Interfaces Free Bound. 25 (2023), 155–191
DOI 10.4171/IFB/488

Stability of self-similar solutions to geometric flows

Hengrong Du and Nung Kwan Yip

Abstract. We show that self-similar solutions for the mean curvature flow, surface diffusion, and Willmore flow of entire graphs are stable upon perturbations of initial data with small Lipschitz norm. Roughly speaking, the perturbed solutions are asymptotically self-similar as time tends to infinity. Our results are built upon the global analytic solutions constructed by Koch and Lamm in 2012, the compactness arguments adapted by Asai and Giga in 2014, and the spatial equi-decay properties on certain weighted function spaces. The proof for all of the above flows are achieved in a unified framework by utilizing the estimates of the linearized operator.

1. Introduction

We analyze in this paper the long-time asymptotics of various geometric flows, in particular the stability of self-similar solutions. From the point of view of calculus of variations, many geometric flows can be seen as the negative gradient flows of some geometric functionals with respect to a certain underlying metric. Heuristically, the gradient descent nature of the flows evolves general initial data toward a critical point of the corresponding functional. These evolutions are often modeled by nonlinear parabolic partial differential equations. The long-time asymptotics of the solution is one of the key questions to be investigated. For instance, in the celebrated work [35] of Leon Simon, the asymptotics of a large class of such geometric evolution equations are studied by an infinite-dimensional version of the Łojasiewicz inequalities combined with the Liapunov–Schmidt reduction. It is also worth pointing out that in [13] Eells and Sampson used the long-time limit of heat flows to construct harmonic mappings between Riemannian manifolds under certain curvature assumptions.

The geometric flows studied in this paper are of curvature-driven type, which arise from the energy minimization of the surface area functional. This naturally leads to evolutions involving mean curvature which is the first variation of the surface area. These motions appear often in the modeling of materials science, such as phase transitions and grain growth [1, 30]. It is also used in describing the bending of membranes in red blood cells [20, 34]. The underlying equations are related to mean curvature flows (MCF), surface diffusion (SD) and Willmore flows (WF), which are the three equations analyzed in this paper.

2020 Mathematics Subject Classification. Primary 53E10; Secondary 35K93, 35K30, 35K25.
Keywords. Self-similar solutions, geometric flows, mean curvature.

One mathematical point to note is that the equations to be analyzed include fourth-order flows, which are much harder to handle than their second-order counterparts due to the lack of maximum or comparison principle. On the other hand, these flows enjoy a certain invariant property, leading to the existence of self-similar solutions. The main goal of the current paper is to analyze the stability of these solutions. More precisely, under fairly general initial conditions, we will show that the solutions to these equations converge to some self-similar form. In order to take advantage of a general unified approach, we restrict ourselves to entire graph solutions relying very much on linearized analysis.

One can also interpret this phenomena of self-similarity using the *renormalization group* method as in [4]. The key idea is that after rescaling or zooming out in the spatial variable, suppose the initial data converges to a scale-invariant function which is determined by the behavior of the data at infinity; then, the solution will converge to a scale-invariant solution, or so-called *self-similar solution*. In other words, the long-time asymptotics are determined by the rescaling limit of the initial data. Hence, we expect that if the initial data is perturbed without changing the scaling limit, then the corresponding solution will increasingly look like the self-similar solution corresponding to the unperturbed scale-invariant initial data. There is also huge literature where such phenomena is proved for semilinear heat equations—see, for example, [6, 19, 23, 29], to name just a few. Another technique extensively used in the case of MCF is the *monotonicity formula*. It has been used in this case to characterize the form of self-similar solutions and the convergence to them [11, 22]. This is also the precursor to the more recent entropy method to characterize self-similar shrinkers [10].

In this paper, we will investigate the stability of self-similar solutions corresponding to MCF, SD, and WF. Note that global-in-time existence of classical solutions to these geometric flows with general initial data does not hold due to the possibility of finite-time blow-ups. On the other hand, in the case of graph setting, it is possible to have long-time solutions. For MCF, this is comprehensively analyzed in [11, 12]. In a very interesting paper [24], Koch and Lamm have constructed a unique global-in-time solution to these geometric flows under a *small Lipschitz norm* assumption on the initial data. This is in contrast to those existence results of classical solutions making use of the maximal regularity property of elliptic operators where the initial data are required to be $C^{1,\alpha}$ or $C^{2,\alpha}$ (depending on the order of the equation)—see [15, 16, 36] for examples of such results. The main technique of [24], which originated from Koch–Tataru [25] for incompressible Navier–Stokes equations, is a fixed point argument on some scale-invariant function spaces. Even though it can only handle the case of graphs, all the above geometric flows in general dimensions can be tackled in a unified framework. In addition, the approach does not rely on a maximum principle which only works for second-order scalar PDEs. Thus, it is applicable for PDE systems and higher-order equations.

Another relevant work is Asai–Giga [2], which establishes a stability result for self-similar solutions to a one-dimensional surface diffusion with bounded initial data. It uses a compactness argument in some Hölder spaces. The earlier work [3] proves a similar result, but it seems the technique is only applicable to the one-dimensional curve case. From an application point of view, these two works touch upon the celebrated model called *thermal grooving*, first described by Mullins [30]. Combining the techniques of [24] and [2], we are able to show a local-in-space stability result (Theorem 2.2) and also a global-in-space result (Theorem 2.4). The latter is achieved in the setting of some weighted function spaces. Qualitatively, we have extended the result of [2, 3] to higher dimensions with unbounded initial data.

This paper is organized as follows: In Section 2, we introduce the geometric flows, the definition of self-similar solutions, and the statement of our main results. Then, we outline the strategy of proof. Section 3 is devoted to the proof of Theorem 2.2, which asserts the local-in-space convergence of the perturbed solution. Next, in Section 4, we prove our global-in-space convergence result (Theorem 2.4) under a spatial decaying assumption on the initial perturbation. We make a remark in Section 5 on the generalization to polyharmonic flows. The proofs of technical Lemmas 4.4 and 4.7 are put in Appendices A and B, respectively.

Before getting into the technical details, we introduce one notation to be used throughout this paper. We write for any two positive quantities that $A \lesssim B$ if there is a universal constant C such that $A \leq CB$. The value of the constant is not relevant in the argument and can change from one line to the other.

2. Geometric flows

Let Σ be a closed hypersurface in $\mathbb{R}^{n+1}$. The area functional of Σ is given by

$$A(\Sigma) = \int_{\Sigma} 1 d\mu_g, \tag{2.1}$$

where g is the induced metric from the immersion and $d\mu_g$ is the corresponding area element. The aim of this paper is to investigate the (L^2- and H^{-1}-) negative gradient flows of (2.1). More precisely, we consider a time-dependent hypersurface Σ_t given by immersions $f : \Sigma \times \mathbb{R}_+ \to \mathbb{R}^{n+1}$ which evolves according to

(1) Mean curvature flow (MCF):

$$\partial_t f = \mathcal{H} := -\nabla_{L^2} A, \tag{2.2}$$

(2) Surface diffusion (SD):

$$\partial_t f = -\Delta_g \mathcal{H} =: -\nabla_{H^{-1}} A, \tag{2.3}$$

where $\mathcal{H}$ represents the mean curvature vector and Δ_g is the Laplace–Beltrami operator with respect to the induced metric g. Note that MCF and SD can be recast as the negative gradient flows of A with respect to the L^2 and H^{-1} metric, respectively—see [5, 41] for more details about the derivation.

We will also consider the following Willmore functional for two-dimensional surfaces ($n = 2$) in $\mathbb{R}^3$:

$$W(f) = \frac{1}{4}\int_{\Sigma} |\mathcal{H}|^2 d\mu_g. \tag{2.4}$$

The negative L^2 gradient flow of (2.4) is then given as follows:

(3) Willmore flow (WF):

$$f_t^{\perp} = -\Delta_g \mathcal{H} - \frac{1}{2}\mathcal{H}^3 + 2\mathcal{H}\mathcal{K} =: -\nabla_{L^2} W, \tag{2.5}$$

where $\mathcal{K}$ is the Gauss curvature of Σ. We refer the reader to [26] for details of the derivation.

As mentioned earlier, in this paper we consider the case that Σ_t is given by an entire graph, that is, there exists a function $u : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ such that

$$\Sigma_t = \big\{(x, u(x,t)) \mid x \in \mathbb{R}^n,\ t \in \mathbb{R}_+\big\}.$$

For concreteness, we write down the graph equations for (2.2), (2.3), and (2.5):

$$\text{MCF:} \quad \frac{\partial u}{\partial t} = \sqrt{1+|\nabla u|^2}\,\mathrm{div}\Big(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\Big), \tag{2.6}$$

$$\text{SD:} \quad \frac{\partial u}{\partial t} = -\,\mathrm{div}\bigg[\sqrt{1+|\nabla u|^2}\Big(I - \frac{\nabla u \otimes \nabla u}{1+|\nabla u|^2}\Big)\nabla\,\mathrm{div}\Big(\frac{\nabla u}{\sqrt{1+|\nabla u|^2}}\Big)\bigg], \tag{2.7}$$

$$\text{WF:} \quad \frac{\partial u}{\partial t} = -w\,\mathrm{div}\bigg[\frac{1}{w}\Big(\Big(I - \frac{\nabla u \otimes \nabla u}{w^2}\Big)\nabla(w\mathcal{H}) - \frac{1}{2}\mathcal{H}^2\nabla u\Big)\bigg]. \tag{2.8}$$

In the above, we have used the following notations and representation:

$$w = \sqrt{1+|\nabla u|^2} \quad \text{and} \quad \mathcal{H} = \mathrm{div}\Big(\frac{\nabla u}{w}\Big).$$

To simplify the above equations, we borrow the contraction operator $\star$ from [24] for all possible contractions between derivatives of u; for example, we use $\nabla^2 u \star \nabla u \star \nabla u$ to indicate any expression of the form $\nabla_{ij} u \nabla_k u \nabla_l u$ with $1 \le i, j, k, l \le n$. They are all treated equally in terms of analysis. Moreover, we use $P_k(\nabla u)$ to denote some k-th power contraction of ∇u, that is,

$$P_k(\nabla u) = \underbrace{\nabla u \star \cdots \star \nabla u}_{k \text{ times}} = \prod_{j=1}^{k} \nabla_{i_j} u \quad \text{for some } 1 \le i_j \le n.$$

As derived in [24], we can rewrite equations (2.2), (2.3), and (2.5) using the above convention as follows:

$$\text{MCF:} \quad \partial_t u - \Delta u = w^{-2}\nabla^2 u \star P_2(\nabla u), \tag{2.9}$$

$$\text{SD:} \quad \partial_t u + \Delta^2 u = \nabla_i f_1^i[u] + \nabla_{ij} f_2^{ij}[u], \tag{2.10}$$

$$\text{WF:} \quad \partial_t u + \Delta^2 u = f_0[u] + \nabla_i f_1^i[u] + \nabla_{ij} f_2^{ij}[u], \tag{2.11}$$

where

$$f_0[u] = \nabla^2 u \star \nabla^2 u \star \nabla^2 u \star \sum_{k=1}^{4} w^{-2k} P_{2k-2}(\nabla u), \tag{2.12}$$

$$f_1[u] = \nabla^2 u \star \nabla^2 u \star \sum_{k=1}^{4} w^{-2k} P_{2k-1}(\nabla u), \tag{2.13}$$

$$f_2[u] = \nabla^2 u \star \sum_{k=1}^{2} w^{-2k} P_{2k}(\nabla u). \tag{2.14}$$

Under the assumption that $|\nabla u| \lesssim 1$, the following crude bounds for the nonlinear terms play crucial roles in our analysis:

$$|f_0[u]| \lesssim |\nabla^2 u|^3, \quad |f_1[u]| \lesssim |\nabla^2 u|^2, \quad \text{and} \quad |f_2[u]| \lesssim |\nabla^2 u|. \tag{2.15}$$

Abstractly, we can write (2.9), (2.10), and (2.11) in the following form:

$$\begin{cases} \partial_t u + Au = N[u], & (x,t) \in \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x), & x \in \mathbb{R}^n, \end{cases} \tag{2.16}$$

where $A = -\Delta$ or Δ^2, and $N[u]$ is the nonlinear term on the right-hand side of equations (2.9), (2.10), or (2.11). In this paper, we will consider *mild solutions* $u(x,t)$ to (2.16), by which we mean that u satisfies the following integral equation:

$$u(x,t) = e^{-At} u_0(x) + \int_0^t e^{-(t-s)A} N[u](x,s) ds, \quad (x,t) \in \mathbb{R}^n \times (0,\infty), \tag{2.17}$$

where e^{-At} is the semigroup generated by $-A$. If the Lipschitz norm of u_0 is small, the global well-posedness of mild solutions to (2.16) is obtained by Koch–Lamm [24]. More specifically, the following result (global well-posedness for initial data with small Lipschitz norm) for (2.16) and the technique to prove it provide a starting point for our investigation (the definition of the function space X_∞ will be given in (3.28) and (3.29)):

Theorem 2.1 (Koch–Lamm [24, Theorems 3.1 and 5.1]). *There exist $\varepsilon > 0$, $C > 0$ such that for every u_0 with $\|\nabla u_0\|_\infty < \varepsilon$ there exists an analytic solution $u \in X_\infty$ of (2.16) with $u(\cdot,0) = u_0$ which satisfies $\|u\|_{X_\infty} \le C\|\nabla u_0\|_{L^\infty(\mathbb{R}^n)}$. The solution is unique in the ball $B^{X_\infty}_{C\varepsilon}(0) := \{u \in X_\infty \mid \|u\|_{X_\infty} \le C\varepsilon\}$. Moreover, there exist $R > 0$, $c > 0$ such that for every $k \in \mathbb{N}_0$ and multi-index $\gamma \in \mathbb{N}_0^n$, we have the estimate*

$$\sup_{x\in\mathbb{R}^n} \sup_{t>0} |(t^{\frac{1}{\alpha}} \nabla)^\gamma (t\partial_t)^k \nabla u(x,t)| \le c\|\nabla u_0\|_{L^\infty(\mathbb{R}^n)} R^{|\gamma|+k} (|\gamma|+k)!. \tag{2.18}$$

Furthermore, u depends analytically on u_0.

Note that even though the estimate resembles those coming from linear parabolic equations and is consistent with the *parabolic scaling*, it is highly nontrivial to establish for nonlinear equations. The fact that the estimates are expressed in terms of the Lipschitz norm of the initial data is particularly useful, as self-similar initial data is necessarily only Lipschitz. Furthermore, note that the following gradient bound for the solution ($\gamma = 0, k = 0$) implies that the smallness of the Lipschitz norm is preserved in time:

$$\|\nabla u(t)\|_{L^\infty(\mathbb{R}^n)} \leq C\|\nabla u_0\|_{L^\infty(\mathbb{R}^n)}. \tag{2.19}$$

This fact is crucial if we want to work in the graph setting because for surface diffusion, it has been shown by [14] that in general the graph property might not be preserved.

One of the most important features of these equations is their *scale-invariant* property. More precisely, for any positive constant λ, if we define $\Sigma_\lambda := \lambda^{-1}\Sigma$, then

$$\mathcal{H}_{\Sigma_\lambda} = \lambda\mathcal{H}_\Sigma, \quad \mathcal{K}_{\Sigma_\lambda} = \lambda^2\mathcal{K}_\Sigma, \quad \text{and} \quad \triangle_{\Sigma_\lambda} = \lambda^2\triangle_\Sigma.$$

In terms of equations, we have the following: let u be a mild solution to (2.16). If we define $u_\lambda(x,t) := \lambda^{-1}u(\lambda x, \lambda^\alpha t)$, where $\alpha = 2$ if $A = -\Delta$ and $\alpha = 4$ if $A = \Delta^2$, then u_λ solves the same PDE but with rescaled initial data, that is,

$$\begin{cases} \partial_t u_\lambda + Au_\lambda = N[u_\lambda], & (x,t) \in \mathbb{R}^n \times (0,\infty), \\ u_\lambda(x,0) = \lambda^{-1}u_0(\lambda x), & x \in \mathbb{R}^n. \end{cases} \tag{2.20}$$

Note that with $y = \lambda x$, we have $\nabla_x u_\lambda = \nabla_y u$, $\nabla_x^2 u_\lambda = \lambda\nabla_y^2 u$, and so forth. The powers of $\nabla^2 u$ in the nonlinear terms f_i are such that $f_i(u_\lambda) = \lambda^{3-i} f_i(u)$ for $i = 0, 1, 2$. Hence, $\nabla^i f_i(u_\lambda) = \lambda^3\nabla^i f_i(u)$. They indeed give the corresponding scale invariance with $\alpha = 4$ for SD and WF. For MCF, we only have the term $f_2(u) \sim \nabla^2 u$, corresponding to $\alpha = 2$.

The above naturally leads to the notion of self-similar solutions v which satisfy $v_\lambda(x,t) = v(x,t)$. Setting $t = 0$, the initial data necessarily has the property that $v(x,0) = \lambda^{-1}v(\lambda x, 0)$. Conversely, let v be the solution of (2.16) with self-similar initial data $v_0(x) = |x|\psi(\frac{x}{|x|})$ for some function $\psi : \mathbb{S}^{n-1} \to \mathbb{R}$ so that v_0 is indeed self-similar (i.e., $v_0(x) = \lambda^{-1}v_0(\lambda x)$). Since v_λ solves the same equation and initial data, by uniqueness of solutions, it holds that $v_\lambda(x,t) = v(x,t)$. Upon introducing $\Psi(y) = v(y,1)$, we then have

$$v(x,t) = v_{t^{-\frac{1}{\alpha}}}(x,t) = t^{\frac{1}{\alpha}}v(xt^{-\frac{1}{\alpha}}, 1) =: t^{\frac{1}{\alpha}}\Psi(xt^{-\frac{1}{\alpha}}). \tag{2.21}$$

The function Ψ is called a *self-similar profile* and it satisfies the following equation:

$$A\Psi(y) + \frac{1}{\alpha}\Psi(y) - \frac{1}{\alpha}y\cdot\nabla\Psi(y) = N(\Psi(y)).$$

The main objective of this paper is to study the stability of self-similar solutions under bounded (and small) perturbations of self-similar initial data. Our main results are given in Theorems 2.2 and 2.4. (Again, see (3.28) and (3.29) for the definition of the function space X_∞.)

Theorem 2.2. *There exists an* $\varepsilon > 0$ *such that if* $u(x,t) \in X_\infty$ *is a global mild solution to* (2.16) *with perturbed self-similar initial data* $u_0(x) = v_0(x) + p(x)$ *that satisfies* $\|p\|_{L^\infty(\mathbb{R}^n)} < \infty$ *and* $\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)}, \|\nabla p\|_{L^\infty(\mathbb{R}^n)} < \varepsilon$*, then for any compact subset* K *of* $\mathbb{R}^n$*, it holds that*

$$\lim_{t\to\infty} \left\| t^{-\frac{1}{\alpha}} u(t^{\frac{1}{\alpha}} x, t) - \Psi(x) \right\|_{C^k(K)} = 0, \quad \forall k \in \mathbb{N}^+. \tag{2.22}$$

By Theorem 2.1, u automatically enjoys analytic regularity estimate (2.18).

The next example demonstrates the validity of Theorem 2.2.

Example 2.3. Consider a spatial shift of self-similar initial data $v_0(x)$ by a vector $a \in \mathbb{R}^n$, that is, $u_0(x) = v_0(x-a)$. In this case, $p(x) = u_0(x) - v_0(x) = v_0(x-a) - v_0(x)$, which satisfies the condition of Theorem 2.2. In fact,

$$\begin{aligned} &\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)} \\ &\quad \le \|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla (v_0(\cdot - a) - v_0(\cdot))\|_{L^\infty(\mathbb{R}^n)} \\ &\quad \le 3\|\nabla v_0\|_{\infty(\mathbb{R}^n)} < 3\varepsilon \end{aligned}$$

and

$$\begin{aligned} \|p(x)\|_{L^\infty(\mathbb{R}^n)} &= \|v_0(x-a) - v_0(x)\|_{L^\infty(\mathbb{R}^n)} \\ &\le \|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} |a| < \infty. \end{aligned}$$

By the uniqueness of mild solutions to (2.16), we have

$$u(x,t) = v(x-a,t) = t^{\frac{1}{\alpha}} \Psi((x-a)t^{-\frac{1}{\alpha}}).$$

Then, Theorem 2.2 gives that

$$\begin{aligned} \lim_{t\to\infty} \left\| t^{-\frac{1}{\alpha}} u(t^{\frac{1}{\alpha}} x, t) - \Psi(x) \right\|_{C^k(K)} &= \lim_{t\to\infty} \left\| \Psi(x - a t^{-\frac{1}{\alpha}}) - \Psi(x) \right\|_{C^k(K)} \\ &= 0, \quad \forall k \in \mathbb{N}^+. \end{aligned}$$

We also have the following result on the global convergence under perturbation with spatial decay:

Theorem 2.4 (Global stability with spatial decay). *There exists an* $\varepsilon > 0$ *such that if* $u(x,t) \in X_\infty$ *is a global mild solution to* (2.16) *with perturbed self-similar initial data* $u_0(x) = v_0(x) + p(x)$ *that satisfies* $\|p\|_{L^\infty(\mathbb{R}^n)} < \infty$ *as well as* $\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|(1 + |x|^\beta)\nabla p\|_{L^\infty(\mathbb{R}^n)} < \varepsilon$ *for some* $\beta > 0$*, then we have*

$$\lim_{t\to\infty} \left\| t^{-\frac{1}{\alpha}} u(t^{\frac{1}{\alpha}} x, t) - \Psi(x) \right\|_{C^1(\mathbb{R}^n)} = 0. \tag{2.23}$$

Again, by Theorem 2.1, u automatically enjoys analytic regularity estimate (2.18).

Remark 2.5. It seems possible to also prove higher-order global-in-space convergence results. The main technical step is to generalize Lemmas 4.4 and 4.7 to higher-order estimates. Paper [24] uses the analytic Banach fixed point theorem to obtain higher-order regularity. For reasons of conciseness and space, we omit this step in this paper.

For the rest of this section, we outline the strategy of the proof of Theorem 2.2. Such an approach is also described in [18, Chapter 1] by M.-H. Giga, Y. Giga, and J. Saal. First, note that upon setting $\lambda = t^{\frac{1}{\alpha}}$, we have $u_\lambda(x,1) = u_{t^{\frac{1}{\alpha}}}(x,1) = t^{-\frac{1}{\alpha}} u(xt^{\frac{1}{\alpha}}, t)$. Hence, (2.22) is equivalent to

$$\begin{aligned}\lim_{\lambda\to\infty} \|u_\lambda(x,1) - v_\lambda(x,1)\|_{C^k(K)} &= \lim_{\lambda\to\infty} \|u_\lambda(x,1) - v(x,1)\|_{C^k(K)} \\ &= 0, \quad \forall k \in \mathbb{N}^+. \end{aligned} \tag{2.24}$$

Thus, all we need is to estimate at time $t = 1$ the difference between the two solutions u_λ and $v_\lambda \equiv v$. Now, let $\Phi_\lambda := u_\lambda - v$. Then, it satisfies

$$\Phi_\lambda(x,t) = e^{-At} p_\lambda(x) + \int_0^t e^{-(t-s)A}(N[v + \Phi_\lambda] - N[v])(x,s)ds, \tag{2.25}$$

where we have used the fact that the difference between the two initial data is given by $u_\lambda(x,0) - v(x,0) = \frac{1}{\lambda} p(\lambda x) := p_\lambda(x)$.

Next, the following estimate from Theorem 2.1 is applicable to both u_λ and v:

$$|\nabla^\gamma \partial_t^k \nabla u(x,t)| \le C t^{-\left(\frac{|\gamma|}{\alpha} + k\right)} \|\nabla u_0\|_{L^\infty(\mathbb{R}^n)}. \tag{2.26}$$

Putting (2.26) and (2.25) together, we can apply the Arzelà–Ascoli compactness theorem to show that there is a subsequence $\{\Phi_{\lambda_k}\}$, $\lambda_k \to \infty$ and $\Phi_\infty \in C^\infty(\mathbb{R}^n \times (0,1])$ such that the following statements hold:

(i) (Convergence.) For any compact subset K of $\mathbb{R}^n$,

$$\lim_{\lambda_k\to\infty} \left\|\Phi_{\lambda_k}(x,1) - \Phi_\infty(x,1)\right\|_{C^k(K)} = 0, \quad \forall k \in \mathbb{N}. \tag{2.27}$$

(ii) (Regularity.) For any $t \in (0,1]$,

$$\left\|\nabla^\gamma \partial_t^k \nabla \Phi_\infty(\cdot,t)\right\|_{L^\infty(\mathbb{R}^n)} \le C t^{-\left(\frac{|\gamma|}{\alpha} + k\right)} \left(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\right). \tag{2.28}$$

(iii) (Integral equation.) The function $\Phi_\infty(x,t)$ solves the following integral equation:

$$\Phi_\infty(x,t) = \int_0^t e^{-(t-s)A}(N[v + \Phi_\infty] - N[v])(x,s)ds, \quad (x,t) \in \mathbb{R}^n \times (0,1]. \tag{2.29}$$

As the last step, we conclude the proof of (2.24) by showing that every solution Φ_∞ of (2.29) satisfying the property $\|\nabla\Phi_\infty\|_{L^\infty(\mathbb{R}^n)} \ll 1$ and regularity estimate (2.28) must be equal to 0.

We would like to emphasize that the above strategy is very simple and robust. See again [18] for a general exposition of this strategy. Despite the fact that the results are restricted to the graph setting, it is applicable to all the geometric evolutions under consideration here. Another advantage is that neither maximum nor comparison principle is used in the current approach; see, for example, the results for MCF [7, 9, 39] that do rely on such principle.

It is worth mentioning that our strategy is also applicable to *graphical MCF* in arbitrary codimension to obtain the stability of self-similar solutions. In the case of Euclidean space, the dynamical equation is given by a system of equations for a function $f : \mathbb{R}^n \times [0, \infty) \to \mathbb{R}^m$:

$$\frac{\partial f^\alpha}{\partial t} = g^{ij} \frac{\partial^2 f^\alpha}{\partial x_i \partial x_j}, \quad \alpha = 1, \dots m, \tag{2.30}$$

where $g^{ij} = (g_{ij})^{-1}$ and

$$g_{ij} = \delta_{ij} + \sum_{\beta=1}^{m} \frac{\partial f^\beta}{\partial x_i} \frac{\partial f^\beta}{\partial x_j}.$$

(Equation (2.30) becomes (2.6) when $m = 1$; this is the hypersurface case.) Note that (2.30) also satisfies the structural form in (2.9). Compared with the celebrated result of Ecker–Huisken [11] for MCF of a hypersurface given by an entire graph on $\mathbb{R}^n$, in order to have existence and convergence of a global-in-time solution for graphical MCF, the initial map is required to satisfy some *area decreasing* property. We refer to [27, 28, 32, 33, 38, 42–44] for a sample of results. For more comprehensive surveys, please see [31, 37, 45].

As a last remark before presenting the proof, note that WF has one more term than SD, namely, $f_0[u]$. Thus, in the current work, we will only consider MCF and WF for simplicity.

3. Stability result – local version

In this section, we will prove Theorem 2.2. As outlined above, we will first establish uniform estimates and compactness of Φ_λ. In all of the following results, we are working in the regime of the *small Lipschitz norm*; more precisely, there exists an $\varepsilon \ll 1$ such that $\|\nabla u_0\|_{L^\infty(\mathbb{R}^n)}, \|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} \ll \varepsilon$:

3.1. The mean curvature flow (MCF) case

In this case, we have $A = -\Delta$, $\alpha = 2$. Thus, equation (2.25) for Φ_λ becomes

$$\Phi_\lambda(x,t) = e^{t\Delta} p_\lambda(x) + \int_0^t e^{(t-s)\Delta} (N[u_\lambda] - N[v])(x,s) ds. \tag{3.1}$$

The nonlinear term $N[u]$ can be estimated as

$$|N[u]| = (1+|\nabla u|^2)^{-1}\nabla u \star \nabla u \star \nabla^2 u \lesssim |\nabla u|^2 |\nabla^2 u| \lesssim |\nabla^2 u|. \tag{3.2}$$

We also recall the heat kernel and its associated semigroup:

$$h(x,t) := \frac{1}{(4\pi t)^{\frac{n}{2}}} \exp\Big(-\frac{|x|^2}{4t}\Big) \quad \text{and} \quad e^{\Delta t} f(x) := \int_{\mathbb{R}^n} h(x-y,t) f(y) dy. \tag{3.3}$$

3.1.1. Uniform estimates and compactness for $\boldsymbol{\Phi_\lambda}$. We first note several useful facts. By the L^1 bound of the heat kernel, we get

$$\sup_{\lambda>1} \sup_{t\geq 0} \|e^{\Delta t} p_\lambda(x)\|_{L^\infty(\mathbb{R}^n)} \leq \sup_{\lambda>1} \|p_\lambda(x)\|_{L^\infty(\mathbb{R}^n)} < \infty. \tag{3.4}$$

Furthermore, the Lipschitz norm is invariant under the rescaling:

$$\|\nabla p_\lambda\|_{L^\infty(\mathbb{R}^n)} = \|\nabla p\|_{L^\infty(\mathbb{R}^n)}. \tag{3.5}$$

From regularity estimate (2.26), we have

$$\begin{aligned} \sup_{\lambda>1} \big\|\partial_t^k \nabla^\gamma \nabla u_\lambda(\cdot,t)\big\|_{L^\infty(\mathbb{R}^n)} &\lesssim t^{-\frac{|\gamma|}{2}-k} \sup_{\lambda>1} \|\nabla(v_0+p_\lambda)\|_{L^\infty(\mathbb{R}^n)} \\ &\lesssim t^{-\frac{|\gamma|}{2}-k} \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big), \end{aligned} \tag{3.6}$$

and similarly for $v_\lambda = v$,

$$\sup_{\lambda>1} \big\|\partial_t^k \nabla^\gamma \nabla v_\lambda(\cdot,t)\big\|_{L^\infty(\mathbb{R}^n)} \lesssim t^{-\frac{|\gamma|}{2}-k} \|\nabla v_0\|_{L^\infty(\mathbb{R}^n)}. \tag{3.7}$$

Now we estimate

$$\begin{aligned} &\sup_{\lambda>1} \|\Phi_\lambda(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} \\ &\quad\leq \sup_{\lambda>1} \|e^{\Delta t} p_\lambda(\cdot)\|_{L^\infty(\mathbb{R}^n)} + \sup_{\lambda>1} \Big\| \int_0^t e^{-(t-s)\Delta}(N[u_\lambda]-N[v])(\cdot,s) ds \Big\|_{L^\infty(\mathbb{R}^n)} \\ &\quad\lesssim \sup_{\lambda>1} \|e^{\Delta t} p_\lambda(\cdot)\|_{L^\infty(\mathbb{R}^n)} \\ &\qquad + \sup_{\lambda>1} \Big\| \int_0^t \int_{\mathbb{R}^n} |h(\cdot-y,t-s)| \big(|N[u_\lambda]| + |N[v]|\big)(y,s) dy ds \Big\|_{L^\infty(\mathbb{R}^n)} \\ &\quad\lesssim \sup_{\lambda>1} \|p_\lambda\|_{L^\infty(\mathbb{R}^n)} \\ &\qquad + \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big) \Big\| \int_0^t \int_{\mathbb{R}^n} h(\cdot-y,t-s) s^{-\frac{1}{2}} dy ds \Big\|_{L^\infty(\mathbb{R}^n)} \\ &\quad\lesssim \sup_{\lambda>1} \|p_\lambda\|_{L^\infty(\mathbb{R}^n)} + \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big) \\ &\quad< \infty. \end{aligned} \tag{3.8}$$

In the above, we have used the estimate

$$\|N[u_\lambda(\cdot,s)]\|_{L^\infty(\mathbb{R}^n)} \lesssim \|\nabla u_\lambda(\cdot,s)\|^2_{L^\infty(\mathbb{R}^n)} \|\nabla^2 u_\lambda(\cdot,s)\|_{L^\infty(\mathbb{R}^n)} \lesssim s^{-\frac{1}{2}}. \tag{3.9}$$

By higher-order regularity estimates (3.6) and (3.7), we have for any $k \in \mathbb{N}, \gamma \in \mathbb{N}^n$,

$$\begin{aligned}\sup_{\lambda>1} \|\nabla^\gamma \partial_t^k \nabla \Phi_\lambda(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} &\lesssim \sup_{\lambda>1} \|\nabla^\gamma \partial_t^k \nabla u_\lambda(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} + \|\nabla^\gamma \partial_t^k \nabla v(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} \\ &\lesssim \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big) t^{-\frac{|\gamma|}{2}-k}.\end{aligned}$$

With the above uniform estimates for Φ_λ, we can apply the Arzelà–Ascoli theorem to extract a subsequence $\{\Phi_{\lambda_k}\}$ and $\Phi_\infty(x,t) \in C^\infty(\mathbb{R}^n \times (0,1])$ such that for any $\delta > 0$, compact subset $K \subset \mathbb{R}^n$, and $k \in \mathbb{N}$, we have

$$\lim_{\lambda_k \to \infty} \sup_{\delta \le t \le 1} \|\Phi_{\lambda_k} - \Phi_\infty\|_{C^k(K)} = 0. \tag{3.10}$$

Then, (2.27) and (2.28) follow.

3.1.2. Equation for Φ_∞. Here we verify (2.29) by passing the limit $\lambda_k \to \infty$ in (3.1). First note that

$$\lim_{\lambda\to\infty} \sup_{t\ge 0} \|e^{\Delta t} p_\lambda(\cdot)\|_{L^\infty(\mathbb{R}^n)} \le \lim_{\lambda\to\infty} \|p_\lambda\|_{L^\infty(\mathbb{R}^n)} = \lim_{\lambda\to\infty} \frac{1}{\lambda}\|p\|_{L^\infty(\mathbb{R}^n)} = 0. \tag{3.11}$$

Second, from (3.10), we know that for any $\delta > 0$ and any compact subset $K \subset \mathbb{R}^n$,

$$\lim_{\lambda_k\to\infty} \sup_{\delta\le t\le 1} \|N[v+\Phi_{\lambda_k}] - N[v+\Phi_\infty]\|_{C^k(K)} = 0. \tag{3.12}$$

Now note that

$$\begin{aligned}&\Big|\int_0^t e^{(t-s)\Delta}\big(N[\Phi_\lambda + v] - N[\Phi_\infty + v]\big)(x,s)ds\Big| \\ &\quad \le \int_0^t \int_{\mathbb{R}^n} h(t-s,x-y)\big[|N[\Phi_\lambda+v]| + |N[\Phi_\infty+v]|\big](y,s)\,dyds.\end{aligned}$$

By the formula of the heat kernel (see (3.3)) and the estimate for the nonlinear term given in (3.9), the integrand can be estimated as

$$h(t-s,x-y)\big[|N[\Phi_\lambda+v]| + |N[\Phi_\infty+v]|\big](y,s) \lesssim (t-s)^{-\frac{n}{2}} \exp\Big(-\frac{|x-y|^2}{4(t-s)}\Big) s^{-\frac{1}{2}},$$

which is integrable:

$$\int_0^t \int_{\mathbb{R}^n} (t-s)^{-\frac{n}{2}} \exp\Big(-\frac{|x-y|^2}{4(t-s)}\Big) s^{-\frac{1}{2}}\,dy\,ds \lesssim t^{\frac{1}{2}}.$$

Hence, (2.29) follows by the Lebesgue dominated convergence theorem.

3.2. The Willmore flow (WF) case

In this case, we have $A = \Delta^2, \alpha = 4$, and

$$N[u] = f_0[u] + \nabla_i f_1^i[u] + \nabla_{ij}^2 f_2^{ij}[u].$$

First, we introduce the heat kernel of biharmonic operator $b(x,t)$:

$$b(x,t) = t^{-\frac{n}{4}} g\Big(\frac{x}{t^{\frac{1}{4}}}\Big), \quad \text{where} \quad g(\xi) = (2\pi)^{-\frac{n}{2}} \int_{\mathbb{R}^n} e^{i\xi\cdot k - |k|^4} dk, \quad \xi \in \mathbb{R}^n.$$

It satisfies the following decaying estimates (see [17, Chapter 9, Theorem 7], [24]) which play a very important role in this paper:

$$|b(x,t)| \lesssim t^{-\frac{n}{4}} \exp\Big(-C\frac{|x|^{\frac{4}{3}}}{t^{\frac{1}{3}}}\Big), \tag{3.13}$$

$$|\nabla^k b(x,t)| \lesssim t^{-\frac{n+k}{4}} \exp\Big(-C_k\frac{|x|^{\frac{4}{3}}}{t^{\frac{1}{3}}}\Big), \quad \forall k \geq 1, \tag{3.14}$$

The integral equation for mild solutions $u(x,t)$ to (2.11) now reads

$$\begin{aligned} u(x,t) = &\int_{\mathbb{R}^n} b(x-y,t)u_0(y)dy + \int_0^t \int_{\mathbb{R}^n} b(x-y,t-s) f_0[u](y,s)dyds \\ &- \int_0^t \int_{\mathbb{R}^n} \nabla_i b(x-y,t-s) f_1^i[u](y,s)dyds \\ &+ \int_0^t \int_{\mathbb{R}^n} \nabla_{ij}^2 b(x-y,t-s) f_2^{ij}[u](y,s)dyds. \end{aligned} \tag{3.15}$$

Given the uniform bound for $\|\nabla u\|_{L^\infty(\mathbb{R}^n)} \lesssim 1$, we note here the estimates for the nonlinear structures:

$$|f_0[u]| \lesssim |\nabla^2 u|^3 \lesssim t^{-\frac{3}{4}}, \quad |f_1[u]| \lesssim |\nabla^2 u|^2 \lesssim t^{-\frac{2}{4}}, \quad |f_2[u]| \lesssim |\nabla^2 u| \lesssim t^{-\frac{1}{4}}. \tag{3.16}$$

Note also that in order to take advantage of the kernel decay, we perform integration by parts to eliminate the derivatives on f_1 and f_2. With this, we use the following L^1 bound for b:

$$\|\nabla^k b(\cdot,t)\|_{L^1(\mathbb{R}^n)} \lesssim t^{-\frac{k}{4}} \quad \text{for } k = 0, 1, 2. \tag{3.17}$$

3.2.1. Uniform estimates and convergence for $\boldsymbol{\Phi_\lambda}$. Using the estimates for b, we first establish an L^∞ bound for Φ_λ. For $e^{-\Delta^2 t} p_\lambda$, we have

$$\begin{aligned} \sup_{\lambda>1}\sup_{t\geq 0} \big\| e^{-\Delta^2 t} p_\lambda \big\|_{L^\infty} &= \sup_{\lambda>1}\sup_{t\geq 0} \Big\| \int_{\mathbb{R}^n} b(\cdot - y,t)p_\lambda(y)dy \Big\|_{L^\infty} \\ &\leq \sup_{\lambda>1}\sup_{t\geq 0} \|p_\lambda\|_{L^\infty} \|b(\cdot,t)\|_{L^1} \\ &\lesssim \sup_{\lambda>1} \|p_\lambda\|_{L^\infty} < \infty. \end{aligned} \tag{3.18}$$

From regularity estimate (2.26), we have

$$\begin{aligned}\sup_{\lambda>1}\left\|\partial_t^k\nabla^\gamma\nabla u_\lambda(\cdot,t)\right\|_{L^\infty} &\lesssim t^{-\frac{|\gamma|}{4}-k}\sup_{\lambda>1}\|\nabla(v_0+p_\lambda)\|_{L^\infty}\\ &\lesssim t^{-\frac{|\gamma|}{4}-k}\big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)}+\sup_{\lambda>1}\|\nabla p_\lambda\|_{L^\infty(\mathbb{R}^n)}\big)\\ &\lesssim t^{-\frac{|\gamma|}{4}-k}\big(\|v_0\|_{L^\infty(\mathbb{R}^n)}+\|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big) \qquad (3.19)\end{aligned}$$

and similarly for $v_\lambda = v$,

$$\sup_{\lambda>1}\left\|\partial_t^k\nabla^\gamma\nabla v_\lambda(\cdot,t)\right\|_{L^\infty}\lesssim t^{-\frac{|\gamma|}{4}-k}\|v_0\|_{L^\infty(\mathbb{R}^n)}. \qquad (3.20)$$

For the L^∞ estimate for Φ_λ, we combine (3.13), (3.14), (3.19), and (2.15) to give

$$\begin{aligned}\sup_{\lambda>1}\|\Phi_\lambda(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} &\le \sup_{\lambda>1}\left\|e^{-\Delta^2 t}p_\lambda(\cdot)\right\|_{L^\infty(\mathbb{R}^n)}\\ &\quad+\sup_{\lambda>1}\left\|\int_0^t\int_{\mathbb{R}^n} b(\cdot-y,t-s)\big(f_0[u_\lambda]-f_0[v]\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\quad+\sup_{\lambda>1}\left\|\int_0^t\int_{\mathbb{R}^n} \nabla_i b(\cdot-y,t-s)\big(f_1^i[u_\lambda]-f_1^i[v]\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\quad+\sup_{\lambda>1}\left\|\int_0^t\int_{\mathbb{R}^n} \nabla_{ij} b(\cdot-y,t-s)\big(f_2^{ij}[u_\lambda]-f_2^{ij}[v]\big)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\le \sup_{\lambda>1}\left\|e^{-\Delta^2 t}p_\lambda(\cdot)\right\|_{L^\infty(\mathbb{R}^n)}\\ &\quad+\sup_{\lambda>1}\left\|\int_0^t\int_{\mathbb{R}^n} |b(\cdot-y,t-s)|\big(|f_0[u_\lambda]|+|f_0[v]|\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\quad+\sup_{\lambda>1}\left\|\int_0^t\int_{\mathbb{R}^n} |\nabla_i b(\cdot-y,t-s)|\big(|f_1^i[u_\lambda]|+|f_1^i[v]|\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\quad+\sup_{\lambda>1}\left\|\int_0^t\int_{\mathbb{R}^n} |\nabla_{ij} b(\cdot-y,t-s)|\big(|f_2^{ij}[u_\lambda]|+|f_2^{ij}[v]|\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}.\end{aligned}$$

Now we make use of the structure for nonlinear terms (3.16) together with kernel and regularity estimates (3.17), (3.19), and (3.20) to obtain

$$\begin{aligned}&\left\|\int_0^t\int_{\mathbb{R}^n} |b(\cdot-y,t-s)|\big(|f_0[u_\lambda]|+|f_0[v]|\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\qquad\lesssim\int_0^t\int_{\mathbb{R}^n}|b(y,t-s)|s^{-\frac{3}{4}}\,dy\,ds\lesssim\int_0^t s^{-\frac{3}{4}}\,ds\lesssim t^{\frac{1}{4}}; \qquad (3.21)\\ &\left\|\int_0^t\int_{\mathbb{R}^n} |\nabla_i b(\cdot-y,t-s)|\big(|f_1^i[u_\lambda]|+|f_1^i[v]|\big)(y,s)dyds\right\|_{L^\infty(\mathbb{R}^n)}\\ &\qquad\lesssim\int_0^t\int_{\mathbb{R}^n}|\nabla b(y,t-s)|s^{-\frac{2}{4}}\,dy\,ds\\ &\qquad\lesssim\int_0^t(t-s)^{-\frac{1}{4}}s^{-\frac{2}{4}}\,ds=t^{\frac{1}{4}}\int_0^1(1-s)^{-\frac{1}{4}}s^{-\frac{2}{4}}\,ds\lesssim t^{\frac{1}{4}}; \qquad (3.22)\end{aligned}$$

$$\left\| \int_0^t \int_{\mathbb{R}^n} |\nabla_{ij} b(\cdot - y, t-s)| \big(|f_2^{ij}[u_\lambda]| + |f_2^{ij}[v]| \big)(y,s) dy ds \right\|_{L^\infty(\mathbb{R}^n)}$$
$$\lesssim \int_0^t \int_{\mathbb{R}^n} |\nabla^2 b(y, t-s)| s^{-\frac{1}{4}} \, dy \, ds$$
$$\lesssim \int_0^t (t-s)^{-\frac{2}{4}} s^{-\frac{1}{4}} \, ds = t^{\frac{1}{4}} \int_0^1 (1-s)^{-\frac{2}{4}} s^{-\frac{1}{4}} \, ds \lesssim t^{\frac{1}{4}}. \tag{3.23}$$

Hence, we have

$$\sup_{\lambda > 1} \|\Phi_\lambda(\cdot, t)\|_{L^\infty(\mathbb{R}^n)} \lesssim \sup_{\lambda > 1} \|p_\lambda\|_{L^\infty} + t^{\frac{1}{4}} \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)} \big) < \infty.$$

For higher-order regularity estimates, by (3.19), we have

$$\sup_{\lambda > 1} \left\| \nabla^\gamma \partial_t^k \nabla \Phi_\lambda(\cdot, t) \right\|_{L^\infty(\mathbb{R}^n)} \le C t^{-\frac{|\gamma|}{4} - k} \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)} \big). \tag{3.24}$$

As in the MCF case, we apply the Arzelà–Ascoli theorem to extract a subsequence $\{\Phi_{\lambda_k}\}$ and $\Phi_\infty(x,t) \in C^\infty(\mathbb{R}^n \times (0,1])$ such that for any $\delta > 0$ and any compact subset K of $\mathbb{R}^n$,

$$\lim_{\lambda_k \to \infty} \sup_{\delta \le t \le 1} \|\Phi_{\lambda_k}(\cdot, t) - \Phi_\infty(\cdot, t)\|_{C^k(K)} = 0, \quad \forall k \in \mathbb{N}^+, \tag{3.25}$$

and Φ_∞ satisfies regularity estimate (2.28).

3.2.2. Equation for Φ_∞. Here we check that Φ_∞ satisfies (2.29). The strategy is similar to the MCF case.

Recall that Φ_λ satisfies the following identity:

$$\Phi_\lambda(x,t) = e^{-\Delta^2 t} p_\lambda(x) + \int_0^t \int_{\mathbb{R}^n} b(x-y, t-s) \big(f_0[\Phi_\lambda + v] - f_0[v] \big)(y,s) dy ds$$
$$- \int_0^t \int_{\mathbb{R}^n} \nabla_i b(x-y, t-s) \big(f_1^i[\Phi_\lambda + v] - f_1^i[v] \big)(y,s) dy ds$$
$$+ \int_0 \int_{\mathbb{R}^n} \nabla_{ij}^2 b(x-y, t-s) \big(f_2^{ij}[\Phi_\lambda + v] - f_2^{ij}[v] \big) dy ds. \tag{3.26}$$

First, by the L^1 bound of $b(\cdot, t)$, similar to (3.11), we have

$$\lim_{\lambda \to \infty} \|e^{-\Delta^2 t} p_\lambda(\cdot)\|_{L^\infty(\mathbb{R}^n)} \lesssim \lim_{\lambda \to \infty} \|p_\lambda\|_{L^\infty(\mathbb{R}^n)} \lesssim \lim_{\lambda \to \infty} \frac{1}{\lambda} \|p\|_{L^\infty(\mathbb{R}^n)} = 0. \tag{3.27}$$

Second, similar to the previous computations—in particular, the derivations of estimates (3.21), (3.22), and (3.23)—the integrals of the nonlinear terms are all bounded by integrands that are integrable with bounds independent of λ. Hence, (2.29) follows from the Lebesgue dominated convergence theorem. We emphasize here again the crucial use of estimates (3.16) for the nonlinear terms and L^1 bounds (3.17) for the derivatives of the biharmonic heat kernel.

3.3. Proof of $\Phi_\infty = 0$

In this section, we will show that integral equation (2.29) only admits the zero solution among the class of functions with small Lipschitz norm. This follows from a fixed point-type argument.

Motivated by the translation and scaling invariance of the equation, the following function space was introduced in [24] for $0 < T \le \infty$:

(1) For MCF with $\alpha = 2$,

$$X_T := \Big\{ f(x,t) : \mathbb{R}^n \times (0,T) \to \mathbb{R} \mid \|f\|_{X_T} := \sup_{0<t<T} \|\nabla f(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} + \sup_{x\in\mathbb{R}^n} \sup_{0<R^2<T} R^{\frac{2}{n+4}} \|\nabla^2 f\|_{L^{n+4}(B_R(x)\times(R^2/2,R^2))} < \infty \Big\}. \tag{3.28}$$

(2) For WF with $\alpha = 4$,

$$X_T := \Big\{ f(x,t) : \mathbb{R}^n \times (0,T) \to \mathbb{R} \mid \|f\|_{X_T} = \sup_{0<t<T} \|\nabla f(\cdot,t)\|_{L^\infty(\mathbb{R}^n)} + \sup_{x\in\mathbb{R}^n} \sup_{0<R^4<T} R^{\frac{2}{n+6}} \|\nabla^2 f\|_{L^{n+6}(B_R(x)\times(R^4/2,R^4))} < \infty \Big\}. \tag{3.29}$$

Note that the above norms are scale invariant:

$$\|f_\lambda\|_{X_T} = \|f\|_{X_{\lambda^\alpha T}} \quad \text{and} \quad \|f_\lambda\|_{X_\infty} = \|f\|_{X_\infty}.$$

We then have the following estimate:

Lemma 3.1 ([24, Lemmas 3.10 and 5.2]). *For any $0 < T \le \infty$ and $0 < \delta < 1$ there exists $C(\delta) > 0$ such that for every $g_1, g_2 \in B_\delta^{X_T}(0) := \{g \in X_T \mid \|g\|_{X_T} \le \delta\}$, we have*

$$\Big\| \int_0^T e^{-(T-s)A} N[g_1](x,s)ds - \int_0^T e^{-(T-s)A} N[g_2](x,s)ds \Big\|_{X_T} \le C(\delta)\big(\|g_1\|_{X_T} + \|g_2\|_{X_T}\big)\|g_1 - g_2\|_{X_T}. \tag{3.30}$$

The above is established through the linearized estimate

$$\Big\| \int_0^T e^{-(T-s)A} g\, ds \Big\|_{X_T} \le \|g\|_{Y_T}$$

for some appropriate spatial-temporal function space Y_T; see [24, Lemmas 3.11 and 5.3]. We will in fact present the proof of the above result in the setting of weighted function spaces X_T^β and Y_T^β; see Lemmas 4.4 and 4.7.

We apply the above lemma with $T = 1$, $g_1 = \Phi_\infty + v$, and $g_2 = v$. Suppose we can show that $\|g_1\|_{X_T}, \|g_2\|_{X_T} \ll 1$; then, we would have

$$\|\Phi_\infty\|_{X_T} = \Big\| \int_0^T e^{-(T-s)A} (N[\Phi_\infty + v] - N[v])(x,s)ds \Big\|_{X_T} \ll \|\Phi_\infty\|_{X_T},$$

which implies $\|\Phi_\infty\|_{X_T} = 0$. Hence, $\nabla\Phi_\infty \equiv 0$, which leads to $N[\Phi_\infty + v] = N[v]$, as $N(\cdot)$ only involves the derivatives of Φ_∞. From (2.29), we conclude that $\Phi_\infty \equiv 0$.

Hence, we are led to compute the X_T norm of g_1 and g_2 under the regularity estimates given by (2.18) and (2.28).

For MCF, we have

$$\begin{aligned}
&\|\Phi_\infty\|_{X_T} + \|v\|_{X_T}\\
&\quad\lesssim \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big)\\
&\qquad\cdot\Big(1 + \sup_{0<R^2<T} R^{\frac{2}{n+4}}\Big(\int_{B_R(x)\times(R^2/2,R^2)} (t^{-\frac{1}{2}})^{n+4}\,dt\,dy\Big)^{\frac{1}{n+4}}\Big)\\
&\quad\lesssim (\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)})\Big(1 + \sup_{0<R^2<T} R^{\frac{2}{n+4}}\Big(R^n\int_{R^2/2}^{R^2} t^{-\frac{n+4}{2}}\,dt\Big)^{\frac{1}{n+4}}\Big)\\
&\quad\lesssim v(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)})\Big(1 + \sup_{0<R^2<T} R^{\frac{2}{n+4}}\big(R^n R^{-n-2}\big)^{\frac{1}{n+4}}\Big)\\
&\quad\lesssim \|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}.
\end{aligned}$$

For WF, we have,

$$\begin{aligned}
&\|\Phi_\infty\|_{X_T} + \|v\|_{X_T}\\
&\quad\lesssim \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big)\\
&\qquad\cdot\Big(1 + \sup_{0<R^4<T} R^{\frac{2}{n+6}}\Big(\int_{B_R(x)\times(R^4/2,R^4)} (t^{-\frac{1}{4}})^{n+6}\,dt\,dy\Big)^{\frac{1}{n+6}}\Big)\\
&\quad\lesssim \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big)\Big(1 + \sup_{0<R^4<T} R^{\frac{2}{n+6}}\Big(R^n\int_{R^4/2}^{R^4} t^{-\frac{n+6}{4}}\,dt\Big)^{\frac{1}{n+6}}\Big)\\
&\quad\lesssim \big(\|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}\big)\Big(1 + \sup_{0<R^4<T} R^{\frac{2}{n+6}}\big(R^n R^{-n-2}\big)^{\frac{1}{n+6}}\Big)\\
&\quad\lesssim \|\nabla v_0\|_{L^\infty(\mathbb{R}^n)} + \|\nabla p\|_{L^\infty(\mathbb{R}^n)}.
\end{aligned}$$

The above show that in order to obtain the desired result, we just need to take the Lipschitz norms of v_0 and p to be sufficiently small, which is indeed assumed to be the case under the current setting.

4. Equi-decay and global uniform convergence

Here we will tackle Theorem 2.4. In essence, if the gradient of the initial perturbation is assumed to have some spatial decay, then we can obtain a global-in-space convergence result. The idea is to establish the equi-decay property of $\{\Phi_\lambda\}_{\lambda>1}$ via a contraction property of the nonlinear operators in some weighted spaces. For convenience, we recall here the weighted Lipschitz seminorm used in Theorem 2.4:

$$[p]_\beta := \|(1 + |x|^\beta)\nabla p(x)\|_{L^\infty(\mathbb{R}^n)}. \tag{4.1}$$

4.1. The mean curvature flow (MCF) case

For the mean curvature flow case, we introduce the following function space which is the spatially weighted version of X_T:

Definition 4.1. For every $0 < T \leq \infty$, we define the function space X_T^β by

$$X_T^\beta = \{u \mid \|u\|_{X_T^\beta} := \sup_{0<t<T} \sup_{x\in\mathbb{R}^n} (1+|x|^\beta)|\nabla u(t,x)| + \sup_{x\in\mathbb{R}^n} \sup_{0<R^2<T} (1+|x|^\beta) R^{\frac{2}{n+4}} \|\nabla^2 u\|_{L^{n+4}(Q_R(x))} < \infty\}, \quad (4.2)$$

where

$$Q_R(x) := B_R(x) \times (R^2/2, R^2).$$

Then, we have the following linear estimate:

Lemma 4.2. *For $k \geq 0$ and $0 < t < T$,*

$$\left\| t^{\frac{k}{2}} \nabla^k e^{t\Delta} p(x) \right\|_{X_T^\beta} \lesssim [p]_\beta. \quad (4.3)$$

For the analysis of the nonlinear part, we introduce the weighted function spaces Y_T^β as follows:

Definition 4.3. For every $0 < T \leq \infty$, we define the function space Y_T^β by

$$Y_T^\beta = \{g \mid \|g\|_{Y_T}^\beta = \sup_{x\in\mathbb{R}^n} \sup_{0<R^2<T} (1+|x|^\beta) R^{\frac{2}{n+4}} \|g\|_{L^{n+4}(Q_R(x))} < \infty\}.$$

Now we define

$$Sg(x,t) := \int_0^t \int_{\mathbb{R}^n} h(x-y, t-s) g(y,s) dy ds. \quad (4.4)$$

The following is the key technical estimate concerning S:

Lemma 4.4. *For $0 < t < T < \infty$,*

$$\sup_{0<t<T} \|(1+|x|^\beta) Sg(x,t)\|_{L^\infty(\mathbb{R}^n)} + \|Sg\|_{X_T^\beta} \lesssim \|g\|_{Y_T^\beta}.$$

With the above, we then have the following result for the nonlinear functional:

Lemma 4.5. *For every $0 < T < \infty$,*

$$\|N[u] - N[v]\|_{Y_T^\beta} \lesssim \left(\|u\|_{X_T}^2 + \|v\|_{X_T}^2\right) \|u - v\|_{X_T^\beta}. \quad (4.5)$$

In particular, there exist $\varepsilon > 0$ and $q < 1$ such that for all $[v_0] + [p]_\beta < \varepsilon$,

$$\left\| \int_0^t e^{(t-s)\Delta} (N[u] - N[v])(x,s) ds \right\|_{X_T^\beta} \leq q \|u - v\|_{X_T^\beta}. \quad (4.6)$$

We will give the proofs of Lemmas 4.2 and 4.5 here, but the proof of Lemma 4.4 is given in Appendix A due to its length and technical nature.

Proof of Lemma 4.2. It suffices to show that there exists a $C > 0$ depending only on T, n, β, and k such that if $[p]_\beta \leq 1$, then $\|e^{t\Delta}p(x)\|_{X_T^\beta} \leq C$. From the definition of $\|\cdot\|_{X_T^\beta}$, we need to estimate two terms.

First, consider

$$
\begin{aligned}
&\left|t^{\frac{k}{2}}\nabla^k\nabla e^{t\Delta}p(x)\right| \\
&= \frac{1}{(4\pi t)^{\frac{n}{2}}}\left|\int_{\mathbb{R}^n} t^{\frac{k}{2}}\nabla_x^k\nabla_x e^{-\frac{|x-y|^2}{4t}}p(y)dy\right| \\
&= \frac{1}{(4\pi t)^{\frac{n}{2}}}\left|\int_{\mathbb{R}^n} t^{\frac{k}{2}}\nabla_y^k e^{-\frac{|x-y|^2}{4t}}\nabla_y p(y)dy\right| \\
&\leq \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n}\left|t^{\frac{k}{2}}\nabla_y^k e^{-\frac{|x-y|^2}{4t}}\right||\nabla_y p(y)|dy \\
&\leq \frac{1}{(4\pi t)^{\frac{n}{2}}}\Big(\int_{\left\{y:|y-x|\leq\frac{\sqrt{t}}{2\sqrt{T}}|x|\right\}}+\int_{\left\{y:|y-x|\geq\frac{\sqrt{t}}{2\sqrt{T}}|x|\right\}}\left|\mathcal{P}_k\Big(\frac{x-y}{\sqrt{t}}\Big)\right|e^{-\frac{|x-y|^2}{4t}}\frac{1}{1+|y|^\beta}\Big)dy \\
&=: \mathrm{I}+\mathrm{II},
\end{aligned}
$$

where $\mathcal{P}_k$ is some polynomial of degree k. For I, $|y-x| \leq \frac{\sqrt{t}}{2\sqrt{T}}|x|$ implies that $|y| \geq \frac{|x|}{2}$ for $0 < t < T$. Hence,

$$
\frac{1}{1+|y|^\beta} \leq \frac{1}{1+|x/2|^\beta} = \frac{2^\beta}{2^\beta+|x|^\beta} \leq \frac{2^\beta}{1+|x|^\beta},
$$

so that

$$
\begin{aligned}
\mathrm{I} &\lesssim \frac{1}{1+|x|^\beta}\int_{\left\{y:|y-x|\leq\frac{\sqrt{t}}{2\sqrt{T}}|x|\right\}}\frac{1}{(4\pi t)^{\frac{n}{2}}}\left|\mathcal{P}_k\Big(\frac{x-y}{\sqrt{t}}\Big)\right|e^{-\frac{|x-y|^2}{4t}}dy \\
&\lesssim \frac{1}{1+|x|^\beta}\int_{\mathbb{R}^n}|\mathcal{P}_k(z)|e^{-|z|^2}dz \lesssim \frac{1}{1+|x|^\beta},
\end{aligned}
$$

while for II, when $|y-x| \geq \frac{\sqrt{t}}{2\sqrt{T}}|x|$, we have

$$
e^{-\frac{|x-y|^2}{4t}} = e^{-\frac{|x-y|^2}{8t}}e^{-\frac{|x-y|^2}{8t}} \leq e^{-\frac{|x|^2}{32T}}e^{-\frac{|x-y|^2}{8t}},
$$

so that

$$
\begin{aligned}
\mathrm{II} &\leq e^{-\frac{|x|^2}{32T}}\int_{\left\{y:|y-x|\geq\frac{\sqrt{t}}{2\sqrt{T}}|x|\right\}}\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x-y|^2}{8t}}\left|\mathcal{P}_k\Big(\frac{x-y}{2\sqrt{t}}\Big)\right|dy \\
&\lesssim e^{-\frac{|x|^2}{32T}}\int_{\mathbb{R}^n}|\mathcal{P}_k(z)|e^{-\frac{|z|^2}{2}}dz \lesssim e^{-\frac{|x|^2}{32T}} \lesssim \frac{1}{1+|x|^\beta}.
\end{aligned}
$$

Combining I and II, we have

$$\left|t^{\frac{k}{2}}\nabla^k\nabla e^{t\Delta}p(x)\right| \lesssim \frac{1}{1+|x|^\beta}. \tag{4.7}$$

Second, we estimate

$$\sup_{x\in\mathbb{R}^n}\sup_{0<R^2<T}(1+|x|^\beta)R^{\frac{2}{n+4}}\left\|t^{\frac{k}{2}}\nabla^k\nabla^2 e^{t\Delta}p(x)\right\|_{L^{n+4}(Q_R(x))}. \tag{4.8}$$

Note that

$$\begin{aligned}
&\|t^{\frac{k}{2}}\nabla^k\nabla^2 e^{t\Delta}p(x)\|^{n+4}_{L^{n+4}(Q_R(x))}\\
&\quad=\int_{R^2/2}^{R^2}\int_{B_R(x)}\left[t^{\frac{k}{2}}\nabla^k\nabla^2\int\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|y-z|^2}{4t}}p(z)\,dz\right]^{n+4}dy\,dt\\
&\quad=\int_{R^2/2}^{R^2}\int_{B_R(x)}\left[t^{\frac{k}{2}}\int\frac{1}{(4\pi t)^{\frac{n}{2}}}\nabla^{k+1}e^{-\frac{|y-z|^2}{4t}}\nabla p(z)\,dz\right]^{n+4}dy\,dt\\
&\quad\lesssim\int_{R^2/2}^{R^2}\int_{B_R(x)}\left[t^{\frac{k}{2}}\int\frac{1}{(4\pi t)^{\frac{n}{2}}}t^{-\frac{k+1}{2}}e^{-\frac{|y-z|^2}{4t}}\mathcal{P}_{k+1}\Big(\frac{y-z}{\sqrt{t}}\Big)\frac{1}{1+|z|^\beta}\,dz\right]^{n+4}dy\,dt\\
&\quad\lesssim\int_{R^2/2}^{R^2}\int_{B_R(x)}\left[t^{-\frac{1}{2}}\int\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|y-z|^2}{4t}}\mathcal{P}_{k+1}\Big(\frac{y-z}{\sqrt{t}}\Big)\frac{1}{1+|z|^\beta}\,dz\right]^{n+4}dy\,dt\\
&\quad\lesssim\int_{R^2/2}^{R^2}\int_{B_R(x)}\left[\frac{t^{-\frac{1}{2}}}{1+|y|^\beta}\right]^{n+4}dy\,dt\\
&\quad\lesssim\int_{R^2/2}^{R^2}t^{-\frac{1}{2}(n+4)}dt\int_{B_R(x)}\frac{1}{(1+|y|^\beta)^{n+4}}dy\\
&\quad\lesssim R^{-(n+2)}|B_R(x)|\frac{1}{(1+|x|^\beta)^{n+4}}\\
&\quad\lesssim\frac{R^{-2}}{(1+|x|^\beta)^{n+4}},
\end{aligned}$$

which leads to (4.8) $\lesssim 1$.

The above two parts combined give

$$\|t^{\frac{k}{2}}\nabla^k e^{t\Delta}p(x)\|_{X_T^\beta}\le C. \qquad \blacksquare$$

Proof of Lemma 4.5. Recall form (2.9) for the nonlinear term $N(u)$. First note that

$$|(1+|\nabla u|^2)^{-1}-(1+|\nabla v|^2)^{-1}|\le\frac{(|\nabla u|+|\nabla v|)|\nabla(u-v)|}{(1+|\nabla u|^2)(1+|\nabla v|^2)}.$$

Then, we have

$$\begin{aligned}
&|N[u]-N[v]|\\
&\quad=|(1+|\nabla u|^2)^{-1}\nabla u\star\nabla u\star\nabla^2u-(1+|\nabla v|^2)^{-1}\nabla v\star\nabla v\star\nabla^2 v|\\
&\quad\lesssim(|\nabla u|+|\nabla v|)(|\nabla^2 u|+|\nabla^2 v|)|\nabla(u-v)|+(|\nabla u|+|\nabla v|)^2|\nabla^2(u-v)|.
\end{aligned}$$

Then, estimate (4.5) follows from

$$\begin{aligned}
&\|N[u]-N[v]\|_{Y_T^\beta} \\
&\quad = \sup_{x\in\mathbb{R}^n}\sup_{0<R^2<T}(1+|x|^\beta)R^{\frac{2}{n+4}}\|N[u]-N[v]\|_{L^{n+4}(Q_R(x))} \\
&\quad \lesssim \sup_{0<t<T}(\|\nabla u\|_{L^\infty(\mathbb{R}^n)}+\|\nabla v\|_{L^\infty(\mathbb{R}^n)}) \\
&\qquad\qquad \cdot\Big(\sup_{x\in\mathbb{R}^n}\sup_{0<R^2<T}R^{\frac{2}{n+4}}\big(\|\nabla^2 u\|_{L^{n+4}(Q_R(x))}+\|\nabla^2 v\|_{L^{n+4}(Q_R(x))}\big)\Big) \\
&\qquad\qquad \cdot\Big(\sup_{0<t<T}\sup_{x\in\mathbb{R}^n}(1+|x|^\beta)|\nabla(u-v)|\Big) \\
&\qquad + \sup_{0<t<T}\big(\|\nabla u\|_{L^\infty(\mathbb{R}^n)}+\|\nabla v\|_{L^\infty(\mathbb{R}^n)}\big)^2 \\
&\qquad\qquad \cdot\Big(\sup_{x\in\mathbb{R}^n}\sup_{0<R^2<T}(1+|x|^\beta)R^{\frac{2}{n+4}}\|\nabla^2(u-v)\|_{L^{n+4}(Q_R(x))}\Big) \\
&\quad \lesssim \big(\|u\|_{X_T}+\|v\|_{X_T}\big)^2\|u-v\|_{X_T^\beta}.
\end{aligned}$$

For (4.6), using Lemma 4.4, we have that

$$\begin{aligned}
\|S(N[u]-N[v])\|_{X_T^\beta} &\lesssim \|N[u]-N[v]\|_{Y_T^\beta} \\
&\lesssim \big(\|u\|_{X_T}^2+\|v\|_{X_T}^2\big)\|u-v\|_{X_T^\beta} \\
&\lesssim \big([p]^2+[v_0]^2\big)\|u-v\|_{X_T^\beta} \lesssim \varepsilon^2\|u-v\|_{X_T^\beta}.
\end{aligned}$$

Note that we have used Theorem 2.1, which deals with the unweighted case, to estimate the $\|\cdot\|_{X_T}$ norms by the initial data. Hence, (4.6) holds if we take ε to be sufficiently small. ■

4.2. The Willmore flow (WF) case

The strategy here is similar to the MCF case. We again introduce the following weighted function space:

$$\begin{aligned}
X_T^\beta = \Big\{u \mid \|u\|_{X_T^\beta} := &\sup_{0<t<T}\sup_{x\in\mathbb{R}^n}(1+|x|^\beta)|\nabla u(x,t)| \\
&+ \sup_{x\in\mathbb{R}^n}\sup_{0<R^4<T}(1+|x|^\beta)R^{\frac{2}{n+6}}\|\nabla^2 u\|_{L^{n+6}(Q_R(x))} < \infty\Big\},
\end{aligned}$$

where $Q_R(x) := B_R(x)\times(R^4/2, R^4)$.

Lemma 4.6. *For $k \geq 0$,*

$$\Big\|t^{\frac{k}{4}}\nabla^k e^{-t\Delta^2}p(x)\Big\|_{X_T^\beta} \lesssim [p]_\beta. \tag{4.9}$$

Anticipating the forms of the nonlinear terms in (2.11), we introduce the weighted function spaces $Y_{0,T}^\beta$, $Y_{1,T}^\beta$ and $Y_{2,T}^\beta$, where

$$\|g_0\|_{Y_{0,T}^\beta} = \sup_{x\in\mathbb{R}^n}\sup_{0<R^4<T}(1+|x|^\beta)R^{\frac{6}{n+6}}\|g_0\|_{L^{\frac{n+6}{3}}(Q_R(x))},$$

$$\|g_1\|_{Y^\beta_{1,T}} = \sup_{x\in\mathbb{R}^n}\sup_{0<R^4<T}(1+|x|^\beta)R^{\frac{4}{n+6}}\|g_1\|_{L^{\frac{n+6}{2}}(Q_R(x))},$$
$$\|g_2\|_{Y^\beta_{2,T}} = \sup_{x\in\mathbb{R}^n}\sup_{0<R^4<T}(1+|x|^\beta)R^{\frac{2}{n+6}}\|g_2\|_{L^{n+6}(Q_R(x))}.$$

Now consider the following operator:

$$Sg(x,t) := \int_0^t e^{-(t-s)\Delta^2} g\, ds = \int_0^t \int_{\mathbb{R}^n} b(x-y,t-s)g(y,s)dyds. \tag{4.10}$$

The key estimate is the following lemma:

Lemma 4.7. *For every $0<t<T<\infty$,*

$$\sum_{l=0}^2\Big(\sup_{0<t<T}\big\|(1+|x|^\beta)\nabla^l Sg_l(\cdot,t)\big\|_{L^\infty(\mathbb{R}^n)} + \|\nabla^l Sg_l\|_{X^\beta_T}\Big) \lesssim \sum_{l=0}^2\|g_l\|_{Y^\beta_{l,T}}. \tag{4.11}$$

Lemma 4.8. *For every $0<T<\infty$,*

$$\sum_{l=0}^2\|(f_l[u]-f_l[v])\|_{Y^\beta_{l,T}} \lesssim \big(\|u\|_{X_T}+\|v\|_{X_T}\big)\|u-v\|_{X^\beta_T}. \tag{4.12}$$

(Recall forms (2.12)–(2.14) for the f_l's.) In particular, there exist $\varepsilon>0$ and $q<1$ such that for all $[v_0]+[p]<\varepsilon$,

$$\sum_{l=0}^2\Big\|\int_0^t e^{-(t-s)\Delta^2}\big(\nabla^l f_l(u)-\nabla^l f_l(v)\big)\,ds\Big\|_{X^\beta_T} \le q\|u-v\|_{X^\beta_T}. \tag{4.13}$$

Proof of Lemma 4.6. It suffices to show that there exists a $C>0$ depending only on T,n,β, and k such that if $[p]_\beta\le 1$, then $\|e^{-t\Delta^2}p(x)\|_{X^\beta_T}\le C$. Again, we need to estimate two terms.

First, by estimate (3.14) for the biharmonic kernel b, for any $k\in\mathbb{N}_+$, there exists $c_k>0$ such that

$$\begin{aligned}
\big|t^{\frac{k}{4}}\nabla^k\nabla e^{-t\Delta^2}p(x)\big| &= \Big|\int_{\mathbb{R}^n} t^{\frac{k}{4}}\nabla_x^k\nabla_x b(t,x-y)p(y)dy\Big| \\
&\le \int_{\mathbb{R}^n}\big|t^{\frac{k}{4}}\nabla_y^k b(x-y,t)\big|\big|\nabla_y p(y)\big|dy \\
&\lesssim \Big(\int_{\{y:|y-x|\le\frac{t^{\frac14}}{2T^{\frac14}}|x|\}} + \int_{\{y:|y-x|\ge\frac{t^{\frac14}}{2T^{\frac14}}|x|\}} t^{-\frac{n}{4}}e^{-c_k|(x-y)t^{-\frac14}|^{\frac43}}\frac{1}{1+|y|^\beta}dy\Big) \\
&=: \mathrm{I}+\mathrm{II}
\end{aligned}$$

where, similar to the MCF case, we have

$$\mathrm{I} \lesssim \frac{1}{1+|x|^\beta}\int_{\{y:|y-x|\le\frac{t^{\frac14}}{2T^{\frac14}}|x|\}} t^{-\frac{n}{4}}e^{-c_k|(x-y)t^{-\frac14}|^{\frac43}}dy \lesssim \frac{1}{1+|x|^\beta},$$

$$\mathrm{II} \lesssim e^{-C|x|^{\frac{4}{3}}} \int_{\left\{y:|y-x|\geq \frac{t^{\frac{1}{4}}}{2T^{\frac{1}{4}}}|x|\right\}} t^{-\frac{n}{4}} e^{-\frac{c_k}{2}|(x-y)t^{-\frac{1}{4}}|^{\frac{4}{3}}}\, dy \lesssim \frac{1}{1+|x|^\beta},$$

so that

$$\sup_{0<t<T} \sup_{x\in\mathbb{R}^2} (1+|x|^\beta)\left|t^{\frac{k}{4}}\nabla^k \nabla e^{-t\Delta^2} p(x)\right| \lesssim 1. \tag{4.14}$$

Second, we compute

$$\begin{aligned}
&\left\|t^{\frac{k}{4}}\nabla^k\nabla^2 e^{-t\Delta^2} p(x)\right\|_{L^{n+6}(Q_R(x))} \\
&\quad= \int_{R^4/2}^{R^4}\int_{B_R(x)} \left[t^{\frac{k}{4}}\nabla^{k+1} b(y-z,t)\nabla p(z)\, dz\right]^{n+6} dy\, dt \\
&\quad\lesssim \int_{R^4/2}^{R^4}\int_{B_R(x)} \left[t^{-\frac{1}{4}} t^{-\frac{n}{4}} e^{-c_k|(y-z)t^{-\frac{1}{4}}|^{\frac{4}{3}}} \frac{1}{1+|z|^\beta}\, dz\right]^{n+6} dy\, dt \\
&\quad\lesssim \int_{R^4/2}^{R^4}\int_{B_R(x)} \left[\frac{t^{-\frac{1}{4}}}{1+|y|^\beta}\, dz\right]^{n+6} dy\, dt \\
&\quad\lesssim \int_{R^4/2}^{R^4} t^{-\frac{n+6}{4}}\, dt \int_{B_R(x)} \frac{1}{(1+|y|^\beta)^{n+6}}\, dy \\
&\quad\lesssim R^{-2}\frac{1}{(1+|x|^\beta)^{n+6}},
\end{aligned}$$

which implies that

$$\sup_{x\in\mathbb{R}^n} \sup_{0<R^4<T} (1+|x|^\beta) R^{\frac{2}{n+6}} \left\|t^{\frac{k}{4}}\nabla^k\nabla^2 e^{-t\Delta^2} p\right\|_{L^{n+6}(Q_R(x))} \lesssim 1. \tag{4.15}$$

Combining (4.14) and (4.15) then gives Lemma 4.6. ■

Proof of Lemma 4.8. The proof is similar to that of Lemma 4.5. We will just highlight some key computations, though mostly at the symbolic level.

Recall that the form of f_0 is $f_0(u) = (\nabla^2 u)^3 \mathcal{P}(\nabla u)$, for some polynomial $\mathcal{P}$. Then,

$$\begin{aligned}
f_0(u) - f_0(v) &= \left((\nabla^2 u)^3 - (\nabla v)^3\right)\mathcal{P}(\nabla u) + (\nabla^2 v)^3\left(\mathcal{P}(\nabla u) - \mathcal{P}(\nabla v)\right) \\
&\approx \mathcal{P}(\nabla u)\left((\nabla^2 u)^2 + (\nabla v)^2\right)(\nabla^2(u-v)) + (\nabla^2 v)^3\mathcal{P}'(\nabla u)(\nabla(u-v)),
\end{aligned}$$

so that

$$\begin{aligned}
&\|f_0(u) - f_0(v)\|_{L^{\frac{n+6}{3}}(Q_R(x))} \\
&\quad\lesssim \|\mathcal{P}(\nabla u)\|_{L^\infty(\mathbb{R}^n)} \left\|\left((\nabla^2 u)^2 + (\nabla v)^2\right)\nabla^2(u-v)\right\|_{L^{\frac{n+6}{3}}(Q_R(x))} \\
&\qquad+ \|(\nabla^2 v)^3\|_{L^{\frac{n+6}{3}}(Q_R(x))} \|\mathcal{P}'(\nabla u)\|_{L^\infty(\mathbb{R}^n)} \|\nabla(u-v)\|_{L^\infty(\mathbb{R}^n)} \\
&\quad\lesssim \|\mathcal{P}(\nabla u)\|_{L^\infty(\mathbb{R}^n)} \left(\|\nabla^2 u\|^2_{L^{n+6}(Q_R(x))} + \|\nabla^2 v\|^2_{L^{n+6}(Q_R(x))}\right) \|\nabla^2(u-v)\|_{L^{n+6}(Q_R(x))} \\
&\qquad+ \|\nabla^2 u\|^3_{L^{n+6}(Q_R(x))} \|\mathcal{P}'(\nabla v)\|_{L^\infty(\mathbb{R}^n)} \|\nabla(u-v)\|_{L^\infty(\mathbb{R}^n)},
\end{aligned}$$

and hence,

$$\|f_0(u)-f_0(v)\|_{Y^\beta_{0,T}} \lesssim \big(\|u\|^2_{X_T}+\|v\|^2_{X_T}\big)\|u-v\|_{X^\beta_T}.$$

Similarly, for $f_1(u)=(\nabla^2 u)^2\mathcal{P}(\nabla u)$ and $f_2(u)=(\nabla^2 u)\mathcal{P}(\nabla u)$, we have

$$\begin{aligned}
&\|f_1(u)-f_1(v)\|_{L^{\frac{n+6}{2}}(Q_R(x))}\\
&\quad\lesssim \|\mathcal{P}(\nabla u)\|_{L^\infty(\mathbb{R}^n)}\big(\|\nabla^2 u\|_{L^{n+6}(Q_R(x))}+\|\nabla^2 v\|_{L^{n+6}(Q_R(x))}\big)\|\nabla^2(u-v)\|_{L^{n+6}(Q_R(x))}\\
&\qquad+\|\nabla^2 u\|^2_{L^{n+6}(Q_R(x))}\|\mathcal{P}'(\nabla v)\|_{L^\infty(\mathbb{R}^n)}\|\nabla(u-v)\|_{L^\infty(\mathbb{R}^n)}
\end{aligned}$$

and

$$\begin{aligned}
\|f_2(u)-f_2(v)\|_{L^{n+6}(Q_R(x))} &\lesssim \|\mathcal{P}(\nabla u)\|_{L^\infty(\mathbb{R}^n)}\|\nabla^2(u-v)\|_{L^{n+6}(Q_R(x))}\\
&\quad+\|\nabla^2 u\|_{L^{n+6}(Q_R(x))}\|\mathcal{P}'(\nabla v)\|_{L^\infty(\mathbb{R}^n)}\|\nabla(u-v)\|_{L^\infty(\mathbb{R}^n)}
\end{aligned}$$

so that

$$\|f_1(u)-f_1(v)\|_{Y^\beta_{1,T}},\quad \|f_2(u)-f_2(v)\|_{Y^\beta_{2,T}} \lesssim \big(\|u\|^2_{X_T}+\|v\|^2_{X_T}\big)\|u-v\|_{X^\beta_T},$$

thus completing the proof of (4.12). ■

Again, we postpone the proof of Lemma 4.7 to Appendix B due to its technicality.

4.3. Conclusion of the proof of Theorem 2.4

For simplicity, we just write down the steps for WF as it involves more terms. Recall the equation for Φ_λ:

$$\Phi_\lambda = e^{-\Delta^2 t}p_\lambda + \sum_{l=0}^{2}(\mathcal{N}_l[v+\Phi_\lambda]-\mathcal{N}_l[v]), \tag{4.16}$$

where

$$\mathcal{N}_l(g)=\int_0^t e^{-\Delta^2(t-s)}\nabla^l f_l(g)\,ds.$$

First, taking the X^β_T norm of both sides of the equation, by Lemma 4.7 and (4.13) of Lemma 4.8, we get

$$\begin{aligned}
\|\Phi_\lambda\|_{X^\beta_T} &\le \big\|e^{-\Delta^2 t}p_\lambda\big\|_{X^\beta_T} + \sum_{l=0}^{2}\|\mathcal{N}_l[\Phi_\lambda+v]-\mathcal{N}_l[v]\|_{X^\beta_T}\\
&\le \big\|e^{-\Delta^2 t}p_\lambda\big\|_{X^\beta_T} + \sum_{l=0}^{2}\|f_l(\Phi_\lambda+v)-f_l[v]\|_{Y^\beta_{l,T}}\\
&\le \big\|e^{-\Delta^2 t}p_\lambda\big\|_{X^\beta_T} + q\|\Phi_\lambda\|_{X^\beta_T}.
\end{aligned}$$

Hence, upon choosing $[v_0]$, $[p]$ small enough, we will have $q < 1$, which implies a uniform bound for Φ_λ in X_T^β. More precisely,

$$\|\Phi_\lambda\|_{X_T^\beta} \lesssim \left\|e^{-\Delta^2 t} p_\lambda\right\|_{X_T^\beta} \lesssim [p_\lambda]_\beta. \tag{4.17}$$

Second, from Lemmas 4.7 and 4.8 again, we have that

$$\begin{aligned}\sum_{l=0}^{2} \sup_{0<t<T} &\|(1 + |x|^\beta)(\mathcal{N}_l[\Phi_\lambda + v] - \mathcal{N}_l[v])\|_{L^\infty(\mathbb{R}^n)} \\ &\lesssim \sum_{l=0}^{2} \| f_l(\Phi_\lambda + v) - f_l(v)\|_{Y_{l,T}^\beta} \lesssim \|\Phi_\lambda\|_{X_T^\beta} \lesssim [p_\lambda]_\beta.\end{aligned}$$

When $\lambda > 1$, we have $[p_\lambda]_\beta \leq [p]_\beta$. Hence,

$$\sup_{\lambda>1} \sum_{l=0}^{2} \|(1 + |x|^\beta)(\mathcal{N}_l[\Phi_\lambda + v] - \mathcal{N}_l[v])(x, T)\|_{L^\infty(\mathbb{R}^n)} \lesssim [p]_\beta. \tag{4.18}$$

With the above, we can prove the global C^1-convergence. Upon setting $T = 1$ in equation (4.18), we have that the set $\{\Phi_\lambda(\cdot, 1) - e^{-\Delta^2} p_\lambda(\cdot) = \sum_{l=0}^{2} \mathcal{N}_l(\Phi_\lambda + v) - \mathcal{N}_l(v)\}_{\lambda>1}$ satisfies the equi-decay property, that is,

$$\lim_{R\to\infty} \sup_{\lambda>0} \sup_{|x|<R} \left|\Phi_\lambda(x, 1) - e^{-\Delta^2} p_\lambda(x, 1)\right| = 0.$$

From (3.24) (with $\gamma = k = 0$) and (3.18) (with the latter applied to ∇p_λ), we have

$$\begin{aligned}&\left\|\nabla\big(\Phi_\lambda(\cdot, 1) - e^{-\Delta^2} p_\lambda(\cdot, 1)\big)\right\|_{L^\infty(\mathbb{R}^n)} \\ &\quad\leq \|\nabla\Phi_\lambda(\cdot, 1)\|_{L^\infty(\mathbb{R}^n)} + \left\|e^{-\Delta^2}\nabla p_\lambda(\cdot, 1)\right\|_{L^\infty(\mathbb{R}^n)} < \infty.\end{aligned}$$

Finally, recall estimate (3.27). By the Arzelà–Ascoli theorem, we can then conclude that $\Phi_{\lambda_j} \to \Phi_\infty$ in $C^0(\mathbb{R}^n)$ for a subsequence $\lambda_j \to \infty$. The proof of $\Phi_\infty \equiv 0$ is the same as in Section 3.3 for the spatially unweighted case.

For the convergence of $\nabla\Phi_\lambda$, by (4.17), we have that $\nabla\Phi_\lambda$ has the equi-decay property, that is,

$$\lim_{R\to\infty} \sup_{\lambda>0} \sup_{|x|<R} \left|\nabla\Phi_\lambda(x, 1)\right| = 0.$$

From (3.24) (with $\gamma = 1, k = 0$), we further have

$$\sup_{\lambda>0} \left\|\nabla^2\Phi_\lambda(\cdot, 1)\right\|_{L^\infty(\mathbb{R}^n)} < \infty.$$

Hence, we deduce that $\nabla\Phi_{\lambda_j} \to \nabla\Phi_\infty \equiv 0$ uniformly in $\mathbb{R}^n$.

The overall C^1-convergence of $u_\lambda = \Phi_\lambda + v$ to v is thus established.

5. Generalization to polyharmonic flows

As a future perspective and direction, we use this section to illustrate the robustness of the current approach and outline an abstract framework for the stability of self-similar solutions to possible higher-order polyharmonic flows. Suppose the polyharmonic flow, in the graphical setting, takes the following form:

$$\begin{cases} \partial_t u + Au = N[u] & \text{on } \mathbb{R}^n \times (0,\infty), \\ u(x,0) = u_0(x) & \text{in } \mathbb{R}^n, \end{cases} \tag{5.1}$$

where $A = (-\Delta)^m$, $m \geq 2$, and $N[u]$ is the nonlinear term—see [21] for an example of the form of N. Furthermore, assume that (5.1) is invariant under the rescaling

$$u_\lambda := \frac{1}{\lambda} u(\lambda x, \lambda^{2m} t). \tag{5.2}$$

Then, for the self-similar initial data $v_0(x) = \lambda^{-1} v_0(\lambda x)$ with small Lipschitz norm, we expect the existence of a self-similar solution $v(x,t)$ to (5.1), that is,

$$v(x,t) = v_{t^{-\frac{1}{2m}}}(x,t) = t^{\frac{1}{2m}} v(xt^{-\frac{1}{2m}}, 1) =: t^{\frac{1}{2m}} \Psi(xt^{-\frac{1}{2m}}).$$

One could follow Koch–Lamm's method to find a unique analytic solution to (5.1) with initial data of small Lipschitz norm in the following scale-invariant function space:

$$\begin{aligned} X_T := \Big\{ f(x,t) : \mathbb{R}^n \times (0,T) \to \mathbb{R} \mid \|f\|_{X_T} := \sum_{k=0}^{m-2} \sup_{0<t<T} t^{\frac{k}{2m}} \|\nabla^k \nabla f(x,t)\|_{L^\infty(\mathbb{R}^n)} \\ + \sup_{x\in\mathbb{R}^n} \sup_{0<R^{2m}<T} R^{\frac{(m-1)p-n-2m}{p}} \|\nabla^m f\|_{L^p(B_R(x)\times(R^{2m}/2,R^{2m}))} < \infty \Big\} \end{aligned} \tag{5.3}$$

for some $p > n + 2m$. We anticipate that an approach similar to this paper can show the stability of the self-similar solution v under bounded (and small) perturbations; more specifically, for $u_0 = v_0(x) + p(x)$ with $\|p\|_{L^\infty(\mathbb{R}^n)} < \infty$ and $\|\nabla p\|_{L^\infty(\mathbb{R}^n)} < \varepsilon$, it holds that

$$\lim_{t\to\infty} \left\| t^{-\frac{1}{2m}} u(t^{\frac{1}{2m}} x, t) - \Psi(x) \right\|_{C^k_{\mathrm{loc}}(\mathbb{R}^n)} = 0, \quad \forall k \in \mathbb{N}^+. \tag{5.4}$$

Moreover, by putting the difference $u - v$ in the weighted space

$$\begin{aligned} X_T^\beta := \Big\{ f(x,t) : \mathbb{R}^n \times (0,T) \to \mathbb{R} \mid \\ \|f\|_{X_T^\beta} := \sum_{k=0}^{m-2} \sup_{0<t<T} t^{\frac{k}{2m}} \|(1+|x|^\beta)\nabla^k \nabla f(x,t)\|_{L^\infty(\mathbb{R}^n)} \\ + \sup_{x\in\mathbb{R}^n} \sup_{0<R^{2m}<T} (1+|x|^\beta) R^{\frac{(m-1)p-n-2m}{p}} \|\nabla^m f\|_{L^p(B_R(x)\times(R^{2m}/2,R^{2m}))} < \infty \Big\}, \end{aligned} \tag{5.5}$$

we can similarly gain the equi-decay property which leads to the global convergence

$$\lim_{t\to\infty}\left\|t^{-\frac{1}{2m}}u(t^{\frac{1}{2m}}x,t)-\Psi(x)\right\|_{C^k(\mathbb{R}^n)}=0, \tag{5.6}$$

provided the initial perturbation is small in some appropriate weighted space, for example, $\|(1+|x|^\beta)\nabla p\|_{L^\infty(\mathbb{R}^n)} \ll 1$.

A. Proof of Lemma 4.4

Before the proof, we first recall some L^p estimates concerning the heat kernel $h(x,t)$ given in (3.3): for $0<t<\infty$,

$$\|h\|_{L^p(\mathbb{R}^n\times(0,t))} \lesssim t^{\frac{(n+2)-pn}{2p}} \qquad \text{for } 1\le p<\frac{n+2}{n}, \tag{A.1}$$

$$\|\nabla h\|_{L^p(\mathbb{R}^n\times(0,t))} \lesssim t^{\frac{(n+2)-(n+1)p}{2p}} \qquad \text{for } 1\le p<\frac{n+2}{n+1}, \tag{A.2}$$

and

$$\begin{aligned}\Big\|\int_0^t\int_{\mathbb{R}^n}\nabla^2 h(z-y,t-s)g(y,s)dyds\Big\|_{L^p(\mathbb{R}^n\times\mathbb{R}^+)}&\\ \lesssim \|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^+)} \quad \text{for } 1\le p<\infty,&\end{aligned} \tag{A.3}$$

where the last is from the theory of singular integrals [40]. The following pointwise estimate will also be used: for all $(z,s)\in\mathbb{R}^n\times(0,t)\setminus B_{\sqrt{t}}(0)\times(0,\frac{t}{2})$, it holds that

$$|h(z,s)|+\sqrt{t}|\nabla h(z,s)|+t|\nabla^2 h(z,s)|\le Ct^{-\frac{n}{2}}\exp\Big(-c\frac{|z|}{\sqrt{t}}\Big), \tag{A.4}$$

which follows from the scaling property of the heat kernel.

Proof of Lemma 4.4. It suffices to show that if $\|g\|_{Y_T^\beta}\le 1$, then

$$\sup_{0<t<T}\|(1+|x|^\beta)Sg(x,t)\|_{L^\infty(\mathbb{R}^n)}+\|Sg\|_{X_T^\beta}\lesssim 1.$$

For this purpose, we need to estimate $|Sg(x,t)|$, $|\nabla Sg(x,t)|$, and $\|\nabla^2 Sg\|_{L^{n+4}(Q_R(x))}$. We recall the notation $Q_R(x)=B_R(x)\times(\frac{R^2}{2},R^2)$ and further define $Q'_R(x):=B_R(x)\times(0,\frac{R^2}{2})$. Without loss of generality, we fix $T=1$.

Estimate for Sg. We decompose

$$\begin{aligned}|Sg(x,t)|&=\Big|\int_0^t\int_{\mathbb{R}^n}h(x-y,t-s)g(y,s)dyds\Big|\\ &\le\int_{Q_{\sqrt{t}}(x)}+\int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt{t}}(x)}|h(x-y,t-s)g(y,s)|dyds\end{aligned}$$

$$:= I_1 + I_2.$$

For I_1, by the Hölder inequality and heat kernel estimate (A.1) with $p = \frac{n+4}{n+3} < \frac{n+2}{n}$, we have

$$\begin{aligned} I_1 &\leq \|h\|_{L^{\frac{n+4}{n+3}}(Q'_{\sqrt{t}}(0))} \|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))} \leq \|h\|_{L^{\frac{n+4}{n+3}}(\mathbb{R}^n\times(0,t/2))} \|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))} \\ &\lesssim t^{\frac{6+n}{8+2n}} \|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))} = t^{\frac{1}{2}} t^{\frac{1}{n+4}} \|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))} \\ &\lesssim \frac{t^{\frac{1}{2}}}{1+|x|^{\beta}}. \end{aligned} \tag{A.5}$$

We estimate I_2 as follows:

$$\begin{aligned} I_2 &= \int_{\mathbb{R}^n\times(0,t)\backslash Q_{\sqrt{t}}(x)} |h(x-y, t-s)g(y,s)| dy ds \\ &\lesssim \sum_{m=0}^{\infty} \sum_{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n} \int_{2^{-m-1}t}^{2^{-m}t} \int_{B_{2^{-\frac{m}{2}}\sqrt{t}}(z)} t^{-\frac{n}{2}} e^{-c\frac{|x-y|}{\sqrt{t}}} |g(y,s)| dy ds \\ &= \sum_{m=0}^{\infty} \sum_{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n} \int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)} t^{-\frac{n}{2}} e^{-c\frac{|x-y|}{\sqrt{t}}} |g(y,s)| dy ds \\ &\lesssim \Bigg(\sum_{m=0}^{\infty} \sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n \\ |z-x|\leq \frac{\sqrt{t}|x|}{2}}} + \sum_{m=0}^{\infty} \sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n \\ |z-x|\geq \frac{\sqrt{t}|x|}{2}}} \Bigg) \int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)} t^{-\frac{n}{2}} e^{-c\frac{|x-y|}{\sqrt{t}}} |g(y,s)| dy ds \\ &:= I_{21} + I_{22}. \end{aligned}$$

To estimate I_{21}, we compute

$$\begin{aligned} I_{21} &\lesssim \sum_{m=0}^{\infty} \sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n \\ |z-x|\leq \frac{\sqrt{t}|x|}{2}}} e^{-c\frac{|x-z|}{\sqrt{t}}} \int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)} t^{-\frac{n}{2}} |g(y,s)| dy ds \\ &\lesssim \sum_{m=0}^{\infty} \Bigg(\sup_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n \\ |z-x|\leq \frac{\sqrt{t}|x|}{2}}} \int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)} t^{-\frac{n}{2}} |g(y,s)| dy ds \Bigg) \Bigg(\sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n \\ |z-x|\leq \frac{\sqrt{t}|x|}{2}}} e^{-c\frac{|z-x|}{\sqrt{t}}} \Bigg), \end{aligned}$$

where we have used the estimate $|\sum_z a(z)b(z)| \leq \sup_z |a(z)| \sum_z |b(z)|$. Note that

$$\sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n \\ |z-x|\leq \frac{\sqrt{t}|x|}{2}}} e^{-c\frac{|z-x|}{\sqrt{t}}} \lesssim \sum_{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n} e^{-c\frac{|z|}{\sqrt{t}}} = \sum_{\tilde{z}\in\mathbb{Z}^n} e^{-c|\tilde{z}|2^{-\frac{m}{2}}} \lesssim \int_{\mathbb{R}^n} e^{-c|\tilde{z}|2^{-\frac{m}{2}}} d^n\tilde{z} \approx 2^{\frac{mn}{2}},$$

while

$$\begin{aligned}
&\sup_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\le\frac{\sqrt{t}|x|}{2}}}\int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)} t^{-\frac{n}{2}}|g(y,s)|dyds\\
&\le t^{-\frac{n}{2}}\sup_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\le\frac{\sqrt{t}|x|}{2}}}\|1\|_{L^{\frac{n+4}{n+3}}(Q_{2^{-\frac{m}{2}}\sqrt{t}}(z))}\|g\|_{L^{n+4}(Q_{2^{-\frac{m}{2}}\sqrt{t}}(z))}\\
&\le t^{-\frac{1}{2}}2^{\frac{m(2-(n+2)(n+3))}{2(n+4)}}\sup_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\le\frac{\sqrt{t}|x|}{2}}}\left(2^{-\frac{m}{2}}\sqrt{t}\right)^{\frac{2}{n+4}}\|g\|_{L^{n+4}(Q_{2^{-\frac{m}{2}}\sqrt{t}}(z))}\\
&\le t^{\frac{1}{2}}2^{\frac{m(2-(n+2)(n+3))}{2(n+4)}}\sup_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\le\frac{\sqrt{t}|x|}{2}}}\frac{1}{1+|z|^\beta}\lesssim\frac{t^{\frac{1}{2}}2^{\frac{m(2-(n+2)(n+3))}{2(n+4)}}}{1+|x|^\beta}.
\end{aligned}$$

Hence,

$$I_{21}\le\frac{t^{\frac{1}{2}}}{1+|x|^\beta}\sum_{m=0}^{\infty}2^{\frac{m(2-(n+2)(n+3))}{2(n+4)}}2^{\frac{mn}{2}}=\frac{t^{\frac{1}{2}}}{1+|x|^\beta}\sum_{m=0}^{\infty}2^{-\frac{m}{2}}\lesssim\frac{t^{\frac{1}{2}}}{1+|x|^\beta}. \tag{A.6}$$

We estimate I_{22} as

$$\begin{aligned}
I_{22}&\lesssim\sum_{m=0}^{\infty}\sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\ge\frac{\sqrt{t}|x|}{2}}}e^{-\frac{c}{4}|x|}\int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)}t^{-\frac{n}{2}}e^{-\frac{c}{2}\frac{|x-y|}{\sqrt{t}}}|g(y,s)|dyds\\
&\lesssim e^{-\frac{c}{4}|x|}\sum_{m=0}^{\infty}\sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\ge\frac{\sqrt{t}|x|}{2}}}\int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)}t^{-\frac{n}{2}}e^{-\frac{c}{2}\frac{|x-y|}{\sqrt{t}}}|g(y,s)|dyds\\
&\lesssim e^{-\frac{c}{4}|x|}\sum_{m=0}^{\infty}\left(\sup_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\ge\frac{\sqrt{t}|x|}{2}}}\int_{Q_{2^{-\frac{m}{2}}\sqrt{t}}(z)}t^{-\frac{n}{2}}|g(y,s)|dyds\right)\left(\sum_{\substack{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n\\ |z-x|\ge\frac{\sqrt{t}|x|}{2}}}e^{-\frac{c}{2}\frac{|z-x|}{\sqrt{t}}}\right).
\end{aligned}$$

Then, similar to the computation for I_{21}, we arrive at

$$I_{22}\lesssim e^{-\frac{c}{4}|x|}t^{\frac{1}{2}}\sum_{m=0}^{\infty}\left(\sup_{z\in 2^{-\frac{m}{2}}\sqrt{t}\mathbb{Z}^n}\frac{1}{1+|z|^\beta}\right)2^{-\frac{m}{2}}\lesssim e^{-\frac{c}{4}|x|}t^{\frac{1}{2}}\lesssim\frac{t^{\frac{1}{2}}}{1+|x|^\beta}. \tag{A.7}$$

Combining (A.5), (A.6), and (A.7), we obtain

$$\sup_{0<t<T}\|(1+|x|^\beta)Sg(x,t)\|_{L^\infty(\mathbb{R}^n)}\lesssim t^{\frac{1}{2}}\lesssim 1. \tag{A.8}$$

We restate the estimate I_2 here for future usage:

$$I_2 = \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt{t}}(x)} |h(x-y,t-s)g(y,s)|dyds$$
$$\lesssim \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt{t}}(x)} t^{-\frac{n}{2}} e^{-c\frac{|x-y|}{\sqrt{t}}} |g(y,s)|dyds \lesssim \frac{t^{\frac{1}{2}}}{1+|x|^\beta}. \tag{A.9}$$

Estimate for ∇Sg.

$$|\nabla Sg(x,t)| = \Big|\int_0^t \int_{\mathbb{R}^n} \nabla h(x-y,t-s)g(y,s)dyds\Big|$$
$$\le \int_{Q_{\sqrt{t}}(x)} + \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt{t}}(x)} |\nabla h(x-y,t-s)g(y,s)|dyds$$
$$:= J_1 + J_2.$$

For J_1, by the Hölder inequality, using heat kernel estimate (A.2) with $p=\frac{n+4}{n+3}<\frac{n+2}{n+1}$, we can derive

$$J_1 \le \|\nabla h\|_{L^{\frac{n+4}{n+3}}(Q'_{\sqrt{t}}(0))}\|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))} \le \|\nabla h\|_{L^{\frac{n+4}{n+3}}(\mathbb{R}^n\times(0,t/2))}\|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))}$$
$$\lesssim \sqrt{t}^{\frac{2}{n+4}}\|g\|_{L^{n+4}(Q_{\sqrt{t}}(x))} \lesssim \frac{1}{1+|x|^\beta}. \tag{A.10}$$

For J_2, we can follow the derivation of (A.9) exactly. The only change is the appearance of $t^{-\frac{1}{2}}$ due to the pointwise estimate of ∇h in (A.4):

$$J_2 = \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt{t}}(x)} |\nabla h(x-y,t-s)g(y,s)|dyds$$
$$\lesssim \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt{t}}(x)} t^{-\frac{1}{2}}t^{-\frac{n}{2}} e^{-c\frac{|x-y|}{\sqrt{t}}} |g(y,s)|dyds \lesssim \frac{1}{1+|x|^\beta}. \tag{A.11}$$

Combining (A.10) and (A.11), we have

$$\sup_{0<t<1}\sup_{x\in\mathbb{R}^n}(1+|x|^\beta)|\nabla Sg(x,t)| \lesssim 1. \tag{A.12}$$

Estimate for $\nabla^2 Sg$. For this, we need to show

$$\sup_{0<R^2<1} R^{\frac{2}{n+4}}\|\nabla^2 Sg\|_{L^{n+4}(Q_R(x))} \lesssim \frac{1}{1+|x|^\beta}. \tag{A.13}$$

For this purpose, we compute

$$R^{\frac{2}{n+4}}\|\nabla^2 Sg(z,t)\|_{L^{n+4}(Q_R(x))}$$
$$= R^{\frac{2}{n+4}}\Big\|\int_0^t\int_{\mathbb{R}^n}\nabla^2 h(z-y,t-s)g(y,s)dyds\Big\|_{L^{n+4}(Q_R(x))}$$

$$= R^{\frac{2}{n+4}} \Big\| \int_{R^n\times(0,t)\setminus B_{2R}(x)\times(R^2/4,R^2)} + \int_{B_{2R}(x)\times(R^2/4,R^2)} \nabla^2 h(z-y,t-s)g(y,s)dyds \Big\|_{L^{n+4}(Q_R(x))}$$
$$= R^{\frac{2}{n+4}} \Big\| \int_{R^n\times(0,t)\setminus B_{2R}(x)\times(R^2/4,R^2)} \nabla^2 h(z-y,t-s)g(y,s)dyds \Big\|_{L^{n+4}(Q_R(x))}$$
$$+ R^{\frac{2}{n+4}} \Big\| \int_{B_{2R}(x)\times(R^2/4,R^2)} \nabla^2 h(z-y,t-s)g(y,s)dyds \Big\|_{L^{n+4}(Q_R(x))}$$
$$:= K_1 + K_2.$$

For K_1, we have

$$K_1 = R^{\frac{2}{n+4}} \Big\| \int_{R^n\times(0,t)\setminus B_{2R}(x)\times(R^2/4,R^2)} \nabla^2 h(z-y,t-s)g(y,s)dyds \Big\|_{L^{n+4}(Q_R(x))}$$
$$\lesssim R^{\frac{2}{n+4}} \Big\| \frac{t^{-\frac{1}{2}}}{1+|z|^\beta} \Big\|_{L^{n+4}(Q_R(x))} \lesssim \frac{1}{1+|x|^\beta}, \tag{A.14}$$

where we have used again estimate (A.9) for I_2 but with h replaced by $\nabla^2 h$. The $t^{-\frac{1}{2}}$ factor is due to the pointwise estimate for $\nabla^2 h$ from (A.4). Note also that $\frac{R^2}{2} < t < R^2$.

For K_2,

$$K_2 := R^{\frac{2}{n+4}} \Big\| \int_{B_{2R}(x)\times(R^2/4,R^2)} \nabla^2 h(z-y,t-s)g(y,s)dyds \Big\|_{L^{n+4}(Q_R(x))}$$
$$\lesssim R^{\frac{2}{n+4}} \|\chi_{B_{2R}(x)\times(R^2/4,R^2)} g(z,t)\|_{L^{n+4}(\mathbb{R}^n\times\mathbb{R}_+)}$$
$$\lesssim R^{\frac{2}{n+4}} \|g\|_{L^{n+4}(B_{2R}(x)\times(R^2/4,R^2))} \lesssim \frac{1}{1+|x|^\beta}, \tag{A.15}$$

where the second inequality is due to (A.3).

Hence, (A.13) holds upon combining (A.14) and (A.15). ∎

B. Proof of Lemma 4.7

The strategy here is very similar to Lemma 4.4. The main difference is the usage of the estimates of the biharmonic kernel b and also the fact that we need to deal with g_l for $l = 0, 1, 2$. For $0 \le k \le 3$ and $t > 0$, we have from (3.14) that

$$\|\nabla^k b\|_{L^p(\mathbb{R}^n\times(0,t))} \le C t^{\frac{(n+4)-p(n+k)}{4p}} \quad \text{for } 1 \le p < \frac{n+4}{n+k}, \tag{B.1}$$

while for $k = 4$, the following comes from the theory of singular integrals [40]:

$$\Big\| \int_0^t \int_{\mathbb{R}^n} \nabla^4 b(z-y,t-s)g(y,s)\,dyds \Big\|_{L^p(\mathbb{R}^n\times\mathbb{R}^+)} \lesssim \|g\|_{L^p(\mathbb{R}^n\times\mathbb{R}^+)}. \tag{B.2}$$

Furthermore, from the scaling property of the kernel, the following pointwise estimate holds:

$$\sum_{k=0}^{4}|(\sqrt[4]{t}\nabla)^k b(z,s)| \lesssim t^{-\frac{n}{4}} \exp\Big(-c\frac{|z|}{\sqrt[4]{t}}\Big), \quad \forall (y,s) \in \mathbb{R}^n \times (0,t) \setminus Q'_{\sqrt[4]{t}}(0), \tag{B.3}$$

where we recall the notation $Q_R(x) = B_R(x) \times (\frac{R^4}{2}, R^4)$ and $Q'_R(x) = B_R(x) \times (0, \frac{R^4}{2})$.

Proof of Lemma 4.7. The proof is similar to the one given in the previous section. It suffices to show that there exists a $C > 0$ such that if $\sum_{l=0}^{2}\|g_l\|_{Y^\beta_{l,T}} \le 1$, then

$$\sum_{l=0}^{2} \sup_{0<t<T} \|(1+|x|^\beta)\nabla^l S g_l(x,t)\|_{L^\infty(\mathbb{R}^n)} + \|\nabla^l S g_l\|_{X^\beta_T} \le C.$$

Without loss of generality, we fix $T = 1$. Note also that $Q_{\sqrt[4]{t}}(x) := B_{\sqrt[4]{t}}(x) \times (t/2, t)$ and $Q'_{\sqrt[4]{t}}(x) := B_{\sqrt[4]{t}}(x) \times (0, t/2)$. Now we estimate the relevant quantities.

Estimate for Sg_l ($l = 0, 1, 2$). We compute

$$\begin{aligned}
|\nabla^l S g_l(x,t)| &= \Big|\int_0^t \int_{\mathbb{R}^n} \nabla^l b(x-y,t-s) g_l(y,s)dyds\Big| \\
&\le \Big(\int_{Q'_{\sqrt[4]{t}}(0)} + \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt[4]{t}}(x)}\Big)|\nabla^l b(x-y,t-s)g_l(y,s)|dyds \\
&= I_1 + I_2.
\end{aligned}$$

For I_1, by the Hölder inequality, using kernel estimate (B.1) with $p = \frac{n+6}{n+3+l} < \frac{n+4}{n+l}$, we arrive at

$$\begin{aligned}
I_1 &\le \|\nabla^l b\|_{L^{\frac{n+6}{n+3+l}}(Q'_{\sqrt[4]{t}}(0))} \|g_l\|_{L^{\frac{n+6}{3-l}}(Q_{\sqrt[4]{t}}(x))} \le \|\nabla^l b\|_{L^{\frac{n+6}{n+3+l}}(\mathbb{R}^n\times(0,t))} \|g_l\|_{L^{\frac{n+6}{3-l}}(Q_{\sqrt[4]{t}}(x))} \\
&\lesssim t^{\frac{n+4-(n+l)p}{4p}} t^{-\frac{1}{4}\left(\frac{6-2l}{n+6}\right)} \Big(t^{\frac{1}{4}\frac{6-2l}{n+6}} \|g_l\|_{L^{\frac{n+6}{3-l}}(Q_{\sqrt[4]{t}}(x))}\Big) \lesssim \frac{t^{\frac{1}{4}}}{1+|x|^\beta}.
\end{aligned} \tag{B.4}$$

For I_2, we make use of (B.3) and compute

$$\begin{aligned}
I_2 &\lesssim \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt[4]{t}}(x)} |\nabla^l b(x-y,t-s)g_l(y,s)|dyds \\
&\lesssim \sum_{m=0}^{\infty} \sum_{z\in 2^{-\frac{m}{4}}\sqrt[4]{t}\mathbb{Z}^n} \int_{Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)} t^{-\frac{n+l}{4}} e^{-c\frac{|x-y|}{\sqrt[4]{t}}} |g_l(y,s)|dyds \\
&\lesssim \Bigg(\sum_{m=0}^{\infty} \sum_{\substack{z\in 2^{-\frac{m}{4}}\sqrt[4]{t}\mathbb{Z}^n \\ |z-x|\le\frac{\sqrt[4]{t}|x|}{2}}} + \sum_{m=0}^{\infty} \sum_{\substack{z\in 2^{-\frac{m}{4}}\sqrt[4]{t}\mathbb{Z}^n \\ |z-x|\ge\frac{\sqrt[4]{t}|x|}{2}}}\Bigg) \int_{Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(x)} t^{-\frac{n+l}{4}} e^{-c\frac{|x-y|}{\sqrt[4]{t}}} |g_l(y,s)|dyds
\end{aligned}$$

$$:= I_{21} + I_{22}.$$

Again, similar to the previous section, we have

$$\begin{aligned}
I_{21} &= \sum_{m=0}^{\infty} \sum_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| \le \frac{\sqrt[4]{t}|x|}{2}}} \int_{Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)} t^{-\frac{n+l}{4}} e^{-c\frac{|x-y|}{\sqrt[4]{t}}} |g_l(y,s)| dy ds \\
&\le t^{-\frac{n+l}{4}} \sum_{m=0}^{\infty} \Bigg(\sum_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| \le \frac{\sqrt[4]{t}|x|}{2}}} e^{-c\frac{|x-z|}{\sqrt[4]{t}}} \Bigg) \Bigg(\sup_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| \le \frac{\sqrt[4]{t}|x|}{2}}} \int_{Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)} |g_l(y,s)| dy ds \Bigg) \\
&\lesssim t^{-\frac{n+l}{4}} \sum_{m=0}^{\infty} 2^{\frac{mn}{4}} \Bigg(\sup_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| \le \frac{\sqrt[4]{t}|x|}{2}}} \|1\|_{L^{\frac{n+6}{n+3+l}}\left(Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)\right)} \|g_l\|_{L^{\frac{n+6}{3-l}}\left(Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)\right)} \Bigg) \\
&\lesssim t^{\frac{1}{4}} \sum_{m=0}^{\infty} \frac{2^{-\frac{m(1+l)}{4}}}{1+|x|^{\beta}} \lesssim \frac{t^{\frac{1}{4}}}{1+|x|^{\beta}},
\end{aligned} \tag{B.5}$$

while for I_{22},

$$\begin{aligned}
I_{22} &= \sum_{m=0}^{\infty} \sum_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| > \frac{\sqrt[4]{t}|x|}{2}}} \int_{Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)} t^{-\frac{n+l}{4}} e^{-c\frac{|x-y|}{\sqrt[4]{t}}} |g_l(y,s)| dy ds \\
&\le t^{-\frac{n+l}{4}} e^{-\frac{c}{4}|x|} \sum_{m=0}^{\infty} \Bigg(\sum_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| > \frac{\sqrt[4]{t}|x|}{2}}} e^{-c\frac{|x-z|}{\sqrt[4]{t}}} \Bigg) \Bigg(\sup_{\substack{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n \\ |z-x| > \frac{\sqrt[4]{t}|x|}{2}}} \int_{Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)} |g_l(y,s)| dy ds \Bigg) \\
&\lesssim t^{-\frac{n+l}{4}} e^{-\frac{c}{4}|x|} \sum_{m=0}^{\infty} 2^{\frac{mn}{4}} \Bigg(\sup_{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n} \|1\|_{L^{\frac{n+6}{n+3+l}}\left(Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)\right)} \|g_l\|_{L^{\frac{n+6}{3-l}}\left(Q_{2^{-\frac{m}{4}}\sqrt[4]{t}}(z)\right)} \Bigg) \\
&\lesssim t^{\frac{1}{4}} e^{-\frac{c}{4}|x|} \sum_{m=0}^{\infty} 2^{-\frac{m(1+l)}{4}} \sup_{z \in 2^{-\frac{m}{4}} \sqrt[4]{t}\mathbb{Z}^n} \frac{1}{1+|z|^{\beta}} \lesssim \frac{t^{\frac{1}{4}}}{1+|x|^{\beta}}.
\end{aligned} \tag{B.6}$$

Combining (B.4), (B.5), and (B.6) leads to

$$\sup_{0<t<1} \sup_{x \in \mathbb{R}^n} (1+|x|^{\beta}) \sum_{l=0}^{2} |\nabla^l S g_l(x,t)| \lesssim t^{\frac{1}{4}} \sum_{l=0}^{2} \|g_l\|_{Y^{\beta}_{l,1}} \lesssim \sum_{l=0}^{2} \|g_l\|_{Y^{\beta}_{l,1}}.$$

Estimate for $\nabla S g_l$ $(l = 0, 1, 2)$. The same computation leads to

$$\sup_{0<t<1} \sup_{x \in \mathbb{R}^n} (1+|x|^{\beta}) \sum_{l=0}^{2} |\nabla^l \nabla S g_l(x,t)| \lesssim \sum_{l=0}^{2} \|g_l\|_{Y^{\beta}_{l,1}}.$$

This is essentially the same as going from (A.8) to (A.12). Hence, we just outline the key computation:

$$
\begin{aligned}
|\nabla\nabla^l Sg_l(x,t)| &= \Big|\int_0^t \int_{\mathbb{R}^n} \nabla\nabla^l b(x-y,t-s)g_l(y,s)dyds\Big| \\
&\le \Big(\int_{Q'_{\sqrt[4]{t}}(0)} + \int_{\mathbb{R}^n\times(0,t)\setminus Q_{\sqrt[4]{t}}(x)}\Big)|\nabla\nabla^l b(x-y,t-s)g_l(y,s)|dyds \\
&:= J_1 + J_2.
\end{aligned}
$$

To estimate J_1, by applying the Hölder inequality, we use kernel estimate (B.1) with $p = \frac{n+6}{n+3+l} < \frac{n+4}{n+l+1}$, $l = 0,1,2$, to derive

$$
\begin{aligned}
J_1 &\le \|\nabla^{l+1}b\|_{L^{\frac{n+6}{n+3+l}}(Q'_{\sqrt[4]{t}}(0))}\|g_l\|_{L^{\frac{n+6}{3-l}}(Q_{\sqrt[4]{t}}(x))} \\
&\le \|\nabla^{l+1}b\|_{L^{\frac{n+6}{n+3+l}}(\mathbb{R}^n\times(0,t))}\|g_l\|_{L^{\frac{n+6}{3-l}}(Q_{\sqrt[4]{t}}(x))} \\
&\lesssim \sqrt[4]{t}^{\frac{6-2l}{n+6}}\|g_l\|_{L^{\frac{n+6}{3-l}}(Q_{\sqrt[4]{t}}(x))} \lesssim \frac{1}{1+|x|^\beta}.
\end{aligned}
$$

For J_2, the computation is similar. The extra factor $t^{-\frac{1}{4}}$ coming from $\nabla^{l+1}b$ is absorbed by the $t^{\frac{1}{4}}$ in (B.4), (B.5), and (B.6).

Estimate for $\nabla^2 Sg_l$ ($l = 0,1,2$). For this, we need to show

$$
\sup_{0<R^4<1} R^{\frac{2}{n+6}}\|\nabla^{2+l}Sg_l(z,t)\|_{L^{n+6}(Q_R(x))} \lesssim \frac{1}{1+|x|^\beta} \quad \text{for } l = 0,1,2. \tag{B.7}
$$

We first compute

$$
\begin{aligned}
&\|\nabla^{2+l}Sg_l\|_{L^{n+6}(Q_R(x))} \\
&\quad = \Big\|\int_{\mathbb{R}^n\times(0,t)\setminus B_{2R(x)}\times(R^4/4,R^4)} \\
&\qquad + \int_{B_{2R(x)}\times(R^4/4,R^4)} \nabla^{2+l}b(z-y,t-s)g_l(y,s)dyds\Big\|_{L^{n+6}(Q_R(x))} \\
&\quad \le \Big\|\int_{\mathbb{R}^n\times(0,t)\setminus B_{2R(x)}\times(R^4/4,R^4)} \nabla^{2+l}b(z-y,t-s)g_l(y,s)dyds\Big\|_{L^{n+6}(Q_R(x))} \\
&\qquad + \Big\|\int_{B_{2R(x)}\times(R^4/4,R^4)} \nabla^{2+l}b(z-y,t-s)g_l(y,s)dyds\Big\|_{L^{n+6}(Q_R(x))} \\
&\quad := K_1 + K_2.
\end{aligned}
$$

For K_1, using the same arguments as those for K_1 in the previous section, we get the pointwise bound

$$
\int_{\mathbb{R}^n\times(0,t)\setminus B_{2R(x)}\times(R^4/4,R^4)} |\nabla^{2+l}b(z-y,t-s)g(y,s)|dyds \lesssim \frac{t^{-\frac{1}{4}}}{1+|z|^\beta},
$$

so that

$$R^{\frac{2}{n+6}} \left\| \frac{t^{-\frac{1}{4}}}{1+|z|^\beta} \right\|_{L^{n+6}(Q_R(x))} \lesssim R^{\frac{2}{n+6}} \frac{t^{-\frac{1}{4}}}{1+|x|^\beta} (R^n R^4)^{\frac{1}{n+6}} \approx \frac{1}{1+|x|^\beta}, \tag{B.8}$$

where we have used the fact that $\frac{R^4}{2} < t < R^4$.

For K_2, we can focus on the L^{n+6} estimate for $\nabla^{2+l} S g_l$ with g_l supported in $Q_{2R}(x)$. First, we recall the Young inequality:

$$\|f * g\|_{L^m(\mathbb{R}^n\times\mathbb{R}^+)} \le C\|f\|_{L^p(\mathbb{R}^n\times\mathbb{R}^+)}\|g\|_{L^q(\mathbb{R}^n\times\mathbb{R}^+)},$$

where $0 \le p, q, m \le \infty$ and $p^{-1} + q^{-1} = 1 + m^{-1}$. Applying the inequality with $m = n+6$, $p = \frac{n+6}{n+4} < \frac{n+4}{n+2}$, $q = \frac{n+6}{3}$, and $m = n+6$, $p = \frac{n+6}{n+5} < \frac{n+4}{n+3}$, $q = \frac{n+6}{2}$, we get

$$\|\nabla^2 S g_0\|_{L^{n+6}(\mathbb{R}^n\times(0,1))} \lesssim \|g_0\|_{L^{\frac{n+6}{3}}(\mathbb{R}^n\times\mathbb{R}^+)} = \|g_0\|_{L^{\frac{n+6}{3}}(B_{2R}(x)\times(R^4/4,R^4))}$$

$$\text{and} \quad \|\nabla^3 S g_1\|_{L^{n+6}(\mathbb{R}^n\times(0,1))} \lesssim \|g_1\|_{L^{\frac{n+6}{2}}(\mathbb{R}^n\times\mathbb{R}^+)} = \|g_1\|_{L^{\frac{n+6}{2}}(B_{2R}(x)\times(R^4/4,R^4))},$$

respectively. Hence,

$$R^{\frac{2}{n+6}}\|\nabla^2 S g_0\|_{L^{n+6}(\mathbb{R}^n\times(0,1))}, \quad R^{\frac{2}{n+6}}\|\nabla\nabla^2 S g_1\|_{L^{n+6}(\mathbb{R}^n\times(0,1))} \lesssim \frac{1}{1+|x|^\beta}. \tag{B.9}$$

For the L^{n+6} norm of $\nabla^4 S g_2$, by singular integral estimate (B.2) with $p = n+6$, we have that

$$\begin{aligned} &R^{\frac{2}{n+6}} \left\| \int_{B_{2R}(x)\times(R^4/4,R^4)} \nabla^4 b(z-y, t-s) g_2(y,s)\,dy\,ds \right\|_{L^{n+6}(Q_R(x))} \\ &\quad \lesssim R^{\frac{2}{n+6}} \|g_2\|_{L^{n+6}(B_{2R}(x)\times(R^4/4,R^4))} \lesssim \frac{1}{1+|x|^\beta}. \end{aligned} \tag{B.10}$$

Combining (B.8), (B.9), and (B.10) gives (B.7), thus completing the proof. ∎

Acknowledgements. The authors thank Changyou Wang for useful discussions.

References

[1] S. Allen and J. Cahn, A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. *Acta Metallurgica* **27** (2014), no. 6, 1085–1095.

[2] T. Asai and Y. Giga, On self-similar solutions to the surface diffusion flow equations with contact angle boundary conditions. *Interfaces Free Bound.* **16** (2014), no. 4, 539–573 Zbl 1332.35061 MR 3292121

[3] P. Baras, J. Duchon, and R. Robert, Évolution d'une interface par diffusion de surface. *Comm. Partial Differential Equations* **9** (1984), no. 4, 313–335 Zbl 0542.35078 MR 740093

[4] J. Bricmont, A. Kupiainen, and G. Lin, Renormalization group and asymptotics of solutions of nonlinear parabolic equations. *Comm. Pure Appl. Math.* **47** (1994), no. 6, 893–922 Zbl 0806.35067 MR 1280993

[5] J. W. Cahn and J. E. Taylor, Overview no. 113: surface motion by surface diffusion. *Acta metallurgica et materialia* **42** (1994), no. 4, 1045–1063.

[6] T. Cazenave, F. Dickstein, M. Escobedo, and F. B. Weissler, Self-similar solutions of a nonlinear heat equation. *J. Math. Sci. Univ. Tokyo* **8** (2001), no. 3, 501–540 Zbl 0996.35031 MR 1855457

[7] A. Cesaroni, H. Kröner, and M. Novaga, Anisotropic mean curvature flow of Lipschitz graphs and convergence to self-similar solutions. *ESAIM, Control Optim. Calc. Var.* **27** (2021), paper no. 97 Zbl 1483.53107 MR 4323027

[8] R. Chill, E. Fašangová, and R. Schätzle, Willmore blowups are never compact. *Duke Math. J.* **147** (2009), no. 2, 345–376 Zbl 1175.35079 MR 2495079

[9] J. Clutterbuck and O. C. Schnürer, Stability of mean convex cones under mean curvature flow. *Math. Z.* **267** (2011), no. 3–4, 535–547 Zbl 1216.53058 MR 2776047

[10] T. H. Colding and W. P. II Minicozzi, Generic mean curvature flow I: generic singularities. *Ann. of Math. (2)* **175** (2012), no. 2, 755–833 Zbl 1239.53084 MR 2993752

[11] K. Ecker and G. Huisken, Mean curvature evolution of entire graphs. *Ann. of Math. (2)* **130** (1989), no. 3, 453–471 Zbl 0696.53036 MR 1025164

[12] K. Ecker and G. Huisken, Interior estimates for hypersurfaces moving by mean curvature. *Invent. Math.* **105** (1991), no. 3, 547–569 Zbl 0707.53008 MR 1117150

[13] J. Eells, Jr. and J. H. Sampson, Harmonic mappings of Riemannian manifolds. *Amer. J. Math.* **86** (1964), 109–160 Zbl 0122.40102 MR 164306

[14] C. M. Elliott and S. Maier-Paape, Losing a graph with surface diffusion. *Hokkaido Math. J.* **30** (2001), no. 2, 297–305 Zbl 0993.35042 MR 1844821

[15] J. Escher, U. F. Mayer, and G. Simonett, The surface diffusion flow for immersed hypersurfaces. *SIAM J. Math. Anal.* **29** (1998), no. 6, 1419–1433 Zbl 0912.35161 MR 1638074

[16] J. Escher and G. Simonett, The volume preserving mean curvature flow near spheres. *Proc. Amer. Math. Soc.* **126** (1998), no. 9, 2789–2796 Zbl 0909.53043 MR 1485470

[17] A. Friedman, *Partial differential equations of parabolic type*. Courier Dover Publications, NY, 2008

[18] M.-H. Giga, Y. Giga, and J. Saal, Self-similar solutions for various equations. In *Nonlinear partial differential equations*. Prog. Nonlinear Differ. Equ. Appl. 79, Birkhäuser, Boston, MA, 2010 Zbl 1215.35001 MR 2656972

[19] A. Gmira and L. Véron, Large time behaviour of the solutions of a semilinear parabolic equation in $\mathbf{R}^N$. *J. Differential Equations* **53** (1984), no. 2, 258–276 Zbl 0529.35041 MR 748242

[20] W. Helfrich, Elastic properties of lipid bilayers: Theory and possible experiments. *Z. Naturforsch., C* **28** (1973), no. 11, 693–703.

[21] T. Huang and C. Wang, Well-posedness for the heat flow of polyharmonic maps with rough initial data. *Adv. Calc. Var.* **4** (2011), no. 2, 175–193 Zbl 1227.35152 MR 2793836

[22] G. Huisken, Asymptotic behavior for singularities of the mean curvature flow. *J. Differential Geom.* **31** (1990), no. 1, 285–299 Zbl 0694.53005 MR 1030675

[23] O. Kavian, Remarks on the large time behaviour of a nonlinear diffusion equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **4** (1987), no. 5, 423–452 Zbl 0653.35036 MR 921547

[24] H. Koch and T. Lamm, Geometric flows with rough initial data. *Asian J. Math.* **16** (2012), no. 2, 209–235 Zbl 1252.35159 MR 2916362

[25] H. Koch and D. Tataru, Well-posedness for the Navier-Stokes equations. *Adv. Math.* **157** (2001), no. 1, 22–35 Zbl 0972.35084 MR 1808843

[26] E. Kuwert and R. Schätzle, The Willmore functional. In *Topics in modern regularity theory*, pp. 1–115, CRM Ser. (N.S.) 13, Edizioni della Normale, Pisa, 2012 Zbl 1322.53002 MR 2882586
[27] F. Lubbe, Mean curvature flow of contractions between Euclidean spaces. *Calc. Var. Partial Differential Equations* **55** (2016), no. 4, paper no. 104 Zbl 1355.53060 MR 3530212
[28] F. Lubbe, Evolution of area-decreasing maps between two-dimensional Euclidean spaces. *J. Geom. Anal.* **28** (2018), no. 4, 3928–3949 Zbl 1407.53067 MR 3881996
[29] Y. Meyer, Large-time behavior and self-similar solutions of some semilinear diffusion equations. In *Harmonic analysis and partial differential equations (Chicago, IL, 1996)*, pp. 241–261, Chicago Lectures in Math., University of Chicago Press, Chicago, IL, 1999 Zbl 0945.35014 MR 1743866
[30] W. Mullins, Theory of thermal grooving. *J. Appl. Phys.* **28** (1957), no. 3, 333–339.
[31] A. Savas-Halilaj, Graphical mean curvature flow. In *Nonlinear analysis, differential equations, and applications*, pp. 493–577, Springer Optim. Appl. 173, Springer, Cham, 2021 Zbl 1476.53015 MR 4367387
[32] A. Savas-Halilaj and K. Smoczyk, Homotopy of area decreasing maps by mean curvature flow. *Adv. Math.* **255** (2014), 455–473 Zbl 1288.53064 MR 3167489
[33] A. Savas-Halilaj and K. Smoczyk, Evolution of contractions by mean curvature flow. *Math. Ann.* **361** (2015), no. 3–4, 725–740 Zbl 1398.53078 MR 3319546
[34] U. Seifert, Configurations of fluid membranes and vesicles. *Advances in Physics* **46** (1997), no. 1, 13–137.
[35] L. Simon, Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems. *Ann. of Math. (2)* **118** (1983), no. 3, 525–571 Zbl 0549.35071 MR 727703
[36] G. Simonett, The Willmore flow near spheres. *Differ. Integral Equ.* **14** (2001), no. 8, 1005–1014 Zbl 1161.35429 MR 1827100
[37] K. Smoczyk, Mean curvature flow in higher codimension: introduction and survey. In *Global differential geometry*, pp. 231–274, Springer Proc. Math. 17, Springer, Heidelberg, 2012 Zbl 1247.53004 MR 3289845
[38] K. Smoczyk, M.-P. Tsui, and M.-T. Wang, Curvature decay estimates of graphical mean curvature flow in higher codimensions. *Trans. Amer. Math. Soc.* **368** (2016), no. 11, 7763–7775 Zbl 1347.53054 MR 3546783
[39] N. Stavrou, Selfsimilar solutions to the mean curvature flow. *J. Reine Angew. Math.* **499** (1998), 189–198 Zbl 0895.53039 MR 1631112
[40] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*. Princeton Math. Ser. 43, Princeton University Press, Princeton, NJ, 1993 Zbl 0821.42001 MR 1232192
[41] J. E. Taylor, J. W. Cahn, and C. A. Handwerker, Overview No. 98: I–Geometric models of crystal growth. *Acta Metallurgica et Materialia* **40** (1992), no. 7, 1443–1474.
[42] M.-P. Tsui and M.-T. Wang, Mean curvature flows and isotopy of maps between spheres. *Comm. Pure Appl. Math.* **57** (2004), no. 8, 1110–1126 Zbl 1067.53056 MR 2053760
[43] M.-T. Wang, Long-time existence and convergence of graphic mean curvature flow in arbitrary codimension. *Invent. Math.* **148** (2002), no. 3, 525–543 Zbl 1039.53072 MR 1908059
[44] M.-T. Wang, The Dirichlet problem for the minimal surface system in arbitrary dimensions and codimensions. *Comm. Pure Appl. Math.* **57** (2004), no. 2, 267–281 Zbl 1071.35050 MR 2012810

[45] M.-T. Wang, Lectures on mean curvature flows in higher codimensions. In *Handbook of geometric analysis. No. 1*, pp. 525–543, Adv. Lect. Math. (ALM) 7, International Press, Somerville, MA, 2008 Zbl 1167.53058 MR 2483374

Received 30 August 2021; revised 22 June 2022.

Hengrong Du
Department of Mathematics, Vanderbilt University, 1326 Stevenson Center, Station B 407807, Nashville, TN 37240, United States; hengrong.du@vanderbilt.edu

Nung Kwan Yip
Department of Mathematics, Purdue University, 150 N. University Street, West Lafayette, IN 47906, United States; yip@math.purdue.edu