \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\overline{\mathbf{y}},4s)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} + Cs^{k+2}\widehat{\beta}^2_{\mathcal{P},k}(\overline{\mathbf{y}}, 8\cdot 10^4 s),
\end{aligned}$$

where we used (4.31) and boundedness of intersection number in the covering (4.30) in the second line (see Remark 4.16(2)), and we used Lemma 4.15 twice in the last two inequalities. On the other hand,

$$\begin{aligned}
\int_{P^V(\mathbf{x},r)} \int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \left(\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\overline{\mathbf{y}},4s)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2}\right) \frac{ds}{s^{k+3}} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) &\le \int_{P^V(\mathbf{x},r)} \int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \left(\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},4r)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2}\right) \frac{ds}{s^{k+3}} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \\
&\le C \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},4r)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} \int_{P^V(\mathbf{z},2\mathbf{r}_{\mathbf{z}})} \int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \frac{ds}{s^{k+3}} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \\
&\le C \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},4r)} \widehat{\mathbf{r}}_{\mathbf{z}}^{k+2} \int_{P^V(\mathbf{z},2\mathbf{r}_{\mathbf{z}})} \frac{1}{\widehat{\mathbf{r}}_{\mathbf{z}}^{k+2}} \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})
\end{aligned}$$

$$\leq C \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},4r)} \widehat{\mathbf{r}}_{\mathbf{z}}^k \leq Cr^k,$$

where we used (4.30) in the second line, (4.31) in the third line and Ahlfors regularity (4.28) in the last line. We can use Ahlfors regularity because $\widehat{\mathbf{r}}_{\mathbf{x}} \leq \widehat{\mathbf{r}}_{\mathbf{y}} + C\delta r \leq 4r$ if $\delta \leq \overline{\delta}$. Similarly,

$$\begin{aligned}
\int_{P^V(\mathbf{x},r)} \int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \kappa^2(\mathbf{y},s)\,\frac{ds}{s} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) &\leq C \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},r)} \int_{P^V(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} \int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \kappa^2(\mathbf{y},s)\,\frac{ds}{s} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\
&\leq Cr^k + \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},r)} \int_{P^V(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} \int_{\widehat{\mathbf{r}}_{\mathbf{y}}}^r \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{z},\cdot 10^5 s)\,\frac{ds}{s} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\
&\leq Cr^k + \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},r)} \widehat{\mathbf{r}}_{\mathbf{z}}^k \int_{\frac{1}{2}\widehat{\mathbf{r}}_{\mathbf{z}}}^r \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{z},10^5 s)\,\frac{ds}{s}\\
&\leq Cr^k + C\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{x},r)} \int_{P^V(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} \int_{\frac{1}{4}\widehat{\mathbf{r}}_{\mathbf{w}}}^r \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{w},10^5 s)\,d\widehat{\mu}(\mathbf{w})\frac{ds}{s}\\
&\leq Cr^k + C\int_{P^V(\mathbf{x},2r)} \int_{\frac{1}{4}\widehat{\mathbf{r}}_{\mathbf{w}}}^r \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{w},10^5 s)\,d\widehat{\mu}(\mathbf{w})\frac{ds}{s}.
\end{aligned}$$

Now we estimate $I_1$. Suppose $\mathbf{y} \in V \setminus \widehat{\mathcal{C}}_0$. For $s \in (0, \widehat{\mathbf{r}}_{\mathbf{y}}]$ and $\mathbf{w} \in P^V(\mathbf{y}, s)$, using (4.30), we choose $\mathbf{y'} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{0}, \frac{9}{5})$ such that $\mathbf{w} \in P(\mathbf{y'}, 2\widehat{\mathbf{r}}_{\mathbf{y'}})$. Then using Lemma 4.15, Lemma 4.17(3) and (4.31), we get

$$\tag{4.37}\begin{aligned}
|\nabla^2 F(\mathbf{w})| \leq &\left|\sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} (\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}}(\mathbf{w}) - \ell_{\mathbf{y'},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{y'}}}(\mathbf{w}))\nabla^2\psi_{\mathbf{z}}(\mathbf{w})\right|\\
&+ \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+} |\nabla\psi_{\mathbf{z}}(\mathbf{w})| \cdot |\nabla(\ell_{\mathbf{z},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{z}}} - \ell_{\mathbf{y'},\gamma^{-1}\widehat{\mathbf{r}}_{\mathbf{y'}}})(\mathbf{w})|\\
\leq &C\widehat{\mathbf{r}}_{\mathbf{w}}^{-1}\delta^{\frac{2}{k+2}}.
\end{aligned}$$

Fix $\mathbf{y} \in P^V(\mathbf{x}, r)$ and $s \in (0, \mathbf{r}_{\mathbf{y}})$. Suppose $\mathbf{y} \in P^V(\mathbf{z}, 2\widehat{\mathbf{r}}_{\mathbf{z}})$ for some $\mathbf{z} \in \widehat{\mathcal{C}}_+ \cap P^V(\mathbf{0}, \frac{9}{5})$. Then $\mathbf{r}_{\mathbf{y}} \leq 2\mathbf{r}_{\mathbf{z}}$. Therefore,

$$\int_{P^V(\mathbf{x},r)} \int_0^{\widehat{\mathbf{r}}_{\mathbf{y}}} \kappa^2(\mathbf{y},s)\,\frac{ds}{s} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \leq C \sum_{\mathbf{z}\in\widehat{\mathcal{C}}_+\cap P^V(\mathbf{0},\frac{9}{5})} \int_{P^V(\mathbf{z},2\widehat{\mathbf{r}}_{\mathbf{z}})} \int_0^{\widehat{\mathbf{r}}_{\mathbf{y}}} \kappa^2(\mathbf{y},s)\,\frac{ds}{s} d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}).$$

If $\ell$ is the linear component of the Taylor expansion at any $\mathbf{y} \in P^V(\mathbf{z}, 2\widehat{\mathbf{r}}_{\mathbf{z}})$, then we know that

$$\sup_{\mathbf{w}\in P^V(\mathbf{y},s)} |F(\mathbf{w}) - \ell(\mathbf{w})|^2 \leq Cs^4 \sup_{\mathbf{w}\in P^V(\mathbf{y},s)} (|\nabla^2 F(\mathbf{w})|^2 + |\partial_t F(\mathbf{w})|^2) \leq Cs^2\delta^{\frac{2}{k+2}}.$$

Thus,

$$\kappa^2(\mathbf{y},s) \leq \frac{1}{s^{k+2}} \int_{P^V(\mathbf{y},s)} |F(\mathbf{w}) - \ell(\mathbf{w})|^2\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{w}) \leq C\frac{s^2}{\widehat{\mathbf{r}}_{\mathbf{z}}^2}\delta^{\frac{2}{k+2}}.$$

Therefore,

$$\int_0^{2\widehat{\mathbf{r}}_{\mathbf{z}}} \int_{P^V(\mathbf{z},\widehat{\mathbf{r}}_{\mathbf{z}})} \kappa^2(\mathbf{y},s)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\frac{ds}{s} \leq C\int_0^{2\widehat{\mathbf{r}}_{\mathbf{z}}} \int_{P^V(\mathbf{z},\widehat{\mathbf{r}}_{\mathbf{z}})} \frac{s^2}{\widehat{\mathbf{r}}_{\mathbf{z}}^2}\delta^{\frac{2}{k+2}}\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\frac{ds}{s} \leq C\widehat{\mathbf{r}}_{\mathbf{z}}^k\delta^{\frac{2}{k+2}}.$$

Now summing over $\mathbf{z}$ we get the desired estimate.

Second we consider arbitrary $\mathbf{x} \in P^V(\mathbf{0}, \frac{3}{2})$ and $r \in (0, \frac{1}{10}]$. If $r \leq \widehat{\mathbf{r}}_{\mathbf{x}}$, then the claim follows by an estimate similar to that for $I_1$. Suppose instead $\widehat{\mathbf{r}}_{\mathbf{x}} < r$. Then choose $\overline{\mathbf{x}} \in P(\mathbf{0}, \frac{9}{5})$ such that $\mathbf{x} \in P^V(\overline{\mathbf{x}}, \widehat{\mathbf{r}}_{\overline{\mathbf{x}}})$. Then $P(\mathbf{x}, r) \subset P(\overline{\mathbf{x}}, 4r)$ because $\widehat{\mathbf{r}}_{\overline{\mathbf{x}}} \leq 2\widehat{\mathbf{r}}_{\mathbf{x}} < 2r$. Thus, we are back to the case $\mathbf{x} \in P(\mathbf{0}, \frac{3}{2}) \cap \widehat{\mathcal{C}}$. ☐

To ensure the hypothesis of Proposition B.1, we now prove a Carleson estimate for the $\beta^2_{\mathcal{P},k}$ term appearing in the right-hand side of (4.36). The proof is inspired by [CJN21, (5.10)].

**Lemma 4.20.** *For any $\mathbf{x} \in P(\mathbf{0}, \frac{3}{2})$ and $r > 0$ such that $P(\mathbf{x}, 20r) \subset P(\mathbf{0}, \frac{7}{4})$,*

$$\int_{P^V(\mathbf{x},10r)} \int_{\frac{1}{8}\widehat{\mathbf{r}}_{\mathbf{y}}}^{r} \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{y}, 10^5 s)\, \frac{ds}{s} d\widehat{\mu}(\mathbf{y}) \leq C\delta r^k.$$

*Proof.* Set $r_i := 2^{-i}$ for $i \in \mathbb{N}$. Applying Fubini's theorem, Lemma 3.14, and (4.28), we obtain

$$\begin{aligned}
&\int_{P(\mathbf{x},10r)} \int_{\frac{1}{8}\widehat{\mathbf{r}}_{\mathbf{y}}}^{r} \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{y}, 10^5 s)\, \frac{ds}{s} d\widehat{\mu}(\mathbf{y}) \\
&\leq C \int_{P(\mathbf{x},10r)} \sum_{\frac{1}{8}\widehat{\mathbf{r}}_{\mathbf{y}} \leq r_i \leq 10^5 r} \widehat{\beta}^2_{\mathcal{P},k}(\mathbf{y}, r_i) d\widehat{\mu}(\mathbf{y}) \\
&\leq C \int_{P(\mathbf{x},10r)} \sum_{i\in\mathbb{N}} \frac{\mathbb{1}_{\frac{1}{8}\{\widehat{\mathbf{r}}_{\mathbf{y}} \leq r_i \leq 10^5 r\}}}{r_i^k} \left( \int_{P(\mathbf{y},r_i)} \mathcal{E}_{20r_i}(\mathbf{z}) d\widehat{\mu}(\mathbf{z}) \right) d\widehat{\mu}(\mathbf{y}) \\
&\leq C \sum_{i\in\mathbb{N}} \frac{1}{r_i^k} \int_{P(\mathbf{x},20r)} \mathbb{1}_{\{\frac{1}{16}\widehat{\mathbf{r}}_{\mathbf{z}} \leq r_i \leq 10^5 r\}} \mathcal{E}_{20r_i}(\mathbf{z}) \left( \int_{P(\mathbf{z},r_i)} d\widehat{\mu}(\mathbf{y}) \right) d\widehat{\mu}(\mathbf{z}) \\
&\leq C \sum_{i\in\mathbb{N}} \int_{P(\mathbf{x},20r)} \mathbb{1}_{\{\frac{1}{16}\widehat{\mathbf{r}}_{\mathbf{z}} \leq r_i \leq 10^5 r\}} \mathcal{E}_{20r_i}(\mathbf{z}) d\widehat{\mu}(\mathbf{z}) \leq C \int_{P(\mathbf{x},20r)} \left( N_{\mathbf{z}}(10^{12} r^2) - N_{\mathbf{z}}(\widehat{\mathbf{r}}^2_{\mathbf{z}}) \right) d\widehat{\mu}(\mathbf{z}) \\
&\leq C\delta \widehat{\mu}(P(\mathbf{x}, 20r)) \leq C\delta r^k,
\end{aligned}$$

where we used $\widehat{\mathbf{r}}_{\mathbf{x}} \leq \widehat{\mathbf{r}}_{\mathbf{y}} + 10\delta r \leq 10r$ and Ahlfors regularity (4.28) in the last line. ☐

*Proof of Theorem 4.14.* By translation and rescaling, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $r = 1$. Choose $\varphi \in C^\infty_c(P^V(\mathbf{0}, \frac{5}{4}))$ such that $\varphi \equiv 1$ in $P^V(\mathbf{0}, \frac{6}{5})$ and $0 \leq \varphi \leq 1$ and estimates on $|\partial_t^i \nabla^j \varphi|$ for $2i + j \leq 2$. We claim that $G := \varphi F$ is the desired function $G$. Note that (4.26) follows from Lemma 4.18(1) and the fact that $F|_{P^V(\mathbf{0},\frac{6}{5})} = G|_{P^V(\mathbf{0},\frac{6}{5})}$. Because $\mathbf{0} \in \mathcal{C}$, we know $F(\mathbf{0}) = \mathbf{0}$, so from Lemma 4.18(2), it follows that

$$\sup_{P^V(\mathbf{0},\frac{7}{4})} |F| \leq C\delta^{\frac{2}{k+2}}.$$

Again using Lemma 4.18(2), we conclude that $G$ is $C\delta^{\frac{2}{k+2}}$-Lipschitz. Thus, it remains to prove a BMO estimate for $\partial_t^{\frac{1}{2}} G$.

In light of Proposition B.1, it suffices to establish a Carleson estimate for $\kappa^2_k(\mathbf{x}, r; G)$:

$$\int_0^r \int_{V\cap P(\mathbf{x},r)} \kappa^2(\mathbf{y}, s; G)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \frac{ds}{s} < C\delta^{\frac{2}{k+2}} r^k. \tag{4.38}$$

for any $\mathbf{x} \in V$ and $r \in (0, 100]$. To see (4.38), we break the integral depending on location, scale and resolution. In the support of $\varphi$ at small scales, we use Taylor expansion of $\varphi$ to find a comparison linear function to estimate $\kappa^2$. Moreover, there is no contribution from outside of the support of $\varphi$

at small scales. Finally, at larger scales, we can just use a crude global bound on $G$. More precisely, we have

$$\begin{aligned}\int_{P^V(\mathbf{x},r)}\int_0^r \kappa^2(\mathbf{y},s;G)\,\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) &= \int_{P^V(\mathbf{x},r)\cap P(\mathbf{0},\frac{4}{3})}\int_0^{\min\{r,\frac{1}{100}\}} \kappa^2(\mathbf{y},s;G)\,\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\ &\quad+\int_{P^V(\mathbf{x},r)\setminus P(\mathbf{0},\frac{4}{3})}\int_0^{\min\{r,\frac{1}{100}\}} \kappa^2(\mathbf{y},s;G)\,\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\ &\quad+\int_{P^V(\mathbf{x},r)}\int_{\min\{r,\frac{1}{100}\}}^r \kappa^2(\mathbf{y},s;G)\,\frac{ds}{s}d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\ &=: J_1+J_2+J_3.\end{aligned}$$

Note that $J_2=0$. In fact, for any $\mathbf{y}\in P^V(\mathbf{x},r)\setminus P(\mathbf{0},\frac{4}{3})$ and $s\le\min\{r,\frac{1}{100}\}$, we note that $P(\mathbf{y},s)\cap P(\mathbf{0},\frac{5}{4})=\emptyset$. Thus, $G\equiv 0$ on $P(\mathbf{y},s)$. Because $\kappa^2(\mathbf{y},s;G)\le C\delta^{\frac{2}{k+2}}$ when $s\ge\frac{1}{100}$, we also obtain $J_3\le C\delta^{\frac{2}{k+2}}r^k$.

To estimate $J_1$, let $\mathbf{y}\in P^V(\mathbf{0},\frac{4}{3})$ and $s\in(0,\frac{1}{100}]$. Define

$$\bar{\ell}\in\arg\min_{\ell}\frac{1}{s^k}\int_{P^V(\mathbf{y},s)}\left(\frac{|F(\mathbf{w})-\ell(\mathbf{w})|}{s}\right)^2\mathcal{H}^k_{\mathcal{P}}(\mathbf{w}),$$

where $\ell$ is a function $\ell\colon V\to\mathbb{R}^{n+2-k}$ satisfying $\partial_t\ell\equiv 0$. Define

$$\begin{aligned}\widehat{\ell}(\mathbf{w})&\coloneqq\varphi(\mathbf{y})\bar{\ell}(\mathbf{w})+F(\mathbf{y})\nabla\varphi(\mathbf{y})\cdot(\mathbf{w}-\mathbf{y}),\\ q(\mathbf{w})&\coloneqq\varphi(\mathbf{w})-\varphi(\mathbf{y})-\nabla\varphi(\mathbf{y})\cdot(\mathbf{w}-\mathbf{y}).\end{aligned}$$

Using that $\varphi$ is Lipschitz, $|F|\le C\delta^{\frac{2}{k+2}}$, $0\le\varphi\le 1$, and $|q(\mathbf{w})|\le C|\mathbf{w}-\mathbf{y}|^2$, we obtain

$$\begin{aligned}\kappa_k^2(\mathbf{y},s;G)&\le\frac{1}{s^k}\int_{P^V(\mathbf{y},s)}\left(\frac{|\varphi(\mathbf{w})F(\mathbf{w})-\widehat{\ell}(\mathbf{w})|}{s}\right)^2 d\mathcal{H}^k_{\mathcal{P}}(\mathbf{w})\\ &\le\frac{1}{s^{k+2}}\int_{P^V(\mathbf{y},s)}2\Big(|\varphi(\mathbf{y})(F(\mathbf{w})-\bar{\ell}(\mathbf{w}))|^2+|q(\mathbf{w})F(\mathbf{w})|^2\\ &\qquad\qquad+|(F(\mathbf{w})-F(\mathbf{y}))\nabla\varphi(\mathbf{y})\cdot(\mathbf{w}-\mathbf{y})|^2\Big)\,d\mathcal{H}^k_{\mathcal{P}}(\mathbf{w})\\ &\le 2\kappa_k^2(\mathbf{y},s;F)+\frac{2}{s^{k+2}}s^k\cdot C\delta^{\frac{2}{k+2}}\cdot Cs^4\\ &\le 2\kappa_k^2(\mathbf{y},s;F)+C\delta^{\frac{2}{k+2}}s^2.\end{aligned}\tag{4.39}$$

If $J_1\ne 0$, then $P^V(\mathbf{x},\frac{1}{100})\cap P(\mathbf{0},\frac{4}{3})\ne\emptyset$, hence $\mathbf{x}\in P(\mathbf{0},\frac{3}{2})$. We may therefore combine (4.39), Lemma 4.19, Lemma 4.20 to obtain $J_1\le C\delta^{\frac{2}{k+2}}r^k$. Thus, (4.38) holds.

Therefore, the BMO estimate for $\partial_t^{\frac{1}{2}}G$ follows by using Proposition B.1 for each component of $G$. □

## 5. Neck Decomposition

In this section, we prove Theorem 5.1, Theorem 5.2, and Theorem 5.3, which decompose $P(\mathbf{0},1)$ into $k$-neck regions, $(k+1)$-symmetric balls, and a residual set. The strategy for these decomposition comes from [JN21, Section 7] (see also [NV19]).

In the first version of the neck decomposition, the residual set has $k$-Hausdorff measure zero. Furthermore, the neck regions and $(k+1)$-symmetric balls have summable Minkowski content.

**Theorem 5.1** (Neck Decomposition)**.** *Suppose $u$ is a caloric function satisfying $N(\gamma^{-6}) \leq \Lambda$. For any $\epsilon, \eta \in (0,1]$, there exists a decomposition*

$$P(\mathbf{0},1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_b P(\mathbf{x}_b, r_b) \cup \left( \widetilde{\mathcal{C}} \cup \bigcup_a \mathcal{C}_{a,0} \right),$$

*where the following hold:*

*(a) $\mathcal{N}^a = P(\mathbf{x}_a, 2r_a) \setminus \overline{P}(\mathcal{C}_a, \mathbf{r}_a)$ is an $(m_a, k, \epsilon, C^{-\Lambda}\eta)$-neck region for some positive integer $m_a \leq C\Lambda$, where $\mathcal{C}_a$ is the center set associated to $\mathcal{N}^a$;*

*(b) $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b) \leq \eta$ for some $\mathbf{y}_b \in P(\mathbf{x}_b, 4r_b)$;*

*(c) A $k$-dimensional parabolic Minkowski content estimate holds:*

$$\sum_a r_a^k + \sum_b r_b^k \leq C^{\Lambda^2}(\epsilon\eta)^{-C\Lambda};$$

*(d) Further, $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}) = 0$.*

In the second version of the neck decomposition, given a "resolution" $r_*$, the residual set is of scale commensurable to $r_*$. Commensurability will be crucial in proving the volume estimates for the quantitative nodal and singular sets at resolution $r_*$. Such a consideration is necessary because we cannot approximate our solutions $u$ by solutions whose neck regions all satisfy $\mathcal{C}_0 = \emptyset$ unlike the settings of [JN21, NV19].

**Theorem 5.2** (Finite-Resolution Neck Decomposition)**.** *Suppose $u$ is a caloric function satisfying $N(\gamma^{-6}) \leq \Lambda$. For any $\epsilon, \eta, r_* \in (0,1]$, there exists a decomposition*

$$P(\mathbf{0},1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_b P(\mathbf{x}_b, r_b) \cup \bigcup_f P(\mathbf{x}_f, r_f),$$

*where the following hold:*

*(a) $\mathcal{N}^a = P(\mathbf{x}_a, 2r_a) \setminus \overline{P}(\mathcal{C}_a, \mathbf{r}_a)$ is an $(m_a, k, \epsilon, C^{-\Lambda}\eta)$-neck region for some positive integer $m_a \leq C\Lambda$, where $\mathcal{C}_a$ is the center set associated to $\mathcal{N}^a$ and $\mathbf{r}_\bullet \geq r_*$;*

*(b) $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b) \leq \eta$ for some $\mathbf{y}_b \in P(\mathbf{x}_b, 4r_b)$;*

*(c) A $k$-dimensional parabolic Minkowski content estimate holds:*

$$\sum_a r_a^k + \sum_b r_b^k + \sum_f r_f^k \leq C^{\Lambda^2}(\epsilon\eta)^{-C\Lambda},$$

*such that $r_a, r_b \geq r_*$ and $r_f \in [r_*, C^\Lambda \eta^{-\frac{2}{n}} \epsilon^{-20n^2} r_*]$.*

In the third version, we strengthen the neck decomposition to include strong neck regions when $k \in \{n, n+1\}$. This version is useful in proving stronger regularity of the nodal and singular sets.

**Theorem 5.3** (Refined Neck Decomposition)**.** *Let $k \in \{n, n+1\}$. Suppose $u$ is a caloric function satisfying $N_{\mathbf{0}}(\gamma^{-100}) \leq \Lambda$. For any $\epsilon \in (0,1]$ and $\eta \leq \overline{\eta}(\Lambda)$, there exists a decomposition*

$$P(\mathbf{0},1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_g P(\mathbf{x}_g, r_g) \cup \left( \widetilde{\mathcal{C}} \cup \bigcup_a \mathcal{C}_{a,0} \right),$$

*where the following hold:*

*(a) $\mathcal{N}^a = P(\mathbf{x}_a, 2r_a) \setminus \overline{P}(\mathcal{C}_a, \mathbf{r}_a)$ is an $(m_a, k, \epsilon, C(\Lambda)^{-1}\eta)$-strong neck region of scale $r_a$ for some integer $m_a \in [m_*, C\Lambda]$, where*

$$m_* := \begin{cases} 1 & \text{if } k = n+1, \\ 2 & \text{if } k = n. \end{cases} \tag{5.1}$$

*Here, $\mathcal{C}_a$ is the center set associated to $\mathcal{N}^a$. Further, $\mathcal{N}^a = \widehat{\mathcal{N}}^a \cap P(\mathbf{x}_a, 2r_a)$ where $\widehat{\mathcal{N}}^a$ is an $(m_a, k, \epsilon, C(\Lambda)^{-1}\eta)$ strong neck region of scale $2r_a$;*

(b) *For all $g$,*

$$\sup_{\mathbf{x}\in P(\mathbf{x}_g, r_g)} N_{\mathbf{x}}(\epsilon^{-2} r_g^2) \le m_* - 1 + C(\Lambda)^{-1}\epsilon;$$

(c) *A $k$-dimensional parabolic Minkowski content estimate holds:*

$$\sum_a r_a^k + \sum_g r_g^k \le C(\Lambda)(\epsilon\eta)^{-C(\Lambda)};$$

(d) *Further, $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}) = 0$.*

We prove all neck decomposition theorems by an argument involving double induction. Namely, we first induct on scales keeping the frequency fixed. Then we induct on the frequency, which drops by 1 at each step (cf. Lemma 2.11). Eventually, the frequency of the remaining balls is sufficiently small so that $u$ is very symmetric.

**Definition 5.4.** For $\delta > 0$, $m \in \mathbb{N}_0$, and $r \in (0, 1]$, define the *pinched set*

$$\mathcal{V}^u_{\delta,m,r}(\mathbf{x}) \coloneqq \left\{\mathbf{y} \in P(\mathbf{x}, 4r) \colon |N^u_{\mathbf{y}}(s^2) - m| \le \delta \text{ for any } s \in [\delta r, \delta^{-1} r]\right\}.$$

To carry out the induction procedure, given $\mathbf{x} \in P(\mathbf{0}, 1)$, $m, k \in \mathbb{N}_0$, and $r, \epsilon, \eta, \alpha > 0$, we classify $P(\mathbf{x}, r)$ into the following categories:

(a) A ball $P(\mathbf{x}_a, r_a)$ is associated with an $(m, k, \epsilon, \eta)$-neck region $\mathcal{N}^a = P(\mathbf{x}_a, 2r_a) \setminus \overline{P}(\mathcal{C}^a, \mathbf{r}^a)$;

(b) A ball $P(\mathbf{x}_b, r_b)$ such that $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b) \le \eta$ for some $\mathbf{y}_b \in P(\mathbf{x}_b, 4r_b)$;

(c) A ball $P(\mathbf{x}_c, r_c)$ is not a (b)-ball, and $\mathcal{V}_{\delta,m,r_c}(\mathbf{x}_c)$ is $(k, \alpha r_c)$-independent;

(d) A ball $P(\mathbf{x}_d, r_d)$ is not a (b)-ball, and $\mathcal{V}_{\delta,m,r_d}(\mathbf{x}_d) \neq \emptyset$ is not $(k, \alpha r_d)$-independent;

(e) A ball $P(\mathbf{x}_e, r_e)$ is not a (b)-ball, and $\mathcal{V}_{\delta,m,r_e}(\mathbf{x}_e) = \emptyset$.

5.1. **Covering of (c)-balls.** We first show that any (c)-ball can be decomposed into a neck region, (b)-balls, (d)-balls, and (e)-balls. Our proof is similar to [JN21, Section 7.3], except that using Lemma 3.12, we can show that the approximating planes in our inductive covering can be taken independent of scale and location.

**Proposition 5.5** (Covering of (c)-balls)**.** *For any $\alpha, \epsilon, \eta, \Lambda > 0$, the following statements hold when $\delta \le C(\alpha)^{\Lambda}\eta^{\frac{1}{n}}\epsilon^{10n^2}$. Let $u$ be a caloric function satisfying*

$$\sup_{\mathbf{x}\in P(\mathbf{0}, 20r)} N^u_{\mathbf{x}}(\delta^{-2} r^2) \le m + \delta, \tag{5.2}$$

*for some non-negative integer $m \le \Lambda$. Suppose $\mathcal{V} \coloneqq \mathcal{V}^u_{\delta,m,r}(\mathbf{0})$ is a $(k, \alpha r)$-independent set and that $\mathcal{E}^{k+1,1}_{100r}(\mathbf{y}) > \eta$ for all $\mathbf{y} \in P(\mathbf{0}, 2r)$. Then the following hold:*

(1) *There exists a decomposition*

$$P(\mathbf{0}, r) \subseteq (\mathcal{C}_0 \cup \mathcal{N}) \cup \bigcup_{b\in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{d\in D} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e),$$

*where $\mathcal{N}$ is an $(m, k, \epsilon, C(\alpha)^{-\Lambda}\eta)$-neck of scale $10r$, and we have a $k$-dimensional content estimate*

$$\mathcal{H}^k_{\mathcal{P}}(\mathcal{C}_0) + \sum_{b\in B} r_b^k + \sum_{d\in D} r_d^k + \sum_{e\in E} r_e^k \le C r^k.$$

(2) *For any $r_* \in (0, r]$, there exists a decomposition*

$$P(\mathbf{0}, r) \subseteq \mathcal{N} \cup \bigcup_{b\in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{d\in D} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e) \cup \bigcup_{f\in F} P(\mathbf{x}_f, r_f),$$

*where $\mathcal{N}$ is an $(m,k,\epsilon,C(\alpha)^{-\Lambda}\eta)$-neck of scale $10r$ with $\mathbf{r}_\bullet \geq r_*$, and we have a $k$-dimensional content estimate*

$$\sum_{b\in B} r_b^k + \sum_{d\in D} r_d^k + \sum_{e\in E} r_e^k + \sum_{f\in F} r_f^k \leq Cr^k,$$

*$r_e, r_b, r_d \geq r_*$, and $r_f \in [r_*, \gamma^{-1}r_*]$.*

*Remark* 5.6. In Proposition 5.5(1), $\mathcal{C} = \mathcal{C}_0 \cup \{\mathbf{x}_b\} \cup \{\mathbf{x}_d\} \cup \{\mathbf{x}_e\}$ and $\mathbf{r}_{\mathbf{x}_b} = r_b$, $\mathbf{r}_{\mathbf{x}_d} = r_d$, $\mathbf{r}_{\mathbf{x}_e} = r_e$. A similar description exists for the center set in Proposition 5.5(2).

*Proof.* By parabolic rescaling, we may assume $r = 1$. We will first construct a sequence $(\mathcal{N}^j)_{j\in\mathbb{N}}$ of $(m,k,\epsilon,C(\alpha)^{-\Lambda}\eta)$-neck regions inductively.

(1) Because $\mathcal{V} \neq \emptyset$, we can choose $\mathbf{x}_a \in \mathcal{V}$. Set $s_0 := 10^4$. We claim that $u$ is not $(k+1, C^{-\Lambda}\eta^2, s_0)$ symmetric at $\mathbf{x}_a$ if $\delta \leq C^{-\Lambda}\eta^2$. Suppose not. Then we may apply Lemma 3.12(2) with $\mathcal{C} \leftarrow \{\mathbf{x}_a\}$, $\delta \leftarrow C^{-\Lambda}\eta^2$, $s_0 \leftarrow s_0$, $r_0 \leftarrow s_0$, $\kappa \leftarrow \gamma$, $\mathbf{r}_{\mathbf{x}_a} \leftarrow 1$, and $s \leftarrow 100$ so that $\mathcal{E}_{100}^{k+1,1}(\mathbf{x}_a) < \eta$, which is a contradiction.

Since $\mathcal{V}$ is $(k,\alpha)$-independent, $\mathcal{E}_{s_0}^{k,s_0^{-1}\alpha}(\mathbf{x}_a) \leq \delta$. By Theorem 3.6(1), $u$ is $(k, C(\alpha)^\Lambda\delta, s_0)$-symmetric at $\mathbf{x}_a$. From this and (5.2), we can apply Lemma 3.12 with $\mathcal{C} \leftarrow \{\mathbf{x}_a\}$, $r_0 \leftarrow s_0$, $\kappa \leftarrow 10^3\gamma^{-1}\delta$, $\delta \leftarrow C(\alpha)^\Lambda\delta$, and $\mathbf{r}_\bullet \leftarrow \gamma$ to obtain $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ such that:

(1.A) for all $s \in [\gamma, \gamma^{-5}]$, $V \cap P(\mathbf{x}_a, 10s) \subseteq \{\mathbf{y} \in P(\mathbf{x}_a, 10s) \colon |N_{\mathbf{y}}(10^6\delta^2 s^2) - m| < C(\alpha)^\Lambda\sqrt{\delta}\}$;

(1.B) for all $s \in [\gamma, 1]$, for any $\zeta \geq C(\alpha)^\Lambda\sqrt{\delta}$, we have

$$\{\mathbf{y} \in P(\mathbf{x}_a, 10s) : |N_{\mathbf{y}}(s^2) - m| < \zeta\} \subseteq P(V, C(\alpha)^\Lambda\eta^{-\frac{1}{n}}\zeta^{\frac{1}{n}}s).$$

Let $\{\mathbf{x}_i^1\}_{i=1}^{N_1}$ be a maximal subset of $V \cap P(\mathbf{x}_a, 20)$ satisfying $\mathbf{x}_a \in \{\mathbf{x}_i^1\}_{i=1}^{N_1}$ and $|\mathbf{x}_i^1 - \mathbf{x}_j^1| \geq 2\gamma^3$ when $i \neq j$. Each ball $P(\mathbf{x}_i^1, \gamma)$ must be of the type (b), (c), (d) or (e). Let $\mathcal{C}^1 := \{\mathbf{x}_i^1\}_{i=1}^{N_1}$, and set $\mathbf{r}_{\mathbf{x}}^1 := \gamma$ for all $\mathbf{x} \in \mathcal{C}^1$. Defining $\mathcal{N}^1 := P(\mathbf{x}_a, 20) \setminus \overline{P}(\mathcal{C}^1, \mathbf{r}^1)$, we then have

$$P(\mathbf{0},1) \subseteq \mathcal{N}^1 \cup \bigcup_{b\in B^1} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in C^1} P(\mathbf{x}_c, r_c) \cup \bigcup_{d\in D^1} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E^1} P(\mathbf{x}_e, r_e),$$

where $r_b = r_c = r_d = r_e = \gamma$.

**Claim 5.6.1.** *If $\delta \leq C(\alpha)^{-\Lambda}\min\{\epsilon^2, \eta\}$, $\mathcal{N}^1$ is an $(m,k,\epsilon,C(\alpha)^{-\Lambda}\eta)$-neck region of scale* 10.

*Proof of Claim 5.6.1.* Properties (n1) and (n4) hold by construction. Property (n2) holds by (1.A) if $\delta \leq C(\alpha)^{-\Lambda}\epsilon^2$. Suppose $\mathbf{x} \in \mathcal{C}^1$ and $s \in [\gamma, \gamma^{-3}]$. Applying Lemma 3.12(1) with $\mathcal{C} \leftarrow \mathcal{C}^1$, $r_0 \leftarrow 1$, $\kappa \leftarrow \gamma^3$, $\delta \leq C(\alpha)^\Lambda\sqrt{\delta}$, and $\mathbf{r}_\bullet \leftarrow \gamma$ yields that $u$ is $(k, C(\alpha)^\Lambda\sqrt{\delta}, s)$-symmetric at $\mathbf{x}$ with respect to $V$. We claim that $u$ is not $(k+1, C(\alpha)^{-\Lambda}\eta, s)$-symmetric at $\mathbf{x}$. Suppose not. Then Lemma 3.12(1) with $\mathcal{C} \leftarrow \mathcal{C}^1$, $r_0 \leftarrow 1$ $\kappa \leftarrow \gamma^3$, $s_0 \leftarrow s$, $\delta \leftarrow C(\alpha)^{-\Lambda}\eta$, $\mathbf{r}_\bullet \leftarrow \gamma$ implies that $u$ is $(k+1, C(\alpha)^\Lambda\delta, s_0)$-symmetric at $\mathbf{x}_a$, which is a contradiction. Thus (n3) holds. ☐

Set $\epsilon' := C(\alpha)^\Lambda\eta^{-\frac{1}{n}}\delta^{\frac{1}{2n}}$, $\eta' := C(\alpha)^{-\Lambda}\eta$. Suppose by induction that we have defined collections of centers

$$\mathcal{C}^\ell := \{\mathbf{x}_b\}_{b\in B^\ell} \cup \{\mathbf{x}_c\}_{c\in C^\ell} \cup \{\mathbf{x}_d\}_{d\in D^\ell} \cup \{\mathbf{x}_e\}_{e\in E^\ell}$$

for $\ell = 1, \ldots, j$ with

$$P(\mathbf{0},1) \subseteq \mathcal{N}^\ell \cup \bigcup_{b\in B^\ell} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in C^\ell} P(\mathbf{x}_c, r_c) \cup \bigcup_{d\in D^\ell} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E^\ell} P(\mathbf{x}_e, r_e),$$

where $\mathcal{N}^\ell = P(\mathbf{x}_a, 20) \setminus \overline{P}(\mathcal{C}^\ell, \mathbf{r}^\ell)$ is an $(m,k,\epsilon,\eta')$-neck region with corresponding plane $V \in \mathrm{Gr}_{\mathcal{P}}(k)$, such that additionally the following hold for $1 \leq \ell \leq j$:

(j.i) for $b \in B^\ell \setminus B^{\ell-1}$, $c \in C^\ell$, $d \in D^\ell \setminus D^{\ell-1}$, and $e \in E^\ell \setminus E^{\ell-1}$, we have $r_b = r_c = r_d = \gamma^\ell$, (j.ii) there exist $\mathbf{y}_c \in \mathcal{V}_{\delta,m,r_c}(\mathbf{x}_c)$ for each $c \in C^{\ell-1}$ such that $\mathbf{x}_c \in P(\mathbf{y}_c + V, \epsilon'\gamma^\ell)$ and

$$\mathcal{C}^\ell \setminus \mathcal{C}^{\ell-1} \subseteq \bigcup_{c\in C^{\ell-1}} ((\mathbf{y}_c + V) \cap P(\mathbf{x}_c, r_c)) \subseteq \bigcup_{\mathbf{z}\in\mathcal{C}^\ell} P(\mathbf{z}, 2\gamma^2\mathbf{r}^\ell_{\mathbf{z}}),$$

(j.iii) if $\mathbf{x} \in \mathcal{C}^\ell \setminus \mathcal{C}^{\ell-1}$, then $\mathbf{r}^\ell_{\mathbf{x}} = \gamma^\ell$; otherwise, we have $\mathbf{r}^\ell_{\mathbf{x}} = \mathbf{r}^{\ell-1}_{\mathbf{x}}$,
(j.iv) $|N_{\mathbf{x}}(s^2) - m| < \epsilon'$ for $\mathbf{x} \in \mathcal{C}^\ell$ and $s \in [10^3\gamma^{-1}\delta\mathbf{r}^\ell_{\mathbf{x}}, \gamma^{-5}]$,
(j.v) for any $\mathbf{x} \in \mathcal{C}^j$, and $\ell < j$, there exists $c_\ell \in C^\ell$ such that $P(\mathbf{x}, \gamma^j) \subseteq P(\mathbf{x}_{c_\ell}, \frac{3}{2}\gamma^\ell)$.

Note that (1.i), (1.ii), (1.iii), and (1.v) are satisfied by construction. Further, (1.iv) holds by (1.A).

Given the information at $j$th step, we proceed to the $(j+1)$th step as follows. For each $c \in C^j$, choose $\mathbf{y}_c \in \mathcal{V}_{\delta,m,\gamma^j}(\mathbf{x}_c)$, which is non-empty by assumption. Because $u$ is $(k, C(\alpha)^\Lambda\delta, s_0)$-symmetric with respect to $V$ at $\mathbf{x}_a$, but not $(k+1, C^{-\Lambda}\eta, s_0)$-symmetric at $\mathbf{x}_a$, we can apply (5.2) and Lemma 3.12 with $\mathcal{C} \leftarrow \{\mathbf{y}_c, \mathbf{x}_a\}$, $\mathbf{r}_{\mathbf{x}_a} \leftarrow \gamma$, $\kappa \leftarrow 10^3\gamma^{-1}\delta$, $s_0 \leftarrow s_0$, $\mathbf{r}_{\mathbf{y}_c} \leftarrow \gamma^j$, $r_0 \leftarrow s_0$, $\delta \leftarrow C(\alpha)^\Lambda\delta$, $\eta \leftarrow C^{-\Lambda}\eta$ to conclude the following for all $c \in C^j$:

(j.A) for all $s \in [\gamma^j, \gamma^{-5}]$, $(\mathbf{y}_c + V) \cap P(\mathbf{y}_c, 10s) \subseteq \{\mathbf{y} \in P(\mathbf{y}_c, 10s) \colon |N_{\mathbf{y}}(10^6\delta^2 s^2) - m| < C(\alpha)^\Lambda\sqrt{\delta}\}$;
(j.B) for all $s \in [\gamma^j, 1]$ and $\zeta \in [C(\alpha)^\Lambda\sqrt{\delta}, 1]$, we have

$$\{\mathbf{y} \in P(\mathbf{y}_c, 10s) : |N_{\mathbf{y}}(s^2) - m| < \zeta\} \subseteq P(\mathbf{y}_c + V, C(\alpha)^\Lambda\eta^{-\frac{1}{n}}\zeta^{\frac{1}{n}}s).$$

Let $\{\mathbf{x}^{j+1}_i\}_{i=1}^{N_{j+1}}$ be a maximal subset of $\cup_{c\in C^j}((\mathbf{y}_c + V) \cap P(\mathbf{x}_c, \gamma^j))$ such that

$$|\mathbf{x}^{j+1}_{i_1} - \mathbf{x}^{j+1}_{i_2}| \geq 2\gamma^{(j+1)+2} \text{ for } 1 \leq i_1 \neq i_2 \leq N_{j+1},$$
$$|\mathbf{x}^{j+1}_i - \mathbf{x}| \geq \gamma^2(\gamma^{j+1} + \mathbf{r}^j_{\mathbf{x}}) \text{ for } 1 \leq i \leq N_{j+1} \text{ and } \mathbf{x} \in \mathcal{C}^j \setminus \{\mathbf{x}_c\}_{c\in C^j}.$$

Define $\mathcal{C}^{j+1} := (\mathcal{C}^j \setminus \{\mathbf{x}_c\}_{c\in C^j}) \cup \{\mathbf{x}^{j+1}_i\}_{i=1}^{N_{j+1}}$, and set

$$\mathbf{r}^{j+1}_{\mathbf{x}} := \begin{cases} \mathbf{r}^j_{\mathbf{x}} & \mathbf{x} \in \mathcal{C}^j \setminus \{\mathbf{x}_c\}_{c\in C^j}, \\ \gamma^{j+1} & \text{otherwise.} \end{cases}$$

Categorize $\{P(\mathbf{x}^{j+1}_i, \gamma^{j+1})\}_{i=1}^{N_{j+1}}$ into balls of type (b), (c), (d) and (e). Let $B^{j+1}, D^{j+1}, E^{j+1}$ be the union of $B^j, D^j, E^j$ with the new balls of type (b), (d), (e), respectively. Let $C^{j+1}$ be the new balls of type (c). Define $r_b$, $r_c$, $r_d$ so that induction hypothesis (j+1.i) holds.

**Claim 5.6.2.** *The statements* (j+1.i)-(j+1.v) *hold.*

*Proof of Claim 5.6.2.* Statements (j+1.i), (j+1.iii) and (j+1.v) hold by construction. For $c \in C^{j+1}$, the statement (j+1.iv) for $\mathbf{x} = \mathbf{x}_c$ follows from (j.A) and $\mathbf{x}_c \in \cup_{c'\in C^j}((\mathbf{y}_{c'} + V) \cap P(\mathbf{y}_{c'}, 5\gamma^j))$. Finally, we prove (j+1.ii). By our construction, there exists $c' \in C^{j-1}$ such that

$$\mathbf{x}_c \in (\mathbf{y}_{c'} + V) \cap P(\mathbf{x}_{c'}, \gamma^{j-1}) \subset (\mathbf{y}_{c'} + V) \cap P(\mathbf{y}_{c'}, 2\gamma^{j-1}).$$

By (j$-$1.A) with $s \leftarrow \gamma^{j-1}$, we thus have $|N_{\mathbf{x}_c}(\gamma^{2j}) - m| < C^\Lambda\sqrt{\delta}$. Using this, and $\mathbf{x}_c \in P(\mathbf{y}_c, 10\gamma^j)$, (j.B) with $s \leftarrow \gamma^j$ and $\zeta \leftarrow C(\alpha)^\Lambda\sqrt{\delta}$ gives $\mathbf{x}_c \in P(\mathbf{y}_c + V, C(\alpha)^\Lambda\eta^{-\frac{1}{n}}\delta^{\frac{1}{2n}}\gamma^j)$. The other claims of (j+1.ii) follow by construction. □

**Claim 5.6.3.** *For any* $\ell \leq j+1$ *and* $\mathbf{x} \in \mathcal{C}^\ell$*, there exists* $\mathbf{y} \in \mathcal{C}^{j+1}$ *such that* $|\mathbf{x} - \mathbf{y}| < 2\gamma^2(\gamma^\ell + \mathbf{r}^{j+1}_{\mathbf{y}})$.

*Proof of Claim 5.6.3.* We prove by induction on $\ell = j+1$ down to 1. If $\ell = j+1$, we take $\mathbf{y} = \mathbf{x}$. Suppose the claim holds for $\ell + 1$, where $\ell \leq j$, and consider $\mathbf{x} \in \mathcal{C}^\ell$. If $\mathbf{x} \in \mathcal{C}^{j+1}$, then we can take $\mathbf{y} = \mathbf{x}$. Otherwise, we have $\mathbf{x} = \mathbf{x}_c$ for some $c \in C^\ell$. By $(\ell + 1.\text{ii})$, we have $\mathbf{x}_c \in P(\mathbf{y}_c + V, \epsilon'\gamma^{\ell+1})$. Again using $(\ell + 1.\text{ii})$, we have $(\mathbf{y}_c + V) \cap P(\mathbf{x}_c, r_c) \subseteq P(\mathbf{z}, 2\gamma^2\mathbf{r}^\ell_{\mathbf{z}})$ for some $\mathbf{z} \in \mathcal{C}^{\ell+1}$. Hence, by the triangle inequality, $|\mathbf{x} - \mathbf{z}| < \epsilon'\gamma^{\ell+1} + 2\gamma^2\mathbf{r}^{\ell+1}_{\mathbf{z}}$. If $\mathbf{z} \in \mathcal{C}^{j+1}$, then we take $\mathbf{y} = \mathbf{z}$, and we are done. Otherwise, $\mathbf{z} \in \{\mathbf{x}_{c'}\}_{c' \in C^{\ell+1}}$, so that $\mathbf{r}_{\mathbf{z}}^{\ell+1} = \gamma^{\ell+1}$, and by the induction hypothesis, there is $\mathbf{y} \in \mathcal{C}^{j+1}$ such that $|\mathbf{z} - \mathbf{y}| < 2\gamma^2(\gamma^{\ell+1} + \mathbf{r}_{\mathbf{y}}^{j+1})$. By the triangle inequality,

$$|\mathbf{x} - \mathbf{y}| \le |\mathbf{x} - \mathbf{z}| + |\mathbf{z} - \mathbf{y}| \le \epsilon' \gamma^{\ell+1} + 2\gamma^2\gamma^{\ell+1} + 2\gamma^2(\gamma^{\ell+1} + \mathbf{r}_{\mathbf{y}}^{j+1}) < 2\gamma^2(\gamma^{\ell} + \mathbf{r}_{\mathbf{y}}^{j+1}),$$

if $\epsilon' < \gamma^2$ and we have proved the claim. □

**Claim 5.6.4.** *If* $\delta < C(\alpha)^{-\Lambda} \min\{\eta^{\frac{1}{n}} \epsilon^{8n^2}, \eta'\}$, *then*

$$\mathcal{N}^{j+1} := P(\mathbf{x}_a, 20) \setminus P(\mathcal{C}^{j+1}, \mathbf{r}^{j+1})$$

*is an* $(m, k, \epsilon, \eta')$*-neck region.*

*Proof of Claim 5.6.4.* Condition (n1) holds because $\mathcal{N}^j$ is an $(m, k, \epsilon, \eta')$-neck region, and by our construction of $\mathcal{C}^{j+1}$. For any $\mathbf{x} \in \mathcal{C}^j \cap \mathcal{C}^{j+1}$, (n2) and (n3) hold by the fact that $\mathcal{N}^j$ is an $(m, k, \epsilon, \eta')$-neck region. If instead $\mathbf{x} \in \mathcal{C}^{j+1} \setminus \mathcal{C}^j$, then (n2) holds by (j+1.iv). We now prove (n3). Let $\mathbf{x} \in \mathcal{C}^j \cap \mathcal{C}^{j+1}$. Using $\mathbf{x} \in \cup_{c \in C^j}((\mathbf{y}_c + V) \cap P(\mathbf{y}_c, 2\gamma^j))$ and (j.A) we know that

$$|N_{\mathbf{x}}(\gamma^{18}s^2) - m| < C(\alpha)^{\Lambda}\sqrt{\delta}$$

for any $s \in [\gamma^j, s_0]$. We now apply Lemma 3.12(1) with $\mathcal{C} \leftarrow \{\mathbf{x}, \mathbf{x}_a\}$, $\delta \leftarrow C(\alpha)^{\Lambda}\sqrt{\delta}$, $\mathbf{r} \leftarrow \mathbf{r}^{j+1}$, $r_0 \leftarrow \gamma^{-5}$, $s_0 \leftarrow s_0$, $\kappa \leftarrow \gamma^7$, $\mathbf{x} \leftarrow \mathbf{x}$, and $\mathbf{x}_0 \leftarrow \mathbf{x}_a$ so that for any $s \in [\gamma^{j+8}, \gamma^{-5}]$, $u$ is $(k, C^{\Lambda}(\alpha)\sqrt{\delta}, s)$-symmetric at $\mathbf{x}$. We claim that $u$ is not $(k+1, \eta', s)$-symmetric at $\mathbf{x}$. Suppose not. Then we apply Lemma 3.12(1) with $\mathcal{C} \leftarrow \{\mathbf{x}, \mathbf{x}_a\}$, $\delta \leftarrow \eta'$, $\mathbf{r} \leftarrow \mathbf{r}^{j+1}$, $r_0 \leftarrow \gamma^{-5}$, $s_0 \leftarrow s$, $\kappa \leftarrow \gamma^7$, $\mathbf{x}_0 \leftarrow \mathbf{x}$, $\mathbf{x} \leftarrow \mathbf{x}_a$, and $s \leftarrow s_0$ so that $u$ is $(k+1, C^{\Lambda}\eta', s_0)$-symmetric which is a contradiction since $C^{\Lambda}\eta' < \eta$.

We now prove (n4.a). Fix $\mathbf{x} \in \mathcal{C}^{j+1}$ and $s \in [\mathbf{r}_{\mathbf{x}}^{j+1}, \gamma^{-3}]$ such that $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_a, 20)$. Suppose $\mathbf{z} \in \mathcal{C}^{j+1} \cap P(\mathbf{x}, s)$. By (n1), we then have $\mathbf{r}_{\mathbf{z}}^{j+1} \le \gamma^{-2}|\mathbf{x} - \mathbf{z}| \le \gamma^{-2}s$, hence $|N_{\mathbf{z}}(s^2) - m| < \epsilon'$ and $|N_{\mathbf{x}}(s^2) - m| < \epsilon'$ by (j+1.iv). Choose $\ell \le j$ such that $s \in [\gamma^{\ell+1}, \gamma^{\ell}]$. First assume that $\ell \ge 0$. By (j+1.v), there exists $c \in C^{\ell}$ such that $\mathbf{x} \in P(\mathbf{y}_c, 2\gamma^{\ell})$. This, along with $|\mathbf{x} - \mathbf{z}| < s \le \gamma^{\ell}$ and (5.2) implies

$$\mathbf{x}, \mathbf{z} \in \{\mathbf{y} \in P(\mathbf{y}_c, 10\gamma^{\ell}) \colon |N_{\mathbf{y}}(\gamma^{2\ell}) - m| < \epsilon'\}.$$

By ($\ell$.B) with $s \leftarrow \gamma^{\ell}$, and $\zeta \leftarrow C^{\Lambda}\sqrt{\delta}$, we thus have $\mathbf{x}, \mathbf{z} \in P(\mathbf{y}_c + V, C(\alpha)^{\Lambda}\eta^{-\frac{2}{n}}\delta^{\frac{1}{4n^2}}s)$, hence $\mathbf{x} - \mathbf{z} \in P(V, C(\alpha)^{\Lambda}\eta^{-\frac{2}{n}}\delta^{\frac{1}{4n^2}}s)$. If $-3 \le \ell \le -1$, then we can take $\mathbf{y}_c = \mathbf{x}_a$ and the same argument works.

We now prove (n4.b). Again fix $\mathbf{x} \in \mathcal{C}^{j+1}$ and $s \in [\mathbf{r}_{\mathbf{x}}^{j+1}, \gamma^{-3}]$ such that $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_a, 10)$, and now suppose that $\mathbf{z} \in (\mathbf{x} + V) \cap P(\mathbf{x}, s)$. Choose $\ell \le j$ such that $s \in [\gamma^{\ell+1}, \gamma^{\ell}]$. Recall that $\mathbf{x} \in (\mathbf{y}_c + V) \cap P(\mathbf{x}_c, \gamma^j)$ for some $c \in C^j$. Therefore, $\mathbf{z} \in (\mathbf{y}_c + V) \cap P(\mathbf{y}_c, 10\gamma^{\ell})$. Then applying (j.A) with $s \leftarrow \gamma^{\ell}$ yields $|N_{\mathbf{z}}(\gamma^{6+2\ell}) - m| \le C(\alpha)^{\Lambda}\sqrt{\delta}$. Now we use (1.B) with $\zeta \leftarrow C(\alpha)^{\Lambda}\sqrt{\delta}$, $s \leftarrow 1$ to get $\mathbf{z} \in P(V, C(\alpha)^{\Lambda}\eta^{-\frac{1}{n}}\delta^{\frac{1}{2n}})$. From (1.ii) and $|\mathbf{z}| \le 2$, there exists $\mathbf{x}_1 \in \mathcal{C}^1$ such that $\mathbf{z} \in P(\mathbf{x}_1, 3\gamma^3)$. If $P(\mathbf{x}_1, \gamma)$ is not a (c)-ball, then $\mathbf{x}_1 \in \mathcal{C}^{j+1}$ and $\mathbf{r}_{\mathbf{x}_1}^{j+1} = \gamma$, hence $\mathbf{z} \in P(\mathbf{x}_1, \gamma(\mathbf{r}_{\mathbf{x}_1}^{j+1} + s))$ and the claim follows in this case. Suppose instead that $P(\mathbf{x}_1, \gamma)$ is a (c)-ball, so that $\mathbf{x}_1 = \mathbf{x}_{c_1}$ for some $c_1 \in C^1$. Then (1.B) gives $\mathbf{z} \in P(\mathbf{y}_{c_1} + V, C(\alpha)^{\Lambda}\eta^{-\frac{1}{n}}\delta^{\frac{1}{2n}}\gamma) \cap P(\mathbf{x}_{c_1}, \gamma)$, hence by (2.ii), there exists $\mathbf{x}_2 \in \mathcal{C}^2$ such that $\mathbf{z} \in P(\mathbf{x}_2, 3\gamma^2\mathbf{r}_{\mathbf{x}_2}^2)$. If $P(\mathbf{x}_2, \gamma^2)$ is not a (c)-ball, then $\mathbf{x}_2 \in \mathcal{C}^{j+1}$ and $\mathbf{r}_{\mathbf{x}_2}^{j+1} = \mathbf{r}_{\mathbf{x}_2}^2 = \gamma^2$, hence $\mathbf{z} \in P(\mathbf{x}_2, \gamma(s + \mathbf{r}_{\mathbf{x}_2}^{j+1}))$, and claim follows in this case. We can repeat this argument to obtain $\mathbf{x}_{\ell} \in \mathcal{C}^{\ell}$ such that $\mathbf{z} \in P(\mathbf{x}_{\ell}, 3\gamma^2\mathbf{r}_{\mathbf{x}_{\ell}}^{\ell})$. If $\mathbf{x}_{\ell} \in \mathcal{C}^{j+1}$, then we are done by reasoning as above. Otherwise, we have $\mathbf{x}_{\ell} \in \{\mathbf{x}_{c'}\}_{c' \in C^{\ell}}$, hence $\mathbf{r}_{\mathbf{x}_{\ell}}^{\ell} = \gamma^{\ell}$. By Claim 5.6.3, there exists $\mathbf{y} \in \mathcal{C}^{j+1}$ such that $|\mathbf{x}_{c_{\ell}} - \mathbf{y}| < 2\gamma^2(\gamma^{\ell} + \mathbf{r}_{\mathbf{y}}^{j+1})$. It follows that

$$|\mathbf{z} - \mathbf{y}| \le |\mathbf{z} - \mathbf{x}_{c_{\ell}}| + |\mathbf{x}_{c_{\ell}} - \mathbf{y}| < 3\gamma^2\mathbf{r}_{\mathbf{x}_{\ell}}^{\ell} + 2\gamma^2(\gamma^{\ell} + \mathbf{r}_{\mathbf{y}}^{j+1}) \le 3\gamma(s + \mathbf{r}_{\mathbf{y}}^{j+1}). \tag{5.3}$$

Thus (n4.b) holds. □

For any $\mathbf{x} \in \mathcal{C}^\ell$, we have $|\mathbf{x} - \mathbf{x}_c| < \gamma^{\ell-1}$ for some $c \in C^{\ell-1}$, so that

$$\mathcal{C}^\ell \subseteq P(\mathcal{C}^{\ell-1}, \gamma^{\ell-1}).$$

By Lemma 5.8, there is a closed subset $\mathcal{C} \subseteq \overline{P}(\mathbf{x}_a, 20)$ such that $\lim_{\ell\to\infty} d_{\mathcal{P},H}(\mathcal{C}, \mathcal{C}^\ell) = 0$. For $\mathbf{x} \in \mathcal{C} \setminus \cup_\ell \mathcal{C}^\ell$, set $\mathbf{r}_\mathbf{x} := 0$, and set $\mathcal{N} := P(\mathbf{x}_a, 20) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$. Setting $B := \cup_\ell B^\ell$, $D := \cup_\ell D^\ell$, and $E := \cup_\ell E^\ell$, we thus have

$$P(\mathbf{0}, 1) \subseteq P(\mathcal{C} \setminus \mathcal{C}^\ell, \gamma^\ell) \cup \mathcal{N} \cup \bigcup_{b\in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{d\in D} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e).$$

Taking the intersection over all $\ell \in \mathbb{N}$ gives

$$P(\mathbf{0}, 1) \subseteq (\mathcal{N} \cup \mathcal{C}_0) \cup \bigcup_{b\in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{d\in D} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e).$$

We claim that $\mathcal{N}$ is an $(m, k, 2\epsilon, C(\alpha)^{-\Lambda}\eta)$-neck region of scale 10. For any $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}_+$, there exists $j \in \mathbb{N}$ such that $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}^j$ and $\mathbf{r}_{\mathbf{x}_1} = \mathbf{r}^j_{\mathbf{x}_1}$, $\mathbf{r}_{\mathbf{x}_2} = \mathbf{r}^j_{\mathbf{x}_2}$, so $P(\mathbf{x}_1, \gamma^2 \mathbf{r}_{\mathbf{x}_1}) \cap P(\mathbf{x}_2, \gamma^2 \mathbf{r}_{\mathbf{x}_2}) = \emptyset$ by the fact that $\mathcal{C}^j$ is an $(m, k, \epsilon)$-neck region. Thus (n1) holds. To prove (n2), (n3), it suffices to consider the case where $\mathbf{x} \in \mathcal{C}_0$. In this case, there is a sequence $\mathbf{x}_{c_j}$ with $c_j \in C^j$, $\mathbf{x}_{c_j} \to \mathbf{x}$. For any $s \in (0, \gamma^{-3}]$, $\mathbf{x}_{c_j} \in \mathcal{C}^j$ implies that for sufficiently large $j \in \mathbb{N}$, we have $|N_{\mathbf{x}_{c_j}}(s^2) - m| < \epsilon$. By the continuity of $\mathbf{y} \mapsto N_\mathbf{y}(s^2)$, we can therefore take $j \to \infty$ to obtain (n2). Similarly, $\mathbf{x}_{c_j} \in \mathcal{C}^j$ and the continuity of the integrals in Definition 3.5 as a function of basepoint yield (n3).

Suppose $\mathbf{x} \in \mathcal{C}_+$ and $s \in [\mathbf{r}_\mathbf{x}, \gamma^{-3}]$ satisfies $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_a, 20)$. Then $\mathcal{C}^j \cap P(\mathbf{x}, s) \subseteq P(\mathbf{x} + V, \epsilon s)$ for all $j \in \mathbb{N}$, so the Hausdorff convergence $\mathcal{C}^j \to \mathcal{C}$ implies that $\mathcal{C} \cap P(\mathbf{x}, s) \subseteq P(\mathbf{x} + V, \frac{3}{2}\epsilon s)$, hence (n4.a) holds with $\epsilon$ replaced by $\frac{3}{2}\epsilon$. The Hausdorff convergence $\mathcal{C}^j \to \mathcal{C}$ gives $\mathcal{C}^j \subseteq P(\mathcal{C}, \frac{1}{8}\gamma s)$ for sufficiently large $j \in \mathbb{N}$, so that if we choose $j$ large enough such that $\gamma^j \le \frac{1}{8}\gamma s$, then

$$(\mathbf{x} + V) \cap P(\mathbf{x}, s) \subseteq \bigcup_{\mathbf{y}\in\mathcal{C}^j} P(\mathbf{y}, 3\gamma(s + \mathbf{r}^j_\mathbf{y})) \subseteq \bigcup_{\mathbf{y}\in\mathcal{C}} P(\mathbf{y}, 10\gamma(s + \mathbf{r}_\mathbf{y})),$$

where we used (5.3) for the first inclusion, and for the last inclusion we used that $\mathbf{r}^j_\mathbf{y} = \mathbf{r}_\mathbf{y}$ for $\mathbf{y} \in \mathcal{C}^j \cap \mathcal{C}$, and $\mathbf{r}_\mathbf{y} \le \gamma^j \le \frac{\gamma}{4}s$ for any $\mathbf{y} \in \mathcal{C}^j \setminus \mathcal{C}$. Thus (n4.b) holds for all $\mathbf{x} \in \mathcal{C}_+$.

Suppose instead $\mathbf{x} \in \mathcal{C}_0$, and fix $s \in (0, \gamma^{-3}]$. As before, choose $c_i \in C^i$ such that $\mathbf{x}_{c_i} \to \mathbf{x}$ as $i \to \infty$. Then the arguments of the previous paragraph and the Hausdorff convergence $\mathcal{C}^j \to \mathcal{C}$ give

$$\mathcal{C} \cap P(\mathbf{x}, s) \subseteq P(\mathcal{C}^j \cap P(\mathbf{x}_{c_i}, s), 10^{-2}\epsilon s) \subseteq P(\mathbf{x}_c + V, \tfrac{5}{3}\epsilon s) \subseteq P(\mathbf{x} + V, 2\epsilon s),$$

hence (n4.a) holds, and a similar argument yields (n4.b).

The $k$-dimensional content estimate then follows from Theorem 4.6.

(2) Choose $\ell \in \mathbb{N}_0$ such that $\gamma^{\ell+1} \le r_* \le \gamma^\ell$. Set

$$\mathcal{C} := \mathcal{C}^\ell = \{\mathbf{x}_b\}_{b\in B^\ell} \cup \{\mathbf{x}_c\}_{c\in C^\ell} \cup \{\mathbf{x}_d\}_{d\in D^\ell} \cup \{\mathbf{x}_e\}_{e\in E^\ell}$$

as in the proof of (1). Set $B := B^\ell$, $D := D^\ell$, $E := E^\ell$, and $F := C^\ell$. Note that $r_c = \gamma^\ell$ when $c \in C^\ell$. □

*Remark* 5.7. By applying ($\infty$.iv), we can assume that $|N_\mathbf{x}(s^2) - m| \le C(\alpha)^\Lambda \eta^{-\frac{1}{n}} \delta^{\frac{1}{2n}}$ for all $\mathbf{x} \in \mathcal{C}$ and $s \in [10^3\gamma^{-1}\delta \mathbf{r}_\mathbf{x}, \gamma^{-10} r]$. We also observe that $\mathbf{r}_\bullet \le \gamma$ from our construction.

In the proof of Proposition 5.5, we used the result concerning Hausdorff convergence of almost-monotone sequences of sets.

**Lemma 5.8.** *Suppose $(X, d)$ is a locally compact metric space, and $(\mathcal{C}_j)$ is a collection of closed subsets of $X$ satisfying*

$$\mathcal{C}_{j+1} \subseteq B(\mathcal{C}_j, 2^{-j}).$$

*Define*

$$\mathcal{C} := \bigcap_{k} \bigcup_{j\geq k} B(\mathcal{C}_j, 2^{-j}),$$

*so that, for any $x \in \mathcal{C}$, there exists a sequence $x_i \in \mathcal{C}_{j(i)}$ with $\lim_{i\to\infty} j(i) = \infty$ and $d(x, x_i) < 2^{-j(i)}$. Then $\mathcal{C}_j \to \mathcal{C}$ in the Hausdorff sense.*

*Proof.* We need to show that $\mathcal{C}_{k+1} \subseteq B(\mathcal{C}, \epsilon)$ and $\mathcal{C} \subseteq B(\mathcal{C}_k, \epsilon)$ for sufficiently large $k \in \mathbb{N}$. For all $j \geq k$, we have $\mathcal{C}_j \subseteq B(\mathcal{C}_{j-1}, 2^{-(j-1)}) \subseteq B(\mathcal{C}_{j-2}, 2^{-(j-2)} + 2^{-(j-1)}) \subseteq \cdots \subseteq B(\mathcal{C}_k, 2^{-k})$, hence $\bigcup_{j\geq k} B(\mathcal{C}_j, 2^{-j}) \subseteq B(\mathcal{C}_k, 2^{-(k-2)})$ for all $k \in \mathbb{N}$. In particular, for all $k \in \mathbb{N}$ satisfying $2^{-(k-2)} < \epsilon$, we have

$$\mathcal{C} \subseteq \bigcup_{j\geq k} B(\mathcal{C}_j, 2^{-j}) \subseteq B(\mathcal{C}_k, \epsilon).$$

Because $\left(\bigcup_{j\geq k} \overline{B}(\mathcal{C}_j, 2^{-j})\right)_{k\in\mathbb{N}}$ is a decreasing sequence of compact sets whose intersection is $\mathcal{C}$, it follows that

$$\lim_{k\to\infty} d_H\left(\mathcal{C}, \bigcup_{j\geq k} \overline{B}(\mathcal{C}_j, 2^{-j})\right) = 0.$$

That is, there exists $k_0 \in \mathbb{N}$ such that $\bigcup_{j\geq k} B(\mathcal{C}_j, 2^{-j}) \subseteq B(\mathcal{C}, \epsilon)$ when $k \geq k_0$. In particular, $B(\mathcal{C}_j, 2^{-j}) \subseteq B(\mathcal{C}, \epsilon)$ when $j \geq k_0$. □

5.2. **Refined covering of (c)-balls.** In this section, we show how to refine the neck regions constructed in Proposition 5.5 to arrive at strong neck regions, which is an essential step in applying Theorem 4.14. The strategy for this covering comes from [JN21, Section 7.4].

Our refinement of neck regions requires proving the following $\epsilon$-regularity for the frequency based on the existence of symmetry.

**Lemma 5.9.** *Suppose $u$ is a caloric function with $N(10^5 r^2) \leq \Lambda$. If $u$ is $(k, \epsilon, 2r)$ symmetric at $\mathbf{x}_0$, then*

$$N_{\mathbf{x}_0}(r^2) \leq \begin{cases} 2\epsilon & \text{if } k = n+2, \\ 1 + C^{\Lambda}\epsilon & \text{if } k = n+1. \end{cases}$$

*Proof.* By translating and parabolic rescaling, we assume that $\mathbf{x}_0 = \mathbf{x}_a$ and $r = 1$. If $u$ is $(n+2, \epsilon, 2)$-symmetric at $\mathbf{0}$ then

$$\int_{\mathbb{R}^n} |\nabla u|^2 \, d\nu_{-4} \leq \epsilon \int_{\mathbb{R}^n} u^2 \, d\nu_{-4},$$

which implies that $N(1) \leq N(4) \leq 2\epsilon$.

Now suppose $u$ is $(n+1, \epsilon, 2)$-symmetric with respect to $V = L \times \mathbb{R}$ for some $(n-1)$-plane $L$ which we can suppose to be the span of the first $n-1$ coordinates of $\mathbb{R}^n$. We use the normalization $\int_{\mathbb{R}^n} u^2 \, d\nu_{-1} = 1$. Then

$$\begin{aligned} \int_{\mathbb{R}^n} |\nabla^2 u|^2 \, d\nu_{-1} &\leq \int_{\mathbb{R}^n} |\partial_n \partial_n u|^2 \, d\nu_{-1} + 2\sum_{i\neq n} \int_{\mathbb{R}^n} |\nabla \partial_i u|^2 \, d\nu_{-1} \\ &\leq 2\int_{\mathbb{R}^n} |\Delta u|^2 \, d\nu_{-1} + 4\sum_{i\neq n} \int_{\mathbb{R}^n} |\nabla \partial_i u|^2 \, d\nu_{-1} \\ &\leq C\int_{\mathbb{R}^n} |\partial_t u|^2 \, d\nu_{-4} + C\sum_{i\neq n} \int_{\mathbb{R}^n} |\partial_i u|^2 \, d\nu_{-4} < C^{\Lambda}\epsilon. \end{aligned} \tag{5.4}$$

On the other hand, integrating the $f$-Bochner formula (2.5) and integrating by parts, we get

$$\begin{aligned} C^{\Lambda}\epsilon &> 2\int_{\mathbb{R}^n} |\nabla^2 u|^2 \, d\nu_{-1} = -2\int_{\mathbb{R}^n} \langle \nabla u, \nabla \Delta_f u\rangle \, d\nu_{-1} - \int_{\mathbb{R}^n} |\nabla u|^2 \, d\nu_{-1} \\ &= 2\int_{\mathbb{R}^n} (\Delta_f u)^2 + u\Delta_f u \, d\nu_{-1} + \int_{\mathbb{R}^n} |\nabla u|^2 \, d\nu_{-1} \\ &= 2\int_{\mathbb{R}^n} (\Delta_f u + \frac{1}{2}u)^2 \, d\nu_{-1} - \frac{1}{2}\int_{\mathbb{R}^n} u^2 \, d\nu_{-1} + \frac{N(1)}{2} \\ &\geq \frac{1}{2}(N(1)-1). \end{aligned}$$

□

We now upgrade the decomposition in Proposition 5.5 to include strong neck regions.

**Proposition 5.10** (Strong covering of (c)-balls)**.** *For any $\alpha, \epsilon \in (0,1]$, $\Lambda > 0$, the following holds if $\delta \leq C(\alpha,\Lambda)^{-1}(\eta\epsilon)^C$ and $\eta \leq C(\alpha)^{-\Lambda}$. Let $u$ be a caloric function satisfying*

$$\sup_{\mathbf{x}\in P(\mathbf{0},20r)} N^u_{\mathbf{x}}(\delta^{-2}r^2) \leq m + \delta, \tag{5.5}$$

*for some integer $m_* \leq m \leq \Lambda$, where $m_*$ is defined in (5.1). Assume that $\mathcal{V} := \mathcal{V}^u_{\delta,m,r}(\mathbf{0})$ is a $(k,\alpha r)$-independent set and that $\mathcal{E}^{k+1,1}_{100r}(\mathbf{y}) > \eta$ for all $\mathbf{y} \in P(\mathbf{0},2r)$. Then there is a decomposition*

$$P(\mathbf{0},r) \subseteq (\mathcal{C}_0 \cup \mathcal{N}) \cup \bigcup_{c\in C} P(\mathbf{x}_c, r_c) \cup \bigcup_{d\in D} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e),$$

*where $\mathcal{N} = P(\mathbf{x}_a, 5r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$ is a strong $(m,k,\epsilon,C(\alpha,\Lambda)^{-1}\eta)$-neck of scale $\frac{5}{2}r$, and $\mathcal{N} = \widehat{\mathcal{N}} \cap P(\mathbf{x}_a, 5r)$, where $\widehat{\mathcal{N}}$ is a strong $(m,k,\epsilon,C(\alpha,\Lambda)^{-1}\eta)$-neck region of scale* 10. *Furthermore,*

$$\mathcal{H}^k_{\mathcal{P}}(\mathcal{C}_0) + \sum_{d\in D} r_d^k + \sum_{e\in E} r_e^k \leq Cr^k, \qquad \sum_{c\in C} r_c^k \leq C\epsilon r^k.$$

*Proof.* After parabolic rescaling, we may assume that $r = 1$. Take $\delta' > 0$ to be determined. Set $\delta = C(\alpha)^{\Lambda}\eta^{\frac{1}{n}}(\delta')^{10n^2}$. Thus Proposition 5.5(1) yields a cover

$$P(\mathbf{0},1) \subseteq (\widetilde{\mathcal{C}}_0 \cup \widetilde{\mathcal{N}}) \cup \bigcup_{b\in \widetilde{B}} P(\mathbf{x}_b, r_b) \cup \bigcup_{d\in \widetilde{D}} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in \widetilde{E}} P(\mathbf{x}_e, r_e),$$

where $\widetilde{\mathcal{N}} = P(\mathbf{x_a}, 20) \setminus \bigcup_{\mathbf{x}\in\widetilde{\mathcal{C}}} \overline{P}(\mathbf{x}, \widetilde{\mathbf{r}}_{\mathbf{x}})$ is a $(m,k,\delta',C(\alpha)^{-\Lambda}\eta)$-neck region of scale 10, and

$$\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}_0) + \sum_{b\in\widetilde{B}} r_b^k + \sum_{d\in\widetilde{D}} r_d^k + \sum_{e\in\widetilde{E}} r_e^k \leq C. \tag{5.6}$$

For $\mathbf{x} \in P(\mathbf{x}_a, 20)$, define

$$\mathbf{r}_{\mathbf{x}} := \begin{cases} \epsilon\gamma^3 \widetilde{\mathbf{r}}_{\mathbf{w}} & \text{if } \mathbf{x} \in P(\mathbf{w}, \gamma^3\widetilde{\mathbf{r}}_{\mathbf{w}}) \text{ for some } \mathbf{w} \in \widetilde{\mathcal{C}}, \\ \epsilon d(\mathbf{x}, \widetilde{\mathcal{C}}) & \text{if } \mathbf{x} \notin \bigcup_{\mathbf{w}\in\widetilde{\mathcal{C}}} P(\mathbf{w}, \gamma^3\widetilde{\mathbf{r}}_{\mathbf{w}}). \end{cases}$$

If $\mathbf{x} \in P(\mathbf{w}, \gamma^3\widetilde{r}_{\mathbf{w}})$ for some $\mathbf{w} \in \widetilde{\mathcal{C}}$, then we use (n1) to deduce that for any $\mathbf{w}' \in \widetilde{\mathcal{C}} \setminus \{\mathbf{w}\}$

$$|\mathbf{x} - \mathbf{w}'| \geq \gamma^2(1-\gamma)\widetilde{\mathbf{r}}_{\mathbf{w}} > \gamma^3\widetilde{\mathbf{r}}_{\mathbf{w}} > |\mathbf{x} - \mathbf{w}|, \qquad |\mathbf{x} - \mathbf{w}'| \geq \gamma^3\widetilde{\mathbf{r}}_{\mathbf{w}'}.$$

As a consequence, $\mathbf{r}_\bullet$ is well-defined and is $\epsilon$-Lipschitz. We also have

$$\mathbf{r}_\bullet \geq \epsilon d_{\mathcal{P}}(\cdot, \widetilde{\mathcal{C}}). \tag{5.7}$$

For each $\mathbf{w} \in P(\mathbf{x}_a, 20)$, we define

$$\widetilde{\mathbf{w}} \in \arg\min_{\widetilde{\mathcal{C}}} d_{\mathcal{P}}(\mathbf{w}, \cdot). \tag{5.8}$$

For $\widetilde{\delta} > 0$ to be determined, define
$$\check{\mathcal{C}} := \left\{ \mathbf{w} \in P(\mathbf{x}_a, 20) \ : \ \mathbf{w} \in P(\widetilde{\mathbf{w}} + V, \widetilde{\delta}^2 \mathbf{r}_{\mathbf{w}}) \cap P(\widetilde{\mathbf{w}}, 2\max\{d_{\mathcal{P}}(\mathbf{w}, \widetilde{\mathcal{C}}), \widetilde{\mathbf{r}}_{\widetilde{\mathbf{w}}}\}) \right\},$$
so that $\widetilde{\mathcal{C}} \subseteq \check{\mathcal{C}}$.

**Claim 5.10.1.** *For all $\mathbf{w} \in \check{\mathcal{C}} \cap P(\mathbf{x}_a, 20)$, we have*
$$\sup_{s \in [\epsilon^{-2}\gamma^{-6}\delta \mathbf{r}_{\mathbf{w}}, \gamma^{-10}]} |N_{\mathbf{w}}(s^2) - m| < C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}}).$$

*Proof.* By $\epsilon$-Lipschitz property of $\mathbf{r}_\bullet$ and (5.7), we get for any $\mathbf{w} \in P(\mathbf{x}_a, 20)$,
$$\widetilde{\mathbf{r}}_{\widetilde{\mathbf{w}}} = \epsilon^{-1}\gamma^{-3}\mathbf{r}_{\widetilde{\mathbf{w}}} \le \epsilon^{-1}\gamma^{-3}(\mathbf{r}_{\mathbf{w}} + \epsilon|\mathbf{w} - \widetilde{\mathbf{w}}|) \le 2\epsilon^{-1}\gamma^{-3}\mathbf{r}_{\mathbf{w}}. \tag{5.9}$$
Thus, Remark 5.7 implies that for any $\widetilde{\mathbf{w}} \in \widetilde{\mathcal{C}} \cap P(\mathbf{x}_a, 20)$
$$\sup_{s \in [2\gamma^{-4}\epsilon^{-1}\delta \mathbf{r}_{\mathbf{w}}, \gamma^{-10}]} |N_{\widetilde{\mathbf{w}}}(s^2) - m| \le C(\alpha)^{\Lambda}\eta^{-\frac{1}{n}}(\delta')^{\frac{1}{2n}}.$$
We apply Lemma 3.12(2) with $\mathcal{C} \leftarrow \{\widetilde{\mathbf{w}}\}$, $r_0 \leftarrow \gamma^{-10}$, $\mathbf{r}_{\widetilde{\mathbf{w}}} \leftarrow \mathbf{r}_{\mathbf{w}}$, $\kappa \leftarrow \epsilon^{-1}\delta\gamma^{-5}$, $s_0 \leftarrow 1$, $s \leftarrow \epsilon^{-1}\mathbf{r}_{\mathbf{w}}$, and $\mathbf{x}_0 \leftarrow \widetilde{\mathbf{w}}$ to obtain
$$(\widetilde{\mathbf{w}} + V) \cap P(\widetilde{\mathbf{w}}, 10\epsilon^{-1}\mathbf{r}_{\mathbf{w}}) \subseteq \{\mathbf{y} \in P(\widetilde{\mathbf{w}}, 10\epsilon^{-1}\mathbf{r}_{\mathbf{w}}) \colon |N_{\mathbf{y}}(\epsilon^{-4}\gamma^{-10}\delta^2\mathbf{r}_{\mathbf{w}}^2) - m| < C(\alpha)^{\Lambda}\eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}}\}.$$
By Lemma 2.6 with $L \leftarrow \mathbb{R}^n$, we can argue as in Theorem 3.6(2) to obtain
$$|N_{\mathbf{w}}(\epsilon^{-4}\gamma^{-12}\delta^2\mathbf{r}_{\mathbf{w}}) - m| < C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}}).$$
Then using (5.5) and the monotonicity of $N$, we obtain the claim. ☐

**Claim 5.10.2.** *For all $\mathbf{x} \in \check{\mathcal{C}} \cap P(\mathbf{x}_a, 20)$, $u$ is $(k, \epsilon, s)$-symmetric at $\mathbf{x}$ but not $(k+1, C^{-\Lambda}\eta, s)$-symmetric for all $s \in [\mathbf{r}_{\mathbf{x}}, \gamma^{-10}]$ if $\widetilde{\delta} \le C(\alpha, \Lambda)^{-1}\epsilon$ and $\delta' \le \eta^2\widetilde{\delta}^{4n}$.*

*Proof.* Fix $\mathbf{x} \in \check{\mathcal{C}}$ and $s \in [\mathbf{r}_{\mathbf{x}}, \gamma^{-3}]$. Taking $\delta \le \gamma^6\epsilon^2$, we can apply Claim 5.10.1 and Lemma 3.12(1) with $\mathcal{C} \leftarrow \check{\mathcal{C}} \cup \{\widetilde{\mathbf{x}}\}$ for some $\mathbf{x}_0 \leftarrow \widetilde{\mathbf{x}} \in \widetilde{\mathcal{C}}$, $r_0 \leftarrow \gamma^{-10}$, $s_0 \leftarrow 1$, $\kappa \leftarrow 1$, $\mathbf{r} \leftarrow \mathbf{r}$, $\delta \leftarrow \max\{\delta', C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}})\}$ and the fact that $u$ is $(k, \delta', 1)$-symmetric at $\widetilde{\mathbf{x}}$, we get that $u$ is $(k, \epsilon, s)$-symmetric at $\mathbf{x}$ if $\max\{\delta', C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}})\} \le C^{-\Lambda}\epsilon$.

Further, we claim that $u$ is not $(k+1, C^{-\Lambda}\eta, s)$-symmetric at $\mathbf{x}$ for any $s \in [\mathbf{r}_{\mathbf{x}}, \gamma^{-10}]$. Suppose not. By Claim 5.10.1, if $\max\{\delta', C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}})\} \le C^{-\Lambda}\eta$, we can use Lemma 3.12(1) $\mathcal{C} \leftarrow \{\mathbf{x}, \widetilde{\mathbf{x}}\}$ with $\mathbf{x}_0 \leftarrow \mathbf{x}$, $s_0 \leftarrow s$, $r_0 \leftarrow \gamma^{-10}$, $\kappa \leftarrow 1$, $\mathbf{r} \leftarrow \mathbf{r}$, $\delta \leftarrow C^{-\Lambda}\eta$ to prove that $u$ is $(k+1, \eta, 1)$-symmetric at $\widetilde{\mathbf{x}}$, which is a contradiction. ☐

Let $\mathcal{C}_+$ be a maximal subset of $\check{\mathcal{C}}$ such that $\widetilde{\mathcal{C}}_+ \subseteq \mathcal{C}_+$ and $\{P(\mathbf{x}, \gamma^2\mathbf{r}_{\mathbf{x}})\}_{\mathbf{x} \in \mathcal{C}_+ \cup \widetilde{\mathcal{C}}_0}$ is pairwise disjoint. Set $\mathcal{C} := \mathcal{C}_+ \cup \widetilde{\mathcal{C}}_0$.

**Claim 5.10.3.** *If $\widetilde{\delta} \le \overline{\overline{\delta}}(\alpha, \epsilon, \eta, \Lambda)$ and $\delta' \le \overline{\delta}'(\alpha, \widetilde{\delta}, \epsilon, \eta, \Lambda)$, then $\mathcal{N} := P(\mathbf{x}_a, 20) \setminus \bigcup_{\mathbf{x} \in \mathcal{C}} \overline{P}(\mathbf{x}, \mathbf{r}_{\mathbf{x}})$ is a strong $(m, k, \epsilon, C^{-\Lambda}\eta)$-neck of scale 10.*

*Proof.* First, (n1) and (n5) hold by construction. Second, since $\mathcal{C} \subset \check{\mathcal{C}}$, (n2) and (n3) hold because of Claim 5.10.1 and Claim 5.10.2 if $\widetilde{\delta} \le C(\alpha, \Lambda)^{-1}\epsilon$ and $\delta' \le \eta^2\widetilde{\delta}^{4n}$.

Now we prove (n4.a). Fix $\mathbf{x} \in \check{\mathcal{C}}$ and $s \in [\mathbf{r}_{\mathbf{x}}, \gamma^{-3}10]$ such that $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_a, 20)$. Now we apply Claim 5.10.1 and Lemma 3.12(3) with $\mathcal{C} \leftarrow \check{\mathcal{C}}$, $\mathbf{x} \leftarrow \mathbf{x}$, $\mathbf{x}_0 \leftarrow \widetilde{\mathbf{x}} \in \widetilde{\mathcal{C}}$, $\delta \leftarrow C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}}(\delta')^{\frac{1}{4n}})$, $\kappa \leftarrow 1$, $s_0 \leftarrow 1$, $r_0 \leftarrow \gamma^{-4}$, $\eta \leftarrow C(\alpha, \Lambda)^{-1}\eta$, $\zeta \leftarrow C(\alpha, \Lambda)(\widetilde{\delta}^{\frac{1}{2}} + \eta^{-\frac{1}{4n}}(\delta')^{\frac{1}{8n}})$, $\mathbf{r}_\bullet \leftarrow \mathbf{r}_\bullet$ to obtain
$$\check{\mathcal{C}} \cap P(\mathbf{x}, s) \subseteq P(\mathbf{x} + V, C(\alpha, \Lambda)\eta^{-\frac{1}{n}}(\widetilde{\delta}^{\frac{1}{2n}} + \eta^{-\frac{1}{4n^2}}(\delta')^{\frac{1}{8n^2}})s).$$
The claim follows if we take $\widetilde{\delta} \le \overline{\overline{\delta}}(\alpha, \epsilon, \eta, \Lambda)$ and $\delta' \le \overline{\delta}'(\alpha, \widetilde{\delta}, \epsilon, \eta, \Lambda)$.

It thus remains to prove (n4.b'). This will follow from proving that for all $\mathbf{x} \in \mathcal{C}$ and $s \in [\mathbf{r_x}, \gamma^{-3}]$ such that $P(\mathbf{x}, s) \subseteq P(\mathbf{x}_a, 20)$, we want to show

$$(\mathbf{x} + V) \cap P(\mathbf{x}, s) \subseteq P(\mathcal{C}, \tfrac{\gamma}{10} s). \tag{5.10}$$

Fix $\mathbf{z} \in (\mathbf{x} + V) \cap P(\mathbf{x}, s)$. Recalling (5.8), if $|\mathbf{z} - \widetilde{\mathbf{z}}| < \frac{\gamma}{10} s$, then because $\widetilde{\mathcal{C}} \subseteq \mathcal{C}$, we are done. Suppose instead $|\mathbf{z} - \widetilde{\mathbf{z}}| \geq \frac{\gamma}{10} s$. By Claim 5.10.1 with $\mathbf{w} \leftarrow \mathbf{x}$, we can apply Lemma 3.12(2) with $\mathcal{C} \leftarrow \check{\mathcal{C}}$, $\mathbf{x}_0 \leftarrow \widetilde{\mathbf{x}}$, $\mathbf{x} \leftarrow \mathbf{x}$, $\mathbf{r}_\bullet \leftarrow \mathbf{r}_\bullet$, $r_0 \leftarrow \gamma^{-10}$, $s_0 \leftarrow 1$, $s \leftarrow s$, $\delta \leftarrow C(\alpha, \Lambda)(\widetilde{\delta} + \eta^{-\frac{1}{2n}} (\delta')^{\frac{1}{4n}})$, $\kappa \leftarrow 1$, we have

$$\sup_{s' \in [s, \gamma^{-10}]} |N_{\mathbf{z}}((s')^2) - m| \leq C(\alpha, \Lambda)((\widetilde{\delta})^{\frac{1}{2}} + \eta^{-\frac{1}{4n}} (\delta')^{\frac{1}{8n}}). \tag{5.11}$$

Using (5.9), the $\epsilon$-Lipschitz property of $\mathbf{r}_\bullet$, and $\mathbf{r_x} \leq s$ we get

$$\widetilde{\mathbf{r}}_{\widetilde{\mathbf{z}}} \leq 2\epsilon^{-1}\gamma^{-3}\mathbf{r_z} \leq 2\epsilon^{-1}\gamma^{-3}(\mathbf{r_x} + \epsilon s) \leq 4\gamma\epsilon^{-1}\gamma^{-3} s. \tag{5.12}$$

Thus, we can apply Lemma 3.12(3) with $\mathcal{C} \leftarrow \widetilde{\mathcal{C}}$, $\mathbf{r}_\bullet \leftarrow \widetilde{\mathbf{r}}_\bullet$, $\eta \leftarrow C(\alpha, \Lambda)^{-1}\eta$, $\mathbf{x}_0 \leftarrow \widetilde{\mathbf{z}}$, $\mathbf{x} \leftarrow \widetilde{\mathbf{z}}$, $r_0 \leftarrow \gamma^{-10}$, $s_0 \leftarrow 1$, $\delta \leftarrow C(\alpha, \Lambda)\eta^{-\frac{1}{n}} (\delta')^{\frac{1}{2n}}$ $\kappa \leftarrow 1$, $s \leftarrow 4\epsilon^{-1}\gamma^{-3} s$, and $\zeta \leftarrow C(\alpha, \Lambda)((\widetilde{\delta})^{\frac{1}{2}} + \eta^{-\frac{1}{4n}} (\delta')^{\frac{1}{8n}})$ to obtain

$$\begin{aligned} &\{\mathbf{y} \in P(\widetilde{\mathbf{z}}, 40\epsilon^{-1}\gamma^{-3}s) \colon \mathcal{E}_{400\gamma^{-3}\epsilon^{-1}s}(\mathbf{y}) < C(\alpha, \Lambda)((\widetilde{\delta})^{\frac{1}{2}} + \eta^{-\frac{1}{4n}} (\delta')^{\frac{1}{8n}})\} \\ &\qquad \subseteq P(\widetilde{\mathbf{z}} + V, C(\alpha, \Lambda)\eta^{-\frac{1}{n}}((\widetilde{\delta})^{\frac{1}{2n}} + \eta^{-\frac{1}{4n^2}} (\delta')^{\frac{1}{8n^2}})\epsilon^{-1} s). \end{aligned}$$

Combining this with (5.11) and $|\mathbf{z} - \widetilde{\mathbf{z}}| \leq 2\epsilon^{-1} s$ (which follows from (5.7) and (5.12)), we conclude the existence of $\mathbf{y} \in \widetilde{\mathbf{z}} + V$ such that

$$|\mathbf{y} - \mathbf{z}| \leq C(\alpha, \Lambda)\eta^{-\frac{1}{n}}((\widetilde{\delta})^{\frac{1}{2n}} + \eta^{-\frac{1}{4n^2}} (\delta')^{\frac{1}{8n^2}})\epsilon^{-1} s \leq \gamma^{10} s \tag{5.13}$$

if we choose $\widetilde{\delta} \leq \overline{\overline{\delta}}(\alpha, \epsilon, \eta, \Lambda)$ and then $\delta' \leq \overline{\delta'}(\alpha, \epsilon, \eta, \Lambda)$. We claim that $\mathbf{y} \in \check{\mathcal{C}}$ or equivalently (because $\mathbf{y} \in \widetilde{\mathbf{z}} + V$)

$$d_{\mathcal{P}}(\widetilde{\mathbf{z}}, \widetilde{\mathbf{y}} + V) \leq \widetilde{\delta}^2 \mathbf{r_y}. \tag{5.14}$$

Using our assumption $|\mathbf{z} - \widetilde{\mathbf{z}}| > \frac{\gamma}{10} s$ and recalling (5.7), we estimate

$$s \leq 10\gamma^{-1} d_{\mathcal{P}}(\mathbf{z}, \widetilde{\mathcal{C}}) \leq 10\gamma^{-1}(d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}}) + |\mathbf{y} - \mathbf{z}|) \leq 10\gamma^{-1}\epsilon^{-1}\mathbf{r_y} + 10\gamma^9 s,$$

to obtain $s \leq 20\epsilon^{-1}\gamma^{-1}\mathbf{r_y}$. We thus also have

$$\begin{aligned} |\widetilde{\mathbf{y}} - \widetilde{\mathbf{z}}| &\leq d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}}) + |\mathbf{y} - \mathbf{z}| + d_{\mathcal{P}}(\mathbf{z}, \widetilde{\mathcal{C}}) \leq 2|\mathbf{y} - \mathbf{z}| + 2d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}}) \\ &\leq 2\gamma^{10} s + 2\epsilon^{-1}\mathbf{r_y} \leq 4\epsilon^{-1}\mathbf{r_y}. \end{aligned} \tag{5.15}$$

By (5.9), we have $\mathbf{r_y} \geq \frac{1}{2}\epsilon\gamma^3 \widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}$, so we can apply Lemma 3.12(3) with $\mathcal{C} \leftarrow \widetilde{\mathcal{C}}$, $r_0 \leftarrow \gamma^{-9}$, $\delta \leftarrow \delta'$ $s_0 \leftarrow 1$, $\mathbf{x}_0 \leftarrow \widetilde{\mathbf{y}}$, $\mathbf{x} \leftarrow \widetilde{\mathbf{y}}$, $s \leftarrow 2\epsilon^{-1}\gamma^{-3}\mathbf{r_y}$, $\zeta \leftarrow C(\Lambda)(\delta')^{\frac{1}{2}}$, $\eta \leftarrow C(\alpha, \Lambda)\eta$, $\kappa \leftarrow 1$ to get

$$\widetilde{\mathbf{z}} \in P(\widetilde{\mathbf{y}} + V, C(\alpha, \Lambda)\eta^{-\frac{1}{n}} (\delta')^{\frac{1}{2n}} \epsilon^{-1}\mathbf{r_y}),$$

where we also used (5.15). If we choose $\delta' \leq \overline{\delta'}(\alpha, \widetilde{\delta}, \epsilon, \eta, \Lambda)$, then (5.14) follows, hence $\mathbf{y} \in \check{\mathcal{C}}$. If $\mathbf{y} \in \widetilde{\mathcal{C}}_0 \subseteq \mathcal{C}$, then we are done by using (5.13). Otherwise, $\check{\mathcal{C}} \subseteq \widetilde{\mathcal{C}}_0 \cup \bigcup_{\mathbf{w} \in \mathcal{C}_+} P(\mathbf{w}, 10\gamma^2 \mathbf{r_w})$ gives the existence of $\mathbf{w} \in \mathcal{C}_+$ such that $|\mathbf{y} - \mathbf{w}| < 10\gamma^2 \mathbf{r_w}$. Because

$$\mathbf{r_w} \leq \mathbf{r_x} + \epsilon(|\mathbf{x} - \mathbf{z}| + |\mathbf{z} - \mathbf{y}| + |\mathbf{y} - \mathbf{w}|) \leq s + \epsilon(s + \gamma^{10} s + 10\gamma^2 \mathbf{r_w})$$

we have $\mathbf{r_w} \leq 4s$, so that

$$|\mathbf{z} - \mathbf{w}| \leq |\mathbf{z} - \mathbf{y}| + |\mathbf{y} - \mathbf{w}| \leq \gamma^{10} s + 40\gamma^2 s \leq \frac{\gamma}{10} s.$$

This proves (5.10), and the claim follows. $\square$

Let $\{P(\mathbf{x}_b, r_b)\}_{b\in B}$ be the set of (b)-balls of $\mathcal{N}$. We claim that there are no (b)-balls in the neck decomposition.

**Claim 5.10.4.** $B = \widetilde{B} = \emptyset$.

*Proof.* Suppose by way of contradiction there exists $b \in \widetilde{B} \cup B$. By Theorem 3.6(1), $u$ is $(k+1, C(\alpha,\Lambda)\eta, 100r_b)$-symmetric at some point $\mathbf{y}_b \in P(\mathbf{x}_b, 4r_b)$. By Lemma 2.3 with $\mathbf{x}_1 \leftarrow \mathbf{y}_b$, $\mathbf{x}_0 \leftarrow \mathbf{x}_b$, $\sigma \leftarrow 1$, $\theta \leftarrow \frac{1}{4}$, $r \leftarrow 4r_b$, $\tau \leftarrow 384r_b^2$, $u$ is $(k+1, C(\Lambda,\alpha)\eta, 20r_b)$-symmetric at $\mathbf{x}_b$, where we also used $(1+\frac{1}{4})\tau \le 10^4$ and the monotonicity of the $L^2$ norms. Thus Lemma 5.9 gives

$$N_{\mathbf{x}_b}(r_b^2) \le \begin{cases} 2\eta & \text{if } k = n+1, \\ 1 + C(\alpha,\Lambda)\eta & \text{if } k = n, \end{cases}$$

which contradicts our assumption on $m$ if $\eta \le \overline{\eta}(\Lambda,\alpha)$, in view of (n2). $\square$

Now it remains to prove the content estimate for the $c$-balls appearing in the strong neck region in Claim 5.10.3:

$$\sum_{c\in C} r_c^k \le C\epsilon.$$

First, we prove the containment of $\check{\mathcal{C}}$ inside a neighborhood of $\widetilde{\mathcal{C}}$.

**Claim 5.10.5.** $\check{\mathcal{C}} \cap P(\mathbf{x}_a, 5) \subseteq \widetilde{\mathcal{C}}_0 \cup \bigcup_{\mathbf{z}\in\widetilde{\mathcal{C}}_+} P(\mathbf{z}, 50\gamma\widetilde{\mathbf{r}}_{\mathbf{z}})$.

*Proof.* Suppose by way of contradiction that $\mathbf{y} \in (\check{\mathcal{C}} \cap P(\mathbf{x}_a, 5)) \setminus \left(\widetilde{\mathcal{C}}_0 \cup \bigcup_{\mathbf{z}\in\widetilde{\mathcal{C}}_+} P(\mathbf{z}, 50\gamma\widetilde{\mathbf{r}}_{\mathbf{z}})\right)$.

First, assume that $d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}}) > 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}$. Note that $d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}}) \le 5$. Therefore, $P(\widetilde{\mathbf{y}}, d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}})) \subset P(\mathbf{x}_a, 20)$ so that we can apply (n4.b) to $\widetilde{\mathcal{C}}$ with $s \leftarrow 2|\mathbf{y} - \widetilde{\mathbf{y}}|$ and $\mathbf{x} \leftarrow \widetilde{\mathbf{y}}$ to obtain

$$(\widetilde{\mathbf{y}} + V) \cap P(\widetilde{\mathbf{y}}, 2|\mathbf{y} - \widetilde{\mathbf{y}}|) \subseteq \bigcup_{\mathbf{z}\in\widetilde{\mathcal{C}}} P(\mathbf{z}, 10\gamma(\widetilde{\mathbf{r}}_{\mathbf{z}} + 2|\mathbf{y} - \widetilde{\mathbf{y}}|)).$$

Thus, for some $\mathbf{z} \in \widetilde{\mathcal{C}}$

$$|\mathbf{y} - \mathbf{z}| \le 10\gamma(\widetilde{\mathbf{r}}_{\mathbf{z}} + 2|\mathbf{y} - \widetilde{\mathbf{y}}|) \le 10\gamma(\widetilde{\mathbf{r}}_{\mathbf{z}} + 2|\mathbf{y} - \mathbf{z}|),$$

which implies that $|\mathbf{y} - \mathbf{z}| < 20\gamma\widetilde{\mathbf{r}}_{\mathbf{z}}$ which contradicts $\mathbf{y} \notin P(\mathbf{z}, 50\gamma\widetilde{\mathbf{r}}_{\mathbf{z}})$.

We now instead assume that $d_{\mathcal{P}}(\mathbf{y}, \widetilde{\mathcal{C}}) \le 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}$. Since $2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}} \le 40\gamma$, we have $P(\widetilde{\mathbf{y}}, 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}) \subset P(\mathbf{x}_a, 20)$. Therefore, we can apply (n4.b) to $\widetilde{\mathcal{C}}$ with $\mathbf{x} \leftarrow \widetilde{\mathbf{y}}$ and $s \leftarrow 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}$ to obtain

$$(\widetilde{\mathbf{y}} + V) \cap P(\widetilde{\mathbf{y}}, 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}) \subseteq \bigcup_{\mathbf{z}\in\widetilde{\mathcal{C}}} P(\mathbf{z}, 10\gamma(\widetilde{\mathbf{r}}_{\mathbf{z}} + 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}})).$$

Therefore, because $\mathbf{y} \in \check{\mathcal{C}}$ and $\mathbf{y} \in P(\widetilde{\mathbf{y}}, 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}})$, there exists $\mathbf{z} \in \widetilde{\mathcal{C}}$ such that

$$|\mathbf{y} - \mathbf{z}| \le 10\gamma(\widetilde{\mathbf{r}}_{\mathbf{z}} + 2\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}) + \widetilde{\delta}^2\mathbf{r}_{\mathbf{y}} \le 10\gamma\widetilde{\mathbf{r}}_{\mathbf{z}} + 20\gamma\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}} + \widetilde{\delta}^2(\mathbf{r}_{\widetilde{\mathbf{y}}} + 2\epsilon\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}) \le 10\gamma\widetilde{\mathbf{r}}_{\mathbf{z}} + 30\gamma\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}},$$

where we used $\epsilon$-Lipschitz property of $\mathbf{r}$ in the second inequality, and chose $\widetilde{\delta} \le \gamma\epsilon$. From $\mathbf{y} \notin P(\mathbf{z}, 50\gamma\widetilde{\mathbf{r}}_{\mathbf{z}})$, we have $50\gamma\widetilde{\mathbf{r}}_{\mathbf{z}} < |\mathbf{y} - \mathbf{z}| \le 10\gamma\widetilde{\mathbf{r}}_{\mathbf{z}} + 30\gamma\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}$ hence $\widetilde{\mathbf{r}}_{\mathbf{z}} \le \widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}}$. It follows that

$$|\mathbf{y} - \widetilde{\mathbf{y}}| \le |\mathbf{y} - \mathbf{z}| \le 40\gamma\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}},$$

contradicting $\mathbf{y} \notin P(\widetilde{\mathbf{y}}, 50\gamma\widetilde{\mathbf{r}}_{\widetilde{\mathbf{y}}})$. $\square$

Now we prove that the content of the new $c$-balls intersecting with balls around an old center is small compared to its scale. More precisely, for $\mathbf{z} \in \widetilde{\mathcal{C}}_+$, we define

$$\mathcal{C}_{\mathbf{z}} := \{\mathbf{x} \in \mathcal{C} \cap P(\mathbf{z}, 50\gamma\widetilde{\mathbf{r}}_{\mathbf{z}}) \colon P(\mathbf{x}, \mathbf{r}_{\mathbf{x}}) \text{ is a } c\text{-ball}\}.$$

**Claim 5.10.6.** *For any* $\mathbf{z} \in \widetilde{\mathcal{C}}_+$,

$$\sum_{\mathbf{x} \in \mathcal{C}_{\mathbf{z}}} \mathbf{r}_{\mathbf{x}}^k \leq C\epsilon \widetilde{\mathbf{r}}_{\mathbf{z}}^k.$$

*Proof.* Note that if $P(\mathbf{z}, \widetilde{\mathbf{r}}_{\mathbf{z}})$ is an (e)-ball, then $\mathcal{C}_{\mathbf{z}} = \emptyset$ because $\mathbf{x} \in \mathcal{C}_{\mathbf{z}}$ is a center of a (c)-ball. By Claim 5.10.4, it therefore suffices to consider the case when $P(\mathbf{z}, \widetilde{\mathbf{r}}_{\mathbf{z}})$ is a (d)-ball.

Any $\mathbf{x} \in \mathcal{C}_{\mathbf{z}}$ satisfies $\mathbf{r}_{\mathbf{x}} \leq \mathbf{r}_{\mathbf{z}} + \gamma\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}} \leq 100\epsilon\gamma\widetilde{\mathbf{r}}_{\mathbf{z}}$, so that $P(\mathbf{x}, 4\mathbf{r}_{\mathbf{x}}) \subseteq P(\mathbf{z}, \widetilde{\mathbf{r}}_{\mathbf{z}})$. Because $P(\mathbf{x}, \mathbf{r}_{\mathbf{x}})$ is a (c)-ball, there exists $\mathbf{y} \in \mathcal{V}_{\delta,m,\mathbf{r}_{\mathbf{x}}}(\mathbf{x}) \subseteq \mathcal{V}_{\delta,m,\widetilde{\mathbf{r}}_{\mathbf{z}}}(\mathbf{z}) \cap P(\mathbf{z}, \widetilde{\mathbf{r}}_{\mathbf{z}})$. Because $|\mathbf{x} - \mathbf{y}| \leq 4\mathbf{r}_{\mathbf{x}}$, by triangle inequality and the previous inclusion

$$P(\mathbf{x}, \epsilon\widetilde{\mathbf{r}}_{\mathbf{z}}) \subseteq P(\mathcal{V}_{\delta,m,\widetilde{\mathbf{r}}_{\mathbf{z}}}(\mathbf{z}), 2\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}}). \tag{5.16}$$

Let $\{\mathbf{x}_i\}_{i=1}^M$ be a maximal subset of $\mathcal{C}_{\mathbf{z}}$ such that $|\mathbf{x}_i - \mathbf{x}_j| \geq 2\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}}$. Using (5.16) and that $P(\mathbf{z}, \widetilde{\mathbf{r}}_{\mathbf{z}})$ is a (d)-ball, there exists $W \in \mathrm{Gr}_{\mathcal{P}}(k-1)$ such that

$$\bigcup_{i=1}^M P(\mathbf{x}_i, \epsilon\widetilde{\mathbf{r}}_{\mathbf{z}}) \subseteq P(W, (2\epsilon + \alpha)\widetilde{\mathbf{r}}_{\mathbf{z}}) \cap P(\mathbf{z}, \widetilde{\mathbf{r}}_{\mathbf{z}}).$$

It follows that

$$M(\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}})^{n+2} \leq C(n)(\epsilon + \alpha)^{n+2-k+1}\widetilde{\mathbf{r}}_{\mathbf{z}}^{n+2}.$$

That is, $M \leq C\epsilon^{-k+1}$. Now by Ahlfors regularity (4.10) and the maximality of $\mathbf{x}_i$

$$\sum_{\mathbf{x} \in \mathcal{C}_{\mathbf{z}}} \mathbf{r}_{\mathbf{x}}^k \leq \sum_{i=1}^M \mu(\mathcal{C}_{\mathbf{z}} \cap P(\mathbf{x}_i, 10\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}})) \leq CM(\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}})^k \leq C\epsilon\widetilde{\mathbf{r}}_{\mathbf{z}}^k.$$

□

From (5.10) and Remark 5.7, we may apply Lemma 4.3 to conclude that $\mathcal{N}' := \mathcal{N} \cap P(\mathbf{x}_a, 5)$ is a strong $(m, k, \delta, C(\alpha, \Lambda)\eta)$-neck region. Moreover, setting

$$\mathcal{C}_c := \{\mathbf{x} \in \mathcal{C} \cap P(\mathbf{x}_a, 5) \colon P(\mathbf{x}, \mathbf{r}_{\mathbf{x}}) \text{ is a (c)-ball}\},$$

we can combine Claim 5.10.5, Claim 5.10.6 and (5.6) to obtain

$$\sum_{\mathbf{x} \in \mathcal{C}_c} r_{\mathbf{x}}^k \leq \sum_{\mathbf{z} \in \widetilde{\mathcal{C}}_+} \sum_{\mathbf{x} \in \mathcal{C}_{\mathbf{z}}} \mathbf{r}_{\mathbf{x}}^k \leq C\epsilon \sum_{\mathbf{z} \in \widetilde{\mathcal{C}}_+} \widetilde{\mathbf{r}}_{\mathbf{z}}^k \leq C\epsilon,$$

from which the claim follows. □

5.3. **Covering of (d)-balls.** In this section, we show that (d)-balls can be decomposed into balls of other types, in such a way that the $k$-content of the resulting (c)-balls is small. The proof closely follows [JN21, Section 7.5].

**Proposition 5.11** (Covering of (d)-balls)**.** *If $\alpha \leq \bar{\alpha}$ and $\delta > 0$, then the following statements hold. Let $u$ be a caloric function satisfying*

$$\sup_{\mathbf{x} \in P(\mathbf{0}, 20r)} N_{\mathbf{x}}^u(\gamma^{-5}r^2) \leq m + \delta, \tag{5.17}$$

*for some integer $m \leq \Lambda$. Suppose $\mathcal{V} := \mathcal{V}_{\delta,m,r}(\mathbf{0}) \neq \emptyset$ is not $(k, \alpha r)$-independent. Then the following hold.*

*(1) There exists a decomposition*

$$P(\mathbf{0}, r) \subseteq \bigcup_{b \in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{c \in C} P(\mathbf{x}_c, r_c) \cup \bigcup_{e \in E} P(\mathbf{x}_e, r_e) \cup \mathcal{S}_d,$$

*where* $\mathcal{H}^k_{\mathcal{P}}(\mathcal{S}_d) = 0$, *and we have the* $k$*-dimensional content estimates:*

$$\sum_{c\in C} r_c^k \le C\alpha r^k, \qquad\qquad \sum_{b\in B} r_b^k + \sum_{e\in E} r_e^k \le C\alpha^{k-n-2} r^k.$$

*(2) For any* $r_* \in (0,1]$, *there exists a decomposition*

$$P(\mathbf{0},r) \subseteq \bigcup_{b\in B} P(\mathbf{x}_b,r_b) \cup \bigcup_{c\in C} P(\mathbf{x}_c,r_c) \cup \bigcup_{e\in E} P(\mathbf{x}_e,r_e) \cup \bigcup_{f\in F} P(\mathbf{x}_f,r_f),$$

*where* $r_b, r_c, r_e \ge r_*$, $r_f \in [r_*, 10\alpha^{-1} r_*]$, *and we have the* $k$*-dimensional content estimates:*

$$\sum_{c\in C} r_c^k \le C\alpha r^k, \qquad\qquad \sum_{b\in B} r_b^k + \sum_{e\in E} r_e^k + \sum_{f\in F} r_f^k \le C\alpha^{k-n-2} r^k.$$

*Proof.* (1) By parabolic rescaling, we may assume $r = 1$. Choose a maximal subset $\{\mathbf{x}_i\}_{i=1}^N$ of $P(\mathbf{0},1)$ satisfying $|\mathbf{x}_i - \mathbf{x}_j| \ge \frac{\alpha}{10}$ when $i \ne j$. Then categorize the balls of this cover into types (b)-(e) in order to obtain

$$P(\mathbf{0},1) \subseteq \bigcup_{b\in B^1} P(\mathbf{x}_b,r_b) \cup \bigcup_{c\in C^1} P(\mathbf{x}_c,r_c) \cup \bigcup_{d\in D^1} P(\mathbf{x}_d,r_d) \cup \bigcup_{e\in E^1} P(\mathbf{x}_e,r_e),$$

where $r_b \coloneqq r_c \coloneqq r_d \coloneqq r_e \coloneqq \frac{\alpha}{10}$. A simple covering argument gives

$$\sum_{b\in B^1} r_b^k + \sum_{e\in E^1} r_e^k \le C\alpha^{k-n-2} =: C_1(\alpha).$$

Because $\mathcal{V} \coloneqq \mathcal{V}_{\delta,m,1}(\mathbf{0}) \ne \emptyset$ is not $(k,\alpha)$-independent, there exists $V \in \mathrm{Aff}_{\mathcal{P}}(k-1)$ such that $\mathcal{V} \subseteq P(V,\alpha) \cap P(\mathbf{0},1)$. By the monotonicity of the frequency and (5.17), it follows that $\mathcal{V}_{\delta,m,r_c}(\mathbf{x}_c) \subseteq \mathcal{V}$ and $\mathcal{V}_{\delta,m,r_d}(\mathbf{x}_d) \subseteq \mathcal{V}$. Thus, $P(\mathbf{x}_c,r_c), P(\mathbf{x}_d,r_d) \subseteq P(V,5\alpha) \cap P(\mathbf{0},1)$ for any $c \in C^1$ and $d \in D^1$, which implies

$$|C^1| + |D^1| \le C\alpha^{1-k}.$$

In particular,

$$\sum_{c\in C^1} r_c^k + \sum_{d\in D^1} r_d^k \le C\alpha.$$

Because $P(\mathbf{x}_d,r_d)$ satisfies the same hypotheses for $d \in D^1$, we may apply the above procedure to each such ball, obtaining a cover

$$P(\mathbf{0},1) \subseteq \bigcup_{b\in B^2} P(\mathbf{x}_b,r_b) \cup \bigcup_{c\in C^2} P(\mathbf{x}_c,r_c) \cup \bigcup_{d\in D^2} P(\mathbf{x}_d,r_d) \cup \bigcup_{e\in E^2} P(\mathbf{x}_e,r_e),$$

where $r_b \coloneqq r_c \coloneqq r_d \coloneqq r_e \coloneqq (\frac{\alpha}{10})^2$ for $b \in B^2 \setminus B^1$, $c \in C^2 \setminus C^1$, $e \in E^2 \setminus E^1$, $d \in D^2$. Further, the following $k$-content estimates hold:

$$\sum_{b\in B^2} r_b^k + \sum_{e\in E^2} r_e^k \le C_1(1+C\alpha), \qquad \sum_{c\in C^2} r_c^k \le C\alpha(1+C\alpha), \qquad \sum_{d\in D^2} r_d^k \le (C\alpha)^2.$$

By iterating this procedure, we therefore obtain a cover

$$P(\mathbf{0},1) \subseteq \bigcup_{b\in B^\ell} P(\mathbf{x}_b,r_b) \cup \bigcup_{c\in C^\ell} P(\mathbf{x}_c,r_c) \cup \bigcup_{d\in D^\ell} P(\mathbf{x}_d,(\tfrac{\alpha}{10})^\ell) \cup \bigcup_{e\in E^\ell} P(\mathbf{x}_e,r_e), \tag{5.18}$$

where we have the $k$-content estimates

$$\sum_{b\in B^\ell} r_b^k + \sum_{c\in C^\ell} r_c^k \le 2C_1, \qquad \sum_{c\in C^\ell} r_c^k \le C\alpha \sum_{i=0}^{\ell} (C\alpha)^i, \qquad \sum_{d\in D^\ell} r_d^k \le (C\alpha)^\ell,$$

if we choose $\alpha < (2C)^{-1}$. If we set $B := \cup_\ell B^\ell$, $C := \cup_\ell C^\ell$, $E := \cup_\ell E^\ell$, it follows that

$$P(\mathbf{0},1) \subseteq \bigcup_{b\in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in C} P(\mathbf{x}_c, r_c) \cup \mathcal{S}_d \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e),$$

where we observe that $\mathcal{S}_d := \bigcap_\ell \bigcup_{d\in D^\ell} P(\mathbf{x}_d, (\frac{\alpha}{10})^\ell)$ satisfies $\mathcal{H}^k_{\mathcal{P}}(\mathcal{S}_d) = 0$.

(2) Fix $\ell \in \mathbb{N}_0$ such that $(\frac{\alpha}{10})^{\ell+1} \leq r_* \leq (\frac{\alpha}{10})^\ell$. Set $B := B^\ell$, $C := C^\ell$, $E := E^\ell$ and $F := D^\ell$ in the decomposition (5.18). □

5.4. **Proof of neck decomposition theorems.** We now apply Proposition 5.5 and Proposition 5.11 inductively to obtain the neck decomposition theorems, following [CJN21, Section 10.4].

*Proof of Theorem 5.1.* By Lemma 2.10, we have

$$\sup_{\mathbf{x}\in P(\mathbf{0},2)} N_{\mathbf{x}}(\gamma^{-5}) \leq m \tag{5.19}$$

for some integer $m \leq C\Lambda$. Let $\alpha, \delta > 0$ be the parameters in (c), (d), and (e) balls to be determined. Choose a Vitali cover of $P(\mathbf{0},1)$ by parabolic balls of radius $\gamma^{10}$, and categorize these as balls of type (b)-(e), obtaining a cover

$$P(\mathbf{0},1) \subseteq \bigcup_{b\in B^0} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in C^0} P(\mathbf{x}_c, r_c) \cup \bigcup_{d\in D^0} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E^0} P(\mathbf{x}_e, r_e)$$

with $r_b = r_c = r_d = r_e = \gamma^{10}$, and

$$\sum_{b\in B^0} r_b^k + \sum_{c\in C^0} r_c^k + \sum_{d\in D^0} r_d^k + \sum_{e\in E^0} r_e^k \leq C.$$

Let $C_0$ be the maximum of $C$ and the constants $C$ appearing in Proposition 5.5 and Proposition 5.11.

For each $c \in C^0$, if $\delta \leq C(\alpha)^\Lambda \eta^{\frac{1}{n}} \epsilon^{10n^2}$, we can apply the (c)-ball covering (Proposition 5.5(1)) to $P(\mathbf{x}_c, r_c)$ for each $c \in C^0$ to obtain

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A^1} (\mathcal{N}^a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in B^1} P(\mathbf{x}_b, r_b) \cup \bigcup_{d\in D^1} P(\mathbf{x}_d, r_d) \cup \bigcup_{e\in E^1} P(\mathbf{x}_e, r_e).$$

Here $\mathcal{N}^a$ is an $(m,k,\epsilon,C(\alpha)^{-\Lambda})$-neck region and

$$\sum_{a\in A^1} r_a^k + \sum_{b\in B^1} r_b^k + \sum_{d\in D^1} r_d^k + \sum_{e\in E^1} r_e^k \leq C_0 + C_0 \sum_{c\in C^0} r_c^k \leq 2C_0^2.$$

For each $d \in D^1$, fixing $\alpha \leq \bar{\alpha}$ to be determined later, we can apply the (d)-ball covering (Proposition 5.11(1)) to obtain

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in \overline{A}^1} (\mathcal{N}^a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in \overline{B}^1} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in \overline{C}^1} P(\mathbf{x}_c, r_c) \cup \bigcup_{e\in \overline{E}^1} P(\mathbf{x}_e, r_e) \cup S_0^1,$$

with the $k$-content estimates

$$\sum_{b\in \overline{B}^1} r_b^k + \sum_{e\in \overline{E}^1} r_e^k \leq 2C_0^2 + C_1 \sum_{d\in D^1} r_d^k \leq 2C_0^2(1+C_1), \qquad \sum_{c\in \overline{C}^1} r_c^k \leq C_0\alpha \sum_{d\in D^1} r_d^k \leq 2\alpha C_0^3.$$

Furthermore, $S_0^1 := \cup_{d\in D^1} \mathcal{S}_d$ satisfies $\mathcal{H}^k_{\mathcal{P}}(S_0^1) = 0$, where $C_1 := C\alpha^{k-n-2}$ appearing in Proposition 5.11(1).

We then apply the (c)-ball covering to $P(\mathbf{x}_c, r_c)$ with $c \in \overline{C}^1$, and then apply the (d)-ball covering to $P(\mathbf{x}_d, r_d)$ with $d \in D^2$. After the $\ell$-th iteration, we obtain

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in \overline{A}^\ell} (\mathcal{N}_a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in \overline{B}^\ell} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in \overline{C}^\ell} P(\mathbf{x}_c, r_c) \cup \bigcup_{e\in \overline{E}^\ell} P(\mathbf{x}_e, r_e) \cup S_0^\ell,$$

where $\mathcal{H}^k_{\mathcal{P}}(S_0^\ell) = 0$, and

$$\sum_{a\in \overline{A}^\ell} \left(r_a^k + \mathcal{H}^k_{\mathcal{P}}(\mathcal{C}_{a,0})\right) + \sum_{b\in\overline{B}^\ell} r_b^k + \sum_{e\in\overline{E}^\ell} r_e^k \le 2C_0^2(1+C_1)\sum_{j=0}^{\ell-1}(\alpha C_0^2)^j, \qquad \sum_{c\in\overline{C}^\ell} r_c^k \le 2C_0(\alpha C_0^2)^\ell.$$

Set $A := \cup_\ell \overline{A}^\ell$, $B := \cup_\ell \overline{B}^\ell$, $E := \cup_\ell \overline{E}^\ell$, and

$$\widetilde{\mathcal{C}}_m := \bigcup_\ell S_0^\ell \cup \bigcap_j \bigcup_{\ell\ge j}\left(\bigcup_{c\in\overline{C}^\ell} P(\mathbf{x}_c, r_c)\right), \tag{5.20}$$

so that $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}_m) = 0$ and

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A}(\mathcal{N}_a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in B} P(\mathbf{x}_b, r_b) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e) \cup \widetilde{\mathcal{C}}_m.$$

If $\alpha < \frac{1}{2C_0^2}$, then we have the $k$-content estimate

$$\sum_{a\in A} r_a^k + \sum_{b\in B} r_b^k + \sum_{e\in E} r_e^k \le C(\alpha).$$

Furthermore, by construction, we have

$$\sup_{e\in E}\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N^u_{\mathbf{x}}(\delta^2 r_e^2) < m - \delta.$$

By Lemma 2.11, we thus have

$$\sup_{e\in E}\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N^u_{\mathbf{x}}(\delta^4 r_e^2) < m - 1 + \delta.$$

We can therefore cover each $P(\mathbf{x}_e, r_e)$ by a set $\{P(\mathbf{x}_{e'}, r_{e'})\}_{e'\in E'_e}$ of $C\delta^{-2(n+2)}$ parabolic balls of radius $r_{e'} = \delta^2 r_e$ which satisfy the hypotheses of the (c)-ball and (d)-ball covering lemmas. For each $e' \in E'_e$, we can repeat this entire procedure in order to obtain a cover

$$P(\mathbf{x}_{e'}, r_{e'}) \subseteq \bigcup_{a\in A_{e'}}(\mathcal{N}_a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in B_{e'}} P(\mathbf{x}_b, r_b) \cup \bigcup_{e''\in E_{e'}} P(\mathbf{x}_{e''}, r_{e''}) \cup \widetilde{\mathcal{C}}_{m-1,e'}$$

satisfying the $k$-content estimate

$$\sum_{a\in A_{e'}} r_a^k + \sum_{b\in B_{e'}} r_b^k + \sum_{e''\in E_{e'}} r_{e''}^k \le C\delta^{2k} r_e^k,$$

as well as $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}_{m-1,e'}) = 0$. By taking the union over the $C\delta^{-2(n+2)}$ balls in this cover and then over all $e \in E$, we obtain a cover

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A^{(1)}}(\mathcal{N}_a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in B^{(1)}} P(\mathbf{x}_b, r_b) \cup \bigcup_{e\in E^{(1)}} P(\mathbf{x}_e, r_e) \cup \widetilde{\mathcal{C}}_m \cup \widetilde{\mathcal{C}}_{m-1}$$

satisfying the $k$-content estimate

$$\begin{aligned}\sum_{a\in A^{(1)}} r_a^k + \sum_{b\in B^{(1)}} r_b^k + \sum_{e\in E^{(1)}} r_e^k &\le \sum_{a\in A} r_a^k + \sum_{b\in B} r_b^k + \sum_{e\in E}\sum_{e'\in E'_e}\left(\sum_{a\in A_{e'}} r_a^k + \sum_{b\in B_{e'}} r_b^k + \sum_{e''\in E_{e'}} r_{e''}^k\right)\\ &\le C + C\delta^{-2(n+2-k)}\sum_{e\in E} r_e^k\\ &\le C\delta^{-2(n+2-k)},\end{aligned}$$

as well as $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}_{m-1}) = 0$, and assertions (a), (b), (d). Moreover, for any $e \in E^{(1)}$, we have $\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N^u_{\mathbf{x}}(\delta^2 r_e^2) < m-1-\delta$, so that $\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N^u_{\mathbf{x}}(\delta^4 r_e^2) < m-2+\delta$. Repeating the entire procedure at most $m \le C\Lambda$, we eventually obtain a cover

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A^{(m+1)}} (\mathcal{N}_a \cup \mathcal{C}_{a,0}) \cup \bigcup_{b\in B^{(m+1)}} P(\mathbf{x}_b, r_b) \cup \bigcup_{e\in E^{(m+1)}} P(\mathbf{x}_e, r_e) \cup \widetilde{\mathcal{C}},$$

which satisfies assertions (a), (b), (d). Moreover, for any $e \in E^{(m+1)}$, we have

$$\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N^u_{\mathbf{x}}(\delta^4 r_e^2) < m-(m+1)+\delta < 0,$$

so that in fact $E^{(m+1)} = \emptyset$. By taking $\delta := C(\alpha)^\Lambda \eta^{\frac{1}{n}} \epsilon^{10n^2}$, it follows that the cover satisfies the content estimate

$$\sum_{a\in A^{(m+1)}} r_a^k + \sum_{b\in B^{(m+1)}} r_b^k \le C^\Lambda \delta^{-C\Lambda} \le C^{\Lambda^2} (\eta\epsilon)^{-C\Lambda}.$$

□

*Proof of Theorem 5.2.* We can argue as in the proof of Theorem 5.1 but using Proposition 5.5(2) and Proposition 5.11(2) instead of Proposition 5.5(1) and Proposition 5.11(1). As a result, we obtain a decomposition for each $\ell \in \mathbb{N}$

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A^\ell} \mathcal{N}_a \cup \bigcup_{b\in B^\ell} P(\mathbf{x}_b, r_b) \cup \bigcup_{c\in C^\ell} P(\mathbf{x}_c, r_c) \cup \bigcup_{e\in E^\ell} P(\mathbf{x}_e, r_e) \cup \bigcup_{f\in F^\ell} P(\mathbf{x}_f, r_f),$$

where $\mathcal{N}_a = P(\mathbf{x}_a, r_a) \setminus P(\mathcal{C}_a, \mathbf{r}_\bullet)$ is an $(m,k,\epsilon,C(\alpha)^{-\Lambda}\eta)$-neck region with $\mathbf{r}_\bullet, r_b, r_c, r_e \ge r_*$, and $r_f \in [r_*, (10\alpha^{-1}+\gamma^{-1})r_*]$. Because $r_c \le \left(\gamma + \frac{\alpha}{10}\right)^\ell$ for any $c \in C^\ell$, we must have $C^\ell = \emptyset$ for sufficiently large $\ell \in \mathbb{N}$. Fix such $\ell$ and set $A := A^\ell$, $B = B^\ell$, $E = E^\ell$, and $F = F^\ell$. We then have

$$\sum_{a\in A} r_a^k + \sum_{b\in B} r_b^k + \sum_{e\in E} r_e^k + \sum_{f\in F} r_f^k \le C(\alpha).$$

Now we proceed as in the proof of Theorem 5.1 to cover (e)-balls for $e \in E$ with radius $\delta^2 r_e$. If $\delta^2 r_e \le r_* \le r_e$ such (e)-balls are considered as $f$-balls. The remaining (e)-balls can be decomposed as above. We repeat this procedure at most $m+1 \le C\Lambda$ times. □

*Proof of Theorem 5.3.* We argue as in the proof of Theorem 5.1 but using Proposition 5.10 instead of instead of Proposition 5.5(1) and we ask $m \ge m_*$ in Proposition 5.11(1). Note that as in Claim 5.10.4, $m \ge m_*$ guarantees that there are no (b)-balls in the decomposition arising from Proposition 5.11(1). Thus

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A_m} (\mathcal{N}_a \cup \mathcal{C}_{a,0}) \cup \bigcup_{e\in E_m} P(\mathbf{x}_e, r_e) \cup \widetilde{\mathcal{C}}_m,$$

where $\mathcal{N}_a = P(\mathbf{x}_a, r_a) \setminus P(\mathcal{C}_a, \mathbf{r}_\bullet)$ are strong $(m,k,\epsilon,C(\alpha,\Lambda)\eta)$-neck regions, $P(\mathbf{x}_e, r_e)$ are (e)-balls, and $\widetilde{\mathcal{C}}_m$ is the residual set, which satisfies $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}_m) = 0$ and arises as in (5.20) when we apply Proposition 5.5(1) and Proposition 5.11(1) iteratively. Moreover, we have the $k$-content estimate

$$\sum_{a\in A_m} r_a^k + \sum_{e\in E} r_e^k \le C(\Lambda)$$

if $\epsilon \le \overline{\epsilon}$, $\alpha \le \overline{\alpha}$. Now we proceed as in the proof of Theorem 5.1 to cover (e)-balls for $e \in E$ with radius $\delta^2 r_e$. We repeat the procedure at most $(m-2) \le C\Lambda$ times so that we are left with a cover

$$P(\mathbf{0},1) \subseteq \bigcup_{a\in A} (\mathcal{N}_a \cup \widetilde{\mathcal{C}}_{a,0}) \cup \bigcup_{e\in E} P(\mathbf{x}_e, r_e) \cup \widetilde{\mathcal{C}},$$

where $\mathcal{N}_a$ are strong $(m_a, k, \epsilon, C(\Lambda)\eta)$-neck regions for some $m_a \geq m_*$,

$$\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N_{\mathbf{x}}(\delta^2 r_e^2) < m_* - \delta.$$

By Lemma 2.11, we thus have

$$\sup_{e\in E}\sup_{\mathbf{x}\in P(\mathbf{x}_e,4r_e)} N^u_{\mathbf{x}}(\delta^4 r_e^2) < m_* - 1 + \delta.$$

For each (e)-ball $P(\mathbf{x}_e, r_e)$, we choose a Vitali cover by $\leq C(\epsilon\delta^2)^{-(n+2)}$ balls $\{P(\mathbf{x}_g, r_g)\}_{g\in G_e}$ of radius $r_g := \delta^2\epsilon r_e$. Set $G := \cup_{e\in E} G_e$. We then have

$$\sum_{a\in A} r_a^k + \sum_{g\in G} r_g^k \leq C(\Lambda)(\eta\epsilon)^{-C\Lambda},$$

so the remaining claim follows from Theorem 4.6. $\square$

## 6. Volume Estimates and Regularity

In this section, we prove our main theorems on the structure of the quantitative nodal and singular sets. The volume estimates will be a consequence of a stronger estimate for quantitative strata defined below, following the strategy of [CJN21, Section 2.4]. Unlike the settings of [CJN21], we must use the finite resolution neck decomposition (Theorem 5.2) to prove Minkowski estimates.

**Definition 6.1** (Quantitative stratum)**.** Given a caloric function $u$, the $(k, \epsilon, r_1, r_2)$ *quantitative stratum* is defined as

$$\mathcal{S}^k_{\epsilon,r_1,r_2} := \mathcal{S}^k_{\epsilon,r_1,r_2}(u) := \Big\{\mathbf{x} \in P(\mathbf{0}, 1)\colon \mathcal{E}^{k+1,1}_r(\mathbf{x}) \geq \epsilon,\ \forall\, r \in [r_1, r_2]\Big\}.$$

We also set $\mathcal{S}^k_{\epsilon,r} := \mathcal{S}^k_{\epsilon,r,1}$, and define $\mathcal{S}^k_\epsilon := \bigcap_{r>0} \mathcal{S}^k_{\epsilon,r}$.

*Remark* 6.2. In light of Theorem 2.13 and Theorem 3.11, the set $\mathcal{S}^k_\epsilon$ roughly consists of points $\mathbf{x} \in P(\mathbf{0}, 1)$ such that the normalized leading term (with respect to parabolic degree) in the Taylor expansion of $u$ at $\mathbf{x}$ is $\epsilon$ away in $L^2(\mathbb{R}^n, d\nu_{-1})$-distance from the subset of caloric polynomials invariant in $k$ spatio-temporal directions (with temporal direction counted as two).

The following theorem ensures that the strata $\mathcal{S}^k_\epsilon$ satisfy uniform estimates on the parabolic $k$-dimensional Minkowski content and are parabolically rectifiable in the following sense.

**Definition 6.3** ([Mat22, Definition 3.7])**.**

1. A set $E \subseteq \mathbb{R}^{n+1}$ is *parabolic* $(k, \ell)$*-rectifiable* if there are $(k, \ell)$-Lipschitz graphs $G_i$, $i \in \mathbb{N}$, such that

   $$\mathcal{H}^k_{\mathcal{P}}\big(E \setminus \cup_i G_i\big) = 0.$$

   A set $E$ is *horizontally* or *vertically parabolic* $(k, \ell)$*-rectifiable* if the graphs are over $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ of the form $W^k \times \{0\}$ or $W^{k-2} \times \mathbb{R}$, respectively.
2. A set $E \subseteq \mathbb{R}^{n+1}$ is *parabolic* $k$*-rectifiable* if $E \subseteq \mathbb{R}^{n+1}$ is parabolic $(k, \ell)$-rectifiable for each $\ell > 0$. *Horizontally* and *vertically parabolic* $k$*-rectifiable* sets are defined analogously.

**Theorem 6.4.** *Let $u$ be a caloric function such that $N(10^5) \leq \Lambda$. Then, for any $k \in \{1, \ldots, n+1\}$, $\epsilon \in (0, 1)$, and $r \in (0, 1]$, we have*

$$\mathcal{H}^{n+2}_{\mathcal{P}}(P(\mathcal{S}^k_{\epsilon,r}, r) \cap P(\mathbf{0}, 1)) \leq C^{\Lambda^2}\epsilon^{-C\Lambda^3} r^{n+2-k}. \tag{6.1}$$

*Moreover, $\mathcal{S}^k_\epsilon$ is parabolic $k$-rectifiable with $\mathcal{H}^k_{\mathcal{P}}(\mathcal{S}^k_\epsilon \cap P(\mathbf{0}, 1)) \leq C(\epsilon)^{\Lambda^4}$. Furthermore, if $k = n+1$ or $n$, then, for $\mathcal{H}^1$-almost every time $t \in (-1, 1)$, $\mathcal{S}^{k,t}_\epsilon := \mathcal{S}^k_\epsilon \cap (\mathbb{R}^n \times \{t\})$ is $(k-2)$-rectifiable with*

$$\mathcal{H}^{k-2}(\mathcal{S}^{k,t}_\epsilon) \leq C^{\Lambda^2}\epsilon^{-C\Lambda^3}, \tag{6.2}$$

*and for $\mathcal{H}^{\frac{1}{2}}$-almost every time $t \in (-1, 1)$, $\mathcal{H}^{k-1}(\mathcal{S}^{k,t}_\epsilon) = 0$.*

We defer the proof to Section 6.3.

6.1. **$\epsilon$-regularity for nodal and singular sets.** In this subsection, using the cone-splitting inequality, we show that the quantitative nodal and singular sets are contained in the quantitative $(n+1)$-stratum and $n$-stratum, respectively. Moreover, they are contained in super-level sets of the frequency function. Therefore, estimating the volume of nodal and singular sets amounts to proving Theorem 6.4. The proof methods in the section are similar to those in [NV17b, Section 3.4].

Recall that $\mathscr{C}^k_r := Z_r$ or $S_r$ and $\mathscr{C}^k := Z$ or $S$ depending on whether $k = n+1$ or $k = n$.

**Proposition 6.5.** *Let $k \in \{n, n+1\}$. Suppose $u$ is a caloric function with $N(10^5 r^2) \le \Lambda$. If $\epsilon \le C^{-\Lambda}$, then, for any $r \in (0,1)$,*

$$\mathscr{C}^k_r(u) \subseteq \mathcal{S}^k_{\epsilon,r}(u). \tag{6.3}$$

In the following two lemmata, we prove *$\epsilon$-regularity* based on the existence of abundant symmetries.

**Lemma 6.6.** *Suppose $u$ is spatially $(n,\epsilon,r)$-symmetric at $\mathbf{x}_0$ with $\epsilon \le \bar{\epsilon}$. Then*

$$\inf_{P(\mathbf{x}_0, \frac{r}{16})} u^2 > \frac{1}{8}\int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_0;t_0-r^2}. \tag{6.4}$$

*Proof.* After a space-time translation and parabolic scaling, we assume that $\mathbf{x}_0 = \mathbf{0}$ and $r = 1$. Additionally, we impose the normalization $\int_{\mathbb{R}^n} u^2 \, d\nu_{-1} = 1$.

Define $\overline{u} := \int_{\mathbb{R}^n} u \, d\nu_{-1}$. Note that $\overline{u}$ is a constant. By applying the interior estimates in Lemma 2.9 and the Poincaré inequality, we have

$$\sup_{\mathbf{x}\in P(\mathbf{0},\frac{1}{16})} |u-\overline{u}|^2 \le C\int_{\mathbb{R}^n} |u-\overline{u}|^2 \, d\nu_{-1} \le C\int_{\mathbb{R}^n} |\nabla u|^2 \, d\nu_{-1} < C\epsilon < \frac{1}{4},$$

provided $\epsilon < \frac{1}{4C}$. Furthermore,

$$\overline{u}^2 = \int_{\mathbb{R}^n} \overline{u}^2 \, d\nu_{-1} = \int_{\mathbb{R}^n} u^2 \, d\nu_{-1} - \int_{\mathbb{R}^n} |u-\overline{u}|^2 \, d\nu_{-1} > \frac{3}{4}.$$

Combining these estimates, we get the conclusion. □

**Lemma 6.7.** *Suppose $u$ is a caloric function with $N(10^5 r^2) \le \Lambda$. If $u$ is $(n+1,\epsilon,2r)$-symmetric at $\mathbf{x}_0$ with $\epsilon \le C^{-\Lambda}$, then*

$$\inf_{P(\mathbf{x}_0,\frac{r}{16})} |u|^2 + 2r^2|\nabla u|^2 > \frac{1}{16}\int_{\mathbb{R}^n} u^2 \, d\nu_{\mathbf{x}_0;t_0-r^2}.$$

*Proof.* After a space-time translation and parabolic scaling, we assume that $\mathbf{x}_0 = \mathbf{0}$, $r = 1$, and we normalize $\int_{\mathbb{R}^n} u^2 \, d\nu_{-1} = 1$. Suppose $u$ is $(n+1,\epsilon,2)$-symmetric with respect to $V = L \times \mathbb{R}$ for some $(n-1)$-plane $L$. We may suppose $L$ to be the span of the first $n-1$ coordinates of $\mathbb{R}^n$.

Set $\overline{u} := \int_{\mathbb{R}^n} u^2 \, d\nu_t$, $\overline{\partial_i u} := \int_{\mathbb{R}^n} \partial_i u \, d\nu_t$, and $\overline{\nabla u} := (\overline{\partial_1 u}, \ldots, \overline{\partial_n u})$. By the Poincaré inequality and the computation in (5.4) we have

$$\int_{\mathbb{R}^n} |\nabla u - \overline{\nabla u}|^2 \, d\nu_{-1} \le \ 2\int_{\mathbb{R}^n} |\nabla^2 u|^2 \, d\nu_{-1} < C^\Lambda \epsilon.$$

Now, using the interior estimates in Lemma 2.9,

$$\sup_{\mathbf{x}\in P(\mathbf{0},\frac{1}{16})} |\nabla u - \overline{\nabla u}|^2 < C^\Lambda \epsilon,$$

$$\sup_{\mathbf{x}\in P(\mathbf{0},\frac{1}{16})} |u - \overline{u} - \overline{\nabla u}\cdot x|^2 \le C\int_{\mathbb{R}^n} |u - \overline{u} - \overline{\nabla u}\cdot x|^2 \, d\nu_{-1} \le C\int_{\mathbb{R}^n} |\nabla u - \overline{\nabla u}|^2 \, d\nu_{-1} < C^\Lambda \epsilon.$$

Applying the Poincaré inequality and the above computations, if $\epsilon \le C^{-\Lambda}$, then

$$\overline{u}^2 + 2|\overline{\nabla u}|^2 = \int_{\mathbb{R}^n} |\overline{u} + \overline{\nabla u}\cdot x|^2\, d\nu_{-1} \ge \frac{1}{2}\int_{\mathbb{R}^n} u^2\, d\nu_{-1} - \int_{\mathbb{R}^n} |u - (\overline{u} + \overline{\nabla u}\cdot x)|^2\, d\nu_{-1} > \frac{7}{16}.$$

It follows that

$$\inf_{\mathbf{x}\in P(\mathbf{0},\frac{1}{16})} |\overline{u} + \overline{\nabla u}\cdot x|^2 + 2|\overline{\nabla u}|^2 \ge \frac{1}{2}\overline{u}^2 + (2 - 2^{-8})|\overline{\nabla u}|^2 > \frac{7}{32}.$$

Therefore, if $\epsilon \le C^{-\Lambda}$,

$$\inf_{\mathbf{x}\in P(\mathbf{0},\frac{1}{16})} |u|^2 + 2|\nabla u|^2 \ge \frac{7}{64} - C\epsilon > \frac{1}{16}.$$

□

*Proof of Proposition 6.5.* Suppose $\mathbf{x} \notin \mathcal{S}^{n+1}_{\epsilon,r}(u)$, so that $\mathcal{E}^{n+2,1}_s(\mathbf{x}) < \epsilon$ for some $s \in [r,1]$. By Theorem 3.6(1), $u$ is $(n+2, C^\Lambda\epsilon, s)$-symmetric at $\mathbf{x}$. By Lemma 6.6, $\mathbf{x} \notin Z_r(u)$ if $\epsilon \le C^{-\Lambda}$. We also have $S_r(u) \subseteq \mathcal{S}^n_{\epsilon,r}(u)$ by the same argument but using Lemma 6.7. □

In the following lemma, we prove a second version of $\epsilon$-regularity that proves the containment of nodal and singular sets in the super-level sets of the frequency function.

**Lemma 6.8.** *Let $k \in \{n, n+1\}$. The following holds if $\epsilon \le C^{-(\Lambda+1)}$. Suppose $u$ is a caloric function with $N_{\mathbf{0}}(10^5) \le \Lambda$. Then, for any $r \in (0,\epsilon]$, we have*

$$\mathscr{C}^k_r(u) \subset \left\{\mathbf{x} \in P(\mathbf{0},1)\colon N_{\mathbf{x}}(\epsilon^{-2}r^2) \ge m_* - \frac{1}{2}\right\},$$

*where $m_*$ is defined as in (5.1).*

*Proof.* By translation, we may assume $\mathbf{x} = \mathbf{0}$. Suppose $N_{\mathbf{0}}(\epsilon^{-2}r^2) < \frac{3}{2}$. By Lemma 2.12, there exists $s \in [4r, \epsilon^{-\frac{1}{20}}r]$ such that $|N_{\mathbf{x}}(4s^2) - N_{\mathbf{x}}(16s^2)| < \epsilon^{\frac{1}{20}}$. By Theorem 2.13 with $\tau_1 \leftarrow \frac{s}{2}$, $\tau_2 \leftarrow 2s$, $\tau \leftarrow s$, $\delta \leftarrow \epsilon^{\frac{1}{20}}$ to obtain

$$s^{2(j+2\ell)}\int_{\mathbb{R}^n} |\partial_t^\ell \nabla^j (u - p_m)|^2\, d\nu_{-s^2} \le \epsilon^{\frac{1}{20}}\int_{\mathbb{R}^n} u^2\, d\nu_{-16s^2} \le C^\Lambda \epsilon^{\frac{1}{20}} \int_{\mathbb{R}^n} u^2 d\nu_{-s^2}$$

for $j + \ell \le 1$, where $p_m \in \mathcal{P}_m$ and $m \le 1$. If $m = 1$, then $p_1$ is $(n+1, 0, s)$-symmetric at $\mathbf{0}$, hence $u$ is $(n+1, C^\Lambda\epsilon^{\frac{1}{20}}, s)$-symmetric at $\mathbf{0}$. By Lemma 6.7, if $\epsilon \le \overline{\epsilon}(\Lambda)$, then $\mathbf{0} \notin S_r(u)$. If $m = 0$, then $p_0$ is $(n+2, 0, s)$-symmetric at $\mathbf{0}$, so $u$ is $(n+2, C^\Lambda\epsilon^{\frac{1}{20}}, s)$-symmetric at $\mathbf{0}$. By Lemma 6.6, we have $\mathbf{0} \notin Z_r(u)$. □

6.2. **Quantitative dimension reduction.** In the following proposition we prove that if $u$ is $k$-symmetric with respect to a plane $V$, then it is $(k+1)$-symmetric at points away from $V$, but only at a smaller scale. We will use the result in Lemma 6.10 in the proof of the volume estimate.

**Proposition 6.9.** *For any $\epsilon, \eta \in (0,1]$, the following holds if $\delta \le \eta^{C\Lambda}\epsilon^{C\Lambda^2}$. Let $u$ be a caloric function such that for some $\mathbf{x}_0 \in \mathbb{R}^{n+1}$ and $r > 0$,*

$$N^u_{\mathbf{x}_0}(10^{10n}r^2) \le \Lambda, \qquad \inf_{\tau\in[10^2r^2, 10^{10n}]} \mathcal{E}_{\sqrt{\tau}}(\mathbf{x}_0) < \delta.$$

*Suppose $u$ is $(k, \delta, 10r)$-symmetric at $\mathbf{x}_0$ with respect to some $V \in \mathrm{Gr}_{\mathcal{P}}(k)$. Then, for any $\mathbf{y} \in P(\mathbf{x}_0, 2r) \setminus P(\mathbf{x}_0 + V, \eta r)$, there exists $\beta = \beta(\mathbf{y}) \ge \eta^2\epsilon^{C\Lambda}$ such that*

$$\mathcal{E}^{k+1,1}_{\beta r}(\mathbf{y}) < \epsilon, \qquad |N^u_{\mathbf{y}}(10^5\beta^2) - N^u_{\mathbf{y}}(10^{-5}\beta^2)| < \epsilon^2.$$

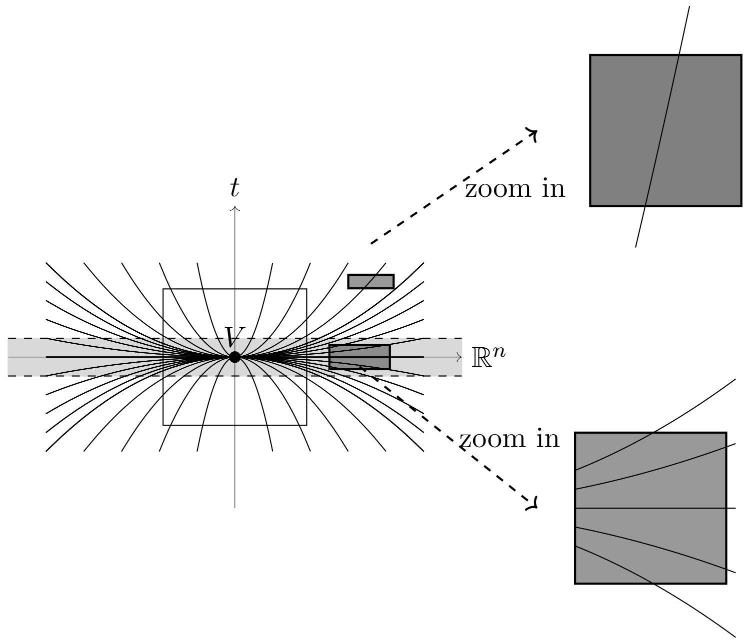


FIGURE 1. Suppose the plane of symmetry is $V = L^k \times \{0\}$. Bottom right: If we are close to the $t = 0$ time slice but away from $V$, then the characteristic curves appear horizontal under parabolic scaling. Top right: If we are away from $t = 0$ and $V$, then the characteristic curves appear vertical under parabolic scaling.

*Proof.* By parabolic rescaling and translation, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $r = 1$. Let $\mathbf{y} = (y, s)$. For simplicity, we assume that $V = L^k \times \{0\}$, as the other case is similar. We may also assume without loss of generality that $\epsilon \leq C^{-\Lambda}$.

First suppose $|s| \leq \eta^2 \epsilon^{C\Lambda}$ so that $|y| \geq \eta$. By Lemma 2.12, there exists $\beta \in [C^{-\Lambda}\eta\epsilon^{4\Lambda+12}, C^{-\Lambda}\eta\epsilon^2]$ such that $N^u_{\mathbf{y}}(10^5\beta^2) - N^u_{\mathbf{y}}(10^{-5}\beta^2) < \epsilon^2$. By Lemma 2.11(3), there exist positive integers $m, m_{\mathbf{y}} \leq C(\Lambda)$ and some $\tau_* \in [10^2, 10^{10n}]$ such that

$$\sup_{s\in[10^{-4}\beta^2, 10^4\beta^2]} |N^u_{\mathbf{y}}(s) - m_{\mathbf{y}}| < 6\epsilon^2, \qquad \sup_{s\in[4^{-1}\tau_*, 4\tau_*]} |N^u(s) - m| < 6\delta^2. \tag{6.5}$$

Set

$$v := 2\tau\Delta_f u + mu, \qquad v_{\mathbf{y}} := 2\tau_{\mathbf{y}}\Delta_{f_{\mathbf{y}}} u + m_{\mathbf{y}} u,$$

where $\tau_{\mathbf{y}} = s - t$ for $t < s$. For any $x \in \mathbb{R}^n$, we then have

$$v(x, s-\beta^2) = 2(\beta^2 - s)\Delta u(x, s-\beta^2) - \nabla u(x, s-\beta^2)\cdot x + mu(x, s-\beta^2),$$
$$v_{\mathbf{y}}(x, s-\beta^2) = 2\beta\Delta u(x, s-\beta^2) - \nabla u(x, s-\beta^2)\cdot(x-y) + m_{\mathbf{y}} u(x, s-\beta^2).$$

Combining equations yields

$$\nabla u(x, s-\beta^2)\cdot y = v_{\mathbf{y}}(x, s-\beta^2) - v(x, s-\beta^2) - 2s\Delta u(x, s-\beta^2) + (m - m_{\mathbf{y}})u(x, s-\beta^2). \tag{6.6}$$

By (6.5) and Lemma 3.8, we have

$$\begin{gathered} \int_{\mathbb{R}^n} v_{\mathbf{y}}^2\, d\nu_{\mathbf{y};s-\beta^2} \leq C^\Lambda \epsilon^2 \int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-\beta^2}, \\ \int_{\mathbb{R}^n} v^2\, d\nu_{-\tau_*} \leq C^\Lambda \delta \int_{\mathbb{R}^n} u^2\, d\nu_{-\tau_*}. \end{gathered} \tag{6.7}$$

By Lemma 2.10, $N_{\mathbf{y}}^u(100) \le C\Lambda$. Using the monotonicity (Lemma 2.5) and comparison of $L^2$-norms at different base points (Lemma 2.3 with $\theta \leftarrow \frac{1}{4}, \tau \leftarrow 24$), we know that

(6.8)
$$\begin{aligned}\int_{\mathbb{R}^n} v^2\, d\nu_{\mathbf{y};s-\beta^2} \le \int_{\mathbb{R}^n} v^2\, d\nu_{\mathbf{y};s-24} \le C\int_{\mathbb{R}^n} v^2\, d\nu_{-100} \le\ C^\Lambda \delta \int_{\mathbb{R}^n} u^2\, d\nu_{-100} \le C^\Lambda\delta \int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-200} \\ \le \left(\frac{C}{\beta}\right)^{C\Lambda} \delta \int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-\beta^2} \le \frac{\delta}{\eta^{C\Lambda}\epsilon^{C\Lambda^2}} \int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-\beta^2}.\end{aligned}$$

On the other hand, by $\beta^4 \ge C^{-\Lambda}\eta^4\epsilon^{20\Lambda}$, $|s| \le \epsilon^{C\Lambda}\eta^2$, and the computation in Lemma 2.9 we know that

$$\int_{\mathbb{R}^n} s^2|\Delta u|^2\, d\nu_{\mathbf{y};s-\beta^2} \le \frac{C^\Lambda s^2}{\beta^4}\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-\beta^2} \le \epsilon^{C\Lambda}\int_{\mathbb{R}^n} u^2 d\nu_{\mathbf{y};s-\beta^2}.$$

Combining this with (6.6), (6.7), and (6.8) yields

$$\begin{aligned}\beta^2 \int_{\mathbb{R}^n} |\nabla u\cdot y|^2\, d\nu_{\mathbf{y};s-\beta^2} &\le C^{-2\Lambda}\epsilon^4\eta^2\left(C^\Lambda\epsilon^2 + \frac{\delta}{\epsilon^{C\Lambda^2}} + \epsilon^{C\Lambda} + 2\Lambda^2\right)\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-\beta^2} \\ &\le C^{-\Lambda}\epsilon^2\eta^2\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-\beta^2}\end{aligned}$$

if $\delta \le \epsilon^{2C\Lambda^2}$. Because $|\pi_L^\perp y| \ge \eta$, it follows that $u$ is $(1, C^{-\Lambda}\epsilon^2, \beta)$-symmetric at $\mathbf{y}$ with respect to $\operatorname{span}(\pi_L^\perp y)$. Because $\partial_t u$ and the components of $\nabla u$ are caloric, we can apply Lemma 2.3 and then Lemma 3.7(1) to conclude that $u$ is also $(k, \epsilon^{-C\Lambda^2}\delta, \beta)$-symmetric with respect to $V$ at $\mathbf{y}$. It follows that $u$ is $(k+1, C^{-\Lambda}\epsilon^2, \beta)$-symmetric with respect to $\operatorname{span}(V, \mathbf{y})$, so by Theorem 3.6(2), we have $\mathcal{E}^{k+1,1}_{\frac{\beta}{10}}(\mathbf{y}) < \epsilon$.

Suppose instead that $|s| \ge \epsilon^{C\Lambda}\eta^2$. Choose $\beta' \in [\epsilon^{4C\Lambda}\eta^2, \epsilon^{3C\Lambda}\eta^2]$ according to Lemma 2.12, so that $N_{\mathbf{y}}(10^5\beta'^2) - N_{\mathbf{y}}(10^{-5}\beta'^2) < \epsilon^2$. By Lemma 2.11(3), there exists a nonnegative integer $m'_{\mathbf{y}} \le C(\Lambda)$ such that $\sup_{s'\in[10^{-4}(\beta')^2, 10^4(\beta')^2]} |N_{\mathbf{y}}^u(s') - m'_{\mathbf{y}}| < \epsilon^2$. By Lemma 3.8, we then have $(x, s-(\beta')^2)$ we have

$$2s\partial_t u(x, s-(\beta')^2) = (v_{\mathbf{y}} - v)(x, s-(\beta')^2) - \nabla u(x, s-(\beta')^2)\cdot y + (m - m'_{\mathbf{y}})u(x, s-(\beta')^2) \tag{6.9}$$

for all $x \in \mathbb{R}^n$. Again appealing to the proof of Lemma 2.9, we get

$$(\beta')^2\int_{\mathbb{R}^n} |\nabla u\cdot y|^2 d\nu_{\mathbf{y};s-(\beta')^2} \le C^\Lambda \int_{\mathbb{R}^n} u^2 d\nu_{\mathbf{y};s-(\beta')^2},$$

so we can combine this with (6.9), (6.7), and (6.8) to obtain

$$\begin{aligned}(\beta')^4\int_{\mathbb{R}^n} |\partial_t u|^2\, d\nu_{\mathbf{y};s-(\beta')^2} &\le \frac{(\beta')^4}{s^2}\left(C^\Lambda\epsilon^2 + \frac{\delta}{\eta^{C\Lambda}\epsilon^{C\Lambda^2}} + \frac{C^\Lambda}{(\beta')^2} + 2\Lambda^2\right)\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-(\beta')^2} \\ &\le C^{-\Lambda}\epsilon^2\int_{\mathbb{R}^n} u^2\, d\nu_{\mathbf{y};s-(\beta')^2}.\end{aligned}$$

We can therefore argue as above to conclude that $u$ is $(k+1, C^{-\Lambda}\epsilon^2, \beta')$-symmetric with respect to $L\times\mathbb{R}$ at $\mathbf{y}$, hence $\mathcal{E}^{k+1,1}_{\beta'}(\mathbf{y}) < \epsilon$. □

### 6.3. **Proof of main theorems in the caloric setting.**

*Proof of Theorem 6.4.* We first prove the content estimate (6.1). Given $\epsilon > 0$, let $\eta \le \bar{\eta}(\Lambda, \epsilon)$ and $\delta \le \bar{\delta}(\Lambda, \epsilon)$ be small constants to be determined later.

By the finite-resolution neck decomposition Theorem 5.2 with $r_* \leftarrow r$, $\epsilon \leftarrow \delta$, and $\eta \leftarrow \eta$, we have

$$P(\mathbf{0},1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_b P(\mathbf{x}_b, r_b) \cup \bigcup_f P(\mathbf{x}_f, r_f), \tag{6.10}$$

where $\mathcal{N}^a = P(\mathbf{x}_a, 2r_a) \setminus \overline{P}(\mathcal{C}^a, \mathbf{r}^a)$ is an $(m_a, k, \delta, C^{-\Lambda}\eta)$-neck region for some $m_a \le C\Lambda$, and $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b) \le \eta$ for some $\mathbf{y}_b \in P(\mathbf{x}_b, 4r_b)$, where we recall that $\gamma = 10^{-10n}$. In addition,

$$\sum_a r_a^k + \sum_b r_b^k + \sum_f r_f^k \le C^{\Lambda^2}(\delta\eta)^{-C\Lambda}, \tag{6.11}$$

$\mathbf{r}^a, r_a, r_b \ge r$ and $r_f \in [r, C^\Lambda \eta^{-\frac{1}{n}} \delta^{-10n^2} r]$.

In the lemmata below, we estimate the contribution from neck regions and (b)-balls with large radius $r_a$ and $r_b$. Because of the multiscale symmetry of a neck region defined in a ball $P(\mathbf{x}_a, r_a)$, we will see that the $k$-th strata won't be far away from the corresponding center set.

**Lemma 6.10.** *For any $\epsilon > 0$, if $m \le \Lambda$, $\delta \le \epsilon^{C\Lambda^2}$, $\chi \le \epsilon^{C\Lambda}$, and $\mathcal{N}^a \subseteq P(\mathbf{x}_a, r_a)$ is an $(m_a, k, \delta, C^{-\Lambda}\eta)$-neck region with $r_a \ge \chi^{-1} r$, then*

$$\mathcal{H}^{n+2}_{\mathcal{P}}\left(P(\mathcal{S}^k_{\epsilon,r} \cap \mathcal{N}^a, r)\right) \le C\epsilon^{-C\Lambda} r_a^k r^{n+2-k}.$$

*Proof.* We first prove that $k$-th strata is not that far from the center set.

**Claim 6.10.1.** *If $\mathbf{y} \in \mathcal{S}^k_{\epsilon,r} \cap \mathcal{N}^a$, then $d_{\mathcal{P}}(\mathbf{y}, \mathcal{C}_a) < \chi^{-1} r$, where $\chi = \epsilon^{C\Lambda}$.*

*Proof.* Let $d := d_{\mathcal{P}}(\mathbf{y}, \mathcal{C}_a) = |\mathbf{x} - \mathbf{y}|$ for some $\mathbf{x} \in \mathcal{C}_a$. By way of contradiction assume $d \ge \chi^{-1} r$. Note that $|\mathbf{y} - \mathbf{x}'| \ge \mathbf{r}_{\mathbf{x}'}$ for any $\mathbf{x}' \in \mathcal{C}$, since $\mathbf{y} \in \mathcal{N}^a$.

Since $d \ge \mathbf{r}_\mathbf{x}$, by (n3), $u$ is $(k, \delta, d)$-symmetric at $\mathbf{x}$ with respect to $V_a \in \mathrm{Gr}(k)$. We claim that $\mathbf{y} \notin P(\mathbf{x} + V_a, \frac{1}{2}d)$. Let $\mathbf{z} \in \mathbf{x} + V_a$ be the closest point to $\mathbf{y}$, and suppose by way of contradiction that $|\mathbf{y} - \mathbf{z}| < \frac{d}{2}$. By (n4), there exists $\mathbf{x}' \in \mathcal{C}^a$ such that $\mathbf{z} \in P(\mathbf{x}', 10\gamma(d + \mathbf{r}_{\mathbf{x}'}))$. Then

$$\mathbf{r}_{\mathbf{x}'} \le |\mathbf{y} - \mathbf{x}'| \le |\mathbf{y} - \mathbf{z}| + |\mathbf{z} - \mathbf{x}'| < \tfrac{1}{2}d + 10\gamma(d + \mathbf{r}_{\mathbf{x}'}),$$

which implies $\mathbf{r}_{\mathbf{x}'} < \frac{3}{4}d$. However,

$$d \le |\mathbf{y} - \mathbf{x}'| < \tfrac{1}{2}d + 10\gamma(d + \mathbf{r}_{\mathbf{x}'}) < \tfrac{1}{2}d + 20\gamma d < d,$$

which is a contradiction. Thus, $\mathbf{y} \notin P(\mathbf{x} + V_a, \frac{1}{2}d)$.

By Proposition 6.9 with $r \leftarrow d$, there is $\underline{\beta} = \epsilon^{C\Lambda}$ and $\beta \ge \underline{\beta}$ such that $\mathcal{E}^{k+1,1}_{\beta d}(\mathbf{y}) < \epsilon$ if $\delta \le \epsilon^{C\Lambda^2}$ which is a contradiction to $\mathbf{y} \in \mathcal{S}^k_{\epsilon,r}$ if we set $\chi = \underline{\beta}$. □

For each $a$, let $\{\mathbf{y}_{a,j}\}_{j=1}^{K_a} \subseteq P(\mathbf{x}_a, r_a)$ be a maximal subset of $\mathcal{C}_a$ with $|\mathbf{y}_{a,i} - \mathbf{y}_{a,j}| \ge 2r$ for $1 \le i \ne j \le K_a$. By Theorem 4.6, if $\delta \le \overline{\delta}$, then there are $V_a \in \mathrm{Gr}_{\mathcal{P}}(k)$ such that the projection $\pi_{V_a} : \mathbb{R}^n \times \mathbb{R} \to V_a$ satisfies $r \le |\pi_{V_a}(\mathbf{y}_{a,i}) - \pi_{V_a}(\mathbf{y}_{a,j})|$ for $1 \le i \ne j \le K_a$, and such that $|\pi_{V_a}(\mathbf{y}_{a,j}) - \pi_{V_a}(\mathbf{x}_a)| \le r_a$ for $1 \le j \le K_a$. Thus,

$$cr^k K_a \le \sum_{j=1}^{K_a} \mathcal{H}^k_{\mathcal{P}}\left(V_a \cap P(\pi_{V_a}(\mathbf{y}_{a,j}), \tfrac{r}{2})\right) \le \mathcal{H}^k_{\mathcal{P}}(V_a \cap P(\pi_{V_a}(\mathbf{x}_a), r_a)) \le Cr_a^k,$$

so that $K_a \le Cr_a^k r^{-k}$. From $\mathcal{C}_a \subseteq \bigcup_{j=1}^{K_a} P(\mathbf{y}_{a,j}, \frac{1}{2}\chi^{-1} r)$ and Claim 6.10.1,

$$\mathcal{S}^k_{\epsilon,r} \cap \mathcal{N}^a \subseteq P(\mathcal{C}^a, \chi^{-1} r) \subseteq \cup_j P(\mathbf{y}_j, \tfrac{3}{2}\chi^{-1} r).$$

Thus,

$$\mathcal{H}^{n+2}_{\mathcal{P}}\left(P(\mathcal{S}^k_{\epsilon,r}\cap\mathcal{N}^a,r)\right)\le\sum_{j=1}^{K_a}\mathcal{H}^{n+2}_{\mathcal{P}}(P(\mathbf{y}_j,2\chi^{-1}r))\le C\chi^{-n-2}r_a^k r^{n+2-k}.$$

□

For the rest of the proof, we set $\delta=\epsilon^{C\Lambda^2}$.

**Lemma 6.11.** *For any $\epsilon\in(0,\frac{1}{10}]$, if $\chi<\epsilon^{C\Lambda}$, $\eta\le C^{-\Lambda}\epsilon^{C\Lambda^2}$, $P(\mathbf{x}_b,r_b)$ is a (b)-ball in the sense that $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b)\le\eta$ for some $\mathbf{y}_b\in P(\mathbf{x}_b,4r_b)$, and $r_b\ge\chi^{-1}r$, then*

$$\mathcal{S}^k_{\epsilon,r}\cap P\left(\mathbf{x}_b,2r_b\right)=\emptyset.$$

*Proof.* It suffices to show that for any $\mathbf{y}\in P(\mathbf{x}_b,4r_b)$, $\mathcal{E}^{k+1,1}_s(\mathbf{y})<\frac{\epsilon}{2}$ for some $s\ge\chi r_b$. By Lemma 2.12, if $\chi<\epsilon^{8\Lambda+20}$, then

$$N_{\mathbf{y}}(\epsilon^{-4}s^2)-N_{\mathbf{y}}(\epsilon^4 s^2)<\epsilon^2 \tag{6.12}$$

for some $s\in[\frac{\chi r_b}{100},\frac{r_b}{100}]$. Because $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b)<\eta$, Theorem 3.6(1) implies that $u$ is $(k+1,C^\Lambda\eta,100r_b)$-symmetric at $\mathbf{y}_b\in P(\mathbf{x}_b,2r_b)$. By Lemma 2.3 applied to $\partial_t u$ and components of $\nabla u$, $u$ is $(k+1,C^\Lambda\eta,r_b)$-symmetric at $\mathbf{y}$. Taking $\eta\le C^{-\Lambda}\chi^{2\Lambda}\epsilon^2$, Lemma 3.7(1) implies $u$ is $(k+1,\epsilon^2,100s)$ symmetric at $\mathbf{y}$. This together with (6.12) and Theorem 3.6(2) gives $\mathcal{E}^{k+1,1}_s(\mathbf{y})<\epsilon$ if $\epsilon\le\frac{1}{10}$. □

Fix $\chi=\epsilon^{C\Lambda}$ to be the minimum of the $\chi$ appearing in Lemma 6.10 and the $\chi$ appearing in Lemma 6.11. Thus by Lemma 6.11, we have

$$\mathcal{S}^k_{\epsilon,r}\subset\bigcup_{r_a>\chi^{-1}r}(\mathcal{S}^k_{\epsilon,r}\cap\mathcal{N}^a)\cup\bigcup_{r_a\le\chi^{-1}r}P(\mathbf{x}_a,r_a)\cup\bigcup_{r_b\le\chi^{-1}r}P(\mathbf{x}_b,r_b)\cup\bigcup_f P(\mathbf{x}_f,r_f),$$

so that

$$\begin{aligned}&P(\mathcal{S}^k_{\epsilon,r},r)\\&\subset\bigcup_{r_a>\chi^{-1}r}P(\mathcal{S}^k_{\epsilon,r}\cap\mathcal{N}^a,r)\cup\bigcup_{r_a\le\chi^{-1}r}P(\mathbf{x}_a,r_a+r)\cup\bigcup_{r_b\le\chi^{-1}r}P(\mathbf{x}_b,r_b+r)\cup\bigcup_f P(\mathbf{x}_f,r_f+r).\end{aligned}$$

By Lemma 6.10 and (6.11), we have

$$\mathcal{H}^{n+2}_{\mathcal{P}}\left(\bigcup_{r_a>\chi^{-1}r}P(\mathcal{S}^k_{\epsilon,r}\cap\mathcal{N}^a,r)\right)\le\sum_{r_a>\chi^{-1}r}C\epsilon^{-C\Lambda}r^{n+2-k}\sum_a r_a^k\le C^{\Lambda^2}(\delta\eta)^{-C\Lambda}r^{n+2-k}.$$

By construction, we have $r_a,r_b\ge r$ and $r_f\in[r,C^\Lambda\eta^{-\frac{1}{n}}\delta^{-10n^2}r]$, so that Theorem 5.2(c) yields

$$\begin{aligned}&\mathcal{H}^{n+2}_{\mathcal{P}}\left(\bigcup_{r_a\le\chi^{-1}r}P(\mathbf{x}_a,r_a+r)\cup\bigcup_{r_b\le\chi^{-1}r}P(\mathbf{x}_b,r_b+r)\cup\bigcup_{f\in F}P(\mathbf{x}_f,r_f+r)\right)\\&\le C\left(\sum_{r_a\le\chi^{-1}r}r_a^{n+2}+\sum_{r_b\le\chi^{-1}r}r_b^{n+2}+\sum_f r_f^{n+2}\right)\\&\le Cr^{n+2-k}\left(\chi^{-(n+2-k)}\sum_a r_a^k+\chi^{-(n+2-k)}\sum_b r_b^k+C^\Lambda\delta^{-10n^3}\eta^{-2}\sum_f r_f^k\right)\\&\le C^{\Lambda^2}(\delta\eta)^{-C\Lambda}r^{n+2-k}.\end{aligned}$$

Combining estimates, taking $\eta = C^{-\Lambda}\epsilon^{C\Lambda^2}$ and recalling that $\delta = \epsilon^{C\Lambda^2}$ yields

$$\mathcal{H}^{n+2}_{\mathcal{P}}(P(\mathcal{S}^k_{\epsilon,r}, r)) \leq C^{\Lambda^3}\epsilon^{-C\Lambda^3} r^{n+2-k}.$$

Next, we prove estimates for $\mathcal{S}^k_\epsilon$. By the neck decomposition Theorem 5.1, there exists a decomposition

$$P(\mathbf{0}, 1) \subseteq \bigcup_a \mathcal{N}^a \cup \bigcup_b P(\mathbf{x}_b, r_b) \cup \left(\widetilde{\mathcal{C}} \cup \bigcup_a \mathcal{C}_{a,0}\right), \tag{6.13}$$

where $\mathcal{N}^a = P(\mathbf{x}_a, 2r_a)\backslash\overline{P}(\mathcal{C}^a, \mathbf{r}^a)$ is an $(m_a, k, \delta)$-neck region for some $m_a \leq C\Lambda$, and $\mathcal{E}^{k+1,1}_{100r_b}(\mathbf{y}_b) \leq \eta$ some $\mathbf{y}_b \in P(\mathbf{x}_b, 4r_b)$, and

$$\sum_a r_a^k \leq C^{\Lambda^2}(\delta\eta)^{-C\Lambda}, \qquad \mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}) = 0.$$

By Claim 6.10.1 in Lemma 6.10, for any $\mathbf{y} \in \mathcal{S}^k_\epsilon$, by choosing $r = \chi d_{\mathcal{P}}(\mathbf{y}, \mathcal{C}_a)$, we have $\mathbf{y} \in \mathcal{S}^k_{\epsilon,r}$ and thus $\mathbf{y} \notin \mathcal{N}^a$. By Lemma 6.11, $P(\mathbf{x}_b, r_b) \cap \mathcal{S}^k_\epsilon \subseteq P(\mathbf{x}_b, r_b) \cap \mathcal{S}^k_{\epsilon,\chi r_b} = \emptyset$. By the covering (6.13), it follows that

$$\mathcal{S}^k_\epsilon \cap P(\mathbf{0}, 1) \subseteq \widetilde{\mathcal{C}} \cup \bigcup_a \mathcal{C}_{a,0}. \tag{6.14}$$

By Theorem 5.1(c) and Theorem 4.6, we thus have

$$\mathcal{H}^k_{\mathcal{P}}(\mathcal{S}^k_\epsilon \cap P(\mathbf{0}, 1)) \leq \sum_a \mathcal{H}^k_{\mathcal{P}}(\mathcal{C}_{a,0}) \leq C\sum_a r_a^k \leq C^{\Lambda^3}\epsilon^{-C\Lambda^3}.$$

Moreover, each $\mathcal{C}_{a,0}$ is parabolic $(k, \ell)$, so that $\mathcal{S}^k_\epsilon$ is $(k, \ell)$-rectifiable if $C\delta < \ell$. Therefore, $\mathcal{S}^k_\epsilon$ is parabolic $k$-rectifiable. When $k \in \{n, n+1\}$, in the neck decomposition (6.13), all neck regions $\mathcal{N}_a = P(\mathbf{x}_a, 2r_a) \setminus \mathcal{C}_a$ must be vertical. In fact, if $k = n$ and $\mathcal{N}_a$ is spatial, then $u$ is in fact $(n+2, C^\Lambda\delta, r_a)$-symmetric at any $\mathbf{x} \in \mathcal{C}_a$, and thus $P(\mathbf{x}_a, r_a)$ is a (b)-ball, a contradiction. Thus, $\mathcal{S}^k_\epsilon$ is vertically parabolic $k$-rectifiable if $k \in \{n, n+1\}$.

Finally, we prove estimates for time slices assuming $k \in \{n, n+1\}$. For all $t \in (-1, 1)$, the restriction $\mathcal{C}^t_{a,0} \coloneqq \mathcal{C}_{a,0} \cap (\mathbb{R}^n \times \{t\})$ is a $(k-2, \delta)$-graph with $\mathcal{H}^{k-2}(\mathcal{C}^t_{a,0}) \leq Cr_a^{k-2}$. By Lemma 6.12 below, $\mathcal{H}^{k-2}(\widetilde{\mathcal{C}}^t) = 0$ for $\mathcal{H}^1$-a.e. $t \in (-1, 1)$, and $\mathcal{H}^{k-1}(\widetilde{\mathcal{C}}^t) = 0$ for $\mathcal{H}^{\frac{1}{2}}$-a.e. $t \in (-1, 1)$, where $\widetilde{\mathcal{C}}^t \coloneqq \widetilde{\mathcal{C}} \cap (\mathbb{R}^n \times \{t\})$. By intersecting (6.14) with $\mathbb{R}^n \times \{t\}$, we have $\mathcal{S}^{k,t}_\epsilon \subseteq \widetilde{\mathcal{C}}^t \cup_a \mathcal{C}^t_{a,0}$, so that

$$\mathcal{H}^{k-2}(\mathcal{S}^{k,t}_\epsilon) \leq \mathcal{H}^{k-2}(\widetilde{\mathcal{C}}^t) + \sum_a \mathcal{H}^{k-2}(\mathcal{C}^t_{a,0}) \leq C\sum_a r_a^k \leq C^{\Lambda^3}\epsilon^{-C\Lambda^3}$$

and

$$\mathcal{H}^{k-1}(\mathcal{S}^{k,t}_\epsilon) = \mathcal{H}^{k-1}(\widetilde{\mathcal{C}}^t) = 0$$

for $\mathcal{H}^{\frac{1}{2}}$-a.e. $t \in (-1, 1)$. □

In the following lemma, we disintegrate the measure of a subset $E \subset P(\mathbf{0}, 1)$ into its measures on time slices.

**Lemma 6.12.** *If $E \subseteq P(\mathbf{0}, 1)$ is Borel and $\sigma \in [0, \min(\frac{k}{2}, 1)]$, then $t \mapsto \mathcal{H}^{k-2\sigma}(E_t)$ is a measurable $[0, \infty]$-valued function, and*

$$\int_{(-1,1)} \mathcal{H}^{k-2\sigma}(E_t)\, d\mathcal{H}^\sigma(t) \leq C\mathcal{H}^k_{\mathcal{P}}(E),$$

*where $E_t = E \cap (\mathbb{R}^n \times \{t\})$, and $\mathcal{H}^j$ denotes the $j$-dimensional Hausdorff measure with respect to the Euclidean metric.*

*Proof.* By [Fed14, §2.10.26], the function $t \mapsto \mathcal{H}^{k-2\sigma}(E_t)$ is measurable. Moreover, appealing to [Fed14, Theorem 2.10.25], taking $X \leftarrow (\mathbb{R}^{n+1}, d_{\mathcal{P}})$, $Y \leftarrow (\mathbb{R}, d_{\mathcal{P}})$, and $f \leftarrow \pi$, where $\pi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ is projection to the second factor, and noting that $C^{-1}\mathcal{H}^{2\sigma} \le \mathcal{H}^{\sigma}_{\mathcal{P}} \le C\mathcal{H}^{2\sigma}$ and $\mathcal{H}^{k-2\sigma}_{\mathcal{P}}(E_t) = \mathcal{H}^{k-2\sigma}(E_t)$, the claim follows.

For completeness, we give the proof in the case $\sigma = 1$ along the lines of [Mat95, Theorem 7.7]. The general argument is roughly the same but requires more technical justifications. We may assume $\mathcal{H}^k_{\mathcal{P}}(E) < \infty$. For $j \in \mathbb{N}$, let $\{P(\mathbf{x}_{j,i}, r_{j,i})\}_{i \in I_j}$ be a covering of $E$ such that $r_{j,i} < \frac{1}{j}$ and

$$\omega_k \sum_{i \in I_j} r^k_{j,i} \le \mathcal{H}^k_{\mathcal{P}, \frac{1}{j}}(E) + \frac{1}{j},$$

where $\omega_k$ is the measure of $k$-dimensional sphere. Here, for $\delta > 0$,

$$\mathcal{H}^k_{\mathcal{P},\delta}(E) \coloneqq \inf \sum_i \omega_k r_i^k,$$

where the infimum is over all covering $E \subseteq \cup_i P(\mathbf{x}_i, r_i)$ with $r_i < \delta$. Further, $\mathcal{H}^k_\delta$ is defined similarly but with coverings by Euclidean balls. By Fatou's Lemma,

$$\begin{aligned}
\int_{-1}^1 \mathcal{H}^{k-2}(E_t)\, d\mathcal{H}^1(t) = \int_{-1}^1 \lim_{j\to\infty} \mathcal{H}^{k-2}_{1/j}(E_t)\, d\mathcal{H}^1(t) &\le \liminf_{j\to\infty} \int_{-1}^1 \sum_{i:\ P(\mathbf{x}_{j,i}, r_{j,i})_t \neq \emptyset} \omega_{k-2} r^{k-2}_{j,i}\, d\mathcal{H}^1(t) \\
&\le \liminf_{j\to\infty} \sum_i \int_{t_{j,i}-r^2_{j,i}}^{t_{j,i}+r^2_{j,i}} \omega_{k-2} r^{k-2}_{j,i}\, d\mathcal{H}^1(t) \le \liminf_{j\to\infty} 2\omega_{k-2} \sum_i r^k_{j,i} \\
&\le C\mathcal{H}^k_{\mathcal{P}}(E),
\end{aligned}$$

where $\mathbf{x}_{j,i} = (x_{j,i}, t_{j,i})$. □

*Proof of Theorem 1.8 and Corollary 1.10.* Combine Proposition 6.5 and Theorem 6.4. □

Below, we prove an analog of Theorem 1.1 in the caloric case, which can be proven using real analyticity and Theorem 1.8. We record it here because the proof in the general setting is similar, while real analyticity fails in that generality.

**Theorem 6.13.** *Let $u$ be a caloric function with $N(10^5) \le \Lambda$. For any $k \in \{n, n+1\}$ and $\epsilon > 0$, there exist $\mathbf{x}_i \in P(\mathbf{0}, \frac{1}{4})$, vertical planes $V_i \in \mathrm{Aff}_{\mathcal{P}}(k)$, $r_i \in (0, 1]$, and $\epsilon$-regular parabolic Lipschitz functions $F_i \colon V_i \to V_i^\perp$ such that*

$$\sum_i r_i^k \le C(\epsilon, \Lambda),$$

*and such that $G_i \coloneqq \{\mathbf{x} + F_i(\mathbf{x}) \colon \mathbf{x} \in V_i \cap P(\mathbf{x}_i, 10r_i)\}$ satisfy*

$$\mathcal{H}^k_{\mathcal{P}}\left( (P(\mathbf{0}, 1) \cap \mathscr{C}^k(u)) \setminus \bigcup_i G_i \right) = 0.$$

*Proof.* Given $\epsilon > 0$, fix $\eta \le \overline{\eta}(\Lambda)$ and $\delta \le C(\eta, \Lambda)^{-1}\epsilon^2$ to be determined later. Let

$$P(\mathbf{0}, 1) \subseteq \bigcup_{a \in A} \mathcal{N}^a \cup \bigcup_{g \in G} P(\mathbf{x}_g, r_g) \cup \left( \widetilde{\mathcal{C}} \cup \bigcup_{a \in A} \mathcal{C}_{a,0} \right),$$

be the decomposition from Theorem 5.3 with $\epsilon \leftarrow \delta$ and $\eta \leftarrow \eta$ if $\eta \le C(\Lambda)^{-1}$. Here $\mathcal{N}^a = \widehat{\mathcal{N}}^a \cap P(\mathbf{x}_a, 2r_a)$ where $\widehat{\mathcal{N}}^a$ is the associated strong neck region of scale $2r_a$.

Recall that $\mathscr{C}^k_r \coloneqq Z_r$ or $S_r$ and $\mathscr{C}^k \coloneqq Z$ or $S$ depending on $k = n+1$ or $n$, respectively. Fix some $\epsilon_0 < C^{-\Lambda}$ so that we can apply Proposition 6.5 to conclude that $\mathscr{C}^k(u) = \bigcap_{r>0} \mathscr{C}^k_r(u) \subset \mathcal{S}^k_{\epsilon_0}$. Now we apply Claim 6.10.1 with $\epsilon \leftarrow \epsilon_0$, $\delta \leftarrow \delta$, $\chi \leftarrow \epsilon_0^{C\Lambda}$ to obtain $\mathscr{C}^k(u) \cap \mathcal{N}^a = \emptyset$ for all $a \in A$.

Further, we claim that $\mathscr{C}^k(u) \cap P(\mathbf{x}_g, r_g) = \emptyset$ for all $g \in G$. In fact,

$$\sup_{\mathbf{x}\in P(\mathbf{x}_g,r_g)} N_{\mathbf{x}}(\delta^{-2}r_g^2) \le m_* - 1 + C(\Lambda)^{-1}\delta < m_* - \frac{3}{4},$$

if $\delta \le C^{-\Lambda}$, where $m_*$ is defined in (5.1). However, Lemma 6.8 implies that $\mathbf{x} \in \mathscr{C}^k(u)$ has $N_{\mathbf{x}}(0) \ge m_* - \frac{1}{2}$. Therefore, $\mathscr{C}^k \subset \widetilde{\mathcal{C}} \cup \bigcup_{a\in A} \mathcal{C}_{a,0}$. Recall that $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}}) = 0$. Therefore, the claim about the structure of $\mathscr{C}^k$ follows from Theorem 4.14 applied to $\widehat{\mathcal{N}}^a$ if $\delta \le C(\eta,\Lambda)^{-1}\epsilon^2$. □

## 7. Parabolic Inequalities with Lipschitz Coefficients

We consider solutions to linear second order parabolic inequality in divergence form with Lipschitz leading coefficients:

$$\begin{aligned}
&|\partial_t u - \mathcal{L}u| \le \lambda(|u| + |\nabla u|) && \text{on } P(\mathbf{0}, 10^2\lambda^{-1})),\\
&(1+\lambda)^{-1} I \le a \le (1+\lambda) I, && |a^{ij}(\mathbf{x}) - a^{ij}(\mathbf{y})| \le \lambda|\mathbf{x}-\mathbf{y}|,
\end{aligned}\tag{7.1}$$

where $\mathcal{L}u := \operatorname{div}(a\cdot\nabla u)$, $\lambda \in (0,1]$, and $a^{ij} = a^{ji}$. Although weak solutions to (7.1) belong to $C^{1+\alpha,\frac{1+\alpha}{2}}_{\mathrm{loc}}(P(\mathbf{0},10^2\lambda^{-1}))$ for any $\alpha \in (0,1)$ (see [Lie96, Theorem 4.8]), neither $\mathcal{L}u$ nor $\partial_t u$ are defined pointwise, so by (7.1), we really mean that $(\partial_t - \mathcal{L})u = g$ in the sense of distributions, where $g \in L^\infty(\mathbb{R}^n\times\mathbb{R})$ satisfies $|g| \le \lambda(|u| + |\nabla u|)$ $\mathcal{H}^{n+2}_{\mathcal{P}}$-a.e. on $P(\mathbf{0}, 10^2\lambda^{-1})$.

We also assume that $u$ satisfies the following doubling condition for all $\mathbf{x}_0 = (x_0,t_0) \in P(\mathbf{0},\lambda^{-1})$ and $r \in (0,\lambda^{-1}]$:

$$\int_{B(x_0,4r)\times[t_0-4r^2,t_0]} u^2 d\mathcal{H}^{n+2}_{\mathcal{P}} \le \Lambda \int_{B(x_0,r)} u^2(x,t_0)dx.\tag{7.2}$$

The condition (7.2) rules out non-trivial functions satisfying (7.1) with support contained in a slab $\mathbb{R}^n \times [a,b]$ (see [DSP05]). Furthermore, solutions to (7.1) which satisfy (7.2) have polynomial growth as made precise in Lemma 7.4 below. The polynomial growth enforces the frequency to be bounded at large scale, see Lemma 7.5.

For the rest of this section, we assume that $u$ is a solution to (7.1) satisfying (7.2).

### 7.1. Generalized frequency.

In this subsection, we will define the energy functionals and prove their almost monotonicity as in Section 2.

Given a solution $u$ to (7.1) satisfying (7.2), we first prove that parabolic rescalings at small scales satisfy an improved version of (7.1), in a sense that will be made precise below.

**Definition 7.1.** For any $\mathbf{x} = (x,t) \in P(\mathbf{0},100)$ and $r \in (0,1]$, the *parabolic rescaling* $u_{\mathbf{x};r}$ is defined by

$$u_{\mathbf{x};r}(y,s) := u(x + ra^{-\frac{1}{2}}(\mathbf{x})y, t + r^2 s).$$

*Remark* 7.2. Given $\mathbf{x}_0 = (x_0,t_0)$ and $r$, we define

$$\widetilde{a}^{ij}(x,t) := a^{-\frac{1}{2}}(\mathbf{x}_0) a(x_0 + ra^{\frac{1}{2}}(\mathbf{x}_0)x, t_0 + r^2 t) a^{-\frac{1}{2}}(\mathbf{x}_0)$$

and $\widetilde{\mathcal{L}}v := \operatorname{div}(\widetilde{a}\cdot\nabla v)$. Then

$$\begin{aligned}
|(\partial_t - \widetilde{\mathcal{L}})u_{\mathbf{x}_0;r}|(x,t) &= r^2|(\partial_t - \mathcal{L})u|(x_0 + ra^{\frac{1}{2}}(\mathbf{x}_0)x, t_0 + r^2 t)\\
&\le r^2\lambda(|u| + |a(\mathbf{x}_0)\nabla u|)(x_0 + ra^{\frac{1}{2}}(\mathbf{x}_0)x, t_0 + r^2 t)\\
&\le \lambda r(r|u_{\mathbf{x}_0;r}| + (1+C)|\nabla u_{\mathbf{x}_0;r}|)(x,t)
\end{aligned}$$

for all $(x,t) \in P(\mathbf{0}, 5r^{-1})$. Moreover, we have $\widetilde{a}^{ij}(\mathbf{0}) = \delta^{ij}$, $(1+\lambda)^{-2}I \le \widetilde{a} \le (1+\lambda)^2 I$, as well as

$$|\widetilde{a}(\mathbf{x}) - \widetilde{a}(\mathbf{y})| \le (1+\lambda)\lambda r \max\{|a^{\frac{1}{2}}(\mathbf{x}_0)(x-y)|, |t-s|^{\frac{1}{2}}\} \le 2\lambda r|\mathbf{x}-\mathbf{y}|$$

for all $\mathbf{x} = (x,t)$ and $\mathbf{y} = (y,s)$ in $P(\mathbf{0}, 5r^{-1})$.

Throughout this section, we apply the global computation in Section 2.2 to functions of the form $\chi u$, where $u$ is a local solution to (7.1) satisfying (7.2) and $\chi$ is a cutoff function defined below. Fix $\xi \in C^\infty(\mathbb{R})$ such that $\xi|_{(-\infty,1]} \equiv 1$ and $\xi|_{[2,\infty)} \equiv 0$, as well as $0 \le \xi \le 1$ and $\xi' \le 0$ everywhere. We define $\chi \in C^\infty(\mathbb{R}^n \times (-\infty, 0))$ as

$$\chi(x,t) \coloneqq \xi\left(\frac{\lambda^{\frac{\Upsilon}{2}}|x|}{|t|^{\Upsilon}}\right),$$

where $\Upsilon \in (0, \frac{1}{2})$ is fixed throughout this section.

We will now define the fundamental almost-monotone quantities which generalize Definition 2.4. Note that these quantities are defined locally.

**Definition 7.3.** For $\mathbf{x}_0 \in P(\mathbf{0}, 5)$ and $r \in (0,1]$, we define

$$\widehat{H}^u_{\mathbf{x}_0}(\tau) \coloneqq \int_{\mathbb{R}^n} (\chi u_{\mathbf{x}_0;1})^2 \, d\nu_{-\tau}, \qquad \widehat{E}^u_{\mathbf{x}_0} \coloneqq 2\tau \int_{\mathbb{R}^n} |\nabla(\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\tau}, \qquad \widehat{N}^u_{\mathbf{x}_0}(\tau) \coloneqq \frac{\widehat{E}_{\mathbf{x}_0}(\tau)}{\widehat{H}_{\mathbf{x}_0}(\tau)}.$$

When $\mathbf{x}_0 = \mathbf{0}$ and $u$ is clear, we write $\widehat{H}(\tau) \coloneqq \widehat{H}^u_{\mathbf{x}_0}(\tau)$ and so on.

In the lemma below, we prove a version of an interior estimate similar to Lemma 2.9.

**Lemma 7.4.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda)$. For any $\mathbf{x}_0 = (x_0, t_0) \in P(\mathbf{0}, 10)$, $r \in (0, \frac{1}{C(\Lambda)\lambda})$, and $\tau \in (0, \frac{1}{C(\Lambda)\lambda}]$, we have*

$$\begin{aligned}\sup_{B(x_0,r)} (|u(\cdot, t_0 - \tau)|^2 + \tau|\nabla u(\cdot, t_0-\tau)|^2) &\le C(\Lambda)\left(1 + \frac{r^2}{\tau}\right)^{C(\Lambda)} \int_{B(0,\sqrt{\tau})} u^2(x,-\tau)dx \\ &\le C(\Lambda)\left(1 + \frac{r^2}{\tau}\right)^{C(\Lambda)} \widehat{H}^u_{\mathbf{x}_0}(\tau).\end{aligned}$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$. If $\lambda \le \overline{\lambda}(\Lambda)$, then, for all $r \in (0, C(\Lambda)^{-1}\lambda^{-1}]$, we use (7.2) and [EFV06, Theorem 3(1)] to get

$$\int_{B(0,r)\times[-\tau-r^2,-\tau]} u^2 \, d\mathcal{H}^{n+2}_{\mathfrak{P}} \le C(\Lambda) r^2 \int_{B(0,r)} u^2(x,-\tau)dx. \tag{7.3}$$

By taking $r \leftarrow 2\sqrt{\tau}$ in (7.3) and combining with interior parabolic estimates (see for instance [Lie96, Theorem 4.8]), we obtain

$$\begin{aligned}\sup_{B(\mathbf{0},\sqrt{\tau})} (u^2 + |\nabla u|^2)(\cdot, -\tau) \le& \frac{C}{\tau^{\frac{n+2}{2}}} \int_{B(0,2\sqrt{\tau})\times[-5\tau,-\tau]} u^2 \, d\mathcal{H}^{n+2}_{\mathfrak{P}} \le \frac{C(\Lambda)}{\tau^{\frac{n}{2}}} \int_{B(0,2\sqrt{\tau})} u^2(x,-\tau)dx \\ \le& C(\Lambda)\widehat{H}^u(\tau),\end{aligned}$$

which implies the claim if $r < \sqrt{\tau}$.

Now suppose $r \in [\sqrt{\tau}, C(\Lambda)^{-1}\lambda^{-1}]$, and choose $k \in \mathbb{N}$ such that $2^k\sqrt{\tau} \le r \le 2^{k+1}\sqrt{\tau}$. By (7.2) and [EFV06, Theorem 2(1)], we have

$$\int_{B(0,2s)} u^2(x,-\tau)dx \le C(\Lambda) \int_{B(0,s)} u^2(x,-\tau)dx$$

for all $s \in (0, \lambda^{-1}]$. Iterating this estimate yields

$$\int_{B(0,r)} u^2(x,-\tau)dx \le \int_{B(0,2^{k+1}\sqrt{\tau})} u^2(x,-\tau)dx \le C(\Lambda)^{k+1} \int_{B(0,\sqrt{\tau})} u^2(x,-\tau)dx,$$

so combining this with (7.3) and interior parabolic estimates yields

$$\begin{aligned}\sup_{B(0,r)}(u^2+r^2|\nabla u|^2)(\cdot,-\tau)\leq&\frac{C}{r^{n+2}}\int_{B(0,2r)\times[-\tau-4r^2,-\tau]}u^2\,d\mathcal{H}^{n+2}_{\mathfrak{P}}\leq\frac{C(\Lambda)}{r^n}\int_{B(0,2r)}u^2(x,-\tau)dx\\ \leq&\frac{C(\Lambda)^{k+1}}{r^n}\int_{B(0,\sqrt{\tau})}u^2(x,-\tau)dx\leq C(\Lambda)^{k+1}\widehat{H}^u(\tau).\end{aligned}$$

The claim then follows from the observation that

$$C(\Lambda)^{k+1}=C(\Lambda)2^{k\log_2 C(\Lambda)}\leq C(\Lambda)\left(\frac{r^2}{\tau}\right)^{\log_2 C(\Lambda)}.$$

□

In the lemma below, we show that the error arising from using a cutoff function while doing global computation is exponentially small. Moreover, we prove that the frequency is bounded.

**Lemma 7.5.** *The following holds if $\lambda\leq\overline{\lambda}(\Lambda,\Upsilon)$. For any $\mathbf{x}_0\in P(\mathbf{0},10)$, $\alpha\in[0,1)$, and $\tau\in(0,1]$, we have*

$$\begin{aligned}&\int_{\mathbb{R}^n}(1_{\mathrm{supp}(\nabla\chi)}+|\partial_t\chi|+|\nabla\chi|+|\Delta\chi|)^2(|u_{\mathbf{x}_0;1}|^2+\tau|\nabla u_{\mathbf{x}_0;1}|^2)e^{\alpha f}\,d\nu_{-\tau}\\ &\qquad\leq C(\Lambda,\alpha,\Upsilon)\widehat{H}^u(\tau)\lambda^\Upsilon e^{-\frac{(1-\alpha)}{8\lambda^\Upsilon|t|^{1-2\Upsilon}}}.\end{aligned}\tag{7.4}$$

*Further,*

$$\int_{\mathbb{R}^n}\left((\chi u_{\mathbf{x}_0;1})^2+\tau|\nabla(\chi u_{\mathbf{x}_0;1})|^2\right)e^{\alpha f}\,d\nu_{-\tau}\leq C(\alpha,\Upsilon,\Lambda)\widehat{H}^u_{\mathbf{x}_0}(\tau),\tag{7.5}$$

*and in particular, $\widehat{N}^u_{\mathbf{x}_0}(\tau)\leq C(\Lambda)$.*

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0=\mathbf{0}$ and $u_{\mathbf{x}_0;1}=u$. We have

$$\mathrm{supp}(|\partial_t\chi|)\cup\mathrm{supp}(|\nabla\chi|)\cup\mathrm{supp}(|\nabla^2\chi|)\subseteq\{(x,t)\in\mathbb{R}^n\times(-\infty,0]\colon |t|^\Upsilon\leq\lambda^{\frac{\Upsilon}{2}}|x|\leq 2|t|^\Upsilon\},$$

as well as

$$|\partial_t\chi|\leq C\lambda^{\frac{\Upsilon}{2}}|t|^{-(\Upsilon+1)},\qquad|\nabla\chi|\leq C\lambda^{\frac{\Upsilon}{2}}|t|^{-\Upsilon},\qquad|\Delta\chi|\leq C\lambda^{\frac{\Upsilon}{2}}|t|^{-2\Upsilon}.$$

Combining these expressions and using Lemma 7.4, we get

$$\begin{aligned}&\int_{\mathbb{R}^n}(1_{\mathrm{supp}(\nabla\chi)}+|\partial_t\chi|+|\nabla\chi|+|\Delta\chi|)^2(|u|^2+\tau|\nabla u|^2)e^{\alpha f}d\nu_t\\ &\quad\leq C(\Lambda)\widehat{H}^u(\tau)\lambda^\Upsilon|t|^{-4\Upsilon-2-\frac{n}{2}}\int_{B(0,2\lambda^{-\frac{\Upsilon}{2}}|t|^\Upsilon)\setminus B(0,\lambda^{-\frac{\Upsilon}{2}}|t|^\Upsilon)}(1+|x|)^{C\Lambda}e^{-(1-\alpha)\frac{|x|^2}{4|t|}}\,dx\\ &\quad\leq C(\Lambda,\alpha,\Upsilon)\widehat{H}^u(\tau)\lambda^\Upsilon e^{-\frac{(1-\alpha)}{8\lambda^\Upsilon|t|^{1-2\Upsilon}}}.\end{aligned}$$

This proves (7.4) To get (7.5), we use Lemma 7.4 and (7.4) as follows:

$$\begin{aligned}&\int_{\mathbb{R}^n}\left((\chi u)^2+\tau|\nabla(\chi u)|^2\right)e^{\alpha f}d\nu_{-1}\\ &\leq\frac{C(\Lambda)\widehat{H}^u(\tau)}{\tau^{\frac{n}{2}}}\int_{\mathbb{R}^n}(1+|y|)^{2\Lambda}e^{-\frac{(1-\alpha)|y|^2}{4\tau}}dy+2\int_{\mathbb{R}^n}\tau|\nabla\chi|^2|u|^2e^{\alpha f}d\nu_t\\ &\leq C(\Lambda)\widehat{H}^u(\tau)\int_{\mathbb{R}^n}(1+|z|^2)^{2\Lambda}e^{-\frac{(1-\alpha)|z|^2}{4}}dz+C(\alpha,\Upsilon,\Lambda)\widehat{H}^u(\tau)\leq C(\alpha,\Upsilon,\Lambda)\widehat{H}^u(\tau).\end{aligned}$$

□

The result below establishes almost monotonicity of $\log \widehat{H}$ when $\lambda$ is small as an analog of Lemma 2.5.

**Lemma 7.6.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$. There exists $C = C(\Lambda)$ such that for any $\mathbf{x}_0 \in P(\mathbf{0}, 10)$ with $0 < \tau_1 < \tau_2 \le 1$, we have*

$$\left(\frac{\tau_1}{\tau_2}\right)^C \widehat{H}^u_{\mathbf{x}_0}(\tau_2) \le \widehat{H}^u_{\mathbf{x}_0}(\tau_1) \le \left(\frac{\tau_2}{\tau_1}\right)^{\lambda^\Upsilon C} \widehat{H}^u_{\mathbf{x}_0}(\tau_2). \tag{7.6}$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$. In light of Lemma 2.5, we estimate the following error to compute $\frac{d}{d\tau} \log \widehat{H}^u$ :

$$\int_{\mathbb{R}^n} (\chi u) \Box (\chi u)\, d\nu_t = \int_{\mathbb{R}^n} (\chi u) \Big( (\partial_t - \mathcal{L}) u + \operatorname{div}\left((a - I) \cdot \nabla(\chi u)\right) \Big)\, d\nu_t.$$

First, we integrate by parts and use Lemma 7.5 to get

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} \chi u \operatorname{div}\left((a - I) \cdot \nabla(\chi u)\right) d\nu_t \right| &= \left| \int_{\mathbb{R}^n} (\nabla(\chi u) - \chi u \nabla f) \cdot ((a - I) \nabla(\chi u))\, d\nu_t \right| \\
&\le C \lambda^{1 - \frac{\Upsilon}{2}} \tau^\Upsilon \int_{\mathbb{R}^n} |\nabla(\chi u)|^2 + |\chi u \nabla f||\nabla(\chi u)|\, d\nu_t \\
&\le C \lambda^{1 - \frac{\Upsilon}{2}} \tau^\Upsilon \frac{\widehat{E}^u(\tau) + \widehat{H}^u(\tau)}{\tau},
\end{aligned}$$

where we used Cauchy–Schwarz in the third line along with Lemma (7.5).

Next, we use

$$(\partial_t - \mathcal{L})(\chi u) = \chi (\partial_t - \mathcal{L}) u + u (\partial_t - \mathcal{L}) \chi - 2(a \nabla u) \cdot \nabla \chi \tag{7.7}$$

to estimate

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} \chi u (\partial_t - \mathcal{L})(\chi u)\, d\nu_t \right| &\le \int_{\mathbb{R}^n} \chi |u| (\chi \lambda (|u| + |\nabla u|) + |u||(\partial_t - \mathcal{L})\chi| + 2(1 + \lambda)|\nabla u||\nabla \chi|)\, d\nu_t \\
&\le C \int_{\mathbb{R}^n} (\lambda \chi^2 + |\nabla \chi|^2 + |(\partial_t - \mathcal{L})\chi|^2)\Big(\frac{1}{\sqrt{\tau}} |u|^2 + \sqrt{\tau} |\nabla u|^2\Big)\, d\nu_t \\
&\le C\Big(\lambda + \lambda^\Upsilon e^{-\frac{1}{8 \lambda^\Upsilon \tau^{1 - 2\Upsilon}}}\Big) \tau^{\frac{1}{2}} \left( \frac{\widehat{E}^u(\tau) + \widehat{H}^u(\tau)}{\tau} \right).
\end{aligned}$$

Therefore,

$$\frac{1}{\tau} (\widehat{N}^u(\tau) - C \lambda^\Upsilon \tau^\Upsilon) \le \frac{d}{d\tau} \log \widehat{H}^u(\tau) \le \frac{1}{\tau} (\widehat{N}^u(\tau) + C \lambda^\Upsilon \tau^\Upsilon). \tag{7.8}$$

Combining this with (7.5) yields (7.6). $\square$

The following result shows that $\nabla^2(\chi u)$ and $\Delta_f(\chi u)$ are square integrable in weighted space-time. Furthermore, we show that $\Box(\chi u)$ is small in a weighted sense.

**Lemma 7.7.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$. Let $\alpha \in [0, \frac{1}{2n}]$ and $\theta \in (0, 1]$. For any $\mathbf{x}_0 \in P(\mathbf{0}, 10)$ and $0 < \theta \tau_2 < \tau_1 \le \tau_2 \le 1$, we have*

$$\int_{\tau_1}^{\tau_2} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\nabla^2 (\chi u_{\mathbf{x}_0;1})|^2 e^{\alpha f}\, d\nu_{-\tau} d\tau \le C(\alpha, \Lambda, \theta, \Upsilon), \tag{7.9a}$$

$$\int_{\tau_1}^{\tau_2} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\Box (\chi u_{\mathbf{x}_0;1})|^2 e^{\alpha f}\, d\nu_{-\tau} d\tau \le C(\alpha, \Lambda, \theta, \Upsilon) \lambda^\Upsilon \tau_2^{2\Upsilon}, \tag{7.9b}$$

$$\int_{\tau_1}^{\tau_2} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\Delta_f (\chi u_{\mathbf{x}_0;1})|^2\, d\nu_{-\tau} d\tau \le C(\Lambda, \theta, \Upsilon). \tag{7.9c}$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$. Note that if $g, h \in C_c^\infty(\mathbb{R}^n \times [-\tau_2, -\tau_1])$, then

$$\frac{d}{dt}\int_{\mathbb{R}^n} gh\, d\nu_t = \int_{\mathbb{R}^n} h\Box g\, d\nu_t - \int_{\mathbb{R}^n} g(\Box^* h + 2\nabla f\cdot\nabla h)\, d\nu_t. \tag{7.10}$$

Further,

$$\Box^* e^{\alpha f} + 2\nabla f \cdot \nabla e^{\alpha f} = (2(1-\alpha)f - n)\,\frac{\alpha}{2\tau} e^{\alpha f}. \tag{7.11}$$

Using (7.10) and (7.11), we compute

$$\begin{aligned}
\frac{d}{dt}\int_{\mathbb{R}^n} \tau|\nabla(\chi u)|^2 e^{\alpha f}\, d\nu_t &= \int_{\mathbb{R}^n} e^{\alpha f}\Box(\tau|\nabla(\chi u)|^2)\, d\nu_t - \tau\int_{\mathbb{R}^n} |\nabla(\chi u)|^2\left(\Box^*(e^{\alpha f}) + 2\nabla f\cdot\nabla e^{\alpha f}\right) d\nu_t \\
&= -2\tau\int_{\mathbb{R}^n} |\nabla^2(\chi u)|^2 e^{\alpha f} d\nu_t + 2\tau\int_{\mathbb{R}^n} e^{\alpha f}\nabla(\chi u)\cdot\nabla\Box(\chi u)\, d\nu_t \\
&\quad + \int_{\mathbb{R}^n} |\nabla(\chi u)|^2\left(\alpha(n - 2(1-\alpha)f) - 1\right) e^{\alpha f} d\nu_t =: I_1 + I_2 + I_3.
\end{aligned}$$

Since $\alpha \le \frac{1}{2n}$, it follows that

$$I_3 \le -\frac{1}{2}\int_{\mathbb{R}^n} |\nabla(\chi u)|^2 e^{\alpha f} d\nu_t. \tag{7.12}$$

We now estimate $I_2$. Using (7.7), we get

$$\begin{aligned}
|\Box(\chi u)| &\le |\nabla(\chi u)\cdot\nabla a| + |(a - I)\nabla^2(\chi u)| + |(\partial_t - \mathcal{L})(\chi u)| \\
&\le C(|u| + |\nabla u|)\cdot(\lambda\chi + |\partial_t\chi| + |\Delta\chi| + |\nabla\chi|) + C\lambda^{1-\frac{\Upsilon}{2}}\tau^{\Upsilon}|\nabla^2(\chi u)|.
\end{aligned} \tag{7.13}$$

We can integrate the Bochner formula for $\Delta_{(1-\alpha)f}|\nabla(\chi u)|^2$ by parts to obtain

$$\tau\int_{\mathbb{R}^n} (\Delta_{(1-\alpha)f}(\chi u))^2 e^{\alpha f} d\nu_t = \int_{\mathbb{R}^n}\left(\tau|\nabla^2(\chi u)|^2 + (1-\alpha)|\nabla(\chi u)|^2\right) e^{\alpha f}\, d\nu_t. \tag{7.14}$$

For $\epsilon > 0$ to be determined, we integrate by parts and use (7.13) and (7.14) to get

$$\begin{aligned}
&\left|2\int_{\mathbb{R}^n} (\nabla(\chi u)\cdot\nabla\Box(\chi u)) e^{\alpha f}\, d\nu_t\right| \\
&= \left|2\int_{\mathbb{R}^n} \Delta_{(1-\alpha)f}(\chi u)\Box(\chi u) e^{\alpha f}\, d\nu_t\right| \le \epsilon\int_{\mathbb{R}^n} (\Delta_{(1-\alpha)f}(\chi u))^2 e^{\alpha f} d\nu_t + \epsilon^{-1}\int_{\mathbb{R}^n} (\Box(\chi u))^2 e^{\alpha f}\, d\nu_t \\
&\le \epsilon\int_{\mathbb{R}^n} |\nabla^2(\chi u)|^2 e^{\alpha f} d\nu_t + \frac{\epsilon}{\tau}\int_{\mathbb{R}^n} |\nabla(\chi u)|^2 e^{\alpha f}\, d\nu_t \\
&\quad + C\epsilon^{-1}\int_{\mathbb{R}^n} (|u|^2 + |\nabla u|^2)(\lambda\chi^2 + |\partial_t\chi| + |\nabla\chi| + |\Delta\chi|)^2 e^{\alpha f}\, d\nu_t + C\epsilon^{-1}\lambda^{1-\frac{\Upsilon}{2}}\int_{\mathbb{R}^n} |\nabla^2(\chi u)|^2 e^{\alpha f} d\nu_t \\
&\le (\epsilon + C\epsilon^{-1}\lambda^{1-\frac{\Upsilon}{2}})\int_{\mathbb{R}^n} |\nabla^2(\chi u)|^2 e^{\alpha f} d\nu_t + (\frac{\epsilon}{\tau} + C\epsilon^{-1}\lambda)\int_{\mathbb{R}^n} (|\chi u|^2 + |\nabla(\chi u)|^2) e^{\alpha f} d\nu_t \\
&\quad + C\epsilon^{-1}\int_{\mathbb{R}^n} (|u|^2 + |\nabla u|^2)(|\partial_t\chi| + |\nabla\chi| + |\Delta\chi|)^2 e^{\alpha f} d\nu_t.
\end{aligned}$$

Therefore, if $\epsilon \le \overline{\epsilon}$ and $\lambda \le \overline{\lambda}(\epsilon, \Lambda)$, we can use Lemma 7.5 to get

$$\begin{aligned}
I_1 + I_2 + I_3 &\le \tau(-2 + \epsilon + C\epsilon^{-1}\lambda^{1-\frac{\Upsilon}{2}})\int_{\mathbb{R}^n} |\nabla^2(\chi u)|^2 e^{\alpha f} d\nu_t \\
&\quad + (\epsilon + C\epsilon^{-1}\lambda)\int_{\mathbb{R}^n} |\chi u|^2 e^{\alpha f} d\nu_t + (-\frac{1}{2} + \epsilon + C\epsilon^{-1}\lambda)\int_{\mathbb{R}^n} |\nabla(\chi u)|^2) e^{\alpha f} d\nu_t
\end{aligned}$$

$$+ C\epsilon^{-1}\int_{\mathbb{R}^n}(|u|^2+\tau|\nabla u|^2)(|\partial_t\chi|+|\nabla\chi|+|\Delta\chi|)^2 e^{\alpha f}d\nu_t$$
$$\le -\tau\int_{\mathbb{R}^n}|\nabla^2(\chi u)|^2 e^{\alpha f}d\nu_t + 2\epsilon\widehat{H}^u(\tau)+C(\alpha,\Upsilon,\Lambda)\lambda^{\Upsilon}e^{-\frac{1-\alpha}{8\lambda^{\Upsilon}|t|^{1-2\Upsilon}}}\widehat{H}^u(\tau).$$

We integrate $I_1+I_2+I_3$ against $\frac{1}{\widehat{H}^u(\tau)}$, integrate by parts, use (7.6), (7.8), and Lemma 7.5 to obtain

$$\begin{aligned}
&\int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}|\nabla^2(\chi u)|^2 e^{\alpha f}d\nu_t dt\\
&\le -\int_{-\tau_2}^{-\tau_1}\frac{1}{\widehat{H}^u(\tau)}\frac{d}{dt}\left(\int_{\mathbb{R}^n}\tau|\nabla(\chi u)|^2 e^{\alpha f}d\nu_t\right)dt + C(\alpha,\Upsilon,\Lambda)(\lambda^{\Upsilon}e^{-\frac{1-\alpha}{2\lambda^{\Upsilon}|t_2|^{1-2\Upsilon}}}+2\epsilon)|\tau_2-\tau_1|\\
&\le \frac{\tau_2}{\widehat{H}^u(\tau_1)}\int_{\mathbb{R}^n}|\nabla(\chi u)|^2 e^{\alpha f}d\nu_{t_2} - \int_{-\tau_2}^{-\tau_1}\frac{d}{dt}\log\widehat{H}^u(\tau)\frac{1}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}\tau|\nabla(\chi u)|^2 e^{\alpha f}d\nu_t dt\\
&\quad + C(\alpha,\Upsilon,\Lambda)(\lambda^{\Upsilon}e^{-\frac{1-\alpha}{8\lambda^{\Upsilon}|t_2|^{1-2\Upsilon}}}+2\epsilon)|\tau_2-\tau_1|\\
&\le C(\alpha,\Upsilon,\Lambda)|t_2-t_1| + C(\Lambda,\alpha)\log\left(\frac{\tau_2}{\tau_1}\right).
\end{aligned}$$

This proves (7.9a). We now integrate (7.13) against $\tau\widehat{H}^u(\tau)^{-1}$ and apply Lemma 7.5 and (7.9a) to obtain

$$\begin{aligned}
&\int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}|\Box(\chi u)|^2 e^{\alpha f}d\nu_t dt\\
&\le C\lambda^2\int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}(|(\chi u)|^2+|\nabla(\chi u)|^2)e^{\alpha f}d\nu_t dt\\
&\qquad + C\int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}(|u|^2+|\nabla u|^2)(|\partial_t\chi|^2+|\Delta\chi|^2+|\nabla\chi|^2)e^{\alpha f}d\nu_t dt\\
&\qquad + C\lambda^{2-\Upsilon}\tau_2^{2\Upsilon}\int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}|\nabla^2(\chi u)|^2 d\nu_t dt\\
&\le C(\alpha,\Upsilon,\Lambda)\lambda^{\Upsilon}|\tau_2-\tau_1| + C\lambda^{2-\Upsilon}\tau_2^{2\Upsilon}\left(C(\alpha,\Upsilon,\Lambda)|\tau_2-\tau_1|+C(\Lambda,\alpha)\log\left(\frac{\tau_2}{\tau_1}\right)\right).
\end{aligned}$$

This proves (7.9b).

Finally, we use (7.14) with $\alpha\leftarrow 0$, (7.9a) and (7.5) to get

$$\begin{aligned}
\int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}(\Delta_f(\chi u))^2 d\nu_t &= \int_{-\tau_2}^{-\tau_1}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}\left(|\nabla^2(\chi u)|^2+\frac{1}{\tau}|\nabla(\chi u)|^2\right)d\nu_t dt\\
&\le C(\Lambda,\Upsilon)|\tau_2-\tau_1| + C(\Lambda,\alpha)\log\left(\frac{\tau_2}{\tau_1}\right).
\end{aligned}$$

□

In the following proposition, we prove almost monotonicity and rigidity of the generalized frequency when the scales are comparable.

**Proposition 7.8.** *The following holds if $\lambda\le\overline{\lambda}(\Lambda,\Upsilon)$. Fix $\theta\in(0,1]$.*

*(1) For all $\mathbf{x}_0\in P(\mathbf{0},10)$ and $0<\theta\tau_2<\tau_1<\tau_2\le 1$, we have*

$$\left|\widehat{N}^u_{\mathbf{x}_0}(\tau_2)-\widehat{N}^u_{\mathbf{x}_0}(\tau_1)-\int_{\tau_1}^{\tau_2}\frac{4\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)}\int_{\mathbb{R}^n}\left(\Delta_f(\chi u_{\mathbf{x}_0;1})+\frac{\widehat{N}^u_{\mathbf{x}_0}(\tau)}{2\tau}(\chi u_{\mathbf{x}_0;1})\right)^2 d\nu_{-\tau}d\tau\right|$$

$$\leq C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}.$$

(2) *For any $\epsilon \in (0,1]$, if $4\tau_1 \leq \tau_2$ and $\widehat{N}^u_{\mathbf{x}_0}(\tau_1) > \widehat{N}^u_{\mathbf{x}_0}(\tau_2) - \epsilon$, then there exists $m \in \mathbb{N}_0$ such that*

$$\sup_{\tau\in[\tau_1,\tau_2]} |\widehat{N}^u_{\mathbf{x}_0}(\tau) - m| + \int_{\tau_1}^{\tau_2} \frac{4\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u_{\mathbf{x}_0;1}) + \frac{m}{2\tau}(\chi u_{\mathbf{x}_0;1})\right)^2 d\nu_{-\tau} d\tau$$
$$< C\epsilon + C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}.$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$.

(1) By Lemma 2.5, we can estimate

$$\left|(\widehat{N}^u)(\tau_2) - (\widehat{N}^u)(\tau_1) - \int_{\tau_1}^{\tau_2} \frac{4\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u) + \frac{\widehat{N}^u(\tau)}{2\tau}(\chi u)\right)^2 d\nu_{-\tau} d\tau\right|$$
$$= \left|\int_{\tau_1}^{\tau_2} \frac{4\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u) + \frac{\widehat{N}^u(\tau)}{2\tau}(\chi u)\right) \Box(\chi u)\, d\nu_{-\tau} d\tau\right|$$
$$\leq 4\left(\int_{\tau_1}^{\tau_2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u) + \frac{\widehat{N}^u(\tau)}{2\tau}(\chi u)\right)^2 d\nu_{0\tau} d\tau\right)^{\frac{1}{2}} \left(\int_{2\tau_1}^{\tau_2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\Box(\chi u)|^2 d\nu_{-\tau} d\tau\right)^{\frac{1}{2}}$$
$$\leq C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}.$$

(2) By (1), we have

$$\int_{\tau_2}^{\tau_1} \frac{4\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u) + \frac{\widehat{N}^u(\tau)}{2\tau}(\chi u)\right)^2 d\nu_{-\tau} d\tau \leq C\epsilon + C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}.$$

Choose $\tau_0 \in [\tau_2, \tau_1]$ such that

$$\frac{4\tau_0}{\widehat{H}^u(\tau_0)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u) + \frac{\widehat{N}^u(\tau_0)}{2\tau_0}(\chi u)\right)^2 d\nu_{-\tau_0} \leq \frac{C\epsilon + C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}}{\tau_2 - \tau_1}.$$

By Lemma 2.8, there exists an integer $m \leq C(\Lambda)$ such that

$$|\widehat{N}^u(\tau_0) - m| \leq \frac{80\tau_0^2}{\widehat{H}^u(\tau_0)} \int_{\mathbb{R}^n} \left(\Delta_f u + \frac{\widehat{N}^u(\tau_0)}{2\tau_0}\right)^2 d\nu_{-\tau_0} \leq \frac{\tau_2}{\tau_2 - \tau_1}\left(C\epsilon + C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}\right).$$

On the other hand, for any $[t_1', t_2'] \subseteq [t_1, t_2]$, we again use (1) to get

$$|\widehat{N}^u(\tau_2') - \widehat{N}^u(\tau_1')| \leq C\epsilon + C(\Lambda,\theta,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}.$$

Therefore, (2) immediately follows. □

Using the fact that the generalized frequency is pinched near integers, we improve its almost monotonicity at two comparable scales to two arbitrary scales.

**Corollary 7.9.** *The following holds if $\lambda \leq \overline{\lambda}(\Lambda,\Upsilon)$. For any $\mathbf{x}_0 \in P(\mathbf{0},10)$ and $0 < \theta\tau_2 < \tau_1 \leq \tau_2 \leq 1$, we have*

$$\widehat{N}^u_{\mathbf{x}_0}(\tau_2) \geq \widehat{N}^u_{\mathbf{x}_0}(\tau_1) - C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}.$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume $\mathbf{x}_0 = 0$. For $B = B(\Lambda,\Upsilon)$ to be determined, suppose by way of contradiction that there exists $\tau_1 \in (0,\tau_2]$ such that $\tau_1 \leq \tau_2$ and $\widehat{N}^u(\tau_2) < \widehat{N}^u(\tau_1) - B\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}$. Choose $\epsilon' \in [\frac{B}{2}\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}, B\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}]$ such that $|\widehat{N}^u(\tau_2) + \epsilon' - m| \geq \frac{B}{4}\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}$ for all

$m \in \mathbb{N}_0$. Choose $\tau_* \in [\tau_1, \tau_2]$ such that $\widehat{N}^u(\tau_*) = \widehat{N}^u(\tau_2) + \epsilon'$ but $\widehat{N}^u(\tau) \leq \widehat{N}^u(\tau_2) + \epsilon'$ for all $\tau \in [\tau_*, \tau_2]$. If $\tau_* \geq \frac{1}{4}\tau_2$, then Proposition 7.8(1) yields a contradiction if we take $B \geq \underline{B}(\Lambda, \Upsilon)$. Thus $\tau_* \leq \frac{1}{4}\tau_2$. If $\lambda \leq \overline{\lambda}(\Lambda, \Upsilon)$, then Proposition 7.8(2) with $\tau_1 \leftarrow \tau_*$, $\tau_2 \leftarrow 4\tau_*$, gives

$$\frac{B}{2}\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon} \leq |\widehat{N}^u(\tau_2) + \epsilon' - m| = |\widehat{N}^u(\tau_*) - m| \leq C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_*^{\Upsilon} \leq C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_2^{\Upsilon}$$

for some $m \in \mathbb{N}_0$, since $\widehat{N}^u(\tau) \leq \widehat{N}^u(\tau_*)$ for all $\tau \in [\tau_*, 4\tau_*]$. If $B \geq \underline{B}(\Lambda, \Upsilon)$, this yields a contradiction, and the claim follows. $\square$

Using a discrete version of (2.8), we prove the following analog of Lemma 2.11 where we prove a definite drop in the frequency at small scale given that the frequency at larger scale is a non-integer.

**Lemma 7.10.** *The following holds if $\lambda \leq \overline{\lambda}(\Lambda, \Upsilon)$. If $\widehat{N}(\tau) \leq m + 1 - 2\epsilon$ and $\lambda^{\frac{\Upsilon}{4}}\tau^{\frac{\Upsilon}{2}} < \epsilon$ for some $\epsilon \in (0, \frac{1}{4})$, then*

$$\widehat{N}\left(\left(\frac{\epsilon}{1-\epsilon}\right)^{60}\tau\right) \leq m + \epsilon.$$

*Proof.* Define $\tau_j = 4^{-j}\tau$. Using Lemma 2.8 and Lemma 7.8(1) there exist a non-negative integer $m_j \leq C(\Lambda)$ and $\tau_j^* \in [\tau_{j+1}, \tau_j]$ such that

$$\begin{aligned}(\widehat{N}(\tau_j^*) - m_j)(m_j + 1 - \widehat{N}(\tau_j^*)) &\leq \frac{20(\tau_j^*)^2}{\widehat{H}^u(\tau_j^*)}\int_{\mathbb{R}^n}\left(\Delta_f(\chi u) + \frac{\widehat{N}(\tau_j^*)}{2\tau_j^*}(\chi u)\right)^2 d\nu_{-\tau_j^*} \\ &\leq \frac{20}{3}\left(\widehat{N}(\tau_j) - \widehat{N}(\tau_{j+1}) + C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_j^{\Upsilon}\right).\end{aligned} \tag{7.15}$$

We can assume that $m_j = m$ and $\widehat{N}(\tau_{j+1}) - m - C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_j^{\Upsilon} > \epsilon$. Using

$$\widehat{N}(\tau_{j+1}) - C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_j^{\Upsilon} \leq \widehat{N}(\tau_j^*) \leq \widehat{N}(\tau_j) + C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_j^{\Upsilon},$$

and setting $\delta_j^2 := C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_j^{\Upsilon}$ and $x_j := \widehat{N}(\tau_j) - m$, (7.15) implies

$$(x_{j+1} - \delta_j^2)(1 - x_j - \delta_j^2) \leq \frac{20}{3}(x_j - x_{j+1}) + \delta_j^2,$$

which implies

$$\frac{1}{x_j}\left(\frac{23}{3} - x_j\right) \leq \frac{20}{3}\frac{1}{x_{j+1}} + \frac{\delta_j^2}{\epsilon x_{j+1}} \leq \frac{20 + \delta_j}{3}\frac{1}{x_{j+1}}.$$

Set $y_j := \frac{1}{x_j}$. Then

$$\frac{23}{20}(y_j - 1) - \frac{\delta_j}{20}\frac{23}{20}(y_j - 1) - \frac{\delta_j}{20} \leq (y_{j+1} - 1),$$

Applying this and $y_j - 1 = \frac{1}{\widehat{N}(\tau_j) - m} - 1 \geq \frac{1}{1-\epsilon} - 1 \geq \epsilon \geq \delta_j$, we get

$$\frac{21}{20}(y_j - 1) \leq y_{j+1} - 1.$$

Using this linear relation, we get

$$\left(\frac{21}{20}\right)^j\frac{\epsilon}{1-\epsilon} + 1 \leq \left(\frac{21}{20}\right)^j(y_0 - 1) + 1 \leq y_{j+1},$$

If $\left(\frac{21}{20}\right)^j\frac{\epsilon}{1-\epsilon} + 1 \geq \frac{1}{\epsilon}$, then $x_{j+1} \leq \epsilon$, which implies $\widehat{N}(\tau_{j+1}) \leq m + \epsilon$. In particular, we have $\tau_j = 4^{-\log_{\frac{21}{20}}\left(\frac{1-\epsilon}{\epsilon}\right)^2}\tau = \left(\frac{\epsilon}{1-\epsilon}\right)^{2\log_{\frac{21}{20}}4}\tau$. Therefore, the claim follows. $\square$

**Lemma 7.11.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$. For any solution to (7.1) satisfying (7.2), and for any $\mathbf{x}_0 \in P(\mathbf{0}, 1)$, $\epsilon \in (0, \frac{1}{10})$, if $0 < r_1 \le \epsilon^{C(\Lambda,\Upsilon)\Lambda} r_2$ and $r_2 \le 1$, then there exists $s \in (r_1, r_2)$ such that*

$$\widehat{N}(\epsilon^{-2}s^2) - \widehat{N}(\epsilon^2 s^2) < \epsilon.$$

*Proof.* This follows from the proof of Lemma 2.12, using Lemma 7.10 instead of Lemma 2.11. ☐

7.2. **Quantitative uniqueness of approximating caloric polynomials.** As an analog of Theorem 2.13, in Theorem 7.13, we prove the quantitative uniqueness of approximating caloric polynomials under the assumption that the frequency is pinched at two scales. We construct the approximating polynomial using dynamical rescaling and quantify the uniqueness using an $L^2$-distance defined below.

Given a function $w$ on $\mathbb{R}^n \times I$, define the *renormalized parabolic dilation* $\widetilde{w}_{\mathbf{x}_0;s}$ at $\mathbf{x}_0 \in \mathbb{R}^n \times \mathbb{R}$ by

$$\widetilde{w}_{\mathbf{x}_0;s}(x) := \frac{w(x_0 + e^{-\frac{s}{2}}x, t_0 - e^{-s})}{\sqrt{H^w_{\mathbf{x}_0}(e^{-\frac{s}{2}})}}.$$

When $\mathbf{x}_0 = \mathbf{0}$, we write $\widetilde{w}_s := \widetilde{w}_{\mathbf{x}_0;s}$. Given $m \in \mathbb{N}_0$, we define the distance between a function and $\mathcal{P}_m$ as follows. We denote $\nu := \nu_{-1}$, and for any $0 \le s_2 < s_1 < \infty$, we define

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s_2,s_1]} := \inf_{p\in\mathcal{P}_m} \|\widetilde{w} - \widehat{p}\|_{[s_2,s_1]} := \inf_{p\in\mathcal{P}_m} \left(\int_{s_2}^{s_1} \|\widetilde{w}_s - \widehat{p}\|^2_{L^2(\mathbb{R}^n,\nu)}\, ds\right)^{\frac{1}{2}},$$

where $\widehat{p}(x) := p(x, -1)$ and $\widehat{\mathcal{P}}_m := \{\widehat{p} \colon p \in \mathcal{P}_m\}$.

*Remark* 7.12. We observe that $\|\cdot - \widehat{\mathcal{P}}_m\|_{[\cdot,\cdot]}$ is sub-additive in the following sense

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s_0,s_0+3A]} \le 10(\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s_0,s_0+2A]} + \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s_0+A,s_0+3A]}).$$

Suppose $\widehat{p}_1 \in \arg\min \|\widetilde{w} - \widehat{p}\|_{[s_0,s_0+2A]}$ and $\widehat{p}_2 \in \arg\min \|\widetilde{w} - \widehat{p}\|_{[s_0+A,s_0+3A]}$. Then

$$\begin{aligned}\left(\int_{\mathbb{R}^n} (\widehat{p}_1 - \widehat{p}_2)^2\, d\nu\right)^{\frac{1}{2}} &= \left(\frac{1}{A}\int_{s_0+A}^{s_0+2A}\int_{\mathbb{R}^n} (\widehat{p}_1 - \widehat{p}_2)^2\, d\nu ds\right)^{\frac{1}{2}} \\ &\le \frac{1}{\sqrt{A}}(\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s_0,s_0+2A]} + \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s_0+A,s_0+3A]}).\end{aligned} \tag{7.16}$$

Therefore, using (7.16), we get

$$\begin{aligned}&\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[s_0,s_0+3A]} \\ &\le \int_{s_0}^{s_0+3A}\int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p}_1)^2\, d\nu ds \le \int_{s_0}^{s_0+2A}\int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p}_1)^2\, d\nu ds + \int_{s_0+A}^{s_0+3A}\int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p}_1)^2\, d\nu ds \\ &\le \|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[s_0,s_0+2A]} + 2\int_{s_0+A}^{s_0+3A}\int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p}_2)^2\, d\nu ds + 2\int_{s_0+A}^{s_0+3A}\int_{\mathbb{R}^n} (\widehat{p}_2 - \widehat{p}_1)^2\, d\nu ds \\ &\le 10\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[s_0,s_0+2A]} + 10\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[s_0+A,s_0+3A]}.\end{aligned}$$

**Notation:** Throughout this subsection, we assume that $u$ is a solution to (7.1) satisfying (7.2). By parabolic rescaling, we can assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$. We will set $w := \chi u$ throughout this section. Further, we will write $s := e^{-\tau}$.

**Theorem 7.13.** *For any $A \in [1, \infty)$, the following holds if $\lambda \le \overline{\lambda}(A, \Lambda)$ and $\delta \le \overline{\delta}(A, \Lambda, \Upsilon)$. Suppose*

$$\sup_{s\in[s_2-8A,s_1+8A]} |\widehat{N}^u(e^{-s}) - m| = \delta \tag{7.17}$$

*for some* $2A < s_2 < s_1 - 2$. *Then there exists* $\widehat{p} \in \widehat{\mathcal{P}}_m$ *such that*

$$\|\widetilde{w} - \widehat{p}\|_{[s-A,s+A]} < C(A,\Lambda,\Upsilon)\delta\left(e^{-\frac{\Upsilon}{4}(s_1-s)} + e^{-\frac{\Upsilon}{4}(s-s_2)}\right) + C\lambda^{\frac{\Upsilon}{4}}\left(e^{-\Upsilon s} + e^{-\frac{\Upsilon}{2}(s_1+s_2)}\right) \tag{7.18a}$$

$$|\widehat{N}^u(e^{-s}) - m| < C(A,\Lambda,\Upsilon)\delta\left(e^{-\frac{\Upsilon}{4}(s_1-s)} + e^{-\frac{\Upsilon}{4}(s-s_2)}\right) + C\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon s} \tag{7.18b}$$

*for all* $s \in [s_2, s_1]$.

As a first step towards proving Theorem 7.13, we prove a weaker version of quantitative uniqueness when the frequency is pinched at comparable scales.

**Proposition 7.14.** *The following holds if* $\lambda \leq \overline{\lambda}(\Lambda)$. *Fix* $\epsilon \in (0,1]$ *and* $B \in [1,\infty)$. *Suppose* $\mathbf{x}_0 \in P(\mathbf{0},10)$ *and* $0 \leq s_1 - 2B \leq s_2 \leq s_1 - 2$ *such that*

$$\widehat{N}^u_{\mathbf{x}_0}(\tfrac{1}{2}e^{-s_1}) > \widehat{N}^u_{\mathbf{x}_0}(2e^{-s_2}) - \epsilon.$$

*Then there exists* $m \in \mathbb{N}_0$ *and* $p \in \mathcal{P}_m$ *such that*

$$\|\widetilde{w} - \widehat{p}\|^2_{[s_2,s_1]} \leq C(B,\Lambda,\Upsilon)\left(\epsilon + \lambda^{\frac{\Upsilon}{2}}e^{-\Upsilon s_2}\right).$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$. We set $s_i := -\log(\tau_i)$ for $i = 1, 2$.

Let $m \in \mathbb{N}_0$ be as in Proposition 7.8(2). Let $h(\cdot,t)$ be the $L^2(\mathbb{R}^n,\nu_t)$-orthogonal projection of $(\chi u)(\cdot,t)$ onto $\mathcal{P}_m$. By Lemma 2.8, we have

$$\int_{\tau_1}^{\tau_2}\frac{1}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}(\chi u - h)^2\,d\nu_{-\tau}\frac{d\tau}{\tau} \leq 20\int_{\tau_1}^{\tau_2}\frac{\tau^2}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}\left(\Delta_f(\chi u) + \frac{\widehat{N}^u(\tau)}{2\tau}\right)^2 d\nu_{-\tau}\frac{d\tau}{\tau}. \tag{7.19}$$

Note that $h$ is not necessarily in $\mathcal{P}_m$. We now construct $p \in \mathcal{P}_m$ that approximates $h$ in an optimal way as explained below. Let $\{p_i\}_{i=1}^N$ be a basis of $\mathcal{P}_m$ satisfying $\int_{\mathbb{R}^n} p_i p_j\,d\nu_t = \delta_{ij}\tau^m$ and set $h_i(t) := \frac{1}{\tau^m}\int_{\mathbb{R}^n}\chi u p_i\,d\nu_t$ so that $h = \sum_{i=1}^N h_i p_i$.

Define for $1 \leq j \leq N$

$$\overline{h}_j(s) := \sqrt{\frac{e^{-ms}}{\widehat{H}^u(e^{-s})}}h_j(-e^{-s}), \qquad a_j := \fint_{s_1}^{s_2}\overline{h}_j(s)\,ds, \qquad p := \sum_{i=1}^N a_i p_i.$$

Note that our choice of $a_j$ minimizes $\sum_{j=1}^N\int_{\tau_1}^{\tau_2}\left(\sqrt{\frac{\tau^m}{\widehat{H}^u(\tau)}}h_j(-\tau) - a_j\right)^2\frac{d\tau}{\tau}$ so that

$$\int_{s_1}^{s_2}(\overline{h}_j(s) - a_j)\,ds = 0. \tag{7.20}$$

Further, a direct computation yields

$$\overline{h}_j'(s) = \frac{1}{2}(\widehat{N}^u(e^{-s}) - m)\overline{h}_j(s) + \frac{e^{\frac{m}{2}s-s}}{\sqrt{\widehat{H}^u(e^{-s})}}\int_{\mathbb{R}^n}\square(\chi u)p_j\,d\nu_{-e^{-s}},$$

from which we can use Hölder's inequality to estimate

$$|\overline{h}_i'(s)| \leq \frac{1}{2}|\widehat{N}(e^{-s}) - m| + \frac{e^{-s}}{\sqrt{\widehat{H}^u(e^{-s})}}\left(\int_{\mathbb{R}^n}|\square(\chi u)|^2\,d\nu_{-e^{-s}}\right)^{\frac{1}{2}}.$$

Integrating in time and combining with Lemma 7.7 yields

$$\begin{aligned}\int_{s_2}^{s_1}|\overline{h}_i'(s)|^2\,ds &\leq \int_{\tau_1}^{\tau_2}\frac{1}{\tau}(\widehat{N}(\tau) - m)^2\,d\tau + 2\int_{\tau_1}^{\tau_2}\frac{\tau}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}|\square(\chi u)|^2\,d\nu_{-\tau}d\tau \\ &\leq \int_{\tau_1}^{\tau_2}\frac{1}{\tau}(\widehat{N}(\tau) - m)^2\,d\tau + C(B,\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_2^{2\Upsilon}.\end{aligned} \tag{7.21}$$

Therefore, using (7.20), Poincaré inequality and (7.21), we get

$$\begin{aligned}
&\int_{\tau_1}^{\tau_2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \left( h - \sqrt{\frac{\widehat{H}^u(\tau)}{\tau^m}} p \right)^2 d\nu_{-\tau} \frac{d\tau}{\tau} \\
&= \sum_{j=1}^N \int_{\tau_1}^{\tau_2} \left( \sqrt{\frac{\tau^m}{\widehat{H}^u(\tau)}} h_j(-\tau) - a_j \right)^2 \frac{d\tau}{\tau} = \sum_{j=1}^N \int_{s_2}^{s_1} (\overline{h}_j(s) - a_j)^2 \, ds \\
&\le C(s_1 - s_2) \sum_{j=1}^N \int_{s_2}^{s_1} (\overline{h}_j'(s))^2 \, ds \le C(B) \int_{\tau_1}^{\tau_2} \frac{1}{\tau} (\widehat{N}(\tau) - m)^2 \, d\tau + C(B, \Lambda, \Upsilon) \lambda^{\Upsilon} \tau_2^{2\Upsilon} \\
&\le C(B)\epsilon^2 + C(B, \Lambda, \Upsilon) \lambda^{\Upsilon} \tau_2^{2\Upsilon}.
\end{aligned}$$

This together with (7.19) gives

$$\begin{aligned}
\|\widetilde{w} - \widehat{p}\|^2_{[s_2, s_1]} &= \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}^n} \left( \frac{\chi u}{\sqrt{\widehat{H}^u(\tau)}} - \frac{p}{\sqrt{\tau^m}} \right)^2 d\nu_{-\tau} \frac{d\tau}{\tau} \\
&\le C\epsilon + C(B, \Lambda, \Upsilon) \lambda^{\frac{\Upsilon}{2}} \tau_2^{\Upsilon} + C(B)\epsilon^2 + C(B, \Lambda, \Upsilon) \lambda^{\Upsilon} \tau_2^{2\Upsilon} \le C(B, \Lambda, \Upsilon) \left( \epsilon + \lambda^{\frac{\Upsilon}{2}} \tau_2^{\Upsilon} \right).
\end{aligned}$$

□

Note that the estimates in Proposition 7.14 worsen as $\frac{\tau_2}{\tau_1} \to \infty$. To improve the estimate and prove Theorem 7.13, we view parabolic rescaling as a dynamical system. We then analyze the stability of caloric approximation of parabolic rescaling.

To this end, we observe that $\widetilde{w}_s$ satisfies the following evolution equation:

$$\partial_s \widetilde{w}_s = \Delta_f \widetilde{w}_s + \frac{m}{2} \widetilde{w}_s + \frac{1}{2} (N(\widetilde{w}_s) - m) \widetilde{w}_s + g_s, \tag{7.22}$$

where $m \in \mathbb{R}$ and

$$g_s := \frac{(\widetilde{\tau \Box w})_s}{\sqrt{H^w(e^{-s})}} - \frac{\widetilde{w}_s}{2H^w(e^{-s})} \int_{\mathbb{R}^n} (\tau w \Box w) d\nu_{-e^{-s}}. \tag{7.23}$$

Moreover, a change of variables implies that

$$\int_{\mathbb{R}^n} \widetilde{w}_s^2 d\nu = \frac{1}{H^w(e^{-\frac{s}{2}})} \int_{\mathbb{R}^n} w^2(e^{-\frac{s}{2}} x, -e^{-s}) \frac{e^{-\frac{|x|^2}{4}}}{(4\pi)^{\frac{n}{2}}} dx = \frac{1}{H^w(e^{-\frac{s}{2}})} \int_{\mathbb{R}^n} w^2 d\nu_{-e^{-s}} = 1. \tag{7.24}$$

In the lemmata below, we show that if $u$ is a solution to (7.1) satisfying (7.2) and is close to a heat polynomial at some scale, then the behavior of the corresponding function $\widetilde{u}$ satisfying (7.22) is modeled on that of the corresponding linearization at $\mathcal{P}_m$:

$$\partial_s \widetilde{w}_s = \Delta_f \widetilde{w}_s + \frac{m}{2} \widetilde{w}_s, \tag{7.25}$$

where $m \in \mathbb{N}_0$.

As a first step, we show that the frequency of a function and that of its caloric approximation are close.

**Lemma 7.15.** *For any $\widehat{p} \in \widehat{\mathcal{P}}_m$ and $w$ satisfying the mild growth condition (2.1) and $\int_{\mathbb{R}^n} w^2 \, d\nu = 1$, we have*

$$|N^w(1) - N^{\widehat{p}}(1)| \le 4 \int_{\mathbb{R}^n} |\nabla (w - \widehat{p})|^2 \, d\nu + 8m \int_{\mathbb{R}^n} (w - \widehat{p})^2 \, d\nu.$$

*Proof.* Let $\widehat{q} \in \widehat{\mathcal{P}}_m$ be the $L^2(\mathbb{R}^n, \nu)$ projection of $w$ onto $\widehat{\mathcal{P}}_m$. Then from

$$\int_{\mathbb{R}^n} (w - \widehat{q})\widehat{q}\, d\nu = 0, \qquad \int_{\mathbb{R}^n} w^2\, d\nu = 1,$$

we can estimate

$$\int_{\mathbb{R}^n} \nabla(w - \widehat{q}) \cdot \nabla\widehat{q}\, d\nu = -\int_{\mathbb{R}^n} (w - \widehat{q})\Delta_f \widehat{q}\, d\nu = \frac{m}{2}\int_{\mathbb{R}^n} (w - \widehat{q})\widehat{q}\, d\nu = 0. \tag{7.26}$$

Furthermore,

$$1 - \int_{\mathbb{R}^n} \widehat{q}^2 d\nu = \int_{\mathbb{R}^n} (w^2 - \widehat{q}^2) d\nu = \int_{\mathbb{R}^n} (w - \widehat{q})^2 d\nu. \tag{7.27}$$

In particular, $\widehat{h} := \frac{\widehat{q}}{\|\widehat{q}\|_{L^2(\mathbb{R}^n,\nu)}}$ satisfies

$$\|\widehat{q} - \widehat{h}\|_{L^2(\mathbb{R}^n,\nu)} = |\|\widehat{q}\|_{L^2(\mathbb{R}^n,\nu)} - 1| = \frac{|\|\widehat{q}\|^2_{L^2(\mathbb{R}^n,\nu)} - 1|}{\|\widehat{q}\|_{L^2(\mathbb{R}^n,\nu)} + 1} \le \|w - \widehat{q}\|^2_{L^2(\mathbb{R}^n,\nu)}. \tag{7.28}$$

Using (7.26), (7.27), and (7.28), we obtain that for any $\widehat{p} \in \widehat{\mathcal{P}}_m$,

$$\begin{aligned}
|N^w(1) - N^{\widehat{p}}(1)| = |N^w(1) - N^{\widehat{h}}(1)| &= 2\left|\int_{\mathbb{R}^n} (|\nabla w|^2 - |\nabla \widehat{h}|^2) d\nu\right| \\
&\le 2\left|\int_{\mathbb{R}^n} (|\nabla w|^2 - |\nabla \widehat{q}|^2) d\nu\right| + 2\left|\int_{\mathbb{R}^n} (|\nabla \widehat{q}|^2 - |\nabla \widehat{h}|^2)\, d\nu\right| \\
&= 2\left|\int_{\mathbb{R}^n} \nabla(w - \widehat{q}) \cdot \nabla(w + \widehat{q}) d\nu\right| + 2\left|\int_{\mathbb{R}^n} \nabla(\widehat{q} - \widehat{h}) \cdot \nabla(\widehat{q} + \widehat{h}) d\nu\right| \\
&= 2\int_{\mathbb{R}^n} |\nabla(w - \widehat{q})|^2 d\nu + m\left|\int_{\mathbb{R}^n} (\widehat{q}^2 - \widehat{h}^2)\, d\nu\right| \\
&\le 4\int_{\mathbb{R}^n} |\nabla(w - \widehat{p})|^2 d\nu + 4\int_{\mathbb{R}^n} |\nabla(\widehat{p} - \widehat{h})|^2 d\nu + m\int_{\mathbb{R}^n} (w - \widehat{q})^2\, d\nu \\
&\le 4\int_{\mathbb{R}^n} |\nabla(w - \widehat{p})|^2 d\nu + 2m\int_{\mathbb{R}^n} (\widehat{p} - \widehat{h})^2 d\nu + m\int_{\mathbb{R}^n} (w - \widehat{q})^2\, d\nu \\
&\le 4\int_{\mathbb{R}^n} |\nabla(w - \widehat{p})|^2 d\nu + 2m\int_{\mathbb{R}^n} (w - \widehat{p})^2 + 2(w - \widehat{q})^2 + (\widehat{q} - h)^2 d\nu\, d\nu \\
&\le 4\int_{\mathbb{R}^n} |\nabla(w - \widehat{p})|^2 d\nu + 8m\int_{\mathbb{R}^n} (w - \widehat{p})^2\, d\nu.
\end{aligned}$$

□

Now we estimate the other terms in (7.22).

**Lemma 7.16.** *For any $A \in [1, \infty)$ and any $\sigma \in (0, 1]$ the following holds if $\lambda \le \overline{\lambda}(\Lambda)$. There exists $C = C(A, \Lambda, \Upsilon)$ such that for any $s_0 \ge 0$:*

$$\int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} g_s^2\, d\nu ds \le C\lambda^\Upsilon e^{-2\Upsilon s_0}, \tag{7.29a}$$

$$\int_{s_0}^{s_0+A} |\widehat{N}^u(e^{-s}) - m|\, ds \le C\sigma^{-1}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[s_0-\sigma, s_0+A]} + C\lambda^\Upsilon e^{-2\Upsilon s_0}. \tag{7.29b}$$

*Proof.* We first prove (7.29a). Suppose $0 < s_2 < s_1$. By combining (7.23), Cauchy's inequality, and Lemma 7.7 we get

$$\int_{s_2}^{s_1} \int_{\mathbb{R}^n} g_s^2\, d\nu ds \le 2\int_{s_2}^{s_1} \frac{1}{H(e^{-s})} \int_{\mathbb{R}^n} (\widetilde{\tau\Box w})_s^2\, d\nu ds + 2\int_{s_2}^{s_1} \frac{1}{H^w(e^{-s})} \left(\int_{\mathbb{R}^n} \widetilde{w}_s (\widetilde{\tau\Box w})_s\, d\nu\right)^2 ds$$

$$\leq C \int_{e^{-s_1}}^{e^{-s_2}} \frac{\tau}{H(\tau)} \int_{\mathbb{R}^n} (\Box w)^2 \, d\nu_{-\tau} d\tau \leq C(A, \Lambda, \Upsilon)\lambda^{\Upsilon} e^{-2\Upsilon s_2}.$$

We now prove (7.29b). Let $\eta \in C^\infty(\mathbb{R})$ be a cutoff function such that $\eta|_{(-\infty, s_0-\sigma]} \equiv 0$, $\eta|_{[s_0,\infty)} \equiv 1$, and $|\eta'| \leq 10\sigma^{-1}$.

**Claim 7.16.1.** *The following holds:*

$$\begin{aligned} &2 \int_{s_0-A}^{s_0+A} \int_{\mathbb{R}^n} \eta |\nabla(\widetilde{w}_s - \widehat{p})|^2 d\nu ds \\ &\quad \leq C(A, \Lambda, \Upsilon) \int_{s_0-A}^{s_0+A} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu ds + \frac{1}{20} \int_{s_0-A}^{s_0+A} \eta |\widehat{N}^u(e^{-s}) - m| ds + C(A, \Lambda, \Upsilon)\lambda^{\Upsilon} e^{-2\Upsilon s_0}. \end{aligned}$$

*Proof.* We use

$$\partial_s(\widetilde{w}_s - \widehat{p}) = \Delta_f(\widetilde{w}_s - \widehat{p}) + \frac{m}{2}(\widetilde{w}_s - \widehat{p}) + \frac{1}{2}(\widehat{N}^u(e^{-s}) - m)\widetilde{w}_s + g_s, \tag{7.30}$$

Hölder's inequality, and (7.24), to estimate

$$\begin{aligned} \frac{d}{ds} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu \leq &- 2 \int_{\mathbb{R}^n} |\nabla(\widetilde{w}_s - \widehat{p})|^2 d\nu + \left(\frac{m}{2} + C(A, \Lambda, \Upsilon)\right) \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu \\ &+ \frac{1}{C(A, \Lambda, \Upsilon)} (\widehat{N}^u(e^{-s}) - m)^2 + \int_{\mathbb{R}^n} g_s^2 \, d\nu. \end{aligned} \tag{7.31}$$

Then integrating (7.31) in time against $\eta$ and applying Lemma 7.16 yields

(7.32)

$$\begin{aligned} &\sup_{s \in [s_0, s_0+A]} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu + 2 \int_{s_0-A}^{s_0+A} \int_{\mathbb{R}^n} \eta |\nabla(\widetilde{w}_s - \widehat{p})|^2 d\nu ds \\ &\quad \leq 2 \int_{s_0-A}^{s_0+A} \left(|\eta'| + \left(\frac{m}{2} + 20\right)\eta\right) \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu ds + \frac{1}{C(A, \Lambda, \Upsilon)} \int_{s_0-A}^{s_0+A} \eta |\widehat{N}^u(e^{-s}) - m|^2 ds \\ &\quad\quad + 2 \int_{s_0-A}^{s_0+A} \int_{\mathbb{R}^n} g_s^2 d\nu ds. \end{aligned}$$

□

For any $\widehat{p} \in \widehat{\mathcal{P}}_m$, we can use Lemma 7.15 with $w \leftarrow \widetilde{w}_s$ to obtain

$$|\widehat{N}^u(e^{-s}) - m| = |N^{\widetilde{w}_s}(1) - N^{\widehat{p}}(1)| \leq 4 \int_{\mathbb{R}^n} |\nabla(\widetilde{w}_s - \widehat{p})|^2 d\nu + 8m \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu. \tag{7.33}$$

Integrating this in time against $\eta$, applying Claim 7.16.1 and rearranging terms yields the desired estimate:

(7.34)

$$\int_{s_0-A}^{s_0+A} \eta(s) |\widehat{N}^u(e^{-s}) - m| ds \leq C(A, \Lambda, \Upsilon)\sigma^{-1} \int_{s_0-\sigma}^{s_0+A} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu ds + C(A, \Lambda, \Upsilon)\lambda^{\Upsilon} e^{-2\Upsilon s_0}.$$

□

In the lemma below, we improve the closeness to a polynomial in Proposition 7.14 up to higher-order derivatives. We need the following definition:

$$\|v\|_{H^{-1}(\mathbb{R}^n,\nu)} := \sup\left\{\int_{\mathbb{R}^n} vw \, d\nu \colon \int_{\mathbb{R}^n} (w^2 + |\nabla w|^2) d\nu \leq 1\right\}.$$

**Lemma 7.17.** *For any $A \in [1, \infty)$ and any $\sigma \in (0, 1]$, the following holds if $\lambda \leq \overline{\lambda}(\Lambda)$.*

*(1) There exists $C = C(A, \Lambda, \Upsilon)$ such that for any $p \in \mathcal{P}_m$ and $s_0 \geq 0$, we have*

$$\sup_{s\in[s_0,s_0+A]} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu + \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} |\nabla(\widetilde{w}_s - \widehat{p})|^2 d\nu ds + \int_{s_0}^{s_0+A} \|\partial_s(\widetilde{w}_s - \widehat{p})\|^2_{H^{-1}(\mathbb{R}^n,\nu)} ds$$
$$\leq C(A, \Lambda, \Upsilon)(\sigma^{-1}\|\widetilde{w} - \widehat{p}\|^2_{[s_0-\sigma,s_0+A]} + \lambda^{\Upsilon} e^{-2\Upsilon s_0}).$$

*(2) For any $\epsilon \in (0, 1]$, if $s_0 \geq 0$ and $\widehat{N}^u_{\mathbf{x}_0}(e^{-(s_0+A)}) > \widehat{N}^u_{\mathbf{x}_0}(e^{-s_0}) - \epsilon$, then there exists $m \in \mathbb{N}_0$ such that for any $\widehat{p} \in \mathcal{P}_m$ we have*

(7.35)
$$\int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} |\partial_s(\widetilde{w}_s - \widehat{p})|^2 + |\nabla^2(\widetilde{w}_s - \widehat{p})|^2 \, d\nu ds \leq C\epsilon + C(A, \Lambda, \Upsilon) \left( \|\widetilde{w} - \widehat{p}\|^2_{[s_0,s_0+A]} + \lambda^{\frac{\Upsilon}{2}} e^{-\Upsilon s_0} \right).$$

*Proof.* Using (7.32), (7.34) and recalling that $\eta \in C^\infty(\mathbb{R})$ is a cutoff function satisfying $\eta|_{(-\infty,s_0-\sigma]} \equiv 0$, $\eta|_{[s_0,\infty)} \equiv 1$, and $|\eta'| \leq 10\sigma^{-1}$, we get the estimate for the first two terms. It remains to bound the third term.

For any $\phi \in C_c^\infty(\mathbb{R}^n)$ with $\int_{\mathbb{R}^n} (\phi^2 + |\nabla\phi|^2) \, d\nu \leq 1$, we use (7.30) to estimate

$$\begin{aligned}
\left| \int_{\mathbb{R}^n} \partial_s(\widetilde{w}_s - \widehat{p})\phi \, d\nu \right| &\leq \left| \int_{\mathbb{R}^n} \nabla(\widetilde{w}_s - \widehat{p}) \cdot \nabla\phi \, d\nu \right| + \frac{1}{2}|\widehat{N}^u(e^{-s}) - m| \left| \int_{\mathbb{R}^n} \widetilde{w}_s \phi \, d\nu \right| \\
&\quad + \frac{m}{2} \left| \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})\phi \, d\nu \right| + \int_{\mathbb{R}^n} g_s \phi \, d\nu \\
&\leq \left( \int_{\mathbb{R}^n} |\nabla(\widetilde{w}_s - \widehat{p})|^2 \, d\nu \right)^{\frac{1}{2}} + \frac{1}{2}|\widehat{N}^u(e^{-s}) - m| \\
&\quad + \frac{m}{2} \left( \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 d\nu \right)^{\frac{1}{2}} + \int_{\mathbb{R}^n} g_s^2 \, d\nu.
\end{aligned}$$

We then take the supremum over all such $\phi$, square both sides, integrate in time, and apply (7.32) and Lemma 7.16 to obtain so that

$$\begin{aligned}
\int_{s_0}^{s_0+A} \|\widetilde{w}_s - \widehat{p}\|^2_{H^{-1}(\mathbb{R}^n,\nu)} \, ds \leq& C \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} |\nabla(\widetilde{w}_s - \widehat{p})|^2 \, d\nu ds + \frac{1}{2} \int_{s_0}^{s_0+A} (\widehat{N}^u(e^{-s}) - m)^2 \, ds \\
&+ C \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} g_s^2 \, d\nu ds + C(\Lambda) \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 \, d\nu ds \\
\leq& C(A, \Lambda, \Upsilon)\sigma^{-1} \int_{s_0-\sigma}^{s_0+A} \int_{\mathbb{R}^n} (\widetilde{w}_s - \widehat{p})^2 \, d\nu ds + C(A, \Lambda, \Upsilon)\lambda^{\Upsilon} e^{-2\Upsilon s_0}.
\end{aligned}$$

To prove (7.35), we use the $f$-Bochner formula (2.5) with $w \leftarrow \widetilde{w} - \widehat{p}$ and Proposition 7.8(2) to get

$$\begin{aligned}
\int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} |\nabla^2(\widetilde{w} - p)|^2 \, d\nu ds &\leq \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} \left( \Delta_f \widetilde{w} + \frac{m}{2}\widehat{p} \right)^2 \, d\nu ds \\
&\leq \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} \left( \Delta_f \widetilde{w} + \frac{m}{2}\widetilde{w} \right)^2 \, d\nu ds + 2m^2 \int_{s_0}^{s_0+A} \int_{\mathbb{R}^n} (\widetilde{w} - \widehat{p})^2 \, d\nu ds \\
&\leq C\epsilon + C(A, \Lambda, \Upsilon) \left( \|\widetilde{w} - \widehat{p}\|^2_{[s_0,s_0+A]} + \lambda^{\frac{\Upsilon}{2}} e^{-\Upsilon s_0} \right).
\end{aligned}$$

Then we combine this with (7.9b) in Lemma 7.7 to get (7.35). $\square$

We now carry out the stability analysis of caloric approximation. Below, we prove that fast growth of closeness to $\mathcal{P}_m$ propagates down the scale but fast decay of closeness to $\mathcal{P}_m$ propagates up the scale. The strategy of our proof is similar to that of [CT94, Lemma 5.31] (see also [Sim85, §II.3]).

**Proposition 7.18** (Growth or decay)**.** *For any $A \in [1, \infty)$, the following holds when $\lambda \leq \overline{\lambda}(A, \Lambda)$ and $\epsilon \leq \overline{\epsilon}(A, \Lambda)$. If*

$$\lambda^{\frac{\Upsilon}{4}} e^{-\Upsilon s} \leq \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+A, s+3A]} < \epsilon$$

*for some $s \geq 0$, then we have the following:*

*(1) If $\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+A,s+3A]} \geq e^{\frac{1}{2}A} \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s,s+2A]}$, then*

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+2A,s+4A]} \geq e^{\frac{1}{2}A} \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+A,s+3A]}. \tag{7.36}$$

*(2) If $\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+2A,s+4A]} \leq e^{-\frac{1}{2}A} \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+A,s+3A]}$, then*

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s+A,s+3A]} \leq e^{-\frac{1}{2}A} \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[s,s+2A]}. \tag{7.37}$$

*(3) Either (7.36) or (7.37) holds.*

*Proof.* First we prove the Proposition in the linear setting.

**Claim 7.18.1.** *For any $A \geq 1$ and $v \in C^\infty(\mathbb{R}^n \times (0, 4A])$, such that*

$$\partial_s v = \Delta_f v + \frac{m}{2} v, \qquad\qquad \int_{\mathbb{R}^n} v(\cdot, s) h \, d\nu = 0 \text{ for all } h \in \widehat{\mathcal{P}}_m, \tag{7.38}$$

*and that $v(\cdot, s) \in L^2(\mathbb{R}^n, \nu)$ for almost every $s \in (0, 4A]$ and $v \in L^2(\mathbb{R}^n \times (0, 4A], d\nu ds)$, the conclusions of Proposition 7.18 hold.*

*Proof of Claim 7.18.1.* Using Lemma A.2, we can decompose solutions $v$ to (7.38) as $v(\cdot, s) \coloneqq \sum_{k \neq m} e^{-\frac{(k-m)}{2}s} \widehat{p}_k$ into homogeneous caloric polynomials $\widehat{p}_k \in \widehat{\mathcal{P}}_k$, hence $\|v(\cdot, s) - \widehat{\mathcal{P}}_m\|_{L^2(\mathbb{R}^n, \nu)} = \|v(\cdot, s)\|_{L^2(\mathbb{R}^n, \nu)}$ for all $s \in [0, 4A]$. Write $v(\cdot, s) = v^+(\cdot, s) + v^-(\cdot, s)$, where

$$v^+(\cdot, s) \coloneqq \sum_{k<m} e^{-\frac{(k-m)}{2}s} \widehat{p}_k, \quad v^-(\cdot, s) \coloneqq \sum_{k>m} e^{-\frac{(k-m)}{2}s} \widehat{p}_k.$$

We first prove (1). Assume that $\int_A^{3A} \int_{\mathbb{R}^n} v^2 \, d\nu ds \geq e^{\frac{1}{2}A} \int_0^{2A} \int_{\mathbb{R}^n} v^2 \, d\nu ds$, so that

$$\begin{aligned} \int_A^{3A} \int_{\mathbb{R}^n} (v^+)^2 \, d\nu ds &\geq e^{\frac{1}{2}A} \int_0^{2A} \int_{\mathbb{R}^n} (v^-)^2 \, d\nu ds - \int_A^{3A} \int_{\mathbb{R}^n} (v^-)^2 \, d\nu ds \\ &\geq (e^{\frac{3}{2}A} - 1) \int_A^{3A} \int_{\mathbb{R}^n} (v^-)^2 \, d\nu ds. \end{aligned}$$

Using this, we can see that

$$\begin{aligned} \int_{2A}^{4A} \int_{\mathbb{R}^n} v^2 \, d\nu ds \geq \int_{2A}^{4A} \int_{\mathbb{R}^n} (v^+)^2 \, d\nu ds &\geq e^A \int_A^{3A} \int_{\mathbb{R}^n} (v^+)^2 \, d\nu ds \\ &\geq e^{\frac{1}{2}A} \int_A^{3A} \int_{\mathbb{R}^n} v^2 \, d\nu ds, \end{aligned}$$

which proves (1). The proof of (2) is similar.

We now prove (3). First consider the case where $\int_A^{3A} \int_{\mathbb{R}^n} (v^+)^2 \, d\nu ds \geq \frac{1}{2} \int_A^{3A} \int_{\mathbb{R}^n} v^2 \, d\nu ds$. Then

$$\begin{aligned} \int_{2A}^{4A} \int_{\mathbb{R}^n} v^2 \, d\nu \, ds \geq \int_{2A}^{4A} \int_{\mathbb{R}^n} (v^+)^2 \, d\nu ds &\geq e^A \int_A^{3A} \int_{\mathbb{R}^n} (v^+)^2 \, d\nu ds \\ &\geq e^{\frac{1}{2}A} \int_A^{3A} \int_{\mathbb{R}^n} v^2 \, d\nu ds, \end{aligned}$$

hence (7.36) holds. If instead $\int_A^{3A}\int_{\mathbb{R}^n}(v^-)^2\,d\nu ds \geq \frac{1}{2}\int_A^{3A}\int_{\mathbb{R}^n} v^2\,d\nu ds$, then we similarly have

$$\int_0^{2A}\int_{\mathbb{R}^n} v^2\,d\nu ds \geq \frac{1}{2}e^A\int_A^{3A}\int_{\mathbb{R}^n} v^2\,d\nu ds,$$

so that (7.37) holds. $\square$

Now we prove the general case. Suppose by way of contradiction that there are $\epsilon_i, \lambda_i \to 0$, solutions $u^i$ to (7.1) (with $\lambda \leftarrow \lambda_i$) satisfying (7.2) so that $(\widetilde{w}^i_s := (\widetilde{\chi u^i})_s)_{s\in[0,\infty)}$ are solutions to (7.22), along with $s_i \in [0,\infty)$ and

$$\lambda_i^{\frac{\Upsilon}{2}} e^{-\Upsilon s_i} \leq \|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+A,s_i+3A]} < \epsilon_i, \tag{7.39}$$

yet one of the following holds for all $i \in \mathbb{N}_0$:

(1) $\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+A,s_i+3A]} \geq e^{\frac{1}{2}A}\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i,s_i+2A]}$, but

$$\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+2A,s_i+4A]} < e^{\frac{1}{2}A}\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+A,s_i+3A]}, \tag{7.40}$$

(2) $\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+2A,s_i+4A]} \leq e^{-\frac{1}{2}A}\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+A,s_i+3A]}$, but

$$\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+A,s_i+3A]} > e^{-\frac{1}{2}A}\|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i,s_i+2A]}, \tag{7.41}$$

(3) Both (7.40) and (7.41) hold.

For each $i \in \mathbb{N}_0$, choose $\widehat{q}_i \in \widehat{\mathcal{P}}_m$ such that

$$\begin{aligned} \kappa_i &:= \|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|_{[s_i+A,s_i+3A]} = \left(\int_{s_i+A}^{s_i+3A}\|\widetilde{w}^i_s - \widehat{q}_i\|^2_{L^2(\mathbb{R}^n,\nu)}ds\right)^{\frac{1}{2}}, \\ &\lim_{i\to\infty}\kappa_i = 0, \quad \lim_{i\to\infty}\kappa_i^{-1}\lambda_i^{\frac{\Upsilon}{2}}e^{-\Upsilon s_i} = 0. \end{aligned} \tag{7.42}$$

Set $v^i_s := \kappa_i^{-1}(\widetilde{w}^i_{s_i+s} - \widehat{q}^i)$, so that

$$\partial_s v^i_s = \Delta_f v^i_s + \frac{m}{2}v^i_s + b^i_s\widetilde{w}^i_{s_i+s} + h^i_s,$$

where

$$b^i_s := \frac{1}{2}\kappa_i^{-1}(N^{\widetilde{w}_{s_i+s}}(1) - m), \quad h^i_s := \kappa_i^{-1}g^i_{s_i+s},$$

as well as

$$\int_A^{3A}\int_{\mathbb{R}^n}(v^i_s)^2\,d\nu ds = \|v^i - \widehat{\mathcal{P}}_m\|_{[A,3A]} = 1. \tag{7.43}$$

If (1) holds, then we have

$$\|v^i - \widehat{\mathcal{P}}_m\|_{[0,2A]} \leq e^{-\frac{1}{2}A}, \qquad \|v^i - \widehat{\mathcal{P}}_m\|_{[2A,4A]} < e^{\frac{1}{2}A}. \tag{7.44}$$

By (7.43), (7.44), and the argument of Remark 7.12 (applied twice)

$$\kappa_i^{-1}\|\widetilde{w}^i - \widehat{q}^i\|_{[s_i,s_i+4A]} = \|v^i\|_{[0,4A]} \leq C(A). \tag{7.45}$$

Therefore, after passing to a subsequence, we can get a weak convergence $v^i \rightharpoonup v$ for some $v \in L^2(\mathbb{R}^n \times [0,4A], d\nu ds)$.

Furthermore, for any $\sigma \in (0,1]$, we can apply (7.45) and Lemma 7.17 with $\widetilde{w} \leftarrow \widetilde{w}^i$, $s_0 \leftarrow s_i + \sigma$, and $A \leftarrow 4A - \sigma$, we obtain

$$
\begin{aligned}
&\sup_{s\in[\sigma,4A]} \|v_s^i\|^2_{L^2(\mathbb{R}^n,\nu)} + \int_\sigma^{4A} \|v_s^i\|^2_{H^1(\mathbb{R}^n,\nu)} ds + \int_\sigma^{4A} \|\partial_s v_s^i\|^2_{H^{-1}(\mathbb{R}^n,\nu)} ds \\
&= \kappa_i^{-2} \sup_{s\in[s_i+\sigma,s_i+4A]} \|\widetilde{w}_s^i - \widehat{q}^i\|^2_{L^2(\mathbb{R}^n,\nu)} + \kappa_i^{-2} \int_{s_i+\sigma}^{s_i+4A} \|\widetilde{w}_s^i - \widehat{q}^i\|^2_{H^1(\mathbb{R}^n,\nu)} ds \\
&\quad + \kappa_i^{-2} \int_{s_i+\sigma}^{s_i+4A} \|\partial_s \widetilde{w}_s^i - \widehat{q}^i\|^2_{H^{-1}(\mathbb{R}^n,\nu)} ds \\
&\leq \kappa_i^{-2} C(A,\Lambda,\Upsilon)(\sigma^{-1}\|\widetilde{w}^i - \widehat{q}^i\|^2_{[s_i,s_i+4A]} + \lambda_i^\Upsilon e^{-2\Upsilon s_i}) \leq C(A,\Lambda,\Upsilon)(\sigma^{-1} + \kappa_i^{-2}\lambda_i^\Upsilon e^{-2\Upsilon s_i}) \\
&\leq C(A,\Lambda,\Upsilon)\sigma^{-1},
\end{aligned}
\tag{7.46}
$$

for sufficiently large $i$, where we used (7.42) for the last inequality. Because $H^1(\mathbb{R}^n,\nu) \hookrightarrow L^2(\mathbb{R}^n,\nu)$ is compact (see [BGL14, Theorem 6.4.2]) and $L^2(\mathbb{R}^n,\nu) \hookrightarrow H^{-1}(\mathbb{R}^n,\nu)$ is continuous, it follows from the Aubin–Lions–Simon lemma (see [BF12, Theorem II.5.16], [Aub63] or [Sim86, Theorem 5]) that $(v^i)_{i\in\mathbb{N}}$ has a subsequence converging (strongly) to $v$ in $L^2([\sigma,4A], L^2(\mathbb{R}^n,\nu))$.

**Claim 7.18.2.** *The limit $v \in C^\infty(\mathbb{R}^n \times (0,4A)) \cap L^2(\mathbb{R}^n \times (0,4A))$ satisfies*

$$\partial_s v_s = \Delta_f v_s + \frac{m}{2} v_s.$$

*Proof of Claim 7.18.2.* By Lemma 7.16, we have

$$\int_0^{4A} \int_{\mathbb{R}^n} (h_s^i)^2 \, d\nu ds = \kappa_i^{-2} \int_{s_i}^{s_i+4A} \int_{\mathbb{R}^n} (g_s^i)^2 \, d\nu ds \leq C(A,\Lambda,\Upsilon)\kappa_i^{-2}\lambda_i^\Upsilon e^{-2\Upsilon s_i},$$

so because of (7.42), we have $\lim_{i\to\infty} \int_0^{4A} \int_{\mathbb{R}^n} h_s^i \varphi \, d\nu ds = 0$ for any $\varphi \in C_c^\infty(\mathbb{R}^n \times (0,4A))$.

Similarly, Lemma 7.16 and (7.45) give

$$
\begin{aligned}
\int_\sigma^{4A} |b_s^i| \, ds &= \frac{1}{2}\kappa_i^{-1} \int_{s_i+\sigma}^{s_i+4A} |\widehat{N}^{u^i}(e^{-s}) - m| \, ds \leq C\kappa_i^{-1} \|\widetilde{w}^i - \widehat{\mathcal{P}}_m\|^2_{[s_i,s_i+4A]} + C\lambda_i^\Upsilon \kappa_i^{-1} e^{-2\Upsilon s_i} \\
&\leq C(A,\Lambda,\Upsilon)\sigma^{-1}\kappa_i^{-1} \|\widetilde{w}^i - \widehat{q}^i\|^2_{[s_i,s_i+4A]} + C\kappa_i^{-1}\lambda_i^\Upsilon e^{-2\Upsilon s_i} \\
&\leq C(A,\Lambda,\Upsilon)\sigma^{-1}\kappa_i + C(A,\Lambda,\Upsilon)\kappa_i^{-1}\lambda_i^\Upsilon e^{-2\Upsilon s_i}
\end{aligned}
$$

for any $\sigma > 0$. By (7.42), we thus have $\lim_{i\to\infty} \int_\sigma^{4A} |b_s^i| \, ds = 0$. Thus recalling that $\int_{\mathbb{R}^n} (\widetilde{w}_s^i)^2 \, d\nu = 1$ (see (7.24)), we have

$$\limsup_{i\to\infty} \int_0^{4A} \int_{\mathbb{R}^n} b_s^i \widetilde{w}_{s_i+s}^i \varphi \, d\nu ds \leq \|\varphi\|_{C^0(\mathbb{R}^n)} \limsup_{i\to\infty} \int_\sigma^{4A} |b_s^i| \, ds = 0$$

for any $\varphi \in C_c^\infty(\mathbb{R}^n \times (0,4A))$.

Therefore, for any $\varphi \in C_c^\infty(\mathbb{R}^n \times (0,4A))$ we have

$$
\begin{aligned}
0 &= \lim_{i\to\infty} \int_0^{4A} \int_{\mathbb{R}^n} v_s^i \partial_s \varphi + v_s^i \Delta_f \varphi + \left(\frac{m}{2} v_s^i + b_s^i \widetilde{w}_{s_i+s}^i + h_s^i\right) \varphi \, d\nu ds \\
&= \int_0^{4A} \int_{\mathbb{R}^n} v_s \partial_s \varphi + v_s \Delta_f \varphi + \frac{m}{2} v_s \varphi \, d\nu ds.
\end{aligned}
$$

In particular, $v$ is a weak solution. By parabolic regularity, we know that $v$ is a strong solution and is smooth. ∎

**Claim 7.18.3.** *The limit $v$ satisfies*

$$\operatorname{esssup}_{s\in[\sigma,4A]}\int_{\mathbb{R}^n} v^2(\cdot,s)\,d\nu \le C(A,\Lambda,\Upsilon)\sigma^{-1}.$$

*Proof of 7.18.3.* After passing to a subsequence, we can assume that $v^i(\cdot,s)\to v(\cdot,s)$ in $L^2(\mathbb{R}^n,\nu)$ for almost-every $s\in[\sigma,4A]$. The claim then follows from (7.46). ☐

Because $v^i\to v$ strongly in $L^2(\mathbb{R}^n\times[A,3A],d\nu ds)$, we also have

$$\|v\|^2_{[A,3A]}=1,\qquad \int_A^{3A}\int_{\mathbb{R}^n} v_s h\,d\nu ds=0 \tag{7.47}$$

for all $h\in\widehat{\mathcal{P}}_m$, hence $\|v-\mathcal{P}_m\|_{[A,3A]}=1$.

**Claim 7.18.4.** *The limit $v$ satisfies*

$$\|v-\widehat{\mathcal{P}}_m\|_{[0,2A]}\le e^{-\frac12 A},\qquad \|v-\widehat{\mathcal{P}}_m\|_{[2A,4A]}\le e^{\frac12 A}.$$

*Proof of Claim 7.18.4.* By (7.44), there exist $\widehat{p}^i\in\widehat{\mathcal{P}}_m$ such that

$$\int_0^{2A}\int_{\mathbb{R}^n}(v^i-\widehat{p}^i)^2 d\nu ds\le e^{-\frac12 A}$$

for all $i\to\infty$. Then $\|\widehat{p}^i\|^2_{L^2(\mathbb{R}^n)}\le A^{-1}\|v^i\|^2_{[A,2A]}+e^{-\frac12 A}\le A^{-1}$, so we can pass to a subsequence to ensure that $\widehat{p}^i\to\widehat{p}\in\widehat{\mathcal{P}}_m$. In particular, $v^i-\widehat{p}^i\to v-\widehat{p}$ weakly in $L^2(\mathbb{R}^n\times[0,2A],\nu ds)$, so

$$e^{-\frac{A}{2}}\ge\limsup_{i\to\infty}\|v^i-\widehat{p}^i\|_{[0,2A]}\ge\|v-\widehat{p}\|_{[0,2A]}\ge\|v-\widehat{\mathcal{P}}_m\|_{[0,2A]}.$$

The proof that $\|v-\widehat{\mathcal{P}}_m\|_{[2A,4A]}\le e^{\frac12 A}$ is similar, so we omit it. ☐

Then Claim 7.18.3, Claim 7.18.2, (7.47), and Claim 7.18.4 contradict Claim 7.18.1. ☐

*Proof of Theorem 7.13.* Choose $\epsilon\le\overline{\epsilon}(A,\Lambda)$ and $\lambda\le\overline{\lambda}(A,\epsilon,\Lambda)$ as in Proposition 7.18. It suffices to prove the claim when $s_1=k_1A$ and $s_2=k_2A$ for some $k_1,k_2\in\mathbb{N}_0$.

Suppose $s\in[kA,(k+1)A]\subset[k_2A,k_1A]$. Then Proposition 7.8 implies that

$$\begin{aligned}|\widehat{N}^u(e^{-s})-m| &= (\widehat{N}^u(e^{-s})-m)_+ + (\widehat{N}^u(e^{-s})-m)_-\\ &\le (\widehat{N}^u(e^{-kA})-m)_+ + (\widehat{N}^u(e^{-(k+1)A})-m)_- + C(A,\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}e^{-\Upsilon s}\\ &\le |\widehat{N}^u(e^{-kA})-m| + |\widehat{N}^u(e^{-(k+1)A})-m| + C(A,\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}e^{-\Upsilon s}.\end{aligned}$$

By this and Remark 7.12, it suffices to prove the claim when $s=kA$, $s_1=k_1A$, $s_2=k_2A$ for some $k,k_1,k_2\in\mathbb{N}_0$.

**Claim 7.18.5.** *For all integers $k\in[k_2,k_1]$, we have*

$$\|\widetilde{w}-\widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]}<C(A,\Lambda,\Upsilon)\delta\left(e^{-\frac{\Upsilon}{2}A(k_1-k)}+e^{-\frac{\Upsilon}{2}A(k-k_2)}\right)+\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-2)A}.$$

*Proof of Claim 7.18.5.* Suppose, for contradiction, that for some $k\in[k_2,k_1]$, there exists $B(A,\Lambda,\Upsilon)$ to be determined such that

$$\|\widetilde{w}-\widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]}\ge B\delta\left(e^{-\frac{\Upsilon}{2}A(k_1-k)}+e^{-\frac{\Upsilon}{2}A(k-k_2)}\right)+\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-2)A}\ge\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-2)A}. \tag{7.48}$$

By (7.17), we have $\widehat{N}^u(e^{-(\ell-1)A})-\widehat{N}^u(e^{-(\ell+1)A})\le 2\delta$ for any integer $\ell\in[k_2,k_1]$. Thus Proposition 7.14 gives

$$\|\widetilde{w}-\widehat{\mathcal{P}}_m\|_{[(\ell-1)A,(\ell+1)A]}<C(A,\Lambda,\Upsilon)\left(\delta+\lambda^{\frac{\Upsilon}{2}}e^{-\ell\Upsilon A}\right)<\epsilon \tag{7.49}$$

if we take $\lambda \le \overline{\lambda}(A, \epsilon, \Upsilon)$ and $\delta \le \overline{\delta}(A, \epsilon, \Upsilon)$. Therefore, using (7.48) and (7.49) and Proposition 7.18 with $s \leftarrow (k-2)A$, we can infer that either

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[kA,(k+2)A]} \ge e^{\frac{1}{2}A}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]} \tag{7.50}$$

or else

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]} \le e^{-\frac{1}{2}A}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-2)A,kA]}. \tag{7.51}$$

First assume (7.50) holds. We claim that

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k+\ell)A,(k+\ell+2)A]} \ge e^{\frac{A}{2}}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k+\ell-1)A,(k+\ell+1)A]} \tag{7.52}$$

for any integer $\ell \in [0, k_1 - k]$. Note that the claim is true when $\ell = 0$ by (7.50). Suppose we have shown (7.52) for all $\ell < \ell_0$, where $\ell_0 < k_1 - k$. Then we use (7.48) to get

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k+\ell_0-1)A,(k+\ell_0+1)A]} \ge e^{\frac{A}{2}\ell_0}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]} \ge e^{\frac{A}{2}\ell_0}\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-2)A} \ge \lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k+\ell_0)A}.$$

We then apply Proposition 7.18(1) with $s \leftarrow (k + \ell_0 - 2)A$ to get

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k+\ell_0)A,(k+\ell_0+2)A]} \ge e^{\frac{A}{2}}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k+\ell_0-1)A,(k+\ell_0+1)A]}. \tag{7.53}$$

Thus (7.52) holds by induction.

Applying (7.52) for all integers $\ell \in [0, k_1 - k]$ and (7.48) yield

$$\begin{aligned}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k_1-1)A,(k_1+1)A]} &\ge e^{\frac{A}{2}(k_1-k)}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]} \\ &\ge B\delta e^{\frac{A}{2}(k_1-k)-\frac{\Upsilon}{2}A(k_1-k)} + \lambda^{\frac{\Upsilon}{4}}e^{\frac{A}{2}(k_1-k)-\Upsilon(k-2)A} \ge B\delta + \lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon k_1 A}.\end{aligned}$$

If we choose $\lambda \le \overline{\lambda}(A, \epsilon, \Lambda, \Upsilon)$ and $B \ge \underline{B}(A, \epsilon, \Lambda, \Upsilon)$, then (7.49) is contradicted. Therefore (7.51) holds.

We claim that

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-\ell-1)A,(k-\ell+1)A]} \le e^{-\frac{A}{2}}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-\ell-2)A,(k-\ell)A]} \tag{7.54}$$

for any integer $\ell \in [0, k - k_2 - 1]$. Note that the claim holds for $\ell = 0$ by (7.51). Suppose we have shown (7.54) for all $\ell < \ell_0$, where $\ell_0 < k - k_2 + 1$. Then we use (7.48) to get

$$\begin{aligned}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-\ell_0-1)A,(k-\ell_0+1)A]} &\ge e^{\frac{A}{2}\ell_0}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]} \ge e^{\frac{A}{2}\ell_0}\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-2)A} \\ &\ge \lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-\ell_0-2)A}.\end{aligned}$$

We may therefore apply Proposition 7.18(2) with $s \leftarrow (k - \ell_0 - 2)A$ and (7.54) with $\ell \leftarrow \ell_0 - 1$ to get

$$\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-\ell_0-1)A,(k-\ell_0+1)A]} \le e^{-\frac{A}{2}}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-\ell_0-2)A,(k-\ell_0)A]}. \tag{7.55}$$

Thus (7.54) holds by induction.

Therefore, (7.54) for all integers $\ell \in [0, k - k_2 - 1]$ and (7.48) yield

$$\begin{aligned}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k_2-1)A,(k_2+1)A]} &\ge e^{\frac{A}{2}(k-k_2)}\|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]} \\ &\ge e^{\frac{A}{2}(k-k_2)}B\delta e^{-\frac{\Upsilon}{2}A(k-k_2)} + \lambda^{\frac{\Upsilon}{4}}e^{\frac{A}{2}(k-k_2)}e^{-\Upsilon(k-2)A} \ge B\delta + \lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon k_2 A}.\end{aligned}$$

This contradicts (7.49) if $B \ge \underline{B}(A, \epsilon, \Lambda, \Upsilon)$ and $\lambda \le \overline{\lambda}(A, \epsilon, \Lambda, \Upsilon)$. Therefore, (7.48) fails and the claim holds. □

For each integer $k \in \{k_2, \ldots, k_1\}$, we choose $\widehat{p}^k \in \widehat{\mathcal{P}}_m$ satisfying $\|\widetilde{w} - \widehat{p}^k\|_{[(k-1)A,(k+1)A]} := \|\widetilde{w} - \widehat{\mathcal{P}}_m\|_{[(k-1)A,(k+1)A]}$. By (7.16) and Claim 7.18.5, we then have

$$\begin{aligned}\|\widehat{p}^k - \widehat{p}^{k+1}\|_{L^2(\mathbb{R}^n,\nu)} &\le \|\widetilde{w} - \widehat{p}^k\|_{[(k-1)A,(k+1)A]} + \|\widetilde{w} - \widehat{p}^k\|_{[kA,(k+2)A]} \\ &< C\delta\left(e^{-\frac{\Upsilon}{2}A(k_1-k)} + e^{-\frac{\Upsilon}{2}A(k-k_2)}\right) + 2\lambda^{\frac{\Upsilon}{4}}e^{-\Upsilon(k-2)A},\end{aligned}$$

where $C = C(A,\Lambda,\Upsilon)$. Fix $k_0 := \lfloor \frac{k_2+k_1}{2} \rfloor$. For any integer $k \in [k_2, k_0]$ we have

$$\begin{aligned} \|\widehat{p}^k - \widehat{p}^{k_0}\|_{L^2(\mathbb{R}^n,\nu)} &\le C \sum_{i=k}^{k_0} \left( \delta \left( e^{-\frac{\Upsilon}{2}A(k_1-i)} + e^{-\frac{\Upsilon}{2}(i-k_2)A} \right) + \lambda^{\frac{\Upsilon}{4}} e^{-\Upsilon(i-2)A} \right) \\ &\le C\delta \left( e^{-\frac{\Upsilon}{4}A(k_1-k_2)} + e^{-\frac{\Upsilon}{2}(k-k_2)A} \right) + C\lambda^{\frac{\Upsilon}{4}} e^{-\Upsilon kA}. \end{aligned} \tag{7.56}$$

On the other hand, for any $k \in [k_0, k_1]$, we have

$$\begin{aligned} \|\widehat{p}^k - \widehat{p}^{k_0}\|_{L^2(\mathbb{R}^n,\nu)} &\le C \sum_{i=k_0}^{k} \left( \delta \left( e^{-\frac{\Upsilon}{2}A(k_1-i)} + e^{-\frac{\Upsilon}{2}(i-k_2)A} \right) + \lambda^{\frac{\Upsilon}{4}} e^{-\Upsilon(i-2)A} \right) \\ &\le C\delta \left( e^{-\frac{\Upsilon}{2}A(k_1-k)} + e^{-\frac{\Upsilon}{4}(k_1-k_2)A} \right) + C\lambda^{\frac{\Upsilon}{4}} e^{-\frac{\Upsilon}{2}(k_1+k_2)A}. \end{aligned} \tag{7.57}$$

Combining the estimates (7.56) and (7.57), for any integer $k \in [k_2, k_1]$, we get

$$\|\widehat{p}^k - \widehat{p}^{k_0}\|_{L^2(\mathbb{R}^n,\nu)} \le C\delta \left( e^{-\frac{\Upsilon}{4}A(k_1-k)} + e^{-\frac{\Upsilon}{4}(k-k_2)A} \right) + C\lambda^{\frac{\Upsilon}{4}} \left( e^{-\Upsilon kA} + e^{-\frac{\Upsilon}{2}(k_1+k_2)A} \right).$$

Therefore,

$$\begin{aligned} \|\widetilde{w} - \widehat{p}^{k_0}\|_{[(k-1)A,(k+1)A]} &\le \|\widetilde{w} - \widehat{p}^{k}\|_{[(k-1)A,(k+1)A]} + \|\widehat{p}^k - \widehat{p}^{k_0}\|_{[(k-1)A,(k+1)A]} \\ &\le C\delta \left( e^{-\frac{\Upsilon}{4}A(k_1-k)} + e^{-\frac{\Upsilon}{4}(k-k_2)A} \right) + C\lambda^{\frac{\Upsilon}{4}} \left( e^{-\Upsilon kA} + e^{-\frac{\Upsilon}{2}(k_1+k_2)A} \right). \end{aligned}$$

This proves (7.18a).

We now prove (7.18b). By Lemma 7.16

$$\begin{aligned} \int_{kA}^{(k+1)A} |\widehat{N}^u(e^{-s}) - m|\, ds &\le C\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[kA-1,(k+1)A]} + C\lambda^{\Upsilon} e^{-2\Upsilon kA} \\ &\le C\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[(k-1)A,(k+1)A]} + C\lambda^{\Upsilon} e^{-2\Upsilon kA}. \end{aligned}$$

Using this along with Proposition 7.8 and Claim 7.18.1, we get

$$\begin{aligned} N(e^{-kA}) - m &\le \fint_{kA}^{(k+1)A} (N(e^{-s}) - m)\, ds + C(A,\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}} e^{-\Upsilon kA} \\ &\le C\|\widetilde{w} - \widehat{\mathcal{P}}_m\|^2_{[(k-1)A,(k+1)A]} + C\lambda^{\frac{\Upsilon}{2}} e^{-\Upsilon kA} \\ &\le C(A,\Lambda,\Upsilon)\delta \left( e^{-\frac{\Upsilon}{2}A(k_1-k)} + e^{-\frac{\Upsilon}{2}A(k-k_2)} \right) + C(A,\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{4}} e^{-\Upsilon kA}. \end{aligned}$$

Similarly,

$$\begin{aligned} -(N(e^{-kA}) - m) &\le -\fint_{(k-1)A}^{kA} (N(e^{-s}) - m)\, ds + C(A,\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}} e^{-\Upsilon kA} \\ &\le C(A,\Lambda,\Upsilon)\delta \left( e^{-\frac{\Upsilon}{2}A(k_1-k)} + e^{-\frac{\Upsilon}{2}A(k-k_2)} \right) + C(A,\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{4}} e^{-\Upsilon kA}. \end{aligned}$$

Thus (7.18b) follows. ☐

7.3. **Symmetry and cone splitting.** Our definitions, methods, and results in this subsection are modifications of those in Section 3. These modifications are needed due to the lack of regularity of solutions to (7.1).

**Definition 7.19** (Frequency pinching)**.** Let $u$ be a solution to (7.1) satisfying (7.2). We define

$$\widehat{\mathcal{E}}_r(\mathbf{x}) := \widehat{N}^u_{\mathbf{x}}(50r^2) - \widehat{N}^u_{\mathbf{x}}(\tfrac{1}{50}r^2) + \lambda^{\frac{\Upsilon}{4}} r^{\Upsilon}, \qquad \widehat{\mathcal{E}}_r(\{\mathbf{x}_i\}_{i=0}^k) := \max_i \widehat{\mathcal{E}}_r(\mathbf{x}_i),$$

so that if $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$, then by Proposition 7.8 we have

$$\widehat{N}^u_{\mathbf{x}}(50r^2) - \widehat{N}^u_{\mathbf{x}}(\tfrac{1}{50}r^2) \ge -\frac{1}{2}\lambda^{\frac{\Upsilon}{4}} r^{\Upsilon}.$$

Define the $(k, \alpha r)$-*pinching* at $\mathbf{x}_0$ to be

$$\widehat{\mathcal{E}}^{k,\alpha}_r(\mathbf{x}_0) := \inf \widehat{\mathcal{E}}_r(\{\mathbf{x}_i\}_{i=0}^K),$$

where the infimum is over $(k, \frac{1}{20}\alpha r)$-independent subsets $\{\mathbf{x}_i\}_{i=0}^K \subseteq P(\mathbf{x}_0, \frac{1}{10}r)$.

**Definition 7.20.** A function $u$ satisfying (7.1) is weakly $(k, \delta, r)$-*symmetric* with respect to $V$ at $\mathbf{x}_0 = (x_0, t_0)$, if one of the following holds:

(1) There is a $k$-plane $L \subseteq \mathbb{R}^n$ such that $V = L \times \{0\}$ and

$$\int_{r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\pi_L \nabla(\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\tau} d\tau + \lambda^{\frac{\Upsilon}{4}} r^{\Upsilon} < \delta,$$

(2) There is a $(k-2)$-plane $L \subseteq \mathbb{R}^n$ such that $V = L \times \mathbb{R}$ and

$$\int_{r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\pi_L \nabla(\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\tau} d\tau + \int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\partial_t(\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\tau} d\tau + \lambda^{\frac{\Upsilon}{4}} r^{\Upsilon} < \delta.$$

7.3.1. *Propagation of symmetry.* As an analogue of Lemma 3.7, we prove that symmetry propagates across scales.

**Proposition 7.21.** *If $\lambda \le \overline{\lambda}(\Lambda)$, $\alpha \in [0, \frac{1}{2}]$, and $\beta, \tau_0 \in (0, 1]$, then there exists $C = C(\beta, \Lambda, \Upsilon)$ such that the following properties hold:*

*(1) For any $k$-plane $L \subset \mathbb{R}^n$ and $\tau \in [\beta^2 \tau_0, \tau_0]$, we have:*

$$\int_{\mathbb{R}^n} \tau |\pi_L \nabla(\chi u_{\mathbf{x}_0;1})|^2 e^{-\alpha f} \, d\nu_{-\tau} \le C \int_{\mathbb{R}^n} \tau_0 |\pi_L \nabla(\chi u_{\mathbf{x}_0;1})|^2 e^{-\alpha f} \, d\nu_{-\tau_0} + C\lambda^{\Upsilon} \tau_0^{2\Upsilon} \widehat{H}_{\mathbf{x}_0}(\tau_0), \tag{7.58a}$$

$$\int_{\beta^2 \tau_0}^{\tau_0} \int_{\mathbb{R}^n} \tau |\partial_t(\chi u_{\mathbf{x}_0;1})|^2 e^{-\alpha f} \, d\nu_{-\tau} \le C \int_{\frac{\tau_0}{2}}^{\tau_0} \int_{\mathbb{R}^n} \tau_0 |\partial_t(\chi u_{\mathbf{x}_0;1})|^2 e^{-\alpha f} \, d\nu_{-\tau_0} + C\lambda^{\Upsilon} \tau_0^{2\Upsilon} \widehat{H}_{\mathbf{x}_0}(\tau_0). \tag{7.58b}$$

*In particular, if $u$ is weakly $(k, \delta, r)$-symmetric with respect to $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ at $\mathbf{x}$, then $u$ is weakly $(k, C\delta, s)$-symmetric with respect to $V$ at $\mathbf{x}$ for all $s \in [\beta r, r]$.*

*(2) Suppose $0 < e^2 r_1 \le r \le 1$,*

$$|\widehat{N}_{\mathbf{x}}(e^{-10} r_1^2) - \widehat{N}_{\mathbf{x}}(e^{10} r^2)| + \lambda^{\frac{\Upsilon}{4}} r^{\Upsilon} < \delta < C(\Lambda)^{-1},$$

*and $u$ is weakly $(k, \delta, r_2)$-symmetric at $\mathbf{x}$ with respect to $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ for some $r_2 \in [r_1, r]$. Then, for all $s \in [r_1, r]$, $u$ is weakly $(k, C\delta, s)$-symmetric at $\mathbf{x}$ with respect to $V$.*

In the following lemma, we show that a solution to (7.1) satisfying (7.2) can be well-approximated near a given scale by a caloric function.

**Lemma 7.22.** *Given $\mathbf{x}_0 \in \mathbb{R}^n \times \mathbb{R}$ and $\tau_0 \in (0, 1]$, let $w \in C^\infty(\mathbb{R}^n \times (-\tau_0, 0]) \cap C(\mathbb{R}^n \times [-\tau_0, 0])$ be the unique solution to*

$$\square w = 0, \qquad w(\cdot, -\tau_0) = \chi(\cdot, -\tau_0) u_{\mathbf{x}_0;1}(\cdot, -\tau_0)$$

*satisfying the mild growth condition (1.5). Then, for any $\beta \in (0, \frac{1}{2})$, we have*

$$\sup_{\tau \in [\beta^2 \tau_0, \tau_0]} \int_{\mathbb{R}^n} (w - \chi u_{\mathbf{x}_0;1})^2 \, d\nu_{-\tau} + \int_{\beta^2 \tau_0}^{\tau_0} \int_{\mathbb{R}^n} |\nabla(w - \chi u_{\mathbf{x}_0;1})|^2 d\nu_{-\tau} d\tau \le C(\beta, \Lambda, \Upsilon) \lambda^{\Upsilon} \tau_0^{2\Upsilon} \widehat{H}_{\mathbf{x}_0}(\tau_0),$$

$$\sup_{\tau\in[\beta^2\tau_0,\frac{1}{2}\tau_0]}\int_{\mathbb{R}^n}\tau|\nabla(w-\chi u_{\mathbf{x}_0;1})|^2d\nu_{-\tau}+\int_{\beta^2\tau_0}^{\frac{1}{2}\tau_0}\tau\int_{\mathbb{R}^n}\left(|\partial_t(w-\chi u_{\mathbf{x}_0;1})|^2+|\nabla^2(w-\chi u_{\mathbf{x}_0;1})|^2\right)d\nu_{-\tau}d\tau$$
$$\leq C(\beta,\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_0^{2\Upsilon}\widehat{H}_{\mathbf{x}_0}(\tau_0).$$

*Proof.* After replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume $\mathbf{x}_0=0$ and $u_{\mathbf{x}_0;1}=u$. We compute:

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^n}(\chi u-w)^2d\nu_t&\leq -2\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2d\nu_t+2\int_{\mathbb{R}^n}|\chi u-w|\cdot|\square(\chi u)|d\nu_t\\&\leq -2\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2d\nu_t+\frac{1}{\tau}\int_{\mathbb{R}^n}(\chi u-w)^2d\nu_t+\tau\int_{\mathbb{R}^n}|\square(\chi u)|^2d\nu_t.\end{aligned}$$

Thus we can integrate:

$$\frac{d}{dt}\left(\tau\int_{\mathbb{R}^n}(\chi u-w)^2d\nu_t\right)+2\tau\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2d\nu_t\leq\tau^2\int_{\mathbb{R}^n}|\square(\chi u)|^2d\nu_t$$

from $t=-\tau_0$ to $-\beta^2\tau_0$ and applying Lemma 7.7 (7.9b) yields

$$\sup_{\tau\in[\beta^2\tau_0,\tau_0]}\int_{\mathbb{R}^n}(\chi u-w)^2d\nu_{-\tau}+\int_{\beta^2\tau_0}^{\tau_0}\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2d\nu_{-\tau}d\tau\leq C(\beta,\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_0^{2\Upsilon}\widehat{H}(\tau_0).$$

Fix a cutoff function $\eta\in C^\infty(\mathbb{R})$, such that $\eta(-\tau_0)=0$, $\eta|_{[-\frac{1}{2}\tau_0,\infty)}\equiv 1$, and $0\leq\eta'\leq 10\tau_0^{-1}$. We then integrate

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^n}\tau\eta(t)|\nabla(\chi u-w)|^2\,d\nu_t=&(-\eta(t)+\tau\eta'(t))\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2\,d\nu_t-2\eta(t)\tau\int_{\mathbb{R}^n}|\nabla^2(\chi u-w)|^2\,d\nu_t\\&-2\tau\eta(t)\int_{\mathbb{R}^n}\Delta_f(\chi u-w)\square(\chi u)\,d\nu_t\\\leq&(-\eta(t)+\tau\eta'(t))\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2\,d\nu_t-2\eta(t)\tau\int_{\mathbb{R}^n}|\nabla^2(\chi u-w)|^2\,d\nu_t\\&+2\tau\eta(t)\int_{\mathbb{R}^n}|\Delta_f(\chi u-w)|^2\,d\nu_t+2\tau\eta(t)\int_{\mathbb{R}^n}|\square(\chi u)|^2\,d\nu_t.\end{aligned}$$

Applying the $f$-Bochner formula (2.5) with $w\leftarrow\chi u-w$ and integrating by parts, we get

$$2\int_{\mathbb{R}^n}(\Delta_f(\chi u-w))^2\,d\nu_t=2\int_{\mathbb{R}^n}\left|\nabla^2(\chi u-w)\right|^2d\nu_t+\frac{1}{\tau}\int_{\mathbb{R}^n}|\nabla(\chi u-w)|^2\,d\nu_t.$$

Therefore, arguing as before,

$$\sup_{\tau\in[\beta^2\tau_0,\tau_0]}\int_{\mathbb{R}^n}\tau\eta(-\tau)|\nabla(\chi u-w)|^2\,d\nu_{-\tau}+\int_{\beta^2\tau_0}^{\tau_0}\eta(-\tau)\tau\int_{\mathbb{R}^n}|\nabla^2(\chi u-w)|^2\,d\nu_{-\tau}d\tau$$
$$\leq C(\beta,\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_0^{2\Upsilon}\widehat{H}(\tau_0).$$

This together with Lemma 7.7 (7.9b) yields,

$$\int_{\beta^2\tau_0}^{\tau_0}\eta(-\tau)\tau\int_{\mathbb{R}^n}|\partial_t(\chi u-w)|^2\,d\nu_{-\tau}d\tau\leq C(\beta,\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_0^{2\Upsilon}\widehat{H}(\tau_0).$$

$\square$

The following will be used to prove a version of Lemma 7.25 in the non-caloric setting.

**Corollary 7.23.** *If $\lambda\leq\overline{\lambda}(\Lambda,\Upsilon)$, then, for any $\mathbf{x}_0\in P(\mathbf{0},10)$ and $\tau_0\in(0,1]$, the following estimates hold for any $B\in[2,10^{10n}]$*

$$\frac{\tau_0}{\widehat{H}^u_{\mathbf{x}_0}(\tau_0)}\int_{\mathbb{R}^n}|\pi_L\nabla(\chi u_{\mathbf{x}_0;1})|^2e^{\frac{1}{8n}f}d\nu_{-\tau_0}\leq C\int_{B\tau_0}^{Be^2\tau_0}\frac{1}{\widehat{H}^u_{\mathbf{x}_0}(\tau)}\int_{\mathbb{R}^n}|\pi_L\nabla(\chi u_{\mathbf{x}_0;1})|^2d\nu_{-\tau}d\tau$$

$$+ C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_0^{\Upsilon},$$

$$\int_{\tau_0}^{e^2\tau_0} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\partial_t(\chi u_{\mathbf{x}_0;1})|^2 e^{\frac{1}{8n}f} d\nu_{-\tau} d\tau \leq C \int_{B\tau_0}^{Be^2\tau_0} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\partial_t(\chi u_{\mathbf{x}_0;1})|^2 d\nu_{-\tau} d\tau + C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau_0^{\Upsilon}.$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume $\mathbf{x}_0 = \mathbf{0}$ and $u = u_{\mathbf{x}_0;1}$. Let $w$ be the caloric function from Lemma 7.22 with $\tau_0 \leftarrow 2B\tau_0$. Then, for any $\tau \in [\tau_0, e^2\tau_0]$, Lemma 7.22 gives

$$\begin{aligned}
\tau \int_{\mathbb{R}^n} |\nabla(\chi u - w)|^2 e^{\frac{1}{8n}f} d\nu_{-\tau} &\leq \lambda^{-\frac{\Upsilon}{2}}\tau^{-\Upsilon}\cdot\tau \int_{\mathbb{R}^n} |\nabla(\chi u - w)|^2 d\nu_{-\tau} \\
&\quad + \lambda^{\frac{\Upsilon}{2}}\tau^{\Upsilon}\cdot\tau \int_{\mathbb{R}^n} (|\nabla(\chi u)|^2 + |\nabla w|^2) e^{\frac{1}{4n}f} d\nu_{-\tau} \\
&\leq C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\Upsilon}\widehat{H}^u(\tau) + C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\Upsilon}\left(\int_{\mathbb{R}^n} |\nabla w|^{\frac{8n}{4n-1}} d\nu_{-\tau}\right)^{\frac{4n-1}{4n}} \\
&\leq C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\Upsilon}\widehat{H}^u(\tau),
\end{aligned}$$

where in the last line we applied Proposition 2.2 and $\widehat{N}^w(2\tau) \leq C(\Lambda,\Upsilon)$. On the other hand, because the components of $\pi_L\nabla w$ are caloric, for any $\tau \in [B\tau_0, Be^2\tau_0]$, we have

$$\tau_0 \int_{\mathbb{R}^n} |\pi_L\nabla w|^2 e^{\frac{1}{8n}f} d\nu_{-\tau_0} \leq C\tau_0 \left(\int_{\mathbb{R}^n} |\pi_L\nabla w|^{\frac{8n}{4n-1}} d\nu_{-\tau_0}\right)^{\frac{4n-1}{4n}} \leq C\tau \int_{\mathbb{R}^n} |\pi_L\nabla w|^2 d\nu_{-B\tau_0}.$$

The first inequality therefore follows by combining expressions, and the second inequality is similar. □

We prove a version of (7.58a) and (7.58b) for caloric functions.

**Lemma 7.24.** *Suppose $w$ is a caloric function with mild growth (2.1). Then, for any $\alpha \in [0, \frac{1}{2}]$ and any $0 < \tau_1 \leq \tau_2$, we have*

$$\int_{\mathbb{R}^n} w^2 e^{-\alpha f}\, d\nu_{-\tau_1} \leq \left(\frac{\tau_2}{\tau_1}\right)^{\frac{n\alpha}{2}} \int_{\mathbb{R}^n} w^2 e^{-\alpha f}\, d\nu_{-\tau_2}.$$

*Proof.* Since $\Delta e^{-(1+\alpha)f} = e^{-(1+\alpha)f}\left(-\frac{(1+\alpha)n}{2\tau} + \frac{(1+\alpha)^2 f}{\tau}\right)$, we have

$$-\mathrm{div}\,(w^2\nabla e^{-(1+\alpha)f}) = 2(1+\alpha)\nabla w\cdot\nabla f w e^{-(1+\alpha)f} + w^2 e^{-(1+\alpha)f}\left(\frac{(1+\alpha)n}{2\tau} - \frac{(1+\alpha)^2 f}{\tau}\right).$$

Integration by parts then yields

$$\begin{aligned}
&\frac{(1+\alpha)}{\tau}\int_{\mathbb{R}^n} w^2 f e^{-\alpha f}\, d\nu_t \\
&\leq \frac{n}{2\tau}\int_{\mathbb{R}^n} w^2 e^{-\alpha f} d\nu_t + \frac{2}{(1+\alpha)}\int_{\mathbb{R}^n} |\nabla w|^2 e^{-\alpha f}\, d\nu_t + \frac{(1+\alpha)}{2}\int_{\mathbb{R}^n} \frac{1}{\tau} w^2 f e^{-\alpha f} d\nu_t.
\end{aligned}$$

Rearranging this gives

$$\int_{\mathbb{R}^n} w^2(2(1+\alpha)f - n)\frac{\alpha}{2\tau} e^{-\alpha f} d\nu_t \leq \frac{4\alpha}{(1+\alpha)}\int_{\mathbb{R}^n} |\nabla w|^2 e^{-\alpha f}\, d\nu_t + \frac{n\alpha}{2\tau}\int_{\mathbb{R}^n} w^2 e^{-\alpha f}\, d\nu_t. \tag{7.59}$$

Therefore, if $\alpha < \frac{1}{2}$, then

$$\frac{d}{dt}\int_{\mathbb{R}^n} w^2 e^{-\alpha f}\, d\nu_t = -2\int_{\mathbb{R}^n} |\nabla w|^2 e^{-\alpha f}\, d\nu_t - \int_{\mathbb{R}^n} w^2(\Box^* e^{-\alpha f} + 2\nabla f\cdot\nabla e^{-\alpha f})\, d\nu_t$$

$$= -2\int_{\mathbb{R}^n} |\nabla w|^2 e^{-\alpha f}\, d\nu_t + \int_{\mathbb{R}^n} w^2(2(1+\alpha)f - n)\frac{\alpha}{2\tau} e^{-\alpha f}\, d\nu_t$$
$$\leq \frac{n\alpha}{2\tau}\int_{\mathbb{R}^n} w^2 e^{-\alpha f}\, d\nu_t.$$

Thus the claim follows by Gronwall's inequality. □

*Proof of Proposition 7.21.* By replacing $u$ with $u_{\mathbf{x};1}$, we can assume that $\mathbf{x} = 0$ and $u_{\mathbf{x};1} = u$.

(1) Suppose $h$ is a caloric approximation of $\chi u$ on $[-2\tau_0, -\beta^2\tau_0]$. Then Lemma 7.22 and Lemma 7.6 imply that

$$\sup_{\tau\in[\beta^2\tau_0,\tau_0]}\int_{\mathbb{R}^n} \tau|\nabla(h-\chi u)|^2 d\nu_{-\tau} + \int_{\beta^2\tau_0}^{\tau_0} \tau\int_{\mathbb{R}^n} |\partial_t(h-\chi u)|^2\, d\nu_{-\tau} d\tau \tag{7.60}$$
$$\leq C(\beta,\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_0^{2\Upsilon}\widehat{H}^u(\tau_0).$$

On the other hand, Lemma 7.24 applied twice with $w \leftarrow \partial_t h$ as well as $w \leftarrow \partial_i h$, we get

$$\sup_{\tau\in[\beta^2\tau_0,\tau_0]}\int_{\mathbb{R}^n} |\pi_L\nabla h|^2 e^{-\alpha f}\, d\nu_{-\tau} \leq \beta^{n\alpha}\int_{\mathbb{R}^n} |\pi_L\nabla h|^2 e^{-\alpha f}\, d\nu_{-\tau_0},$$
$$\sup_{\tau\in[\beta^2\tau_0,\tau_0]}\int_{\mathbb{R}^n} |\partial_t h|^2 e^{-\alpha f}\, d\nu_{-\tau} \leq \beta^{n\alpha}\int_{\mathbb{R}^n} |\partial_t h|^2 e^{-\alpha f}\, d\nu_{-\tau_0}.$$

Combining these with (7.60), we get (7.58a) and (7.58b). As a consequence, if $u$ is weakly $(k,\delta,r)$-symmetric with respect to $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ at $\mathbf{x}$, then $u$ is weakly $(k, C\delta, s)$-symmetric with respect to $V$ at $\mathbf{x}$ for all $s \in [\beta r, r]$.

(2) We will assume that $u$ is weakly temporally symmetric, as the other case is similar. By Theorem 7.13 there exists $p \in \mathcal{P}_m$ such that for any $\tau \in [r_1^2, e^2 r^2]$ we have

$$\|\widetilde{w} - \widehat{p}\|_{[-\log\tau - 1, -\log\tau + 1]} \leq C(\Lambda,\Upsilon)\delta + C\lambda^{\frac{\Upsilon}{4}}\tau^{\frac{\Upsilon}{2}} \leq C(\Lambda,\Upsilon)\delta.$$

Then Lemma 7.17 implies for all $\tau \in [r_1^2, e^2 r^2]$

$$\frac{1}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n} (\chi u - q)^2\, d\nu_{-\tau} + \int_{\frac{1}{2}\tau}^{2\tau}\frac{1}{\widehat{H}^u(s)}\int_{\mathbb{R}^n}\left(|\nabla(\chi u - q)|^2 + s|\partial_t(\chi u - q)|^2\right) d\nu_{-s} ds \leq C(\Lambda,\Upsilon)\delta, \tag{7.62}$$

where we set $q := \sqrt{\frac{\widehat{H}^u(s)}{s^m}}p$. Observe that:

$$\frac{1}{\tau^m}\int_{\mathbb{R}^n} p^2\, d\nu_{-\tau} \geq \frac{1}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}(\chi u)^2\, d\nu_{-\tau} - \frac{1}{\widehat{H}^u(\tau)}\int_{\mathbb{R}^n}(\chi u - q)^2\, d\nu_{-\tau} \geq 1 - C(\Lambda,\Upsilon)\delta, \tag{7.63}$$

and similarly

$$\frac{1}{\tau^m}\int_{\mathbb{R}^n} p^2\, d\nu_{-\tau} \leq 1 + C(\Lambda,\Upsilon)\delta. \tag{7.64}$$

Using (7.8), we get

$$\int_{\mathbb{R}^n}\left|\partial_t q - \sqrt{\frac{\widehat{H}^u(s)}{s^m}}\partial_t p\right|^2 d\nu_{-s} \leq C\left(\frac{d}{ds}\log\widehat{H}^u(s) - \frac{m}{s}\right)^2\int_{\mathbb{R}^n} q^2\, d\nu_{-s} \tag{7.65}$$
$$\leq C(\Lambda,\Upsilon)\lambda^{2\Upsilon}s^{2\Upsilon-2}\frac{\widehat{H}^u(s)}{s^m}\int_{\mathbb{R}^n} p^2\, d\nu_{-s}.$$

Combining (7.17) and (7.63) with $\tau \leftarrow 2r_2^2$ and $p \in \mathcal{P}_m$, we get

$$\begin{aligned}
&\log(2) \int_{\mathbb{R}^n} |\partial_t p|^2 \, d\nu_{-1} \\
&= \int_{\frac{\tau}{2}}^{\tau} \frac{1}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} s \left| \sqrt{\frac{\widehat{H}^u(s)}{s^m}} \partial_t p \right|^2 d\nu_{-s} ds \\
&\leq 4 \int_{\frac{\tau}{2}}^{\tau} \frac{1}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} s \left| \sqrt{\frac{\widehat{H}^u(s)}{s^m}} \partial_t p - \partial_t q \right|^2 d\nu_{-s} ds + 4 \int_{\frac{\tau}{2}}^{\tau} \frac{1}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} s |\partial_t(\chi u) - \partial_t q|^2 \, d\nu_{-s} ds \\
&\quad + 4 \int_{\frac{\tau}{2}}^{\tau} \frac{1}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} s |\partial_t(\chi u)|^2 \, d\nu_{-s} ds \\
&\leq C(\Lambda, \Upsilon) \int_{\frac{\tau}{2}}^{\tau} \lambda^{2\Upsilon} s^{2\Upsilon - 1} \frac{1}{s^m} \int_{\mathbb{R}^n} p^2 \, d\nu_{-s} ds + C(\Lambda, \Upsilon)\delta \leq C(\Lambda, \Upsilon)\delta \int_{\mathbb{R}^n} p^2 d\nu_{-1}.
\end{aligned}$$

Combining this along with $p \in \mathcal{P}_m$, for any $\tau \in [r_1^2, r]$, we have

$$\begin{aligned}
\int_{\tau}^{e^2\tau} \frac{s}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} |\partial_t(\chi u)|^2 d\nu_{-s} ds &\leq 4 \int_{\tau}^{e^2\tau} \frac{s}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} |\partial_t(\chi u - q)|^2 d\nu_{-s} ds \\
&\quad + 4 \int_{\tau}^{e^2\tau} \frac{s}{\widehat{H}^u(s)} \int_{\mathbb{R}^n} \left| \partial_t q - \sqrt{\frac{\widehat{H}^u(s)}{s^m}} \partial_t p \right|^2 d\nu_{-s} ds \\
&\quad + 4 \int_{\tau}^{e^2\tau} \frac{s}{s^m} \int_{\mathbb{R}^n} |\partial_t p|^2 d\nu_{-s} ds \\
&\leq C(\Lambda, \Upsilon)(\delta + \lambda^{2\Upsilon}\tau^{2\Upsilon} + \delta) \leq C(\Lambda, \Upsilon)\delta,
\end{aligned}$$

where we used (7.62), (7.65), (7.64) in the second inequality. □

As an analog of Lemma 2.3, the following result proves that energy functionals based at nearby points are comparable. It allows us to propagate symmetry at nearby points.

**Lemma 7.25.** *The following holds if $\lambda \leq \overline{\lambda}(\Lambda, \Upsilon)$. For any $r \in (0, 1)$ and any $\mathbf{x}_0, \mathbf{x}_1 \in P(\mathbf{0}, 10)$ satisfying $|\mathbf{x}_1 - \mathbf{x}_0| < \frac{1}{10n} r$:*

*(1) If $\tau_0, \tau_1 \in [r^2, e^{100} r^2]$, then*

$$C(\Lambda)^{-1} \widehat{H}^u_{\mathbf{x}_0}(\tau_0) \leq \widehat{H}^u_{\mathbf{x}_1}(\tau_1) \leq C(\Lambda) \widehat{H}^u_{\mathbf{x}_0}(\tau_0).$$

*(2) For any $k$-plane $L \subseteq \mathbb{R}^n$ and $\tau \in [r^2, e^{100} r^2]$:*

$$\frac{\tau}{\widehat{H}^u_{\mathbf{x}_1}(\tau)} \int_{\mathbb{R}^n} |\pi_L \nabla(\chi u_{\mathbf{x}_1;1})|^2 \, d\nu_{-\tau} \leq C(\Lambda, \Upsilon) \frac{4\tau}{\widehat{H}^u_{\mathbf{x}_0}(4\tau)} \int_{\mathbb{R}^n} |\pi_L \nabla(\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-4\tau} + C(\Lambda, \Upsilon) \lambda^{\frac{\Upsilon}{2}} \tau^{\Upsilon}.$$

*(3) Finally,*

$$\begin{aligned}
&\int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_1}(\tau)} \int_{\mathbb{R}^n} |\partial_t(\chi u_{\mathbf{x}_1;1})|^2 \, d\nu_{-\tau} d\tau \\
&\quad \leq C(\Lambda, \Upsilon) \int_{4r^2}^{4e^2 r^2} \frac{\tau}{\widehat{H}^u_{\mathbf{x}_0}(\tau)} \int_{\mathbb{R}^n} |\partial_t(\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\tau} d\tau + C(\Lambda, \Upsilon) \lambda^{\frac{\Upsilon}{2}} r^{2\Upsilon}.
\end{aligned}$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0 = \mathbf{0}$ and $u = u_{\mathbf{x}_0;1}$ and we write $\tau := \tau_0$. Define

$$\chi_1(z, -s) := \chi(a^{\frac{1}{2}}(\mathbf{x}_1)(z - x_1), -t_1 - s).$$

Then, for any $s \in [r^2, e^{100} r^2]$, we have

$$\chi_1(\cdot, s)|_{B(0, \frac{1}{10} s^{\Upsilon} \lambda^{-\frac{\Upsilon}{2}})} \equiv \chi(\cdot, s)|_{B(0, 10 s^{\Upsilon} \lambda^{-\frac{\Upsilon}{2}})} \equiv 1. \tag{7.66}$$

(1) Using (7.66) and Lemma 7.4, we estimate

(7.67)
$$\begin{aligned}
\int_{\mathbb{R}^n} (\chi_1 u - \chi u)^2 e^{\lambda f_1}\, d\nu_{\mathbf{x}_1; t_1 - \tau} &\leq 2 \int_{B(x_1, 10\lambda^{-\frac{\Upsilon}{2}} \tau^{\Upsilon}) \setminus \overline{B}(x_1, 10^{-1} \lambda^{-\frac{\Upsilon}{2}} \tau^{\Upsilon})} u^2 e^{\lambda f_1}\, d\nu_{\mathbf{x}_1; t_1 - \tau} \\
&\leq C(\Lambda, \Upsilon) \left(1 + \frac{\lambda^{-\Upsilon}}{\tau^{1 - 2\Upsilon}}\right)^{C(\Lambda)} \frac{\lambda^{-\frac{n\Upsilon}{2}}}{\tau^{\frac{n}{2}(1 - 2\Upsilon)}} \exp\left(-\frac{\lambda^{-\Upsilon}}{400 \tau^{1 - 2\Upsilon}}\right) \widehat{H}^u_{\mathbf{x}_1}(\tau) \\
&\leq C(\Lambda, \Upsilon) \exp\left(-\frac{1}{e^{60} \lambda^{\Upsilon} \tau^{1 - 2\Upsilon}}\right) \widehat{H}^u_{\mathbf{x}_1}(\tau) \leq \frac{1}{2} \widehat{H}^u_{\mathbf{x}_1}(\tau)
\end{aligned}$$

if $\lambda \leq \overline{\lambda}(\Lambda, \Upsilon)$. On the other hand, we change variables $z = x_1 + a^{\frac{1}{2}}(\mathbf{x}_1) y$ to obtain

$$\begin{aligned}
\widehat{H}^u_{\mathbf{x}_1}(\tau) &= \int_{\mathbb{R}^n} \chi^2(y, -\tau) u^2(x_1 + a^{\frac{1}{2}}(\mathbf{x}_1) y, t_1 - \tau) d\nu_{-\tau}(y) \\
&= \left(\sqrt{\det a^{\frac{1}{2}}(\mathbf{x}_1)}\right)^{-1} \int_{\mathbb{R}^n} \chi_1^2(z, t_1 - \tau) u^2(z, t_1 - \tau) \frac{1}{(4\pi\tau)^{\frac{n}{2}}} e^{-\frac{|a^{-\frac{1}{2}}(\mathbf{x}_1)(z - x_1)|^2}{4\tau}} dz \\
&\leq 2 \int_{\mathbb{R}^n} \chi_1^2 u^2 e^{\lambda f_1}\, d\nu_{\mathbf{x}_1; t_1 - \tau} \leq 4 \int_{\mathbb{R}^n} \chi^2 u^2 e^{\lambda f_1}\, d\nu_{\mathbf{x}_1; t_1 - \tau} + 4 \int_{\mathbb{R}^n} (\chi_1 u - \chi u)^2 e^{\lambda f_1}\, d\nu_{\mathbf{x}_1; t_1 - \tau} \\
&\leq C \int_{\mathbb{R}^n} \chi^2 u^2 e^{\frac{1}{2} f}\, d\nu_{t_1 - \tau} + \frac{1}{2} \widehat{H}^u_{\mathbf{x}_1}(\tau) \leq C \widehat{H}^u(\tau - t_1) + \frac{1}{2} \widehat{H}^u_{\mathbf{x}_1}(\tau) \\
&\leq C(\Lambda, \Upsilon) \widehat{H}^u(\tau) + \frac{1}{2} \widehat{H}^u_{\mathbf{x}_1}(\tau)
\end{aligned}$$

if $\lambda \leq \overline{\lambda}(\Lambda, \Upsilon)$, where we used (7.1) in the third line. Further, in the fourth line, we used Lemma 2.1 with $\mathbf{x}_1 \leftarrow \mathbf{0}$, $\mathbf{x}_0 \leftarrow \mathbf{x}_1$, $\tau \leftarrow \tau - t_1$, $\alpha_0 \leftarrow \lambda$, $\alpha \leftarrow \frac{1}{2}$ and $\sigma \leftarrow \frac{1}{100} r^2$ as well as Lemma 7.5. Finally, in the last line, we used Lemma 7.6. Rearranging the computation and using Lemma 7.6 again, we get

$$\widehat{H}^u_{\mathbf{x}_1}(\tau_1) \leq C(\Lambda) \widehat{H}^u(\tau).$$

By symmetry, we get (1).

(2) Fix $v \in L$ with $|v| = 1$. Using (7.66) and Lemma 7.4, we have

(7.68)
$$\begin{aligned}
\int_{\mathbb{R}^n} |\nabla(\chi_1 u - \chi u)|^2 e^{\frac{1}{8n} f}\, d\nu_{t_1 - \tau} &\leq \frac{C \lambda^{\frac{\Upsilon}{2}}}{\tau^{\Upsilon}} \int_{B(0, 20 \tau^{\Upsilon} \lambda^{-\frac{\Upsilon}{2}}) \setminus \overline{B}(0, \frac{1}{20} \tau^{\Upsilon} \lambda^{-\frac{\Upsilon}{2}})} \left(|u|^2 + |\nabla u|^2\right) e^{\frac{1}{8n} f}\, d\nu_{t_1 - \tau} \\
&\leq C(\Lambda, \Upsilon) e^{-\frac{1}{e^{70} r^{2 - 4\Upsilon} \lambda^{\Upsilon}}} \widehat{H}(\tau),
\end{aligned}$$

where we used Lemma 7.7 and (1) in the last inequality. Using a coordinate change together with Lemma 7.5 and $|v - a^{-\frac{1}{2}}(\mathbf{x}_1) v| \leq \lambda |\mathbf{x}_1| \leq \lambda r$, we estimate

$$\int_{\mathbb{R}^n} |v \cdot \nabla(\chi u_{\mathbf{x}_1;1})|^2 d\nu_{-\tau} = \int_{\mathbb{R}^n} |a^{-\frac{1}{2}}(\mathbf{x}_1) v \cdot \nabla(\chi_1 u)(z, t_1 - \tau)|^2 \frac{e^{-\frac{(z - x_1) \cdot a(\mathbf{x}_1)(z - x_1)}{4\tau}}}{(4\pi\tau)^{\frac{n}{2}} \sqrt{\det a(\mathbf{x}_1)}} dz$$

$$\leq C \int_{\mathbb{R}^n} |v\cdot\nabla(\chi_1 u)|^2 e^{\frac{1}{16n}f_{\mathbf{x}_1}} d\nu_{\mathbf{x}_1;t_1-\tau} + \lambda^2 r^2 \int_{\mathbb{R}^n} |\nabla(\chi_1 u)|^2 e^{\frac{1}{16n}f_{\mathbf{x}_1}} d\nu_{\mathbf{x}_1;t_1-\tau}$$
$$\leq C \int_{\mathbb{R}^n} |v\cdot\nabla(\chi_1 u)|^2 e^{\frac{1}{8n}f} d\nu_{t_1-\tau} + C\lambda^2 r^2 \int_{\mathbb{R}^n} |\nabla(\chi_1 u)|^2 e^{\frac{1}{8n}f} d\nu_{t_1-\tau}.$$

where for the last inequality we applied Lemma 2.1 with $\mathbf{x}_0 \leftarrow \mathbf{x}_1$ $\mathbf{x}_1 \leftarrow \mathbf{0}$, $\tau \leftarrow t_1-\tau$, $\alpha_0 \leftarrow \frac{1}{16n}$, $\alpha \leftarrow \frac{1}{8n}$ and $\sigma \leftarrow \frac{1}{100n^2}r^2$. Combining this estimate with (7.68) gives

$$r^2 \int_{\mathbb{R}^n} |v\cdot\nabla(\chi u_{\mathbf{x}_1;1})|^2 d\nu_{-\tau} \leq C r^2 \int_{\mathbb{R}^n} |v\cdot\nabla(\chi_1 u)|^2 e^{\frac{1}{8n}f} d\nu_{t_1-\tau} + C(\Lambda,\Upsilon)\lambda^2 r^2 \widehat{H}(\tau),$$

and (2) then follows by Corollary 7.23.

(3) Using (7.66) and Lemma 7.4, we have

$$\begin{aligned}
&\int_{e^{-1}r^2}^{e^3r^2} \frac{s}{\widehat{H}^u_{\mathbf{x}_1}(s)} \int_{\mathbb{R}^n} |\partial_t(\chi_1 u - \chi u)|^2 e^{\frac{1}{8n}f}\, d\nu_{-s} ds \\
&\leq \frac{C}{r^4}\int_{e^{-1}r^2}^{e^3r^2} \frac{s e^{-\frac{1}{e^{60}s^{1-2\Upsilon}\lambda^{\Upsilon}}}}{\widehat{H}^u_{\mathbf{x}_1}(s)} \int_{B(0,20s^{\Upsilon}\lambda^{-\frac{\Upsilon}{2}})\setminus \overline{B}(0,\frac{1}{10}s^{\Upsilon}\lambda^{-\frac{\Upsilon}{2}})} \left(|u|^2+|\partial_t u|^2\right) d\nu_{-s} ds \qquad (7.69)\\
&\leq C(\Lambda,\Upsilon) e^{-\frac{1}{e^{70}r^{2-4\Upsilon}\lambda^{\Upsilon}}},
\end{aligned}$$

where we used Lemma 7.7 in the last inequality. On the other hand, (7.1) yields

$$\begin{aligned}
\int_{r^2}^{e^2r^2}\int_{\mathbb{R}^n} |\partial_t(\chi u_{\mathbf{x}_1;1})|^2 d\nu_{-\tau} d\tau &= \int_{r^2-t_1}^{e^2r^2-t_1}\int_{\mathbb{R}^n} |\partial_t(\chi_1 u)(z,-s)|^2 \frac{e^{-\frac{(z-x_1)\cdot a(\mathbf{x}_1)(z-x_1)}{4(s+t_1)}}}{(4\pi(s+t_1))^{\frac{n}{2}}\sqrt{\det a(\mathbf{x}_1)}} dz ds \\
&\leq C\int_{r^2-t_1}^{e^2r^2-t_1}\int_{\mathbb{R}^n} |\partial_t(\chi_1 u)|^2 e^{\frac{1}{16n}f_{\mathbf{x}_1}} d\nu_{\mathbf{x}_1;-s} ds \\
&\leq C\int_{r^2-t_1}^{e^2r^2-t_1}\int_{\mathbb{R}^n} |\partial_t(\chi_1 u)|^2 e^{\frac{1}{8n}f} d\nu_{-s} ds,
\end{aligned}$$

where for the last inequality, we applied Lemma 2.1 with $\mathbf{x}_0 \leftarrow \mathbf{x}_1$ $\mathbf{x}_1 \leftarrow \mathbf{0}$, $\tau \leftarrow s$, $\alpha_0 \leftarrow \frac{1}{16n}$, $\alpha \leftarrow \frac{1}{8n}$ and $\sigma \leftarrow \frac{1}{100}r$. We now combine this with (7.69), apply Corollary 7.23, and then apply Lemma 7.21 to obtain the claim. □

We also have a version of the gradient estimates similar to Lemma 2.6. We define $\widehat{H}_s(\mathbf{x}) := H_s^{\chi u_{\mathbf{x};1}}(\mathbf{x})$. It allows us to propagate symmetry along the plane of symmetry.

**Lemma 7.26.** *The following holds if $\lambda \leq \overline{\lambda}(\Lambda,\Upsilon)$. For any $\beta, r\in(0,1]$, $\mathbf{x}_0 \in P(\mathbf{0},10)$, and any $k$-plane $L\subset\mathbb{R}^n$, there exists $C := C(\beta,\Lambda,\Upsilon)$ such that for all $\mathbf{x}_1 \in ((L\times\mathbb{R})+\mathbf{x}_0)\cap P(\mathbf{x}_0,\frac{1}{10}r)$ and $s\in[\beta^2r^2, r^2]$:*

$$\begin{aligned}
\left|\log\left(\frac{\widehat{H}_s(\mathbf{x}_1)}{\widehat{H}_s(\mathbf{x}_0)}\right)\right|^2 &\leq C\frac{|x_1-x_0|^2}{r^2}\int_{r^2}^{e^2r^2}\frac{1}{\widehat{H}_\tau(\mathbf{x}_0)}\int_{\mathbb{R}^n}|\pi_L\nabla(\chi u_{\mathbf{x}_0;1})|^2\, d\nu_{-\tau}d\tau \\
&\quad + C\frac{|t_1-t_0|}{r^2}\int_{r^2}^{e^2r^2}\frac{\tau}{\widehat{H}_\tau(\mathbf{x}_0)}\int_{\mathbb{R}^n}|\partial_t(\chi u_{\mathbf{x}_0;1})|^2\, d\nu_{-\tau}d\tau + C\lambda^2r^2.
\end{aligned}$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we may assume that $\mathbf{x}_0=\mathbf{0}$ and $u=u_{\mathbf{x}_0;1}$. After a rotation, we can write $v\cdot\nabla w$ as $e_1\cdot\nabla w = \partial_{x_1}w$ for any $v\in L$. At $\mathbf{x}=(x,t)$, we compute

$$\partial_{x_1}\widehat{H}_s(\mathbf{x}) = 2\int_{\mathbb{R}^n}(\chi u_{\mathbf{x};1})(y,-s)\partial_{x_1}(\chi(y,-s)u(x+a^{-\frac{1}{2}}(\mathbf{x})y, t-s))\, d\nu_{-s}(y)$$

$$
\begin{aligned}
&= 2\int_{\mathbb{R}^n} (\chi^2 u_{\mathbf{x};1})(y,-s)\left((e_1\cdot\nabla u)_{\mathbf{x};1}(y,-s) + y\cdot(\partial_{x_1} a^{-\frac{1}{2}}(\mathbf{x}))(\nabla u)_{\mathbf{x};1}(y,-s)\right)\,d\nu_{-s}(y)\\
&= 2\int_{\mathbb{R}^n} (u_{\mathbf{x};1}\chi^2)(y,-s) e_1\cdot(I-a^{-\frac{1}{2}}(\mathbf{x}))(\nabla u)_{\mathbf{x};1}(y,-s)\,d\nu_{-s}(y)\\
&\quad\; - 2\int_{\mathbb{R}^n} u_{\mathbf{x};1}^2(y,-s)\chi(y,s)\partial_{y_1}\chi(y,-s)\,d\nu_{-s}(y)\\
&\quad\; + 2\int_{\mathbb{R}^n} (\chi^2 u_{\mathbf{x};1})(y,-s) y\cdot(\partial_{x_1} a^{-\frac{1}{2}}(\mathbf{x}))(\nabla u)_{\mathbf{x};1}(y,-s)\,d\nu_{-s}(y)\\
&\quad\; + 2\int_{\mathbb{R}^n} (u_{\mathbf{x};1}\chi)(y,-s)\partial_{y_1}(u_{\mathbf{x};1}\chi)(y,-s)\,d\nu_{-s}(y)\\
&=: K_1+K_2+K_3+K_4.
\end{aligned}
$$

Using the $\lambda$-Lipschitz property of $a$ in (7.1) and Lemma 7.5, we get

$$
K_1 \le 2\lambda|\mathbf{x}|\int_{\mathbb{R}^n} (u_{\mathbf{x};1}\chi^2)(y,-s)|(\nabla u)_{\mathbf{x};1}(y,-s)|\,d\nu_{-s}(y) \le C(\Lambda,\Upsilon)\lambda|\mathbf{x}|\widehat{H}_s(\mathbf{x}).
$$

We can use Lemma 7.5 again to get

$$
K_2 \le C(\Lambda,\Upsilon)\lambda^{\Upsilon} e^{-\frac{1}{8\lambda^{\Upsilon} s^{1-2\Upsilon}}}\widehat{H}_s(\mathbf{x}).
$$

Furthermore, we use (7.1) to get

$$
\begin{aligned}
K_3 &\le 2\lambda\sqrt{s}\int_{\mathbb{R}^n} (\chi^2 u_{\mathbf{x};1})(y,-s)\frac{|y|}{\sqrt{s}}|(\nabla u)_{\mathbf{x};1}(y,-s)|\,d\nu_{-s}(y)\\
&\le 2\lambda\sqrt{s}\int_{\mathbb{R}^n} (\chi^2 u_{\mathbf{x};1})(y,-s)|(\nabla u)_{\mathbf{x};1}(y,-s)|e^{\frac{1}{2}f}\,d\nu_{-s}(y) \le C(\Lambda,\Upsilon)\lambda\sqrt{s}\widehat{H}_s(\mathbf{x}),
\end{aligned}
$$

where we use Lemma 7.5 in the last line. Finally, using Hölder's inequality, we get

$$
\begin{aligned}
K_4 &\le 2\left(\int_{\mathbb{R}^n} (u_{\mathbf{x};1}\chi)^2(y,-s)e^{\frac{1}{2}f}\,d\nu_{-s}(y)\right)^{\frac{1}{2}}\left(\int_{\mathbb{R}^n} |\partial_{y_1}(u_{\mathbf{x};1}\chi)(y,-s)|^2 e^{-\frac{1}{2}f}\,d\nu_{-s}(y)\right)^{\frac{1}{2}}\\
&\le C(\Lambda,\Upsilon)\left(\widehat{H}_s(\mathbf{x})\int_{\mathbb{R}^n} |\partial_{y_1}(u_{\mathbf{x};1}\chi)(y,-s)|^2 e^{-\frac{1}{2}f}\,d\nu_{-s}(y)\right)^{\frac{1}{2}},
\end{aligned}
$$

where we use Lemma 7.5 in the last line. On the other hand, using Proposition 7.21(1), Lemma 7.25(1)(2), and Lemma 7.6, we get

$$
\begin{aligned}
&\int_{\mathbb{R}^n} |\partial_{y_1}(\chi u_{\mathbf{x};1})(y,-s)|^2 e^{-\frac{1}{2}f}\,d\nu_{-s}(y)\\
&\le C(\beta,\Lambda,\Upsilon)\int_{\mathbb{R}^n} |\partial_{y_1}(\chi u_{\mathbf{x};1})(y,-r^2)|^2 e^{-\frac{1}{2}f}\,d\nu_{-r^2}(y) + C\lambda^{\Upsilon} r^{4\Upsilon}\widehat{H}_{r^2}(\mathbf{x})\\
&\le C(\beta,\Lambda,\Upsilon)\int_{\mathbb{R}^n} |\partial_{y_1}(\chi u)(y,-r^2)|^2\,d\nu_{-r^2}(y) + C\lambda^{\Upsilon} r^{4\Upsilon}\widehat{H}_{r^2}(\mathbf{0}).
\end{aligned}
$$

Therefore, combining estimates and using Lemma 7.25(1), we get

$$
|\partial_{x_1}\log\widehat{H}_s(\mathbf{x})|^2 \le C(\beta,\Lambda,\Upsilon)\frac{1}{\widehat{H}_{r^2}(\mathbf{0})}\int_{\mathbb{R}^n} |\partial_{y_1}(\chi u)|^2\,d\nu_{-r^2}(y) + C(\Lambda,\Upsilon)\lambda^2 r^2.
$$

Integrating this from $\mathbf{0}$ to $(x_1,0)$ and then applying Lemma 7.25(1) and (7.58a), we obtain

$$
\left|\log\left(\frac{\widehat{H}_s((x_1,0))}{\widehat{H}_s(\mathbf{0})}\right)\right|^2 \le C(\beta,\Lambda,\Upsilon)\frac{|x_1|^2}{\widehat{H}_{r^2}(\mathbf{0})}\int_{\mathbb{R}^n} |\partial_{y_1}(\chi u)|^2\,d\nu_{-r^2}(y) + C(\Lambda,\Upsilon)\lambda^2 r^2
$$

$$\leq C(\beta,\Lambda,\Upsilon)\frac{|x_1|^2}{r^2}\int_{r^2}^{e^2r^2}\frac{1}{\widehat{H}_\tau(\mathbf{0})}\int_{\mathbb{R}^n}|\partial_{y_1}(\chi u)|^2\,d\nu_{-\tau}(y)d\tau+C(\Lambda,\Upsilon)\lambda^2r^2.$$

Next, we estimate how $\widehat{H}_s$ changes in the time direction. Let $\{\mathbf{x}'_j := (x_1,t'_j)\}_{j=0}^M$ with $t'_0=0$ and $t'_M=t_1$, where $M\leq C(\beta)$ and $|t'_j-t'_{j+1}|\leq\frac{\beta^2r^2}{100}$. Then

$$\begin{aligned}
&\widehat{H}_s(\mathbf{x}'_{j-1})-\widehat{H}_s(\mathbf{x}'_j)\\
&=\int_{\mathbb{R}^n}\left((\chi u_{\mathbf{x}'_{j-1};1})^2-(\chi u_{\mathbf{x}'_j;1})^2\right)\,d\nu_{-s}\\
&=\int_{\mathbb{R}^n}\left((\chi u_{\mathbf{x}'_{j-1};1})^2(y,-s)-(\chi u_{\mathbf{x}'_{j-1};1})^2(y,t'_j-t'_{j-1}-s)\right)\,d\nu_{-s}\\
&\quad-\int_{\mathbb{R}^n}\left(\chi^2(y,-s)-\chi^2(y,t'_j-t'_{j-1}-s)\right)u^2_{\mathbf{x}'_{j-1};1}(y,t'_j-t'_{j-1}-s)\,d\nu_{-s}\\
&\quad+\int_{\mathbb{R}^n}\chi^2(y,-s)\left(u^2_{\mathbf{x}'_{j-1};1}(y,t'_j-t'_{j-1}-s)-u^2_{\mathbf{x}'_{j-1};1}(a^{\frac12}(\mathbf{x}'_{j-1})a^{-\frac12}(\mathbf{x}'_j)y,t'_j-t'_{j-1}-s)\right)\,d\nu_{-s}\\
&=: L_1+L_2+L_3.
\end{aligned}$$

We first estimate $L_1$. Using Hölder's inequality, Lemma 7.5, and Lemma 7.6, we get

$$\begin{aligned}
|L_1|&=2\left|\int_{\mathbb{R}^n}\left(\int_{t'_j-t'_{j-1}}^0(\chi u_{\mathbf{x}'_{j-1};1})(y,\overline{s}-s)\partial_t(\chi u_{\mathbf{x}'_{j-1};1})(y,\overline{s}-s)\,d\overline{s}\right)\,d\nu_{-s}\right|\\
&\leq C(\beta,\Lambda,\Upsilon)\left|t_1\widehat{H}_{r^2}(\mathbf{x}'_{j-1})\int_{t'_j-t'_{j-1}}^0\int_{\mathbb{R}^n}\left|\partial_t(\chi u_{\mathbf{x}'_{j-1};1})(y,\overline{s}-s)\right|^2e^{-\frac12 f}\,d\nu_{-s}d\overline{s}\right|^{\frac12}\\
&\leq C(\beta,\Lambda,\Upsilon)\left|t_1\widehat{H}_s(\mathbf{x}'_{j-1})\int_{s-t'_j+t'_{j-1}}^s\int_{\mathbb{R}^n}|\partial_t(\chi u_{\mathbf{x}'_{j-1};1})(y,-\tau)|^2e^{-\frac14 f}\,d\nu_{-\tau}d\tau\right|^{\frac12}\\
&\leq C(\beta,\Lambda,\Upsilon)\left|t_1\widehat{H}_s(\mathbf{x}'_{j-1})\int_{\frac{r^2}{100}}^{\frac{r^2}{10}}\int_{\mathbb{R}^n}|\partial_t(\chi u_{\mathbf{x}'_{j-1};1})|^2e^{-\frac14 f}d\nu_{-\tau}d\tau\right|^{\frac12}\\
&\leq C(\beta,\Lambda,\Upsilon)\left(|t_1|\widehat{H}_s(\mathbf{x}'_{j-1})\int_{\frac{r^2}{20}}^{\frac{r^2}{10}}\int_{\mathbb{R}^n}|\partial_t(\chi u)|^2d\nu_{-\tau}d\tau\right)^{\frac12}\\
&\leq C(\beta,\Lambda,\Upsilon)\widehat{H}_s(\mathbf{x}'_{j-1})\left(\frac{|t_1|}{r^2}\int_{r^2}^{e^2r^2}\frac{\tau}{\widehat{H}_\tau(\mathbf{0})}\int_{\mathbb{R}^n}|\partial_t(\chi u)|^2d\nu_{-\tau}d\tau\right)^{\frac12},
\end{aligned}$$

where we used Lemma 2.1 in the third line. Further, we used Proposition 7.21(1) in the fourth line. Moreover, we used Lemma 7.25(3) in the fifth line. Finally, we used Proposition 7.21(1) and Lemma 7.25(1) in the last line.

We now estimate $L_2$ using Lemma 7.5 and Lemma 7.25(1):

$$\begin{aligned}
L_2&=\int_{\mathbb{R}^n}\left(\int_{t'_j-t'_{j-1}}^0\partial_{\overline{s}}\chi^2(y,\overline{s}-s)\,d\overline{s}\right)u^2(y,t'_j-t'_{j-1}-s)\,d\nu_{-s}\\
&\leq C(\beta,\Lambda,\Upsilon)\lambda^\Upsilon|t_1|e^{-\frac{1}{10\lambda^\Upsilon s^{1-2\Upsilon}}}\widehat{H}_s(\mathbf{x}'_{j-1}).
\end{aligned}$$

Finally, we estimate $L_3$. Define $A(\ell) := \ell a^{\frac{1}{2}}(\mathbf{x}'_{j-1})a^{-\frac{1}{2}}(\mathbf{x}'_j) + (1-\ell)I$ for $\ell = [0,1]$. Note that $|A'(\ell)| \leq \lambda\sqrt{|t|_1}$. Then

$$\begin{aligned}
&\left|\frac{d}{d\ell}\int_{\mathbb{R}^n}\chi^2(y,-s)u^2_{\mathbf{x}'_{j-1};1}(A(\ell)y,t'_j-t'_{j-1}-s)\,d\nu_{-s}\right| \\
&= \left|\int_{\mathbb{R}^n}\chi^2(y,-s)2u_{\mathbf{x}'_{j-1};1}(A(\ell)y,t'_j-t'_{j-1}-s)(\nabla u_{\mathbf{x}'_{j-1};1}(A(\ell)y,t'_j-t'_{j-1}-s))\cdot(A'(\ell)y)\,d\nu_{-s}\right| \\
&\leq C|A'(\ell)|\int_{\mathbb{R}^n}\chi^2(A(\ell)^{-1}y,-s)|u_{\mathbf{x}'_{j-1};1}(y,t'_j-t'_{j-1}-s)||(\nabla u_{\mathbf{x}'_{j-1};1}(y,t'_j-t'_{j-1}-s))||y|e^{\frac{1}{4}f}\,d\nu_{-s} \\
&\leq \lambda\sqrt{|t_1|}C(\beta,\Lambda,\Upsilon)\widehat{H}_s(\mathbf{x}'_{j-1}),
\end{aligned}$$

where we changed variables $y \leftarrow A(\ell)^{-1}y$ in the third line and in the last line we used a computation similar to that in Lemma 7.5. Integration from $\ell = 0$ to $\ell = 1$ yields

$$|L_3| \leq \lambda\sqrt{|t_1|}C(\beta,\Lambda,\Upsilon)\widehat{H}_s(\mathbf{x}'_{j-1}).$$

Combining the estimates for $L_1$, $L_2$, and $L_3$, we get

$$\begin{aligned}
&\left|\frac{\widehat{H}_s(\mathbf{x}'_j)-\widehat{H}_s(\mathbf{x}'_{j-1})}{\widehat{H}_s(\mathbf{x}'_{j-1})}\right| \\
&\leq C(\Lambda,\Upsilon)\left(\frac{|t_1|}{r^2}\int_{r^2}^{e^2r^2}\frac{\tau}{\widehat{H}_\tau(\mathbf{0})}\int_{\mathbb{R}^n}|\partial_t(\chi u)(y,-\tau)|^2\,d\nu_{-\tau}d\tau\right)^{\frac{1}{2}} + C(\beta,\Lambda,\Upsilon)\lambda\sqrt{|t_1|}.
\end{aligned}\tag{7.70}$$

Since $\left|\log\left(\frac{\widehat{H}_s(\mathbf{x}_1)}{\widehat{H}_s(\mathbf{0})}\right)\right| \leq C(\Lambda)$ by Lemma 7.25(1), we can assume that the right-hand side of (7.70) is at most $\frac{1}{4}$. In that case, we can use the following elementary estimate

$$\left|\log\left(\frac{\widehat{H}_s(\mathbf{x}'_j)}{\widehat{H}_s(\mathbf{x}'_{j-1})}\right)\right| = \left|\log\left(1+\frac{\widehat{H}_s(\mathbf{x}'_j)-\widehat{H}_s(\mathbf{x}'_{j-1})}{\widehat{H}_s(\mathbf{x}'_{j-1})}\right)\right| \leq 4\left|\frac{\widehat{H}_s(\mathbf{x}'_{\mathbf{j}})-\widehat{H}_s(\mathbf{x}'_{j-1})}{\widehat{H}_s(\mathbf{x}'_{j-1})}\right|$$

to obtain the claim. □

7.3.2. *Symmetry and pinching.* As a modification of Theorem 3.6, we prove that $\widehat{\mathcal{E}}^{k,\alpha}_r(\mathbf{x})$ is small if and only if $\mathbf{x}$ is frequency-pinched and almost $k$-symmetric at scale $r$.

**Theorem 7.27.** *If $\lambda \leq \overline{\lambda}(\Lambda)$, then the following hold for any $\epsilon \in (0,1]$. Let $u$ satisfy (7.1) and (7.2), $r \in (0,1]$, and $\mathbf{x}_0 \in P(\mathbf{0},10)$.*

*(1) Then $u$ is weakly $(k, C(\Lambda,\Upsilon)\alpha^{-n}\widehat{\mathcal{E}}^{k,\alpha}_r(\mathbf{x}_0), r)$-symmetric at $\mathbf{x}_0$.*

*(2) Suppose $u$ is weakly $(k,\epsilon,10^2r)$-symmetric at $\mathbf{x}$ with respect to $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ and*

$$|\widehat{N}_{\mathbf{x}}(10^2r^2)-\widehat{N}_{\mathbf{x}}(10^{-2}\kappa^2r^2)| + \lambda^{\frac{\Upsilon}{4}}r^\Upsilon \leq \epsilon.$$

*Then*

$$\left|\widehat{N}_{\mathbf{v}}(50r^2)-\widehat{N}_{\mathbf{v}}\left(50^{-1}\kappa^2r^2\right)\right| + \lambda^{\frac{\Upsilon}{4}}r^\Upsilon \leq C(\kappa,\Lambda,\Upsilon)\sqrt{\epsilon}\tag{7.71}$$

*for any $\mathbf{v} \in (\mathbf{x}+V)\cap P(\mathbf{x},10r)$. As a consequence, $\widehat{\mathcal{E}}^{k,1}_s(\mathbf{v}) \leq C(\kappa,\Lambda,\Upsilon)\sqrt{\epsilon}$, and $u$ is weakly $(k, C(\kappa,\Lambda,\Upsilon)\sqrt{\epsilon}, s)$-symmetric at $\mathbf{v}$ with respect to $V$ for all $s \in [\kappa r, r]$.*

*Proof of Theorem 7.27(2).* Using (7.8) and Theorem 7.8(1), we have

$$\log\left(\frac{\widehat{H}_\tau(\mathbf{x})}{\widehat{H}_{\frac{1}{2}\tau}(\mathbf{x})}\right) - C\lambda^\Upsilon\tau^\Upsilon \leq \log(2)\widehat{N}_{\mathbf{x}}(\tau) \leq \log\left(\frac{\widehat{H}_{2\tau}(\mathbf{x})}{\widehat{H}_\tau(\mathbf{x})}\right) + C\lambda^\Upsilon\tau^\Upsilon\tag{7.72}$$

for all $\mathbf{x} \in P(\mathbf{0}, 10)$ and $\tau \in (0, 1]$. This together with Lemma 7.26 implies that for any $\mathbf{v} \in (\mathbf{x} + V) \cap P(\mathbf{x}, 10r)$ and $\tau \in (0, 10^2 r^2]$, we have

$$\begin{aligned}\log(2)\widehat{N}_{\mathbf{x}}(\tau) - C\lambda^{\Upsilon}\tau^{\Upsilon} &\le \log\left(\frac{\widehat{H}_{2\tau}(\mathbf{x})}{\widehat{H}_{\tau}(\mathbf{x})}\right) \le \log\left(\frac{\widehat{H}_{2\tau}(\mathbf{v})}{\widehat{H}_{\tau}(\mathbf{v})}\right) + C\sqrt{\epsilon} + C\lambda r \\ &\le \log(2)\widehat{N}_{\mathbf{v}}(2\tau) + C\sqrt{\epsilon} + C\lambda^{\Upsilon}\tau^{\Upsilon},\end{aligned}$$

so that

$$\widehat{N}_{\mathbf{x}}(\tau) \le \widehat{N}_{\mathbf{v}}(2\tau) + C\sqrt{\epsilon} + C\lambda^{\frac{\Upsilon}{4}} r^{\Upsilon} \tag{7.73}$$

when $\tau \in [50^{-1}\kappa^2 r^2, 50r^2]$. Similarly, we have

$$\widehat{N}_{\mathbf{x}}(\tau) \ge \widehat{N}_{\mathbf{v}}(\frac{1}{2}\tau) - C\sqrt{\epsilon} - C\lambda^{\frac{\Upsilon}{4}} r^{\Upsilon}. \tag{7.74}$$

Combining (7.73) and (7.74), we get (7.71). The remaining claims follow from the proof of (1). □

We will prove Theorem 7.27(1) as a consequence of a cone-splitting inequality in Theorem 7.29. While proving the cone splitting inequality, we will need the following lemma which establishes the equivalence of two weighted $L^2$-norms of polynomials.

**Lemma 7.28.** *Fix $\alpha > 0$. Suppose $p$ is a polynomial on $\mathbb{R}^n$ of degree at most $m$. Then*

$$\int_{\mathbb{R}^n} p^2 \, d\nu_{\mathbf{x}_0;t} \le e^{C\alpha(1+m)} \int_{\mathbb{R}^n} p^2 e^{-\alpha f_{\mathbf{x}_0}} \, d\nu_{\mathbf{x}_0;t}$$

*for all $\mathbf{x}_0 = (x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$ and $t < t_0$.*

*Proof.* By a change of coordinate $x \leftarrow \frac{x - x_0}{\sqrt{t - t_0}}$, we assume that $t = -1$ and $\mathbf{x} = 0$.

Since $p$ has degree at most $m$, we can write $p = \sum_{k=0}^{m} \widehat{p}_k$ where $\widehat{p}_k \in \widehat{\mathcal{P}}_m$. Therefore, integration by parts gives

$$\int_{\mathbb{R}^n} |\nabla p|^2 \, d\nu = -\sum_{k=0}^{m} \int_{\mathbb{R}^n} \widehat{p}_k \Delta_f \widehat{p}_k \, d\nu = \sum_{k=0}^{m} \frac{k}{2} \int_{\mathbb{R}^n} \widehat{p}_k^2 \, d\nu \le \frac{m}{2} \int_{\mathbb{R}^n} p^2 \, d\nu. \tag{7.75}$$

Using (7.75) and (7.59) with $\alpha \leftarrow 0$, we get

$$\int_{\mathbb{R}^n} f p^2 d\nu \le C \int_{\mathbb{R}^n} (p^2 + |\nabla p|^2) \, d\nu \le C(1+m) \int_{\mathbb{R}^n} p^2 \, d\nu. \tag{7.76}$$

Fix $R^2 = 2C(1+m)$. Then

$$\begin{aligned}\int_{\mathbb{R}^n} p^2 \, d\nu &= \int_{\mathbb{R}^n \setminus B_R} p^2 \, d\nu + \int_{B_R} p^2 \, d\nu \le \frac{4}{R^2} \int_{\mathbb{R}^n \setminus B_R} f p^2 \, d\nu + e^{\frac{\alpha R^2}{4}} \int_{B_R} p^2 e^{-\alpha f} \, d\nu \\ &\le \frac{C(1+m)}{R^2} \int_{\mathbb{R}^n} p^2 \, d\nu + e^{\frac{\alpha R^2}{4}} \int_{\mathbb{R}^n} p^2 e^{-\alpha f} \, d\nu.\end{aligned}$$

Rearranging the inequality gives the desired result. □

We now prove the cone splitting inequality as an analog of Proposition 3.10.

**Theorem 7.29.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda)$. Let $u$ satisfy (7.1) and (7.2), and suppose $r \in (0, 1]$. For any $\mathbf{x}_0, \mathbf{x}_1 \in P(\mathbf{0}, 10)$ satisfying $|\mathbf{x}_1 - \mathbf{x}_0| < \frac{1}{10} r$,*

$$\begin{aligned}\int_{e^{-2}r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u_{\mathbf{x}_j}(\tau)} \int_{\mathbb{R}^n} \left(|(x_1 - x_0) \cdot \nabla(\chi u_{\mathbf{x}_j;1})|^2 + |(t_1 - t_0) \cdot \partial_t(\chi u_{\mathbf{x}_j;1})|^2\right) d\nu_{-\tau} \frac{d\tau}{\tau} \\ \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{x}_1, \mathbf{x}_0\})\end{aligned}$$

*for $j = 0, 1$.*

*Proof.* We will prove the statement with $j = 0$. By replacing $u$ with $u_{\mathbf{x}_0;1}$ we can assume that $\mathbf{x}_0 = \mathbf{0}$ and $u_{\mathbf{x}_0;1} = u$. By Lemma 7.17 and Lemma 7.14, there exist $p_i \in \mathcal{P}_{m_i}$ such that

$$\sup_{\tau\in[\frac{1}{40}r^2,40r^2]} \frac{1}{\widehat{H}^u_{\mathbf{x}_i}(\tau)} \int_{\mathbb{R}^n} \left( \chi u_{\mathbf{x}_i;1} - \sqrt{\frac{\widehat{H}^u_{\mathbf{x}_i}(\tau)}{\tau^{m_i}}} p_i \right)^2 d\nu_{-\tau} \tag{7.77}$$
$$\leq C(\Lambda,\Upsilon)(\|\widetilde{w}_{\mathbf{x}_i} - \widehat{p}_i\|_{[-\log(50)-2\log(r),\log(50)-2\log(r)]} + \lambda^{\frac{\Upsilon}{2}} r^{2\Upsilon}) \leq C(\Lambda,\Upsilon)\widehat{\mathcal{E}}_r(\mathbf{x}_i)$$

for all $i \in \{0,1\}$. For ease of notation, we define

$$q_i(x,t) := \sqrt{\frac{\widehat{H}^u_{\mathbf{x}_i}(\tau_i)}{\tau_i^{m_i}}} p_i(a^{\frac{1}{2}}(\mathbf{x}_i)(x-x_i), t-t_i), \qquad \chi_i(x,t) := \chi(a^{\frac{1}{2}}(\mathbf{x}_i)(x-x_i), t-t_i).$$

Note that for all $i,j \in \{0,1\}$ we have

$$\operatorname{supp}\chi_j(\cdot,t) \subset B(x_i, 10\lambda^{-\frac{\Upsilon}{2}}\tau_i^{\Upsilon}), \qquad \chi_j(\cdot,t) \equiv 1 \text{ in } B(x_i, 10^{-1}\lambda^{-\frac{\Upsilon}{2}}\tau_i^{\Upsilon}).$$

Let $\overline{q}_i$ be the $L^2(\mathbb{R}^n, \nu_{\mathbf{x}_i;t_i-\tau})$-projection of $q_i$ onto the $m_i$-eigenspace of $-2\tau_i\Delta_{f_i}$. For $\tau \in [\frac{1}{10}r^2, 10r^2]$, we have

$$\begin{aligned}\int_{\mathbb{R}^n} (q_0 - \overline{q}_i)^2 e^{-\frac{1}{8}f_i} d\nu_{\mathbf{x}_i;-\tau} &\leq C\int_{\mathbb{R}^n} (q_i - \overline{q}_i)^2 d\nu_{\mathbf{x}_i;-\tau} + C\int_{\mathbb{R}^n} (q_i - \chi_i u)^2 e^{-\lambda f_i} d\nu_{\mathbf{x}_i;-\tau} \\ &\quad + C\int_{\mathbb{R}^n} (\chi_i u - \chi u)^2 d\nu_{\mathbf{x}_i;-\tau} + C\int_{\mathbb{R}^n} (\chi u - q_0)^2 e^{-\frac{1}{8}f_i} d\nu_{\mathbf{x}_i;-\tau} \\ &=: I_1 + I_2 + I_3 + I_4.\end{aligned} \tag{7.78}$$

We first estimate $I_1$. Since $p_i$ is a polynomial of degree at most $m_i$, we can argue as in (7.75) to get

$$\int_{\mathbb{R}^n} \tau^2 |\nabla^2 p_i|^2 \, d\nu_{-\tau_i} \leq C(m) \int_{\mathbb{R}^n} p_i^2 \, d\nu_{-\tau_i}. \tag{7.79}$$

Furthermore,

$$\int_{\mathbb{R}^n} |x \otimes \nabla p_i|^2 \, d\nu_{-\tau_i} \leq \int_{\mathbb{R}^n} \tau f |\nabla p|^2 \, d\nu_{-\tau_i} \leq C\int_{\mathbb{R}^n} p_i^2 \, d\nu_{-\tau_i}, \tag{7.80}$$

where in the last inequality we used (7.76) for the components of $\nabla p$ and argued as in (7.75). Since $q_i$ are polynomials of degree at most $m_i$, we can use Lemma 7.28 to get

$$\begin{aligned}&\int_{\mathbb{R}^n} (2\tau_i\Delta_{f_i} q_i + m_i q_i)^2 \, d\nu_{\mathbf{x}_i;t} \\ &\leq C\int_{\mathbb{R}^n} (2\tau_i\Delta_{f_i} q_i + m_i q_i)^2 e^{-10\lambda f_i} \, d\nu_{\mathbf{x}_i;t} \\ &= \frac{C\widehat{H}^u_{\mathbf{x}_i}(\tau_i)}{\tau_i^{m_i}} \int_{\mathbb{R}^n} \left( 2\tau \operatorname{tr}\left( (a(\mathbf{x}_i) - I)\left( \nabla^2 p_i - \frac{x\otimes\nabla p_i}{2\tau} \right) \right) \right)^2 (a^{\frac{1}{2}}(\mathbf{x}_i)(x-x_i), -\tau_i) e^{-10\lambda f_i} \, d\nu_{\mathbf{x}_i;t}(x) \\ &\leq \frac{C\widehat{H}^u_{\mathbf{x}_i}(\tau_i)}{\tau_i^{m_i}} \lambda^2 r^2 \int_{\mathbb{R}^n} \tau^2 \left| \nabla^2 p_i - \frac{x\otimes\nabla p_i}{2\tau} \right|^2 (a^{\frac{1}{2}}(\mathbf{x}_i)(x-x_i), -\tau_i) e^{-10\lambda f_i} \, d\nu_{\mathbf{x}_i;t}(x) \\ &\leq \frac{C\widehat{H}^u_{\mathbf{x}_i}(\tau_i)}{\tau_i^{m_i}} \lambda^2 r^2 \int_{\mathbb{R}^n} \tau^2 \left| \nabla^2 p_i - \frac{x\otimes\nabla p_i}{2\tau} \right|^2 e^{-2\lambda f} \, d\nu_{-\tau_i} \\ &\leq \frac{C\widehat{H}^u_{\mathbf{x}_i}(\tau_i)}{\tau_i^{m_i}} \lambda^2 r^2 \int_{\mathbb{R}^n} \tau^2 \left| \nabla^2 p_i \right|^2 + |x\otimes\nabla p_i|^2 \, d\nu_{-\tau_i} \leq \frac{C\widehat{H}^u_{\mathbf{x}_i}(\tau_i)}{\tau_i^{m_i}} \lambda^2 r^2 \int_{\mathbb{R}^n} p_i^2 \, d\nu_{-\tau_i} \leq C\lambda^2 r^2 \widehat{H}^u_{\mathbf{x}_i}(\tau_i),\end{aligned}$$

where we used (7.79) and (7.80) in the second-to-last inequality. Therefore, Lemma 2.8 implies that

$$I_1 = C \int_{\mathbb{R}^n} (q_i - \overline{q}_i)^2 \, d\nu_{\mathbf{x}_i;-\tau} \le C\lambda^2 r^4 \widehat{H}^u_{\mathbf{x}_i}(\tau_i). \tag{7.81}$$

Second, a coordinate change in (7.77) gives

$$I_2 \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\mathbf{x}_i)\widehat{H}^u_{\mathbf{x}_i}(\tau_i). \tag{7.82}$$

Third, the computation in (7.67) gives

$$I_3 \le C(\Lambda)\exp\left(-\frac{1}{e^{60}\lambda^\Upsilon \tau}\right)\widehat{H}^u_{\mathbf{x}_i}(\tau_i).$$

Finally, by Lemma 2.1, we change bases $e^{-\frac{1}{8}f_i} d\nu_{\mathbf{x}_i;t} \le C d\nu_t$ in (7.77) with $i \leftarrow 0$ to get

$$I_4 \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\mathbf{0})\widehat{H}^u(\tau). \tag{7.83}$$

Combining Lemma 7.28, (7.81), (7.82), (7.67), (7.83), (7.78), and Lemma 7.25(1), we get

$$\int_{\mathbb{R}^n} (q_0 - \overline{q}_i)^2 d\nu_{\mathbf{x}_i;t} \le C \int_{\mathbb{R}^n} (q_0 - \overline{q}_i)^2 e^{-\frac{1}{8}f_i} d\nu_{\mathbf{x}_i;t} \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\})\widehat{H}^u_{\mathbf{x}_i}(\tau).$$

Lemma 7.15 then implies that

$$|N^{p_0}_{\mathbf{x}_i}(\tau_i) - m_i| = |N^{q_0}_{\mathbf{x}_i}(\tau_i) - m_i| \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\}). \tag{7.84}$$

for any $\tau_i \in [\frac{1}{8}r^2, 8r^2]$. Since $p_0$ is caloric, Corollary 2.14 implies that $m_i = m$. Then Proposition 3.10 implies that

$$\int_{\mathbb{R}^n} |x_1 \cdot \nabla p_0|^2 + |t_1 \cdot \partial_t p_0|^2 \, d\nu_{-\tau} \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\}) \int_{\mathbb{R}^n} p_0^2 \, d\nu_{-\tau} \tag{7.85}$$

for any $\tau \in [\frac{1}{8}r^2, 8r^2]$. Using (7.8) together with (7.84) and (7.85), we get

$$\begin{aligned}\int_{\mathbb{R}^n} |t_1 \partial_t q_0|^2 \, d\nu_{-\tau} &\le \frac{C\widehat{H}^u(\tau)}{\tau^m} \int_{\mathbb{R}^n} |t_1 \partial_t p_0|^2 \, d\nu_{-\tau} + |t_1|^2 C\left(\frac{d}{d\tau}\log \widehat{H}^u - \frac{m}{\tau}\right)^2 q_0^2 \, d\nu_{-\tau} \\ &\le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\}) \int_{\mathbb{R}^n} q_0^2 \, d\nu_{-\tau}.\end{aligned}$$

Using this and (7.85) again, we obtain that for any $\tau \in [\frac{1}{8}r^2, 8r^2]$

$$\int_{\mathbb{R}^n} |x_1 \cdot \nabla q_0|^2 + |t_1 \cdot \partial_t q_0|^2 \, d\nu_{-\tau} \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\}) H^{q_0}(\tau) \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\})\widehat{H}^u(\tau),$$

where we used (7.77) in the second inequality. On the other hand, Lemma 7.17 and Lemma 7.14 give

$$\int_{e^{-2}r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla(\chi u - q_0)|^2 + \tau\, |\partial_t(\chi u - q_0)|^2 \, d\nu_{-\tau} d\tau \le C(\Lambda, \Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0}, \mathbf{x}_1\}). \tag{7.86}$$

In particular, we get our claim:

$$\begin{aligned}&\int_{e^{-2}r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |x_1 \cdot \nabla(\chi u)|^2 + |t_1 \partial_t \chi u|^2 \, d\nu_{-\tau} \frac{d\tau}{\tau} \\ &\le C \int_{e^{-2}r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla(\chi u - q_0)|^2 + \tau\, |\partial_t(\chi u - q_0)|^2 \, d\nu_{-\tau} d\tau \\ &\quad + C \int_{e^{-2}r^2}^{e^2 r^2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |x_1 \cdot \nabla q_0|^2 + |t_1 \partial_t q_0|^2 \, d\nu_{-\tau} \frac{d\tau}{\tau}\end{aligned}$$

$$\leq C(\Lambda,\Upsilon)\widehat{\mathcal{E}}_r(\{\mathbf{0},\mathbf{x}_1\}).$$

□

*Proof of Theorem 7.27(1).* The proof follows from Lemma 3.3 and Theorem 7.29 as in the proof of Theorem 3.11. □

7.3.3. *Existence of a plane of symmetry.* As an analog of Lemma 3.12, we prove that there exists a plane of symmetry that is uniform across scales and locations.

**Lemma 7.30.** *Suppose $\Lambda \in [1,\infty)$. The following holds if $\lambda \leq \overline{\lambda}(\Lambda,\Upsilon)$ and $\delta \leq C(\Lambda)^{-1}$. Fix any $\eta,\kappa,r_0 \in (0,1]$ and any solution $u$ of (7.1) satisfying (7.2). Suppose $\mathcal{C} \subseteq P(\mathbf{0},r_0)$ and $\mathbf{r}_\bullet : \mathcal{C} \to [0,r_0]$ is a function satisfying the following for some $m \in \mathbb{N}$:*

*(a) there exists $\mathbf{x}_0 \in \mathcal{C}$ and $s_0 \in [\mathbf{r}_{\mathbf{x}_0}, r_0]$ such that $u$ is weakly $(k,\delta,s_0)$-symmetric with respect to $V$;*

*(b) $\sup_{\mathbf{x}\in\mathcal{C}}\sup_{s\in[10^{-6}\kappa\mathbf{r}_{\mathbf{x}},10^6\kappa^{-1}r_0]}(|\widehat{N}_{\mathbf{x}}(s^2)-m| + \lambda^{\frac{\Upsilon}{4}}s^{\frac{\Upsilon}{2}}) < \delta$.*

*Then the following statements hold for any $\mathbf{x} \in \mathcal{C}$:*

*(1) for any $s \in [\kappa\mathbf{r}_{\mathbf{x}},\kappa^{-1}r_0]$, $u$ is weakly $(k,C(\Lambda,\Upsilon)\delta,s)$-symmetric with respect to $V$ at $\mathbf{x}$;*

*(2) for any $s \in [\mathbf{r}_{\mathbf{x}},r_0]$,*

$$(\mathbf{x}+V)\cap P(\mathbf{x},10s) \subset \{\mathbf{y}\in P(\mathbf{x},10s) \colon |\widehat{N}_{\mathbf{y}}(\kappa^2 s^2)-m| < C(\kappa,\Lambda,\Upsilon)\sqrt{\delta}\};$$

*(3) if in addition we assume that $u$ is not weakly $(k+1,\eta,s_0)$-symmetric at $\mathbf{x}_0$, then, for any $s \in [\mathbf{r}_{\mathbf{x}},10^{-2}r_0]$ and $\zeta \geq C(\Lambda,\Upsilon)\sqrt{\delta}$,*

$$\{\mathbf{y}\in P(\mathbf{x},10s) : \widehat{\mathcal{E}}_{100s}(\mathbf{y}) < \zeta\} \subseteq P(\mathbf{x}+V, C(\Lambda,\Upsilon)\eta^{-\frac{1}{n}}\zeta^{\frac{1}{n}}s).$$

*Proof.* By parabolic rescaling, we will assume that $r_0 = 1$.

(1) By Lemma 7.21(2) with $r \leftarrow 100$, $r_1 \leftarrow \mathbf{r}_{\mathbf{x}_0}$, $r_2 \leftarrow s_0$, $u$ is weakly $(k,C(\Lambda,\Upsilon)\delta,1)$-symmetric at $\mathbf{x}_0$ with respect to $V$. This together with Lemma7.25(2)(3) imply that for any $\mathbf{x} \in \mathcal{C}$

$$\int_{10^3}^{e^2\cdot 5\cdot 10^3}\frac{1}{\widehat{H}^u_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\pi_L\nabla(\chi u_{\mathbf{x};1})|^2 + \tau\,|\partial_t(\chi u_{\mathbf{x};1})|^2 e^{-\frac{1}{4}f}\,d\nu_{-\tau}d\tau \leq C(\Lambda,\Upsilon)\delta. \tag{7.87}$$

We claim that the above estimate holds without the weight $e^{-\frac{1}{4}f}$.

**Claim 7.30.1.** *$u$ is weakly $(k,C(\Lambda,\Upsilon)\delta,10^{\frac{3}{2}})$-symmetric with respect to $V$ at any $\mathbf{x}\in\mathcal{C}$.*

*Proof of Claim 7.30.1.* We consider the case where $V = L\times\mathbb{R}$ for some $k$-plane $L \subseteq \mathbb{R}^n$, as the case $V = L\times\{0\}$ is similar. On the other hand, Lemma 7.17 and Lemma 7.14 imply the existence of $p_{\mathbf{x}} \in \mathcal{P}_m$ for every $\mathbf{x}\in\mathcal{C}$ such that if $q_{\mathbf{x}} := \sqrt{\frac{\widehat{H}_{\mathbf{x}}(\tau)}{\tau^m}}p_{\mathbf{x}}$, then

$$\int_{10^3}^{e^2\cdot 5\cdot 10^3}\frac{1}{\widehat{H}^u_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\nabla(\chi u_{\mathbf{x};1}-q_{\mathbf{x}})|^2 + \tau\,|\partial_t(\chi u_{\mathbf{x};1}-q_{\mathbf{x}})|^2\,d\nu_{-\tau}d\tau \leq C(\Lambda,\Upsilon)\delta. \tag{7.88}$$

This together with (7.87) implies

$$\int_{10^3}^{e^2\cdot 5\cdot 10^3}\frac{1}{\widehat{H}^u_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\pi_L\nabla q_{\mathbf{x}}|^2 + \tau\,|\partial_t q_{\mathbf{x}}|^2 e^{-\frac{1}{4}f}\,d\nu_{-\tau}d\tau \leq C(\Lambda,\Upsilon)\delta.$$

Since $q_{\mathbf{x}}$ is a polynomial of degree at most $m$, we use Lemma 7.28 to get

$$\int_{10^3}^{e^2\cdot 5\cdot 10^3}\frac{1}{\widehat{H}^u_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\pi_L\nabla q_{\mathbf{x}}|^2 + \tau\,|\partial_t q_{\mathbf{x}}|^2\,d\nu_{-\tau}d\tau \leq C(\Lambda,\Upsilon)\delta. \tag{7.89}$$

Thus (7.88) and (7.89) imply the claim. □

The conclusion then follows by applying Claim 7.30.1 and Lemma 7.21(2) to each point $\mathbf{x} \in \mathcal{C}$.

(2) Apply (1) and Theorem 7.27(2).

(3) Suppose $\mathbf{x} \in \mathcal{C}$, $s \in [\mathbf{r_x}, 10^{-2}]$, $\mathbf{y} \in P(\mathbf{x}, 10s)$, and $\widehat{\mathcal{E}}_{100s}(\mathbf{y}) < \zeta$ for some $\zeta \geq C(\Lambda, \Upsilon)\sqrt{\delta}$. Set $\epsilon := C(\Lambda, \Upsilon)\eta^{-\frac{1}{n}}\zeta^{\frac{1}{n}}$, and assume by way of contradiction that $\mathbf{y} \notin P(\mathbf{x} + V, \epsilon s)$. By applying (2) twice with $s \leftarrow s$ and $s \leftarrow 100s$, there exists a $(k, 100s)$-independent set $\{\mathbf{v}_i\}_{i \in I} \subseteq (\mathbf{x}+V) \cap P(\mathbf{x}, 10s)$ containing $\mathbf{x}$ and satisfying $\widehat{\mathcal{E}}_{100s}(\mathbf{v}_i) < C(\Lambda, \Upsilon)\sqrt{\delta} \leq \zeta$ for all $i \in I$. Thus $\widehat{\mathcal{E}}^{k+1,\epsilon}_{100s}(\mathbf{x}) < \zeta$, so by Theorem 7.27(1) and Lemma 7.21(2), $u$ is weakly $(k+1, C(\Lambda, \Upsilon)^{-1}\eta, s)$-symmetric at $\mathbf{x}$. By Lemma 7.21(2), it follows that $u$ is weakly $(k+1, C(\Lambda, \Upsilon)^{-1}\eta, 100)$-symmetric at $\mathbf{x}$, hence by Claim 7.30.1, $u$ is also weakly $(k+1, C(\Lambda, \Upsilon)^{-1}\eta, 1)$-symmetric at $\mathbf{x}_0$. Another application of Lemma 7.21(2) finally yields that $u$ is weakly $(k+1, \eta, s_0)$-symmetric at $\mathbf{x}_0$, a contradiction. □

7.3.4. *$L^2$-subspace approximation.* Let $\mu$ be a probability measure supported on $\overline{P}(\mathbf{0}, 1)$. Recall the definitions of $x_{\mathrm{cm}}$, $Q$ from (3.10).

We now state the analog of Lemma 3.17.

**Lemma 7.31.** *Let $v_j$ be the eigenvector of the covariance matrix $Q$ with eigenvalue $\lambda_j$. Let $u$ be a solution to (7.1) satisfying (7.2). Then, for any $r \in (0, 1]$, we have*

$$\int_{100n^2r^2}^{100e^2n^2r^2} \frac{\lambda_j}{\widehat{H}^u_{\mathbf{x}_0}(100n^2r^2)} \int_{\mathbb{R}^n} (v_j \cdot \nabla(u\chi_{\mathbf{0};1}))^2 \, d\nu_{-\tau} d\tau \leq C(\Lambda, \Upsilon) \int_{P(\mathbf{0},1)} \widehat{\mathcal{E}}_{20nr}(\mathbf{x}) \, d\mu(\mathbf{x}).$$

*Proof.* Taking $u \leftarrow \chi u_{\mathbf{0};1}$ in (3.12) and integrating in time against $(\widehat{H}^u_{\mathbf{0}})^{-1}$ yields

(7.90)
$$\begin{aligned}
&\lambda_j \int_{100n^2r^2}^{100e^2n^2r^2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} (v_j \cdot \nabla(\chi u_{\mathbf{0};1})(z, -\tau))^2 \, d\nu_{-\tau}(z) d\tau \\
&\quad \leq \int_{100n^2r^2}^{100e^2n^2r^2} \frac{1}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \int_{P(\mathbf{0},1)} \int_{P(\mathbf{0},1)} \left(\nabla(\chi u_{\mathbf{0};1})(z, -\tau) \cdot (x - y)\right)^2 \, d\mu(\mathbf{x}) d\mu(\mathbf{y}) d\nu_{-\tau}(z) d\tau.
\end{aligned}$$

Moreover, for any fixed $\mathbf{x}, \mathbf{y} \in P(\mathbf{0}, 1)$, by Lemma 7.25(2) and Proposition 7.29, we obtain

$$\begin{aligned}
&\int_{100n^2r^2}^{100e^2n^2r^2} \frac{1}{\widehat{H}^u_{\mathbf{0}}(\tau)} \int_{\mathbb{R}^n} (\nabla(\chi u_{\mathbf{0};1})(z, -\tau) \cdot (x - y))^2 \, d\nu_{-\tau}(z) d\tau \\
&\leq C(\Lambda, \Upsilon) \int_{400n^2r^2}^{400e^2n^2r^2} \frac{1}{\widehat{H}^u_{\mathbf{x}}(\tau)} \int_{\mathbb{R}^n} (\nabla(\chi u_{\mathbf{x};1})(z, -\tau) \cdot (x - y))^2 \, d\nu_{-\tau}(z) d\tau + C(\Lambda, \Upsilon)\lambda^{\frac{\Upsilon}{2}} r^{2\Upsilon} \\
&\leq C(\Lambda, \Upsilon) \widehat{\mathcal{E}}_{20nr}(\{\mathbf{x}, \mathbf{y}\}).
\end{aligned}$$

The lemma then follows by integrating over $\mathbf{x}, \mathbf{y} \in P(\mathbf{0}, 1)$ and combining with (7.90). □

The following is a modification of Lemma 3.14.

**Lemma 7.32.** *Let $k \in \{n, n+1\}$. For any finite measure $\mu$ supported in $P(\mathbf{0}, 1)$ on $\mathbb{R}^{n+1}$ and any solution $u$ of (7.1) satisfying (7.2), if $u$ is weakly $(k, \epsilon, 3r)$-symmetric but not weakly $(k+1, \eta, r)$-symmetric at $\mathbf{0}$, then, for any $P(\mathbf{x}, r) \subseteq P(\mathbf{0}, 2)$, if $\epsilon \leq C^{-1}(\Lambda, \Upsilon)\eta$, we have*

$$\beta^2_{\mathcal{P},k}(\mathbf{x}, r; \mu) \leq \frac{C^\Lambda}{r^k \eta} \left( \int_{P(\mathbf{x},r)} \widehat{\mathcal{E}}_{20nr}(\mathbf{y}) d\mu(\mathbf{y}) \right).$$

*Proof.* Given Lemma 7.31 and the proof of Lemma 3.14, it remains to consider the case where $k = n$ and the symmetry is spatial. Let $w$ be the caloric function with $w(\cdot, 20r^2) = (\chi u_{\mathbf{0};1})(\cdot, 20r^2)$, so that

$$\sup_{\tau\in[r^2,10r^2]}\int_{\mathbb{R}^n}\tau|\nabla(w-\chi u_{\mathbf{x}_0;1})|^2d\nu_{-\tau}+\int_{r^2}^{10r^2}\tau\int_{\mathbb{R}^n}|\partial_t(w-\chi u_{\mathbf{x}_0;1})|^2\,d\nu_{-\tau}d\tau \tag{7.91}$$
$$\leq C(\Lambda,\Upsilon)\lambda^{\Upsilon}\tau_0^{2\Upsilon}\widehat{H}_{\mathbf{x}_0}(\tau_0).$$

It follows that

$$r^2\int_{\mathbb{R}^n}|\nabla w|^2d\nu_{-9r^2}\leq (C\epsilon+C(\Lambda,\Upsilon)\lambda^{\Upsilon}r^{4\Upsilon})\int_{\mathbb{R}^n}w^2d\nu_{-9r^2},$$

so that integrating $\frac{d}{dt}\int_{\mathbb{R}^n}|\nabla w|^2d\nu_t\leq -2\int_{\mathbb{R}^n}|\nabla^2 w|^2d\nu_t$ from $t=-9r^2$ to $t=-e^2r^2$ yields

$$r^2\int_{-e^2r^2}^{-r^2}\int_{\mathbb{R}^n}|\partial_t w|^2d\nu_t dt\leq Cr^2\int_{-e^2r^2}^{-r^2}\int_{\mathbb{R}^n}|\nabla^2 w|^2d\nu_t dt\leq (C\epsilon+C(\Lambda,\Upsilon)\lambda^{\Upsilon}4^{4\Upsilon})\int_{\mathbb{R}^n}w^2d\nu_{-9r^2}.$$

Along with (7.91), this implies that $u$ is weakly $(k+1, C\epsilon+C(\Lambda,\Upsilon)\lambda^{\Upsilon}4^{4\Upsilon},1)$-symmetric, a contradiction if $\epsilon\leq C(\Lambda,\Upsilon)^{-1}\eta$. □

7.4. **$\epsilon$-regularity.** We now prove $\epsilon$-regularity results in the general setting, analogous to those in Section 6.1.

We define effective versions of nodal and singular sets in the general setting:

**Definition 7.33.** Suppose $u$ is a solution to (7.1) satisfying (7.2). For $r>0$, define

$$\widehat{Z}_r(u):=\left\{\mathbf{x}\in P(\mathbf{0},1)\colon \inf_{P(\mathbf{x},10^{-6}s)}|u|^2\leq\frac{1}{8}\int_{\mathbb{R}^n}|u_{\mathbf{x};1}|^2\,d\nu_{-s^2}\text{ for all } s\in[r,1]\right\},$$
$$\widehat{S}_r(u):=\left\{\mathbf{x}\in P(\mathbf{0},1)\colon \inf_{P(\mathbf{x},10^{-6}s)}\left(|u|^2+2s^2|\nabla u|^2\right)\leq\frac{1}{16}\int_{\mathbb{R}^n}|u_{\mathbf{x};1}|^2\,d\nu_{-s^2}\text{ for all } s\in[r,1]\right\}.$$

We define

$$\widehat{\mathscr{C}}^k(u):=\begin{cases}Z(u) & \text{when } k=n+1,\\ S(u) & \text{when } k=n,\end{cases}$$

and $\widehat{\mathscr{C}}_r^k$ analogously.

**Definition 7.34** (Quantitative stratum)**.** Suppose $u$ is a solution to (7.1) satisfying (7.2). The $(k,\epsilon,r_1,r_2)$ *quantitative stratum* is

$$\widehat{\mathcal{S}}^k_{\epsilon,r_1,r_2}:=\widehat{\mathcal{S}}^k_{\epsilon,r_1,r_2}(u):=\left\{\mathbf{x}\in P(\mathbf{0},1)\colon \widehat{\mathcal{E}}_r^{k+1,1}(\mathbf{x})\geq\epsilon,\ \forall\, r\in[r_1,r_2]\right\}.$$

We also set $\widehat{\mathcal{S}}^k_{\epsilon,r}:=\widehat{\mathcal{S}}^k_{\epsilon,r,1}$, and define $\widehat{\mathcal{S}}^k_{\epsilon}:=\bigcap_{r>0}\widehat{\mathcal{S}}^k_{\epsilon,r}$.

The following is a version of Proposition 6.5, where we prove the containment of $\widehat{\mathscr{C}}_r^k$ in the $k$-th quantitative strata.

**Proposition 7.35.** *The following holds if $\lambda\leq\overline{\lambda}(\Lambda,\Upsilon)$. Suppose $u$ is a solution to (7.1) satisfying (7.2). Let $k\in\{n,n+1\}$. If $\epsilon\leq C(\Lambda)^{-1}$, then, for any $r\in(0,1)$,*

$$\widehat{\mathscr{C}}_r^k(u)\subseteq\widehat{\mathcal{S}}^k_{\epsilon,r}(u). \tag{7.93}$$

As in the caloric setting (Lemma 6.6), we first prove that only constant functions can attain maximal symmetry.

**Lemma 7.36.** *The following holds if $\lambda\leq\overline{\lambda}(\Lambda,\Upsilon)$. Let $u$ be a solution of (7.1) satisfying (7.2). Suppose $u$ is spatially weakly $(n,\epsilon,r)$-symmetric at $\mathbf{x}_0$ with $\epsilon\leq\bar{\epsilon}(\Lambda)$. Then*

$$\inf_{P(\mathbf{x}_0,10^{-6}r)}|u|^2>\frac{1}{8}\int_{\mathbb{R}^n}(\chi u_{\mathbf{x}_0;1})^2\,d\nu_{-r^2}.$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we can assume that $\mathbf{x}_0 = \mathbf{0}$ and $u = u_{\mathbf{x}_0;1}$. Then

$$\log(2)\widehat{N}^u(r^2) \le \log\frac{\widehat{H}(4r^2)}{\widehat{H}(r^2)} + C\lambda^{\Upsilon}r^{2\Upsilon} \le \int_{r^2}^{e^2r^2}\frac{\widehat{N}^u(\tau)}{\tau}\,d\tau + C\lambda^{\Upsilon}r^{2\Upsilon}\tau d\tau < \epsilon,$$

where in the last inequality we use that $u$ is spatially weakly $(n,\epsilon,r)$-symmetric at $\mathbf{0}$. Then, for any $\mathbf{x} \in P(\mathbf{0},10^{-3}r)$, Lemma 7.25(1)(2) and Lemma 7.6 give

$$\int_{\mathbb{R}^n}|\nabla(\chi u_{\mathbf{x};1})|^2 e^{-\frac{1}{4}f}\,d\nu_{-\frac{1}{2}r^2} \le C(\Lambda,\Upsilon)\epsilon\widehat{H}^u_{\mathbf{x}}(\frac{1}{2}r^2).$$

Thus for any $\mathbf{x} \in P(\mathbf{0},10^{-3}r)$ and any $\tau \le \frac{1}{2}r^2$, Corollary 7.9 implies

$$\widehat{N}_{\mathbf{x}}(\tau) \le C(\Lambda,\Upsilon)\sqrt{\epsilon}.$$

Now arguing as in Lemma 6.6, for any $\mathbf{x} \in P(\mathbf{0},10^{-3}r)$, we have

$$C^{-1}\frac{1}{\tau^{\frac{n}{2}}}\int_{B(x,\sqrt{\tau})}|u(y,t-\tau)-1|^2\,dy \le \int_{\mathbb{R}^n}|\chi u_{\mathbf{x};1}-1|^2\,d\nu_{-\tau} \le C(\Lambda,\Upsilon)\sqrt{\epsilon}\widehat{H}(r^2),$$

where in the last inequality we use Lemma 7.25(1) and Lemma 7.6. Taking $\tau \coloneqq 10^{-7}r^2$, $\mathbf{x} \leftarrow (0,t)$ for $t \in (-10^{-6}r^2, 10^{-6}r^2)$, we then integrate over $t$ to obtain

$$\frac{1}{r^{n+2}}\int_{P(\mathbf{0},10^{-4}r)}|u-1|^2 d\mathcal{H}^{n+2}_{\mathcal{P}} \le C(\Lambda,\Upsilon)\sqrt{\epsilon}\widehat{H}(r^2).$$

The claim follows from standard parabolic interior estimates. $\square$

As in Lemma 6.7, we now prove that only linear functions can possess $n+1$ symmetries.

**Lemma 7.37.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda,\Upsilon)$. Let $u$ be a solution to (7.1) satisfying (7.2). Suppose $u$ is weakly $(n+1,\epsilon,r)$-symmetric at $\mathbf{x}_0$ with $\epsilon \le \overline{\epsilon}$. Then*

$$\inf_{P(\mathbf{x}_0,10^{-6}r)}|u|^2 + 2r^2|\nabla u|^2 > \frac{1}{16}\int_{\mathbb{R}^n}(\chi u_{\mathbf{x}_0})^2\,d\nu_{-r^2}.$$

*Proof.* By replacing $u$ with $u_{\mathbf{x}_0;1}$, we can assume that $\mathbf{x}_0 = \mathbf{0}$ and $u = u_{\mathbf{x}_0;1}$. By rotating coordinates, we may assume that $u$ is weakly symmetric with respect to $(\mathbb{R}^{n-1}\times\{0\})\times\mathbb{R}$ so that

$$\sum_{i=1}^{n-1}\int_{r^2}^{e^2r^2}\frac{1}{\widehat{H}(\tau)}\int_{\mathbb{R}^n}|\partial_i(\chi u)|^2\,d\nu_{-\tau}d\tau + \int_{r^2}^{e^2r^2}\frac{\tau}{\widehat{H}(\tau)}\int_{\mathbb{R}^n}|\partial_t(\chi u)|^2\,d\nu_{-\tau}d\tau \le \epsilon.$$

By propagation of symmetry in Lemma 7.21 and change of base point in Lemma 7.25, we know that for any $\mathbf{x} \in P(\mathbf{0},\frac{1}{100}r)$

$$\sum_{i=1}^{n-1}\int_{r^2}^{e^2r^2}\frac{1}{\widehat{H}_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\partial_i(\chi u_{\mathbf{x};1})|^2\,d\nu_{-\tau}d\tau + \int_{r^2}^{e^2r^2}\frac{\tau}{\widehat{H}_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\partial_t(\chi u_{\mathbf{x};1})|^2\,d\nu_{-\tau}d\tau \le C\sqrt{\epsilon}.$$

If $w_{\mathbf{x}}$ is the caloric approximation of $\chi u_{\mathbf{x};1}$, then we know that

$$\sum_{i=1}^{n-1}\int_{r^2}^{e^2r^2}\frac{1}{\widehat{H}_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\partial_i(w_{\mathbf{x}})|^2\,d\nu_{-\tau}d\tau + \int_{r^2}^{e^2r^2}\frac{\tau}{\widehat{H}_{\mathbf{x}}(\tau)}\int_{\mathbb{R}^n}|\partial_t(w_{\mathbf{x}})|^2\,d\nu_{-\tau}d\tau \le C\sqrt{\epsilon}.$$

This implies $w_{\mathbf{x}}$ is $(n+1,C\sqrt{\epsilon},r)$-symmetric. Since $w_{\mathbf{x}}$ is caloric, Lemma 6.7 gives pointwise lower bounds for $w_{\mathbf{x}}$ and using standard parabolic regularity and the argument of Lemma 7.36, we get the pointwise estimate

$$\inf_{\mathbf{x}\in P(\mathbf{x}_0,10^{-6}r)}\left(|u|^2 + 2r^2|\nabla u|^2\right) \ge \frac{1}{16}\int_{\mathbb{R}^n}(\chi u)^2\,d\nu_{-r^2},$$

from which the claim follows. $\square$

*Proof of Proposition 7.35.* The proof is similar to the proof of Theorem 6.5 after we replace Theorem 3.6(1), Lemma 6.6, Lemma 6.7 with Theorem 7.27(1), Lemma 7.36, Lemma 7.37. $\square$

This lemma establishes that nodal and singular sets are contained in the super-level sets of the frequency function, see Lemma 6.8.

**Lemma 7.38.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$. Let $k \in \{n, n+1\}$. If $\epsilon \le C(\Lambda)^{-1}$, then, for any $\mathbf{x}_0 \in P(\mathbf{0}, 10)$ and $r \in (0, \epsilon]$, we have*

$$\widehat{\mathscr{C}}^k_r(u) \subset \left\{ \mathbf{x} \in P(\mathbf{0}, 1) \colon \widehat{N}(\epsilon^{-2} r^2) \ge m_* - \frac{2}{3} \right\},$$

*where $m_*$ is defined as in (5.1).*

*Proof.* We focus on $k = n$, since the case $k = n+1$ is analogous. Suppose $\widehat{N}^u_{\mathbf{x}_0}(\epsilon^{-2} r^2) < \frac{4}{3}$. Let $w$ be the caloric function with $w(\cdot, -2\epsilon^{-2} r^2) = (\chi u_{\mathbf{x}_0;1})(\cdot, -2\epsilon^{-2} r^2)$ of $\chi u_{\mathbf{x}_0;1}$ as in Lemma 7.22. If $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$, we have

$$\begin{aligned}
&2\tau \int_{\mathbb{R}^n} |\nabla w|^2 \, d\nu_{-\epsilon^{-2} r^2} \\
&\le 2\lambda^{-\frac{\Upsilon}{2}} r^{-2\Upsilon} \int_{\mathbb{R}^n} |\nabla (w - \chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\epsilon^{-2} r^2} + (1 + \lambda^{\frac{\Upsilon}{2}} r^{2\Upsilon}) \int_{\mathbb{R}^n} |\nabla (\chi u_{\mathbf{x}_0;1})|^2 \, d\nu_{-\epsilon^{-2} r^2} \\
&\le (C\lambda^{\Upsilon} r^{2\Upsilon} + \frac{4}{3}) \widehat{H}^u_{\mathbf{x}_0}(\epsilon^{-2} r^2) \le (C\lambda^{\Upsilon} r^{2\Upsilon} + \frac{4}{3}) H^w(\epsilon^{-2} r^2) \le \frac{3}{2} H^w(\epsilon^{-2} r^2).
\end{aligned}$$

By the proof of Lemma 6.8, $w$ is $(n+1, C\epsilon^{\frac{1}{C(\Lambda,\Upsilon)}}, s)$-symmetric at $\mathbf{0}$. If $\lambda \le \overline{\lambda}(\Lambda, \Upsilon, \epsilon)$, then $u$ is weakly $(n+1, C\epsilon^{\frac{1}{C(\Lambda,\Upsilon)}}, s)$-symmetric at $\mathbf{x}_0$, so if we choose $\epsilon \le \overline{\epsilon}(\Lambda, \Upsilon)$, then Lemma 7.37 implies the claim. $\square$

The following is an epsilon regularity of the frequency based on symmetry similar to Lemma 5.9.

**Lemma 7.39.** *The following holds if $\lambda \le \overline{\lambda}(\Lambda, \Upsilon)$. If $u$ is weakly $(k, \epsilon, r)$-symmetric at $\mathbf{x}_0$, then*

$$\widehat{N}_{\mathbf{x}_0}(r^2) \le \begin{cases} 4\epsilon & \text{if } k = n+2, \\ 1 + C(\Lambda, \Upsilon)\epsilon & \text{if } k = n+1. \end{cases}$$

*Proof.* When $k = n+2$, the claim follows from the proof of Lemma 7.36.

We now suppose $k = n+1$. By replacing $u$ with $u_{\mathbf{x}_0;1}$, we can assume that $\mathbf{x}_0 = \mathbf{0}$ and $u = u_{\mathbf{x}_0;1}$. Let $w$ be a caloric approximation of $\chi u$. As in the proof of Lemma 7.37, we can prove that $w$ is $(n+1, C\epsilon, r)$-symmetric. Since $w$ is caloric, the proof of Lemma 5.9 implies that

$$\int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla^2 w|^2 \, d\nu_{-\tau} d\tau \le C(\Lambda, \Upsilon)\epsilon.$$

In particular, Lemma 7.22 implies that

$$\begin{aligned}
\int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla^2 (\chi u)|^2 \, d\nu_{-\tau} d\tau &\le \int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla^2 w|^2 \, d\nu_{-\tau} d\tau \\
&\quad + \int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla^2 (w - \chi u)|^2 \, d\nu_{-\tau} d\tau \\
&\le C(\Lambda, \Upsilon)\epsilon.
\end{aligned}$$

Therefore,

$$C(\Lambda, \Upsilon)\epsilon \ge \int_{r^2}^{e^2 r^2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} |\nabla^2 \chi u|^2 \, d\nu_{-\tau} d\tau$$

$$= \int_{r^2}^{e^2r^2} \frac{\tau}{\widehat{H}^u(\tau)} \int_{\mathbb{R}^n} \left(\Delta_f(\chi u) + \frac{1}{2\tau}\chi u\right)^2 d\nu_{-\tau} d\tau + \int_{r^2}^{e^2r^2} \frac{1}{2\tau}\left(\widehat{N}(\tau) - 1\right) d\tau$$
$$\geq \int_{r^2}^{e^2r^2} \frac{1}{2\tau}\left(\widehat{N}(\tau) - 1\right) d\tau \geq \widehat{N}(r^2) - 1.$$

This proves the second claim. □

7.5. **Quantitative dimension reduction.** We prove a result concerning quantitative dimension reduction similar to Proposition 6.9. In particular, $u$ is $k$-symmetric with respect to a plane $V$, then it is $(k+1)$-symmetric at points away from $V$, but at a smaller scale.

**Proposition 7.40.** *For any $\epsilon > 0$, there exists $\beta_0 = \beta_0(\epsilon, \Lambda, \Upsilon) > 0$ such that following holds if $\lambda \leq \overline{\lambda}(\Lambda, \Upsilon)$ and $\delta \leq \overline{\delta}(\epsilon, \Lambda, \Upsilon)$. Let $u$ be a solution of (7.1) satisfying (7.2), such that for some $r \in (0,1]$ and $\mathbf{x}_0 \in P(\mathbf{0}, 10)$, we have*

$$\inf_{\tau \in [10^5 n^2 r^2, 10^8 n^2 r^2]} \widehat{\mathcal{E}}_{\sqrt{\tau}}(\mathbf{x}_0) < \delta. \tag{7.94}$$

*Suppose $u$ is weakly $(k, \delta, 10^5 nr)$-symmetric at $\mathbf{x}_0$ with respect to some $V \in \mathrm{Gr}_{\mathcal{P}}(k)$. Then, for any $\mathbf{y} \in P(\mathbf{x}_0, 2r) \setminus P(\mathbf{x}_0 + V, r)$, there exists $\beta(\mathbf{y}) \geq \beta_0$ such that*

$$\widehat{\mathcal{E}}^{k+1,1}_{\beta(\mathbf{y})r}(\mathbf{y}) < \epsilon.$$

*Proof.* We may assume without loss of generality that $\epsilon \leq \overline{\epsilon}(\Lambda, \Upsilon)$. For $\beta > 0$ to be determined, let $w$ be the caloric approximation to $u$ at $\mathbf{y}$ on scales $[\beta^2 r^2, 10^{12} n^2 r^2]$ from Lemma 7.22, so that

$$\begin{aligned}\sup_{\tau \in [\beta^2 r^2, 10^{11} n^2 r^2]} \sum_{j=0}^{1} \int_{\mathbb{R}^n} \tau^j |\nabla^j (w - \chi u_{\mathbf{y};1})|^2 \, d\nu_{-\tau} + \int_{\beta^2 r^2}^{10^{11} n^2 r^2} \tau \int_{\mathbb{R}^n} |\partial_t (w - \chi u_{\mathbf{y};1})|^2 \, d\nu_{-\tau} d\tau \\ \leq C(\Lambda, \Upsilon) \lambda^{\Upsilon} r^{4\Upsilon} \widehat{H}_{\mathbf{y}}(r^2).\end{aligned} \tag{7.95}$$

Writing $\mathbf{y} = (y, s)$, we set $\mathbf{y}' \coloneqq (y', -s) \coloneqq (-a^{\frac{1}{2}}(\mathbf{y}) y, -s)$ and

$$\widehat{w}(z, -\tau) \coloneqq w(z + y', -\tau - s), \qquad \check{w}(z, -\tau) \coloneqq w(a^{\frac{1}{2}}(\mathbf{y}) a^{-\frac{1}{2}}(\mathbf{0}) z + y', -\tau - s),$$
$$\check{\chi}(z, -\tau) \coloneqq \chi(a^{\frac{1}{2}}(\mathbf{y}) a^{-\frac{1}{2}}(\mathbf{0}) z + y', -\tau - s),$$

so that $H^{\widehat{w}}(\tau) = H^w_{\mathbf{y}'}(\tau)$ for all $\tau \in [r^2, 10^{12} n^2 r^2]$. For any $\tau \in [r^2, 10^9 n^2 r^2]$, we then have

$$\mathrm{supp}((\chi - \check{\chi})(\cdot, -\tau)) \subseteq B(0, 10^8 \lambda^{-\frac{\Upsilon}{2}} r^{2\Upsilon}) \setminus \overline{B}(0, 10^{-8} \lambda^{-\frac{\Upsilon}{2}} r^{2\Upsilon}).$$

Combining this with the change of variables $x = a^{\frac{1}{2}}(\mathbf{y}) a^{-\frac{1}{2}}(\mathbf{0}) z + y'$, the ellipticity condition of (7.1), and Lemma 2.1 with $\alpha_0 \leftarrow 10\lambda$, $\alpha \leftarrow \frac{1}{50n}$, $\sigma \leftarrow 4r^2$, and $\tau \leftarrow \tau \in [10^4 n^2 r^2, 10^{11} n^2 r^2]$, we obtain

(7.96)
$$
\begin{aligned}
\int_{\mathbb{R}^n} |\check{w} - \chi u_{\mathbf{0};1}|^2 d\nu_{-\tau} &= \int_{\mathbb{R}^n} |\check{w}(z,-\tau) - \chi(z,-\tau) u_{\mathbf{y};1}(a^{\frac{1}{2}}(\mathbf{y})a^{-\frac{1}{2}}(\mathbf{0})z + y', -\tau - s)|^2 d\nu_{-\tau}(z) \\
&\le 2\sqrt{\frac{\det a(\mathbf{0})}{\det a(\mathbf{y})}} \int_{\mathbb{R}^n} |w - \chi u_{\mathbf{y};1}|^2 (x, -\tau - s) \frac{e^{-\frac{|a^{\frac{1}{2}}(\mathbf{0})a^{-\frac{1}{2}}(\mathbf{y})(x-y')|^2}{4\tau}}}{(4\pi\tau)^{\frac{n}{2}}} dx \\
&\quad + 2\int_{\mathbb{R}^n} |\chi(z,-\tau) - \check{\chi}(z,-\tau)| u_{\mathbf{0};1}^2(z,-\tau) d\nu_{-\tau}(z) \\
&\le C\int_{\mathbb{R}^n} |w - \chi u_{\mathbf{y};1}|^2 (x,-\tau-s) e^{10\lambda f_{\mathbf{y}'}} d\nu_{\mathbf{y}';-\tau-s}(x) \\
&\quad + 4\int_{B(0,10^8\lambda^{-\frac{\Upsilon}{2}} r^{2\Upsilon})\setminus \overline{B}(0,10^{-8}\lambda^{-\frac{\Upsilon}{2}} r^{2\Upsilon})} u_{\mathbf{0};1}^2(z,-\tau) d\nu_{-\tau}(z) \\
&\le C\int_{\mathbb{R}^n} |w - \chi u_{\mathbf{y};1}|^2 e^{\frac{1}{50n} f} d\nu_{-\tau-s} + C(\Lambda,\Upsilon) e^{-\frac{1}{10^8\lambda^{\Upsilon} r^{2(1-2\Upsilon)}}} \widehat{H}^u_{r^2}(\mathbf{y}).
\end{aligned}
$$

Next, we set $A(\ell) := (1-\ell)a^{\frac{1}{2}}(\mathbf{y})a^{-\frac{1}{2}}(\mathbf{0}) + \ell I$ and estimate

$$
|A'(\ell)| = |a^{\frac{1}{2}}(\mathbf{y})a^{-\frac{1}{2}}(\mathbf{0}) - I| \le 2|a^{-\frac{1}{2}}(\mathbf{y}) - a^{-\frac{1}{2}}(\mathbf{0})| \le C\lambda|\mathbf{y}| \le C\lambda r,
$$

using the ellipticity and Lipschitz assumptions of (7.1). Using this and Lemma 2.1, along with the change of variables $x = A(\ell)z + y'$, we obtain the following for all $\tau \in [10^4 n^2 r^2, 10^9 n^2 r^2]$

$$
\begin{aligned}
&\left| \frac{d}{d\ell} \int_{\mathbb{R}^n} w^2(A(\ell)z + y', -\tau - s) d\nu_{-\tau}(z) \right| \\
&= 2\left| \int_{\mathbb{R}^n} w(A(\ell)z + y', -\tau - s) \nabla w(A(\ell)z + y', -\tau - s) \cdot A'(\ell) z d\nu_{-\tau}(z) \right| \\
&\le \frac{2|A'(\ell)| \cdot |A(\ell)^{-1}|}{\det A(\ell)} \int_{\mathbb{R}^n} |w(x,-\tau-s)| \cdot |\nabla w(x,-\tau-s)| \cdot |x - y'| \frac{e^{-\frac{|A(\ell)^{-1}(x-y')|^2}{4\tau}}}{(4\pi\tau)^{\frac{n}{2}}} dx \\
&= C\lambda r\sqrt{\tau} \int_{\mathbb{R}^n} |w| \cdot |\nabla w| e^{\frac{1}{20n} f_{\mathbf{y}'}} \, d\nu_{\mathbf{y}';-s-\tau} \\
&\le C\lambda r \left( \int_{\mathbb{R}^n} |w|^2 e^{\frac{1}{4n} f_{\mathbf{y}'}} \, d\nu_{\mathbf{y}';-\tau-s} \right)^{\frac{1}{2}} \left( \int_{\mathbb{R}^n} \tau |\nabla w|^2 \, d\nu_{\mathbf{y}';-\tau-s} \right)^{\frac{1}{2}} \\
&\le C(\Lambda,\Upsilon)\lambda r \int_{\mathbb{R}^n} w^2 \, d\nu_{\mathbf{y}';-s-2\tau} \\
&\le C(\Lambda,\Upsilon)\lambda r \int_{\mathbb{R}^n} w^2 \, d\nu_{\mathbf{y}';-s-\tau},
\end{aligned}
$$

where we used Proposition 2.2, and the fact that $N^w_{\mathbf{y}'}(\tau) \le CN^w(10^2\tau) \le C(\Lambda,\Upsilon)$ by (7.95) and Lemma 2.10. We integrate from $\ell = 0$ to $\ell = 1$ to obtain

$$
C(\Lambda)\lambda r \int_{\mathbb{R}^n} w^2 d\nu_{\mathbf{y}';-s-\tau} \ge \left| \int_{\mathbb{R}^n} w^2 d\nu_{\mathbf{y}';-s-\tau} - \int_{\mathbb{R}^n} \check{w}^2 d\nu_{-\tau} \right|,
$$

which implies

$$
(1 - C(\Lambda)\lambda r) \int_{\mathbb{R}^n} w^2 d\nu_{\mathbf{y}';-s-\tau} \le \int_{\mathbb{R}^n} \check{w}^2 d\nu_{-\tau} \le (1 + C(\Lambda)\lambda r) \int_{\mathbb{R}^n} w^2 d\nu_{\mathbf{y}';-s-\tau}. \tag{7.97}
$$

Combining (7.96) and (7.97) yields

$$
(1 - C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\frac{\Upsilon}{2}})\widehat{H}^u(\tau) \le H^w_{\mathbf{y}'}(\tau) \le (1 + C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\frac{\Upsilon}{2}})\widehat{H}^u(\tau)
$$

for all $\tau \in [10^4 n^2 r^2, 10^9 n^2 r^2]$. We can therefore argue as in (7.72) to obtain

$$\widehat{N}^u(\tau) \leq \frac{1}{\log(2)} \log\left(\frac{\widehat{H}^u(2\tau)}{\widehat{H}^u(\tau)}\right) + C(\Lambda,\Upsilon)\lambda^{\Upsilon}\tau^{\Upsilon} \leq \frac{1}{\log(2)} \log\left(\frac{\widehat{H}^w(2\tau)}{\widehat{H}^w(\tau)}\right) + C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\frac{\Upsilon}{2}}$$
$$\leq N^w_{\mathbf{y}'}(2\tau) + C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\frac{\Upsilon}{2}}.$$

Similarly, we have $N^w_{\mathbf{y}'}(\tau) \leq \widehat{N}^u(2\tau) + C(\Lambda,\Upsilon)\lambda^{\frac{\Upsilon}{2}}\tau^{\frac{\Upsilon}{2}}$, so that (7.94) implies

$$\inf_{\tau\in[10^5 n^2 r^2, 10^8 n^2 r^2]} \mathcal{E}_{\sqrt{\tau}}(\mathbf{y}'; w) < C(\Lambda,\Upsilon)(\delta + \lambda^{\frac{\Upsilon}{2}}\tau^{\frac{\Upsilon}{2}}) \leq C(\Lambda,\Upsilon)\delta.$$

Moreover, because $u$ is weakly $(k, \delta, 10^5 nr)$-symmetric at $\mathbf{0}$, it follows from Lemma 7.25 that $u$ is weakly $(k, C(\Lambda,\Upsilon)\delta, 10^3 r)$-symmetric at $\mathbf{y}$ with respect to $V$. Combining this with (7.95), we conclude that $w$ is $(k, C(\Lambda,\Upsilon)\delta, 10^2 r)$-symmetric at $\mathbf{0}$. By Lemma 2.3, it follows that $w$ is $(k, C(\Lambda,\Upsilon)\delta, 10r)$-symmetric at $\mathbf{y}'$ with respect to $V$. Moreover, we have

$$d_{\mathcal{P}}(\mathbf{0}, \mathbf{y}' + V) = |\pi_V^{\perp}(\mathbf{y}')| \geq |\pi_V^{\perp}(\mathbf{y})| - |\pi_V^{\perp}(y - a^{\frac{1}{2}}(\mathbf{y})y, 0)| = (1 - \lambda r) d_{\mathcal{P}}(\mathbf{y}, V).$$

We can therefore apply Proposition 6.9 with $\mathbf{x}_0 \leftarrow \mathbf{y}'$, $\mathbf{y} \leftarrow \mathbf{0}$, $\eta \leftarrow \frac{1}{2}$, $u \leftarrow w$, $\epsilon \leftarrow \epsilon^3$, $\delta \leftarrow C(\Lambda,\Upsilon)\delta$ to obtain $\beta = \beta(\mathbf{y}) \geq \epsilon^{C(\Lambda,\Upsilon)}$ such that

$$\mathcal{E}^{k+1,1}_{\beta(\mathbf{y})r}(\mathbf{0}; w) < \epsilon^3, \qquad |N^w(10^5\beta^2) - N^w(10^{-5}\beta^2)| < \epsilon^6,$$

if we take $\delta \leq \overline{\delta}(\epsilon, \Lambda, \Upsilon)$. From Proposition 3.6(1), it follows that $w$ is $(k+1, C(\Lambda,\Upsilon)\epsilon^3, \beta(\mathbf{y})r)$-symmetric. From (7.95), it follows that $u$ is weakly $(k+1, C(\Lambda,\Upsilon)\epsilon^3, \beta(\mathbf{y})r)$-symmetric at $\mathbf{y}$, and (arguing as in (7.72))

$$|\widehat{N}^u_{\mathbf{y}}(10^4\beta(\mathbf{y})^2) - \widehat{N}^u_{\mathbf{y}}(10^{-4}\beta(\mathbf{y})^2)| < \epsilon^5.$$

Then Proposition 7.27(2) yields $\widehat{\mathcal{E}}^{k+1,1}_{\beta(\mathbf{y})r}(\mathbf{y}) < \epsilon$. □

7.6. **Proof of main theorems in the general setting.** Given the results we have developed thus far, the proof of Theorem 1.4 closely follows that of the caloric setting. Therefore, we only state the necessary modifications.

We first define neck regions corresponding to solutions to (7.1).

**Definition 7.41.** Given a $r \in (0, 1]$, a closed set $\mathcal{C} \subseteq P(\mathbf{x}_0, 2r)$ with $\mathbf{x}_0 \in \mathcal{C}$, and a function $\mathbf{r}_\bullet : \mathcal{C} \to [0, \gamma r]$, we say that

$$\mathcal{N} := P(\mathbf{x}_0, 2r) \setminus \overline{P}(\mathcal{C}, \mathbf{r}_\bullet)$$

is an $(m, k, \delta, \eta)$-*neck region (of scale $r$)* if it satisfies (n1), (n4) of Definition 4.1, as well as the following:

- (n2) $\sup_{s\in[\mathbf{r}_\mathbf{x}, \gamma^{-3}r]}(|\widehat{N}_\mathbf{x}(s^2) - m| + \gamma^{\frac{\Upsilon}{4}}s^{\Upsilon}) < \delta$ for all $\mathbf{x} \in \mathcal{C}$;
- (n3) For all $s \in [\mathbf{r}_\mathbf{x}, \gamma^{-3}r]$, $u$ is weakly $(k, \delta, s)$-symmetric at $\mathbf{x}$ with respect to $V$ but not weakly $(k+1, \eta, s)$-symmetric.

We say $\mathcal{N}$ is a *strong* $(m, k, \delta, \eta)$*-region* if, in addition, it satisfies (n4.b') and (n5) of Definition 4.1.

Similarly, we define the packing measure as well as

$$\widehat{\mathcal{V}}^u_{\delta,m,r}(\mathbf{x}) := \left\{\mathbf{y} \in P(\mathbf{x}, 4r) \colon |\widehat{N}^u_{\mathbf{y}}(s^2) - m| + \lambda^{\frac{\Upsilon}{4}}r^{\Upsilon} \leq \delta \text{ for any } s \in [\delta r, \delta^{-1}r]\right\}.$$

By making appropriate modifications to the definition of different types of balls, we can prove the following neck structure theorems.

**Theorem 7.42** (Neck structure theorems)**.** *Suppose $u$ is a solution to* (7.1) *satisfying* (7.2)*. Then the conclusions of Theorem 4.6 and Theorem 4.14 hold.*

*Proof.* The proof is the same, except that we replace Lemma 4.20 with the following lemma. ◻

Recall the definition of $\widehat{\beta}$ from 4.27. We the Carleson condition for $\widehat{\beta}$ as in Lemma 4.20 taking into the error term in $\widehat{\mathcal{E}}$.

**Lemma 7.43.** *Suppose $P(\mathbf{0},2)\setminus\bigcup_{\mathbf{y}\in\mathcal{C}}\overline{P}(\mathbf{y},\mathbf{r}_{\mathbf{y}})$ is a strong $(m,k,\delta,\eta)$-neck region modeled on a vertical $k$-plane $V\in\mathrm{Gr}_{\mathcal{P}}(k)$. Let $\mu$ be the packing measure, let $\pi:\mathcal{C}\to V$ be the projection map, and set $\widehat{\mu}:=\pi_*\mu$. Then, for any $\mathbf{x}\in P(\mathbf{0},\frac{3}{2})$ and $r>0$ such that $P(\mathbf{x},20r)\subset P(\mathbf{0},\frac{7}{4})$,*

$$\int_{P^V(\mathbf{x},10r)}\int_{\frac{1}{8}\widehat{\mathbf{r}}_{\mathbf{y}}}^{r}\widehat{\beta}^2_{\mathcal{P},k}(\mathbf{y},10^5 s)\,\frac{ds}{s}d\widehat{\mu}(\mathbf{y})\le C\delta r^k.$$

*Proof.* Set $r_i:=2^{-i}$ for $i\in\mathbb{N}$. We argue as in Lemma 4.20 to obtain

$$\begin{aligned}\int_{P(\mathbf{x},10r)}\int_{\frac{1}{8}\widehat{\mathbf{r}}_{\mathbf{y}}}^{r}\widehat{\beta}^2_{\mathcal{P},k}(\mathbf{y},10^5 s)\,\frac{ds}{s}d\widehat{\mu}(\mathbf{y}) &\le C\sum_{i\in\mathbb{N}}\int_{P(\mathbf{x},20r)}\mathbb{1}_{\{\frac{1}{16}\widehat{\mathbf{r}}_{\mathbf{z}}\le r_i\le 10^5 r\}}\widehat{\mathcal{E}}_{20nr_i}(\mathbf{z})d\widehat{\mu}(\mathbf{z})\\
&\le C\sum_{i\in\mathbb{N}}\int_{P(\mathbf{x},20r)}\left(\widehat{N}_{\mathbf{z}}(2\cdot 10^4 r_i^2)-\widehat{N}_{\mathbf{z}}(8r_i^2)+\lambda^{\frac{\Upsilon}{4}}r_i^{\Upsilon}\right)d\widehat{\mu}(\mathbf{z})\\
&\le C\int_{P(\mathbf{x},20r)}\left(\widehat{N}_{\mathbf{z}}(\gamma^{-2}r_i^2)-\widehat{N}_{\mathbf{z}}(\widehat{\mathbf{r}}_{\mathbf{z}}^2)+\lambda^{\frac{\Upsilon}{4}}r^{\Upsilon}\right)d\widehat{\mu}(\mathbf{z}),\end{aligned}$$

where in the first inequality, we use Lemma 7.32. ◻

We also have neck decomposition theorems in the general case.

**Theorem 7.44** (Neck Decomposition)**.** *For any $\epsilon,\eta\in(0,1]$, the following holds if $r\in(0,1]$ and $\lambda\le\overline{\lambda}(\epsilon,\eta,\Lambda,\Upsilon)$. Suppose $u$ is a solution to (7.1) satisfying (7.2). there exists a decomposition*

$$P(\mathbf{0},r)\subseteq\bigcup_a\mathcal{N}^a\cup\bigcup_b P(\mathbf{x}_b,r_b)\cup\left(\widetilde{\mathcal{C}}\cup\bigcup_a\mathcal{C}_{a,0}\right),$$

*where the following hold:*

- *(a) $\mathcal{N}^a=P(\mathbf{x}_a,2r_a)\setminus\overline{P}(\mathcal{C}_a,\mathbf{r}_a)$ is an $(m_a,k,\epsilon,C^{-1}(\Lambda,\Upsilon)\eta)$-neck region for some positive integer $m_a\le C(\Lambda,\Upsilon)$, where $\mathcal{C}_a$ is the center set associated to $\mathcal{N}^a$;*
- *(b) $\widehat{\mathcal{E}}^{k+1,1}_{100r_b}(\mathbf{y}_b)\le\eta$ for some $\mathbf{y}_b\in P(\mathbf{x}_b,4r_b)$;*
- *(c) A $k$-dimensional parabolic Minkowski content estimate holds:*

$$\sum_a r_a^k+\sum_b r_b^k\le C(\epsilon,\eta,\Lambda,\Upsilon)r^k;$$

- *(d) Further, $\mathcal{H}^k_{\mathcal{P}}(\widetilde{\mathcal{C}})=0$.*

*Similarly modified statements of Theorems 5.2 and Theorem 5.3 hold.*

*Proof.* The proof of Theorem 5.1 relied only on covering arguments, Lemma 3.12, and Theorem 3.6. Hence, the proof can be adapted by replacing these with Lemma 7.30 and Theorem 7.27, noting that the hypotheses hold under our assumption $r\le\overline{r}(\epsilon,\eta,\Lambda,\Upsilon)$.

For the last claim, we use Lemma 7.39 instead of Lemma 5.9. ◻

Similarly, we have volume estimates for the quantitative strata.

**Theorem 7.45.** *For any $\epsilon\in(0,1]$, the following holds if $\lambda\le\overline{\lambda}(\epsilon,\Lambda,\Upsilon)$ and $r\in(0,1]$. Suppose $u$ is a solution to (7.1) satisfying (7.2). Then, for any $k\in\{1,\dots,n+1\}$,*

$$\mathcal{H}^{n+2}_{\mathcal{P}}(P(\widehat{\mathcal{S}}^k_{\epsilon,r},r)\cap P(\mathbf{0},1))\le C(\epsilon,\Lambda,\Upsilon)r^{n+2-k}.\tag{7.98}$$

*Moreover, $\widehat{\mathcal{S}}^k_\epsilon$ is parabolic k-rectifiable with $\mathcal{H}^k_{\mathcal{P}}(\widehat{\mathcal{S}}^k_\epsilon \cap P(\mathbf{0},1)) \le C(\epsilon,\Lambda,\Upsilon)$. Furthermore, if $k=n+1$ or $n$, for $\mathcal{H}^1$-almost every time $t\in(-1,1)$, $\widehat{\mathcal{S}}^{k,t}_\epsilon := \widehat{\mathcal{S}}^k_\epsilon \cap (\mathbb{R}^n\times\{t\})$ is $(k-2)$-rectifiable with*

$$\mathcal{H}^{k-2}(\widehat{\mathcal{S}}^{k,t}_\epsilon) \le C(\epsilon,\Lambda,\Upsilon), \tag{7.99}$$

*and for $\mathcal{H}^{\frac12}$-almost every time $t\in(-1,1)$, $\mathcal{H}^{k-1}(\widehat{\mathcal{S}}^{k,t}_\epsilon)=0$.*

*Proof.* The proof is verbatim as in the proof of Theorem 6.4 after we replace Theorem 3.6, Lemma 3.7, Proposition 6.9, Theorem 4.6 Theorem 5.1 with Theorem 7.27, Proposition 7.21, Proposition 7.40, Theorem 7.44 and Theorem 7.42, respectively. □

We now show that for any function satisfying the hypotheses of our main theorems, parabolic rescaling at sufficiently small scales will satisfy the recurring hypotheses of Section 7.

**Lemma 7.46.** *Under the hypotheses of Theorem 1.1 and Theorem 1.4, there exists $\Lambda=\Lambda(M,\Theta)$ such that for any $\lambda>0$, the following holds if $r\le\overline{r}(\lambda,M)$. For any $\mathbf{x}\in P(\mathbf{0},\frac14)$, $u_{\mathbf{x};r}$ satisfies (7.1) and (7.2).*

*Proof.* We can assume that $M\ge1$. By Remark 7.2 with $\lambda\leftarrow M$ and $r\leftarrow\frac{\lambda}{200^2M}$, imply that $u_{\mathbf{x};r}$ satisfies (7.1) if $r\le\frac{\lambda}{200^2M}$. Fix $\mathbf{y}=(y,s)\in P(\mathbf{0},\frac12)$. It remains to verify (7.2). From

$$\int_{B(y,4)\times[s-16,s]} u^2 d\mathcal{H}^{n+2}_{\mathcal{P}} \le \int_{B(0,5)\times[-25,s]} u^2 d\mathcal{H}^{n+2}_{\mathcal{P}} \le \Theta\int_{B(0,\frac12)} u^2(x,s)dx \le \Theta\int_{B(\mathbf{y},1)} u^2(x,s)dx.$$

By [EFV06, Theorem 3], for any $\mathbf{y}_0\in P(\mathbf{0},\frac12)$ and any $r'\le\overline{r}'(M,\Theta)$, we have

$$\int_{B(y_0,4r')\times[s_0-(4r')^2,s_0]} u^2 d\mathcal{H}^{n+2}_{\mathcal{P}} \le C(M,\Theta)(r')^2\int_{B(y_0,r')} u^2(y,s_0)dy. \tag{7.100}$$

In terms of the rescaling $u_{\mathbf{x};r}$, this is

$$\begin{aligned}&\int_{r^{-1}a^{\frac12}(\mathbf{x})(B(y_0,4r')-x)\times r^{-2}[s_0-(4r')^2-t,s_0-t]} u^2_{\mathbf{x};r} d\mathcal{H}^{n+2}_{\mathcal{P}}\\ &\le C(M,\Theta)\left(\frac{r'}{r}\right)^2\int_{r^{-1}a^{\frac12}(\mathbf{x})(B(y_0,r')-x)} u^2_{\mathbf{x};r}(y,r^{-2}(s_0-t))dy.\end{aligned} \tag{7.101}$$

Fix $\mathbf{y}=(y,s)\in P(\mathbf{0},\lambda^{-1})$ and $\rho\in(0,M\lambda^{-1}]$. If $r\le\overline{r}(\lambda,M)$, then we can apply (7.101) with $y_0\leftarrow x+ra^{\frac12}(\mathbf{x})y$, $s_0\leftarrow t+r^2s$, and $r'\leftarrow\rho r$ to obtain

$$\int_{(y+a^{\frac12}(\mathbf{x})B(0,4\rho))\times[s-(4\rho)^2,s]} u^2_{\mathbf{x};r} d\mathcal{H}^{n+2}_{\mathcal{P}} \le C(M,\Theta)\rho^2\int_{y+a^{\frac12}(\mathbf{x})B(0,\rho)} u^2_{\mathbf{x};r}(z,s)dz.$$

If $r\le\overline{r}(\lambda,\Lambda)$, then we can iterate [EFV06, Theorem 2(1)] and use $(2M)^{-1}I\le a(\mathbf{x})\le 2MI$ to obtain

$$\begin{aligned}\int_{(B(y,\frac{4\rho}{M}))\times[s-\left(\frac{4\rho}{M}\right)^2,s]} u^2_{\mathbf{x};r} d\mathcal{H}^{n+2}_{\mathcal{P}} &\le \int_{(y+a^{\frac12}(\mathbf{x})B(0,2M\rho))\times[s-(4\rho)^2,s]} u^2_{\mathbf{x};r} d\mathcal{H}^{n+2}_{\mathcal{P}}\\ &\le C(M,\Theta)\rho^2\int_{y+a^{\frac12}(\mathbf{x})B(0,(2M)^{-1}\rho)} u^2_{\mathbf{x};r}(z,s)dz\\ &\le C(M,\Theta)\rho^2\int_{B(y,\frac{\rho}{M})} u^2_{\mathbf{x};r}(z,s)dz.\end{aligned}$$

The claim then follows by replacing $\rho\leftarrow M\rho$. □

*Proof of Theorem 1.1, Theorem 1.4, Corollary 1.5.* Choose $\epsilon \leq \overline{\epsilon}(\Lambda)$ such that Proposition 7.35 holds. We now choose $\lambda \leq \overline{\lambda}(\epsilon, \Lambda, \Upsilon)$ as in Theorem 7.45. Due to Lemma 7.46, we can apply Theorem 7.45 to $u_{\mathbf{x},r_0}$, where $r_0 \leq \overline{r}_0(\lambda, M)$ for $\mathbf{x} \in P(\mathbf{0}, \frac{1}{4})$. Then by Proposition 7.35, we have

$$\widehat{\mathscr{C}}^k(u) = \widehat{\mathscr{C}}^k(u_{\mathbf{x};r_0}) \subset \widehat{\mathscr{C}}^k_r(u_{\mathbf{x};r_0}) \subset \widehat{\mathcal{S}}^k_{\epsilon,r}(u_{\mathbf{x};r_0}).$$

Therefore, proof of the Theorem 1.4 follows from a covering argument. The remaining statements are similarly deduced as in the Proof of Corollary 1.10 and Theorem 6.13, by making the following replacements:

(1) Theorem 5.2 and Theorem 5.3 with Theorem 7.44,
(2) Proposition 6.5 with Proposition 7.35,
(3) Proposition 6.8 with Proposition 7.38,
(4) Theorem 4.14 with Theorem 7.42.

□

*Proof of Corollary 1.3.* Fix $\mathbf{x}_0 = (x_0, t_0) \in \Omega$ and $r > 0$ such that $P(\mathbf{x}_0, 10r) \subseteq \Omega$. Set

$$\Theta := \sup_{t \in [t_0 - 2r, t_0 + 2r]} \frac{\int_{B(x_0, 5r) \times [-25, t]} u^2 \, d\mathcal{H}^{n+2}_{\mathcal{P}}}{\int_{B(x_0, \frac{1}{2}r)} u^2(x,t)\, dx}.$$

Because $u$ is continuous, it follows that $t \mapsto \int_{B(x_0,5r)} u^2(x,t)dx$ is continuous, so that either $\Theta < \infty$ or else $u(\cdot, t)|_{B(x_0,5r)} \equiv 0$ for some $t \in [t_0 - 2r, t_0 + 2r]$. In the latter case, it follows from either of [AV04, Theorem 1], [Fer03, Theorem 3] that $u(\cdot, t)$ vanishes on the connected component of $\Omega \cap (\mathbb{R}^n \times \{t\})$ containing $x_0$, yielding a contradiction. We may therefore apply Theorem 1.4 to any locally finite cover $\{P(x_i, r_i)\}_{i \in \mathbb{N}}$ of $\Omega$ with $P(x_i, 10r_i) \subseteq \Omega$ in order to obtain the claim. □

## Appendix A. Integration by Parts and Asymptotic Expansions

In this appendix, we provide justification for integration by parts used in Section 2 as well as an $L^2$-expansion used in Claim 7.18.1.

Suppose $u \in C^\infty(\mathbb{R}^n \times (t_0, t_1))$ is a caloric function satisfying

$$C_\epsilon := \sup_{t \in [t_0 + \epsilon, t_1 - \epsilon]} \int_{\mathbb{R}^n} u^2 d\nu_{0,t_1;t} < \infty$$

for all $\epsilon > 0$. For any $t \in [t_0 + \epsilon, t_1 - \epsilon]$ and $\alpha > 0$, we can then estimate

$$C_\epsilon \geq \int_{\mathbb{R}^n} u^2 d\nu_{0,t_1;t} \geq \frac{1}{(4\pi(t_1 - t))^{\frac{n}{2}}} e^{-\frac{(|x|-1)_+^2}{4(t_1-t)}} \int_{B(x,1)} u^2(z,t)dz,$$

so that interior parabolic estimates yield

$$|\partial_t^k \nabla^\ell u(x,t)| \leq C(\alpha, \epsilon, k, \ell) e^{(1+\alpha)\frac{|x|^2}{4(t_1-t)}}. \tag{A.1}$$

An elementary consequence of (A.1) is that for any $k, \ell \in \mathbb{N}$, $\mathbf{y} = (y, s) \in \mathbb{R}^n \times [t_0 + \epsilon, t_1 - \epsilon]$, and $t \in [t_0 + \epsilon, s)$, there exists $\alpha = \alpha(\epsilon) > 0$ such that

$$\int_{\mathbb{R}^n} (1 + |\partial_t^k \nabla^\ell u|)^2 e^{\alpha(\epsilon) f_{\mathbf{y}}} d\nu_{\mathbf{y};t} < C(\alpha, \epsilon, k, \ell).$$

In particular, this justifies differentiation under the integral sign when computing time derivatives of quantities of the form

$$\frac{d}{dt} \int_{\mathbb{R}^n} \Phi \, d\nu_{\mathbf{x}_0;t},$$

where $\Phi$ is any quantity we consider in Section 2 to Section 6. Moreover, we can freely integrate by parts any such quantities and can differentiate the expansion (2.7) termwise by virtue of (A.1) and the following lemma.

**Lemma A.1.** *Suppose a vector field $X \in C^\infty(\mathbb{R}^n, \mathbb{R}^n)$ such that for any $\delta > 0$, and $t_0 < t < s < t_1$*

$$|X|(y,t) + |\nabla X|(y,t) \le e^{\frac{|y|^2}{(4+\delta)(s-t)}}.$$

*Then for all $y \in \mathbb{R}^n$, $\mathbf{y} := (y,s)$ satisfies*

$$\int_{\mathbb{R}^n} \operatorname{div}_{f_{\mathbf{y}}}(X)\, d\nu_{\mathbf{y};t} = 0.$$

*Proof.* We first observe that $\operatorname{div}_{f_{\mathbf{y}}}(X) \in L^1(\mathbb{R}^n, d\nu_{\mathbf{y};t})$ by our assumptions on $X$ and the fact that $|\nabla f_{\mathbf{y}}|$ has a linear growth. Because

$$\begin{aligned}\left|\int_{B(y,r)} \operatorname{div}_{f_{\mathbf{y}}}(X) d\nu_{\mathbf{y};t}\right| &= \frac{1}{(4\pi(s-t))^{\frac{n}{2}}}\left|\int_{\partial B(y,r)}\left\langle X(x), \frac{x-y}{|x-y|}\right\rangle e^{-f_{\mathbf{y}}(x,t)} d\mathcal{H}^{n-1}(x)\right| \\ &\le \frac{1}{(4\pi(s-t))^{\frac{n}{2}}} r^{n-1} e^{-\frac{r^2}{4(s-t)}} e^{\frac{(r+|y|)^2}{4(1+\delta)(s-t)}},\end{aligned}$$

the claim follows by taking $r \to \infty$. $\square$

The lemma below proves that under certain regularity assumption solutions to drift heat equation admit a series expansion in terms of homogeneous caloric polynomials. The following result is well-known, but we provide a proof for completeness.

**Lemma A.2.** *Suppose $u \in C^\infty(\mathbb{R}^n \times (0, 4A])$ satisfies*

$$(\partial_s - \Delta_f)u = \frac{m}{2}u$$

*such that for any $\epsilon > 0$ there exists $C_\epsilon < \infty$ satisfying*

$$\operatorname{esssup}_{s\in[\epsilon,4A]} \int_{\mathbb{R}^n} u^2(\cdot, s)\, d\nu \le C_\epsilon \text{ and } \int_0^{4A}\int_{\mathbb{R}^n} u^2\, d\nu ds \le C.$$

*Then there exist $\widehat{p}_k \in \widehat{\mathcal{P}}_k$ such that*

$$\lim_{N\to\infty}\int_0^{4A}\int_{\mathbb{R}^n}\left|u - \sum_{k=0}^N e^{-\frac{(k-m)}{2}s}\widehat{p}_k\right|^2 d\nu ds = 0.$$

*Proof.* Define a test function $\varphi(x,s) := \psi(s)\zeta(x)$ where $\psi(s) \in C_c^\infty((0,4A))$ and $\zeta \in C_c^\infty(\mathbb{R}^n)$. Then by the weak formulation, we know that

$$\begin{aligned} 0 &= \int_0^{4A}\int_{\mathbb{R}^n} (-u\partial_s\varphi - u\Delta_f\varphi - \frac{m}{2}u\varphi)\, d\nu ds \\ &= -\int_0^{4A} \partial_s\psi\left(\int_{\mathbb{R}^n} u\zeta d\nu\right) ds - \int_0^{4A}\psi\left(\int_{\mathbb{R}^n} u\left(\Delta_f\zeta + \frac{m}{2}\zeta\right) d\nu\right) ds.\end{aligned}$$

Fix any $\widehat{q}_k \in \widehat{\mathcal{P}}_k$. Choose a sequence $\zeta_j \in C_c^\infty(\mathbb{R}^n)$ such that

$$\lim_{j\to\infty}\int_{\mathbb{R}^n}\left(|\zeta_j - \widehat{q}_k|^2 + |\nabla(\zeta_j - \widehat{q}_k)|^2 + |\Delta_f(\zeta_j - \widehat{q}_k)|^2\right) d\nu = 0.$$

Because $\operatorname{esssup}_{s\in \operatorname{supp}(\psi)}\|u(\cdot,s)\|_{L^2(\mathbb{R}^n,\nu)} < \infty$, the dominated convergence theorem then implies

$$0 = -\int_0^{4A}\widehat{a}_k(s)\partial_s\psi(s) ds - \int_0^{4A}\psi(s)\left(\frac{m-k}{2}\widehat{a}_k(s)\right) ds,$$

where $\widehat{a}_k(s) := \int_{\mathbb{R}^n} u\widehat{q}_k \, d\nu$. Thus the following holds in a weak sense

$$\partial_s \widehat{a}_k(s) = \frac{m-k}{2}\widehat{a}_k(s). \tag{A.2}$$

Since $\widehat{a}_k(s) \in L^2((0,4A))$, we get $\partial_s \widehat{a}_k(s) \in L^2((0,4A))$ which implies that $\widehat{a}_k(s) \in H^1((0,4A)) \subset C((0,4A))$. Therefore, (A.2) holds in a classical sense. In particular, $\widehat{a}_k(s) = e^{\frac{m-k}{2}s}\widehat{a}_k(0)$. Thus letting $b_k(\cdot,s) \in \mathcal{P}_k$ denote the $L^2(\nu)$-projection of $u(\cdot,s)$ onto $\widehat{\mathcal{P}}_k$, we have $b_k(\cdot,s) = e^{\frac{m-k}{2}s}b_k(\cdot,0)$. Moreover, $u(\cdot,s) \in L^2(\mathbb{R}^n,\nu)$ implies

$$\lim_{N\to\infty} \int_{\mathbb{R}^n} \left| u(\cdot,s) - \sum_{j=0}^{N} b_j(\cdot,s) \right|^2 d\nu = 0,$$

as well as

$$\int_{\mathbb{R}^n} \left| u(\cdot,s) - \sum_{j=0}^{N} b_j(\cdot,s) \right|^2 d\nu \le \int_{\mathbb{R}^n} u^2(\cdot,s) \, d\nu < \infty.$$

The dominated convergence theorem and our assumption $\int_0^{4A}\int_{\mathbb{R}^n} u^2 \, d\nu ds < \infty$ therefore imply

$$\lim_{N\to\infty} \int_0^{4A} \int_{\mathbb{R}^n} \left| u - \sum_{k=0}^{N} e^{-\frac{(k-m)}{2}s}\widehat{p}_k \right|^2 d\nu ds = 0,$$

where $\widehat{p}_k := b_k(\cdot,0)$. □

## Appendix B. Criterion for Regularity of Parabolic Lipschitz Functions

For the sake of completeness, we provide a well-known sufficient criterion for the graph of a parabolic Lipschitz function to be regular in the sense of Definition 4.13, see [HLN04].

Fix a vertical plane $V \in \mathrm{Gr}_{\mathcal{P}}(k)$. Suppose $f\colon V \to \mathbb{R}$ is a function supported in $P(\mathbf{0},5)$. For $\mathbf{x} \in \mathbb{R}^{n-1} \times \mathbb{R}$ and $r > 0$, recall that

$$\kappa^2(\mathbf{x},r) = \kappa_k^2(\mathbf{x},r;f) = \inf_{\ell} \frac{1}{r^k} \int_{V\cap P(\mathbf{x},r)} \left( \frac{|f(\mathbf{y}) - \ell(\mathbf{y})|}{r} \right)^2 d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}),$$

where the infimum is taken over all affine functions $\ell\colon V \to \mathbb{R}$ satisfying $\partial_t \ell \equiv 0$.

The following proposition shows that a Lipschitz function whose $\kappa^2$ satisfies a Carleson estimate has half-time derivative in BMO, as defined in (4.25). The proof is modeled closely on [HLN03, proof of Theorem 1].

**Proposition B.1.** *Suppose $V \in \mathrm{Gr}_{\mathcal{P}}(k)$ is vertical, $f\colon V \to \mathbb{R}$ is supported in $V \cap P(\mathbf{0},5)$, that $f$ has parabolic Lipschitz constant at most $\delta$ and*

$$\int_0^r \int_{V\cap P(\mathbf{x},r)} \kappa^2(\mathbf{y},s) \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \frac{ds}{s} < \delta r^k$$

*for all $\mathbf{x} \in V$ and all $r \in (0,100]$. Then $\|\partial_t^{\frac{1}{2}} f\|_{\mathrm{BMO}_{\mathcal{P}}(V)} \le C\sqrt{\delta}$.*

*Proof.* By a rotation, we may assume that $V = (\{\mathbf{0}^2\} \times \mathbb{R}^{k-2}) \times \mathbb{R} \cong \mathbb{R}^{k-2} \times \mathbb{R}$. Using this identification, we lose no generality in assuming $n = k-2$. Given $g\colon \mathbb{R}^{k-2} \times \mathbb{R} \to \mathbb{R}$ and $\Omega \subset \mathbb{R}^{k-2} \times \mathbb{R}$, we write

$$g_\Omega := \frac{1}{\mathcal{H}^k_{\mathcal{P}}(\Omega)} \int_\Omega g \, d\mathcal{H}^k_{\mathcal{P}}.$$

Because the assumptions and conclusion of the theorem are invariant under parabolic rescaling, it suffices to show that if $Q \coloneqq P(\mathbf{0},1)$, then

$$\int_Q |\partial_t^{\frac{1}{2}} f - (\partial_t^{\frac{1}{2}} f)_Q| \, d\mathcal{H}^k_{\mathcal{P}} \leq C\sqrt{\delta}. \tag{B.1}$$

Fix $\varphi \in C^\infty(P(\mathbf{0},2))$ such that $\varphi|_{P(\mathbf{0},\frac{3}{2})} \equiv 1$, $0 \leq \varphi \leq 1$, and $|\nabla^\ell \varphi| + |\partial_t^\ell \varphi| \leq c(\ell)$. Write

$$f = f_1 + f_2 =: \varphi(f - f_Q) + ((1-\varphi)f + \varphi f_Q)\,.$$

We first prove (B.1) after mollifying $f_1$. We handle the general case by a limiting argument. To this end, fix a rotationally invariant function $\chi \in C_c^\infty(P(\mathbf{0},1))$ satisfying $\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \chi \, dxdt = 1$. For any $\epsilon \in (0,1)$, set $\chi_\epsilon(x,t) \coloneqq \epsilon^{-k}\chi(\epsilon^{-1}x, \epsilon^{-2}t)$. Letting $H$ denote the Hilbert transform, for any Schwartz functions $\psi_1, \psi_2 \in \mathcal{S}(\mathbb{R})$, we have the following integration-by-parts formula:

$$\begin{aligned}\int_{\mathbb{R}} (\partial_t^{\frac{1}{2}}\psi_1)(t)(\partial_t^{\frac{1}{2}}\psi_2)(t)dt &= \int_{\mathbb{R}} (|\tau|^{\frac{1}{2}}\widehat{\psi}_1(\tau))(|\tau|^{\frac{1}{2}}\widehat{\psi}_2(\tau))d\tau = \int_{\mathbb{R}} (\mathrm{sgn}(\tau)\widehat{\psi}_1(\tau))(\tau\widehat{\psi}_2(\tau))d\tau \\ &= \int_{\mathbb{R}} H(\psi_1)(t)\partial_t\psi_2(t)dt,\end{aligned}$$

where $\widehat{\psi}$ denotes the Fourier transform of $\psi$. By taking $\psi_1 \coloneqq \partial_\lambda(\chi_\lambda * f_1)(x,\cdot)$ and $\psi_2 \coloneqq (\chi_\lambda * f_1)(\cdot)$, and then integrating in $x$ and $\lambda$, we obtain

$$\begin{aligned}\int_{\mathbb{R}^{k-2}\times\mathbb{R}} (\partial_t^{\frac{1}{2}}(\chi_\epsilon * f_1))^2 \, d\mathcal{H}^k_{\mathcal{P}} &= \int_{\mathbb{R}^{k-2}\times\mathbb{R}} (\partial_t^{\frac{1}{2}}(\chi_1 * f_1))^2 \, d\mathcal{H}^k_{\mathcal{P}} \\ &\quad - 2\int_\epsilon^1 \int_{\mathbb{R}^{k-2}\times\mathbb{R}} H\left(\partial_\lambda(\chi_\lambda * f_1)\right)\partial_t(\chi_\lambda * f)\, d\mathcal{H}^k_{\mathcal{P}} d\lambda \\ &= I_1 + I_2.\end{aligned}$$

For any $\mathbf{x} = (x,t) \in P(\mathbf{0},2)$,

$$|f_1(\mathbf{x})| \leq |f(\mathbf{x}) - f_Q| \leq C\left|\fint_Q (f(\mathbf{x}) - f(\mathbf{y}))\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\right| \leq C\delta, \tag{B.2}$$

so that also $|\chi_1 * f_1| \leq C\delta$. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{k-2}\times\mathbb{R}$, we have

$$|f_1(\mathbf{x}) - f_1(\mathbf{y})| \leq |f(\mathbf{x}) - f(\mathbf{y})|\varphi(\mathbf{x}) + |f(\mathbf{y}) - f_Q| \cdot |\varphi(\mathbf{x}) - \varphi(\mathbf{y})| \leq C\delta|\mathbf{x}-\mathbf{y}|,$$

**Claim B.1.1.** *For any $x \in \mathbb{R}^{k-2}$ and $s,t \in \mathbb{R}$, we have*

$$|(\chi_1 * f_1)(x,t_1) - (\chi_1 * f_1)(x,t_2)| \leq C\delta|t_1 - t_2|.$$

*Proof.* If $|t_1 - t_2| \geq \frac{1}{4}$, then

$$|(\chi_1 * f_1)(x,t) - (\chi_1 * f_1)(x,s)| \leq C\delta\sqrt{|t_1 - t_2|} \leq C\delta|t_1 - t_2|.$$

Suppose instead $|t_1 - t_2| \leq \frac{1}{4}$. For any $t \in [t_1, t_2]$,

$$\begin{aligned}\left|\frac{d}{dt}(f_1 * \chi_1)(\mathbf{x})\right| &= \left|\int_{\mathbb{R}} \partial_t\chi(\mathbf{x}-\mathbf{y})(f_1(y,s) - f_1(y,t_1))\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\right| \\ &\leq C\delta,\end{aligned}$$

where we used that $\mathrm{supp}(\chi(x-y,t-\cdot)) \subseteq [t_1 - 2, t_1 + 2]$. Integration from $t_1$ to $t_2$ yields the claim. ∎

Combining (B.2) and Claim B.1.1 yields

(B.3)

$$|\partial_t^{\frac{1}{2}}(\chi_1 * f_1)|(x,t) = c\left|\int_{\mathbb{R}} \frac{(\chi_1 * f_1)(x,t) - (\chi_1 * f_1)(x,s)}{|t-s|^{\frac{3}{2}}}\,ds\right| \leq C\delta \int_{\mathbb{R}} \frac{\min\{|t-s|,1\}}{|t-s|^{\frac{3}{2}}}\,ds \leq C\delta.$$

If in addition we have $|t| \geq 100$, then because $\operatorname{supp}(\chi_1 * f_1) \subseteq P(\mathbf{0}, 4)$, we obtain

$$|\partial_t^{\frac{1}{2}}(\chi_1 * f_1)|(x,t) \leq \frac{C\delta}{|t|^{\frac{3}{2}}}. \tag{B.4}$$

Combining estimates (B.3) and (B.4) gives

$$|\partial_t^{\frac{1}{2}}(\chi_1 * f_1)|(x,t) \leq \frac{C\delta}{(1+|t|)^{\frac{3}{2}}}$$

for all $(x,t) \in \mathbb{R}^{k-2} \times \mathbb{R}$, so that

$$I_1 = \int_Q |\partial_t^{\frac{1}{2}}(\chi_1 * f_1)|^2\,d\mathcal{H}^k_{\mathcal{P}} \leq C\delta.$$

Next, we use the $L^2$-boundedness of the Hilbert transform to get

$$\int_\epsilon^1 \lambda^{-1} \int_{\mathbb{R}^{k-2}\times\mathbb{R}} |H\left(\partial_\lambda(\chi_\lambda * f_1)\right)|^2\,d\mathcal{H}^k_{\mathcal{P}} d\lambda \leq C \int_\epsilon^1 \lambda^{-1} \int_{\mathbb{R}^{k-2}\times\mathbb{R}} |\partial_\lambda(\chi_\lambda * f_1)|^2\,d\mathcal{H}^k_{\mathcal{P}} d\lambda,$$

hence

$$\begin{aligned}
|I_2| \leq& 2\int_\epsilon^1 \int_{\mathbb{R}^{k-2}\times\mathbb{R}} H\left(\partial_\lambda(\chi_\lambda * f_1)\right) \partial_t(\chi_\lambda * f_1)\,d\mathcal{H}^k_{\mathcal{P}} d\lambda \\
\leq& C\left(\int_\epsilon^1 \lambda^{-1} \int_{\mathbb{R}^{k-2}\times\mathbb{R}} |\partial_\lambda(\chi_\lambda * f_1)|^2\,d\mathcal{H}^k_{\mathcal{P}} d\lambda\right)^{\frac{1}{2}} \left(\int_\epsilon^1 \lambda \int_{\mathbb{R}^{k-2}\times\mathbb{R}} |\partial_t(\chi_\lambda * f_1)|^2\,d\mathcal{H}^k_{\mathcal{P}} d\lambda\right)^{\frac{1}{2}}.
\end{aligned}$$

For any $\mathbf{x} = (x,t) \in \mathbb{R}^{k-2} \times \mathbb{R}$ and $\mathbf{y} = (y,s) \in \mathbb{R}^{k-2} \times \mathbb{R}$ we have

$$\begin{aligned}
|\varphi(\mathbf{y}) - \varphi(\mathbf{x}) - \langle \nabla\varphi(\mathbf{x}), y - x\rangle| &\leq \left|\int_t^s \partial_\tau \varphi(y,\tau)\,d\tau\right| + |(\varphi(y,t) - \varphi(\mathbf{x})) - \nabla\varphi(\mathbf{x}) \cdot (y-x)| \\
&\leq C|\mathbf{x} - \mathbf{y}|^2.
\end{aligned}$$

Because $\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_t \chi_\lambda(\mathbf{y})\,d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) = 0$ and $\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_t \chi_\lambda(\mathbf{y}) y\,d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) = 0$ (the second identity follows from the rotational invariance of $\chi_\lambda$), this implies

$$\begin{aligned}
\partial_t(\chi_\lambda * f_1)(\mathbf{x}) =& \int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_t\chi_\lambda(\mathbf{x} - \mathbf{y})(f(\mathbf{y}) - f_Q)\varphi(\mathbf{y})\,d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \\
=& \int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_t\chi_\lambda(\mathbf{x} - \mathbf{y})(f(\mathbf{y}) - f_Q)\left(\varphi(\mathbf{y}) - \varphi(\mathbf{x}) - \nabla\varphi(\mathbf{x}) \cdot (y - x)\right) d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \\
&+ \varphi(\mathbf{x}) \int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_t\chi_\lambda(\mathbf{x} - \mathbf{y}) f(\mathbf{y})\,d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \\
&+ \nabla\varphi(\mathbf{x}) \cdot \int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_t\chi_\lambda(\mathbf{x} - \mathbf{y})(f(\mathbf{y}) - f(\mathbf{x}))(\mathbf{y} - \mathbf{x})\,d\mathcal{H}^k(\mathbf{y}),
\end{aligned}$$

hence

$$|\partial_t(\chi_\lambda * f_1) - \varphi \partial_t(\chi_\lambda * f)|(\mathbf{x}) \leq C\delta\lambda^{-(n+3)} \int_{P(\mathbf{x},\lambda)} |\mathbf{x} - \mathbf{y}|^2\,d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \leq C\delta.$$

Similarly, because $\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_\lambda \chi_\lambda(\mathbf{y})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) = 0$ and $\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_\lambda \chi_\lambda(\mathbf{y}) y\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) = 0$ (where the second identity again follows by rotational invariance), we have

$$\begin{aligned}\partial_\lambda(\chi_\lambda * f_1)(\mathbf{x}) = &\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_\lambda\chi_\lambda(\mathbf{x}-\mathbf{y})(f(\mathbf{y})-f_Q)\left(\varphi(\mathbf{y})-\varphi(\mathbf{x})-\nabla\varphi(\mathbf{x})\cdot(y-x)\right)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\ &+\varphi(\mathbf{x})\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_\lambda\chi_\lambda(\mathbf{x}-\mathbf{y})f(\mathbf{y})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\\ &+\nabla\varphi(\mathbf{x})\cdot\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \partial_\lambda\chi_\lambda(\mathbf{x}-\mathbf{y})(f(\mathbf{y})-f(\mathbf{x}))(y-x)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}).\end{aligned}$$

Using

$$\begin{aligned}|\partial_\lambda\chi_\lambda(x-y,t-s)| \leq &|-(n+1)\lambda^{-(n+2)}\chi(\lambda^{-1}(x-y),\lambda^{-2}(t-s))|\\ &+|\lambda^{-(n+3)}\langle\nabla\chi(\lambda^{-1}(x-y),\lambda^{-2}(t-s)),x-y\rangle|\\ &+|t-s|\cdot|\lambda^{-(n+4)}\partial_t\chi(\lambda^{-1}(x-y),\lambda^{-2}(t-s))|\\ \leq &C\lambda^{-(n+2)},\end{aligned}$$

we get

$$|\partial_\lambda(\chi_\lambda * f_1)-\varphi\partial_\lambda(\chi_\lambda * f)|\leq C\delta\lambda. \tag{B.5}$$

Combining expressions, we thus obtain

$$\begin{aligned}|I_2|\leq C&\left(\int_\epsilon^1\lambda^{-1}\int_{(\mathbb{R}^{k-2}\times\mathbb{R})\cap P(\mathbf{0},10)}|\partial_\lambda(\chi_\lambda * f)|^2\, d\mathcal{H}^k_{\mathcal{P}}d\lambda + C\delta\right)^{\frac{1}{2}}\times\\ &\left(\int_\epsilon^1\lambda\int_{(\mathbb{R}^{k-2}\times\mathbb{R})\cap P(\mathbf{0},10)}|\partial_t(\chi_\lambda * f)|^2\, d\mathcal{H}^k_{\mathcal{P}}d\lambda + C\delta\right)^{\frac{1}{2}}.\end{aligned}$$

**Claim B.1.2.** *The following holds:*

$$\begin{aligned}&\int_0^1\lambda^{-1}\int_{P(\mathbf{0},10)}|\partial_\lambda(\chi_\lambda * f)|^2\, d\mathcal{H}^k_{\mathcal{P}}d\lambda+\int_0^1\lambda\int_{P(\mathbf{0},10)}|\partial_t(\chi_\lambda * f)|^2\, d\mathcal{H}^k_{\mathcal{P}}d\lambda\\ &\leq C\int_0^1\int_{P(\mathbf{0},10)}\kappa^2(\mathbf{y},\lambda)\, d\mathcal{H}^k_{\mathcal{P}}\frac{d\lambda}{\lambda}<C\delta.\end{aligned}$$

*Proof.* For any $\theta\in C_c^\infty(P(\mathbf{0},\lambda))$ satisfying $\int_{\mathbb{R}^{k-2}\times\mathbb{R}}\theta(\mathbf{y})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})=0$ and $\int_{\mathbb{R}^{k-2}\times\mathbb{R}}\theta(\mathbf{y})y\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})=0$, and for any affine function $\ell:\mathbb{R}^{k-2}\times\mathbb{R}\to\mathbb{R}$ with $\partial_t\ell=0$, we have

$$(\theta*\ell)(\mathbf{x})=\int_{\mathbb{R}^n}\theta(\mathbf{y})\ell(\mathbf{x}-\mathbf{y})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})=0$$

for all $\mathbf{x}\in\mathbb{R}^{k-2}\times\mathbb{R}$. Thus

$$\begin{aligned}(\theta*f)^2(\mathbf{x})=&\left(\theta*(f-\ell)\right)^2(\mathbf{x})=\left(\int_{\mathbb{R}^n}\theta(\mathbf{x}-\mathbf{y})(f(\mathbf{y})-\ell(\mathbf{y}))\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\right)^2\\ \leq& C\lambda^{2k}\|\theta\|_\infty^2\left(\frac{1}{\lambda^k}\int_{P(\mathbf{x},\lambda)}(f(\mathbf{y})-\ell(\mathbf{y}))\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y})\right)^2\\ \leq& C\lambda^{2k}\|\theta\|_\infty^2\frac{1}{\lambda^k}\int_{P(\mathbf{x},\lambda)}(f(\mathbf{y})-\ell(\mathbf{y}))^2\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}).\end{aligned}$$

Taking infimum over $\ell$, we obtain

$$(\theta*f)^2(\mathbf{x})\leq C\lambda^{2k+2}\|\theta\|_\infty^2\kappa^2(\mathbf{x},\lambda).$$

Now set $\theta \coloneqq \partial_\lambda \chi_\lambda$. Recalling that $\|\partial_\lambda \chi_\lambda\|_\infty \le C\lambda^{-(k+1)}$, we integrate in spacetime and then against $\frac{d\lambda}{\lambda}$ to obtain

$$\int_0^1 \int_{P(\mathbf{0},10)} ((\partial_\lambda \chi_\lambda) * f)^2(\mathbf{y})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \frac{d\lambda}{\lambda} \le C \int_0^1 \int_{P(\mathbf{0},10)} \lambda^{2k+2}\lambda^{-2(k+1)}\kappa^2(\mathbf{y},\lambda)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \frac{d\lambda}{\lambda}.$$

Similarly, we can take $\theta \coloneqq \partial_t \chi_\lambda$ and recall that $\|\partial_t \chi_t\|_\infty \le C\lambda^{-(k+2)}$ to obtain

$$\int_0^1 \int_{P(\mathbf{0},10)} \lambda((\partial_t \chi_\lambda) * f)^2(\mathbf{y})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) d\lambda \le C \int_0^1 \int_{P(\mathbf{0},10)} \lambda^{-2(k+2)}\lambda^{2k+4}\kappa^2(\mathbf{y},\lambda)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{y}) \frac{d\lambda}{\lambda}.$$

$\square$

Combining Claim B.1.2 with (B.5) yields $|I_2| \le C\delta$, and further combining this with the estimate $|I_1| \le C\delta$ gives

$$\int_{\mathbb{R}^{k-2}\times\mathbb{R}} (\partial_t^{\frac{1}{2}}(\chi_\epsilon * f_1))^2(\mathbf{x})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) \le C\delta$$

for all $\epsilon \in (0,1]$. We can thus find a sequence $\epsilon_j \searrow 0$ such that

$$\partial_t^{\frac{1}{2}}(\chi_{\epsilon_j} * f) \rightharpoonup \zeta$$

weakly in $L^2(\mathbb{R}^{k-2}\times\mathbb{R})$ to some $\zeta \in L^2(\mathbb{R}^{k-2}\times\mathbb{R})$ as $j\to\infty$, which satisfies $\int_{\mathbb{R}^{k-2}\times\mathbb{R}} \zeta^2\, d\mathcal{H}^k_{\mathcal{P}} \le C\delta$. Define $I_{\frac{1}{2}}(s) \coloneqq |s|^{-\frac{1}{2}}$ for $s \neq 0$. Then the Fourier transform of $I_{\frac{1}{2}}$ is $\mathcal{F}I_{\frac{1}{2}}(\tau) = \frac{1}{c}|\tau|^{-\frac{1}{2}}$. For any Schwartz function $g$ on $\mathbb{R}$, we have

$$c\mathcal{F}\left(I_{\frac{1}{2}} * \partial_t^{\frac{1}{2}} g\right)(\tau) = c\mathcal{F}I_{\frac{1}{2}}(\tau)|\tau|^{\frac{1}{2}}\mathcal{F}g(\tau) = \mathcal{F}g(\tau),$$

so that $cI_{\frac{1}{2}} * (\partial_t^{\frac{1}{2}} g) = g$. Applying this identity to $g \coloneqq (\chi_{\epsilon_j} * f_1)(x,\cdot)$ for $x \in \mathbb{R}^{k-2}$ yields $cI_{\frac{1}{2}} * \left(\partial_t^{\frac{1}{2}}(\chi_\epsilon * f)\right) = \chi_\epsilon * f$. For any Schwartz function $g \in \mathscr{S}(\mathbb{R}^{k-2}\times\mathbb{R})$,

$$\begin{aligned}
\int_{\mathbb{R}^{k-2}\times\mathbb{R}} g(\mathbf{x})(\chi_\epsilon * f_1)(\mathbf{x})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) =& c\int_{\mathbb{R}^{k-2}\times\mathbb{R}} g(x,\cdot)\left(I_{\frac{1}{2}} * \left(\partial_t^{\frac{1}{2}}(\chi_\epsilon * f_1)(x,\cdot)\right)\right)(t)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) \\
=& c\int_{\mathbb{R}^{k-2}\times\mathbb{R}} (I_{\frac{1}{2}} * g(x,\cdot))(t)\partial_t^{\frac{1}{2}}(\chi_\epsilon * f_1)(x,\cdot)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}),
\end{aligned}$$

so that taking $\epsilon \to 0$ and using the fact that $\chi_\epsilon * f_1 \to f_1$ in $L^2(\mathbb{R}^{k-2}\times\mathbb{R})$, we obtain

$$\begin{aligned}
\int_{\mathbb{R}^{k-2}\times\mathbb{R}} g(\mathbf{x})f_1(\mathbf{x})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) &= c\int_{\mathbb{R}^{k-2}\times\mathbb{R}} (I_{\frac{1}{2}} * g(x,\cdot))(t)\zeta(\mathbf{x})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) \\
&= c\int_{\mathbb{R}^{k-2}} g(\mathbf{x})(I_{\frac{1}{2}} * \zeta(x,\cdot))(t)\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}).
\end{aligned}$$

That is, $f_1(x,\cdot) = c(I_{\frac{1}{2}} * \zeta(x,\cdot))$. A computation similar to the above yields $\partial_t^{\frac{1}{2}}(I_{\frac{1}{2}} * \zeta) = \zeta$, so that $\partial_t^{\frac{1}{2}} f_1 = \zeta$ exists and satisfies

$$\int_{\mathbb{R}^{k-2}\times\mathbb{R}} (\partial_t^{\frac{1}{2}} f_1)^2(\mathbf{x})\, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) \le C\delta.$$

We now handle $f_2$. Because $f_2$ is constant on $\mathbb{R}^{k-2}\times\mathbb{R}\cap P(\mathbf{0},\frac{3}{2})$, for any $\mathbf{x} = (x,t) \in Q$, we have

$$(\partial_t^{\frac{1}{2}} f_2)(\mathbf{x}) = c\int_{|s|\ge\frac{9}{4}} \frac{f_2(x,s)-f_2(x,t)}{|t-s|^{\frac{3}{2}}} ds.$$

Set

$$\alpha := c \int_{|s|\geq \frac{9}{4}} \frac{f(0,s)-f(0,0)}{|s|^{\frac{3}{2}}} ds,$$

and recall that $|f - f_Q| \leq C\delta$ on $\mathbb{R}^{k-2} \times \mathbb{R} \cap P(\mathbf{0}, 2)$, so that using

$$f_2(\mathbf{x}) = \chi f_Q + (1-\chi) f = \chi(f_Q - f) + f,$$

we obtain

$$\begin{aligned}
|\partial_t^{\frac{1}{2}} f_2(\mathbf{x}) - \alpha| \leq & c \int_{|s|\geq \frac{9}{4}} \left| \frac{f(x,s)-f(x,t)}{|t-s|^{\frac{3}{2}}} - \frac{f(0,s)-f(0,0)}{|s|^{\frac{3}{2}}} \right| ds \\
& + C \left( \sup_{\mathbb{R}^{k-2}\times\mathbb{R}\cap P(\mathbf{0},2)} |f - f_Q| \right) \int_{|s|\geq \frac{9}{4}} \frac{1}{|s|^{\frac{3}{2}}} ds \\
\leq & C \int_{|s|\geq \frac{9}{4}} \left| \frac{f(x,s)-f_Q}{|t-s|^{\frac{3}{2}}} - \frac{f(0,s)-f_Q}{|s|^{\frac{3}{2}}} \right| ds + C\delta \\
\leq & C \int_{|s|\geq \frac{9}{4}} \left| \frac{f(x,s)-f(0,s)}{|t-s|^{\frac{3}{2}}} + (f(0,s) - f_Q) \left( \frac{1}{|t-s|^{\frac{3}{2}}} - \frac{1}{|s|^{\frac{3}{2}}} \right) \right| ds + C\delta \\
\leq & C\delta \int_{|s|\geq \frac{9}{4}} |s| \frac{|t|}{|s|^{\frac{5}{2}}} ds + C\delta \\
\leq & C\delta.
\end{aligned}$$

Combining estimates yields that $\partial_t^{\frac{1}{2}} f$ exists and is integrable on $Q$, and

$$\int_Q |\partial_t^{\frac{1}{2}} f(\mathbf{x}) - \alpha| \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) \leq \int_Q |\partial_t^{\frac{1}{2}} f_1(\mathbf{x})| \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) + \int_Q |\partial_t^{\frac{1}{2}} f_2(\mathbf{x}) - \alpha| \, d\mathcal{H}^k_{\mathcal{P}}(\mathbf{x}) \leq C\sqrt{\delta},$$

so the claim follows. □

## References


[AB23] Vedansh Arya and Agnid Banerjee, *Space-like quantitative uniqueness for parabolic operators*, Journal de Mathématiques Pures et Appliquées **177** (2023), 214–259.

[ABG25] Vedansh Arya, Agnid Banerjee, and Nicola Garofalo, *Sharp order of vanishing for parabolic equations, nodal set estimates and landis type results*, Archive for Rational Mechanics and Analysis **249** (2025), no. 6, 67.

[Ang91] Sigurd Angenent, *Nodal properties of solutions of parabolic equations*, The Rocky Mountain Journal of Mathematics **21** (1991), no. 2, 585–592.

[Aub63] Jean-Pierre Aubin, *Un théoreme de compacité*, CR Acad. Sci. Paris **256** (1963), no. 24, 5042–5044.

[AV04] Giovanni Alessandrini and Sergio Vessella, *Remark on the strong unique continuation property for parabolic operators*, Proc. Amer. Math. Soc. **132** (2004), no. 2, 499–501. MR 2022375

[BF12] Franck Boyer and Pierre Fabrie, *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models*, vol. 183, Springer Science & Business Media, 2012.

[BGL14] Dominique Bakry, Ivan Gentil, and Michel Ledoux, *Analysis and geometry of Markov diffusion operators*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 348, Springer, Cham, 2014. MR 3155209

[BHH⁺23] Simon Bortz, John Hoffman, Steve Hofmann, José Luis Luna García, and Kaj Nyström, *Carleson measure estimates for caloric functions and parabolic uniformly rectifiable sets*, Anal. PDE **16** (2023), no. 4, 1061–1088. MR 4605204

[BHMN25] Simon Bortz, Steve Hofmann, José María Martell, and Kaj Nyström, *Solvability of the* $\mathrm{L}^p$ *Dirichlet problem for the heat equation is equivalent to parabolic uniform rectifiability in the case of a parabolic Lipschitz graph*, Invent. Math. **239** (2025), no. 1, 165–217. MR 4841778

[CJN21] Jeff Cheeger, Wenshuai Jiang, and Aaron Naber, *Rectifiability of singular sets of noncollapsed limit spaces with Ricci curvature bounded below*, Ann. of Math. (2) **193** (2021), no. 2, 407–538. MR 4226910

[CM11] Tobias H. Colding and William P. Minicozzi, II, *Lower bounds for nodal sets of eigenfunctions*, Comm. Math. Phys. **306** (2011), no. 3, 777–784. MR 2825508

[CMI20] Tobias Holck Colding and William P. Minicozzi II, *Parabolic Frequency on Manifolds*, February 2020.

[CN13] Jeff Cheeger and Aaron Naber, *Lower bounds on Ricci curvature and quantitative behavior of singular sets*, Inventiones mathematicae **191** (2013), no. 2, 321–339.

[CNV15] Jeff Cheeger, Aaron Naber, and Daniele Valtorta, *Critical Sets of Elliptic Equations*, Communications on Pure and Applied Mathematics **68** (2015), no. 2, 173–209.

[CT94] Jeff Cheeger and Gang Tian, *On the cone structure at infinity of Ricci flat manifolds with Euclidean volume growth and quadratic curvature decay*, Invent. Math. **118** (1994), no. 3, 493–571. MR 1296356

[DF88] Harold Donnelly and Charles Fefferman, *Nodal sets of eigenfunctions on Riemannian manifolds*, Invent. Math. **93** (1988), no. 1, 161–183. MR 943927

[DF90] ———, *Nodal sets for eigenfunctions of the Laplacian on surfaces*, J. Amer. Math. Soc. **3** (1990), no. 2, 333–353. MR 1035413

[Don92] Rui-Tao Dong, *Nodal sets of eigenfunctions on Riemann surfaces*, J. Differential Geom. **36** (1992), no. 2, 493–506. MR 1180391

[DS91] G. David and S. Semmes, *Singular integrals and rectifiable sets in* $\mathbf{R}^n$*: Beyond Lipschitz graphs*, Astérisque (1991), no. 193, 152. MR 1113517

[DSP05] Daniele Del Santo and Martino Prizzi, *Backward uniqueness for parabolic operators whose coefficients are non-Lipschitz continuous in time*, J. Math. Pures Appl. (9) **84** (2005), no. 4, 471–491. MR 2133125

[EFV06] Luis Escauriaza, FJ Fernández, and Sergio Vessella, *Doubling properties of caloric functions*, Applicable analysis **85** (2006), no. 1-3, 205–223.

[Eva10] Lawrence C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR 2597943

[Fed70] Herbert Federer, *The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension*, Bull. Amer. Math. Soc. **76** (1970), 767–771. MR 260981

[Fed14] ———, *Geometric measure theory*, Springer, 2014.

[Fer03] F. J. Fernandez, *Unique continuation for parabolic operators. II*, Comm. Partial Differential Equations **28** (2003), no. 9-10, 1597–1604. MR 2001174

[FL25a] Hanbing Fang and Yu Li, *Singular sets in noncollapsed ricci flow limit spaces*, 2025.

[FL25b] Hanbing Fang and Yu Li, *Volume estimates for the singular sets of mean curvature flows*, arXiv preprint arXiv:2504.09811 (2025), 1–32.

[FWWZ25] Haotong Fu, Wei Wang, Ke Wu, and Zhifei Zhang, *Stratification and rectifiability of harmonic map flows via tangent measures*, arXiv preprint arXiv:2504.14880 (2025), 1–41.

[Gia25] Panagiotis Gianniotis, *Diameter bounds in 3d type i ricci flows*, 2025.

[GL25] Panagiotis Gianniotis and Konstantinos Leskas, $l^1$ *curvature bounds for type i ricci flows*, 2025.

[Gro75] Leonard Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97** (1975), no. 4, 1061–1083. MR 420249

[HHL98] Qing Han, Robert Hardt, and Fanghua Lin, *Geometric measure of singular sets of elliptic equations*, Comm. Pure Appl. Math. **51** (1998), no. 11-12, 1425–1443. MR 1639155

[HHOHON99] R. Hardt, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof, and N. Nadirashvili, *Critical sets of solutions to elliptic equations*, J. Differential Geom. **51** (1999), no. 2, 359–373. MR 1728303

[HJ23] Yiqi Huang and Wenshuai Jiang, *Volume Estimates for Singular sets and Critical Sets of Elliptic Equations with Hölder Coefficients*, arXiv preprint arXiv:2309.08089 (2023), 1–47.

[HJ24] ———, *The Nodal Sets of Solutions to Parabolic Equations*, arXiv preprint arXiv:2406.05877 (2024), 1–44.

[HJ25] Yiqi Huang and Wenshuai Jiang, *The bounded diameter conjecture and sharp geometric estimates for mean curvature flow*, 2025.

[HL94a] Qing Han and Fang-Hua Lin, *Nodal sets of solutions of parabolic equations. II*, Comm. Pure Appl. Math. **47** (1994), no. 9, 1219–1238. MR 1290401

[HL94b] ———, *On the geometric measure of nodal sets of solutions*, J. Partial Differential Equations **7** (1994), no. 2, 111–131. MR 1280175

[HLN03] Steve Hofmann, John L. Lewis, and Kaj Nyström, *Existence of big pieces of graphs for parabolic problems*, Ann. Acad. Sci. Fenn. Math. **28** (2003), no. 2, 355–384. MR 1996443

[HLN04] ———, *Caloric measure in parabolic flat domains*, Duke Math. J. **122** (2004), no. 2, 281–346. MR 2053754

[HS89] Robert Hardt and Leon Simon, *Nodal sets for solutions of elliptic equations*, J. Differential Geom. **30** (1989), no. 2, 505–522. MR 1010169

[HS12] Hamid Hezari and Christopher D. Sogge, *A natural lower bound for the size of nodal sets*, Anal. PDE **5** (2012), no. 5, 1133–1137. MR 3022851

[JN21] Wenshuai Jiang and Aaron Naber, *$L^2$ curvature bounds on manifolds with bounded Ricci curvature*, Annals of Mathematics **193** (2021), no. 1, 107–222.

[Kuk95] Igor Kukavica, *Hausdorff measure of level sets for solutions of parabolic equations*, Internat. Math. Res. Notices (1995), no. 13, 671–682. MR 1383945

[KZZ22] Carlos E. Kenig, Jiuyi Zhu, and Jinping Zhuge, *Doubling inequalities and nodal sets in periodic elliptic homogenization*, Comm. Partial Differential Equations **47** (2022), no. 3, 549–584. MR 4387203

[Lie96] Gary M Lieberman, *Second order parabolic differential equations*, World scientific, 1996.

[Lin91] Fang-Hua Lin, *Nodal sets of solutions of elliptic and parabolic equations*, Comm. Pure Appl. Math. **44** (1991), no. 3, 287–308. MR 1090434

[LL16] Fanghua Lin and Dan Liu, *On the Betti numbers of level sets of solutions to elliptic equations*, Discrete Contin. Dyn. Syst. **36** (2016), no. 8, 4517–4529. MR 3479524

[LM18] Alexander Logunov and Eugenia Malinnikova, *Nodal sets of Laplace eigenfunctions: estimates of the Hausdorff measure in dimensions two and three*, 50 years with Hardy spaces, Oper. Theory Adv. Appl., vol. 261, Birkhäuser/Springer, Cham, 2018, pp. 333–344. MR 3792104

[LM20] ———, *Review of Yau's conjecture on zero sets of Laplace eigenfunctions*, Current developments in mathematics 2018, Int. Press, Somerville, MA, [2020] ©2020, pp. 179–212. MR 4363378

[LMNN21] A. Logunov, E. Malinnikova, N. Nadirashvili, and F. Nazarov, *The sharp upper bound for the area of the nodal sets of Dirichlet Laplace eigenfunctions*, Geom. Funct. Anal. **31** (2021), no. 5, 1219–1244. MR 4356702

[Log18a] Alexander Logunov, *Nodal sets of Laplace eigenfunctions: polynomial upper estimates of the Hausdorff measure*, Annals of Mathematics **187** (2018), no. 1, 221–239.

[Log18b] ———, *Nodal sets of Laplace eigenfunctions: proof of Nadirashvili's conjecture and of the lower bound in Yau's conjecture*, Ann. of Math. (2) **187** (2018), no. 1, 241–262. MR 3739232

[LS19] Fanghua Lin and Zhongwei Shen, *Nodal sets and doubling conditions in elliptic homogenization*, Acta Math. Sin. (Engl. Ser.) **35** (2019), no. 6, 815–831. MR 3952693

[LTY24] Fang Liu, Long Tian, and Xiaoping Yang, *Measure upper bounds of nodal sets of Robin eigenfunctions*, Math. Z. **306** (2024), no. 1, Paper No. 14, 14. MR 4675263

[Mat95] Pertti Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, Cambridge, 1995, Fractals and rectifiability. MR 1333890

[Mat22] ———, *Parabolic Rectifiability, Tangent Planes and Tangent Measures*, Annales Fennici Mathematici **47** (2022), no. 2, 855–884.

[MPP02] Giorgio Metafune, Diego Pallara, and Enrico Priola, *Spectrum of Ornstein-Uhlenbeck operators in Lp spaces with respect to invariant measures*, Journal of Functional Analysis **196** (2002), no. 1, 40–60.

[Nab18] Aaron Naber, *Lecture notes on rectifiable reifenberg for measures*, 2018.

[NV17a] Aaron Naber and Daniele Valtorta, *Rectifiable-Reifenberg and the regularity of stationary and minimizing harmonic maps*, Ann. of Math. (2) **185** (2017), no. 1, 131–227. MR 3583353

[NV17b] ———, *Volume estimates on the critical sets of solutions to elliptic PDEs*, Communications on Pure and Applied Mathematics **70** (2017), no. 10, 1835–1897.

[NV19] ———, *Energy identity for stationary Yang Mills*, Invent. Math. **216** (2019), no. 3, 847–925. MR 3955711

[NV24] ———, *Energy Identity for Stationary Harmonic Maps*, arXiv preprint arXiv:2401.02242 (2024), 1–101.

[Poo96] Chi-Cheung Poon, *Unique continuation for parabolic equations*, Comm. Partial Differential Equations **21** (1996), no. 3-4, 521–539. MR 1387458

[Sim85] Leon Simon, *Isolated singularities of extrema of geometric variational problems*, Harmonic mappings and minimal immersions (Montecatini, 1984), Lecture Notes in Math., vol. 1161, Springer, Berlin, 1985, pp. 206–277. MR 821971

[Sim86] Jacques Simon, *Compact sets in the space $L^p(O,T;B)$*, Annali di Matematica pura ed applicata **146** (1986), no. 1, 65–96.

[SZ12] Christopher D. Sogge and Steve Zelditch, *Lower bounds on the Hausdorff measure of nodal sets II*, Math. Res. Lett. **19** (2012), no. 6, 1361–1364. MR 3091613

[Won93] Richard Wongkew, *Volumes of tubular neighbourhoods of real algebraic varieties*, Pacific Journal of Mathematics **159** (1993), no. 1, 177–184.

[Zel13] Steve Zelditch, *Eigenfunctions and nodal sets*, Surveys in differential geometry. Geometry and topology, Surv. Differ. Geom., vol. 18, Int. Press, Somerville, MA, 2013, pp. 237–308. MR 3087922

[Zel15] ———, *Hausdorff measure of nodal sets of analytic Steklov eigenfunctions*, Math. Res. Lett. **22** (2015), no. 6, 1821–1842. MR 3507264

Department of Mathematics, Rutgers University
*Email address*: `mh1564@scarletmail.rutgers.edu`

Department of Mathematics, University of California San Diego
*Email address*: `rkoirala@ucsd.edu`

Department of Mathematics, University of Tennessee Knoxville
*Email address*: `zma21@utk.edu`